MODELING DIVERSIFIED EQUITY INDICES

By
Renata Rendek

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MATHEMATICAL SCIENCES AT THE UNIVERSITY OF TECHNOLOGY, SYDNEY

SYDNEY, AUSTRALIA, APR 2013

UNIVERSITY OF TECHNOLOGY, SYDNEY
CERTIFICATE OF AUTHORSHIP/ORIGINALITY

Date: **Apr 2013**

Author: **Renata Rendek**

Title: **Modeling Diversified Equity Indices**

Department: **School of Mathematical Sciences**

Degree: **PhD**

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Author

ACKNOWLEDGMENTS

There are many people I would like to acknowledge for their support, both technical and emotional, during my PhD studies. I list but a few here.

I would like to thank Professor Eckhard Platen, my supervisor, for providing the opportunity to take up a PhD and for continuing advice and enthusiasm throughout the research.

My thanks go to the staff and students of both the School of Mathematical Sciences and the Finance Discipline Group for their support and for friendly atmosphere. I wish to express my gratitude to my supervisor Professor Eckhard Platen for the financial support he arranged during my study. Further, I would like to thank Alex Novikov, Truc Lee, Wolfgang Breymann, Hardy Hulley and Katja Ignatieva for their help throughout my PhD studies.

Moreover, since I have used the Linux operating system and GNU R while writing my PhD thesis, I would like to thank the entire Linux Community and the GNU project for their help and support.

Finally to my friends, my family and to Lubos, thank you all for being there for me.

Contents

List of Tables

List of Figures

Abstract

The objective of this thesis is to study and model the dynamics of aggregate wealth, that is, the dynamics of the market capitalization weighted world stock index in different currency denominations. In order for the considered model to be valid over long time periods, it turns out that it needs to be formulated in a general financial modeling framework, the benchmark approach. In order to visualize and test the proposed aggregate wealth dynamics, exact and almost exact simulation techniques for multi-dimensional stochastic processes have been developed. Moreover, the model specification is preceded by a detailed study of the distribution of log-returns of world stock indices in different currency denominations. Various types of world equity indices are constructed and systematically studied, in particular, equi-weighted indices. When the number of constituents is increasing and the given investment universe is well securitized the Naive Diversification Theorem states that a sequence of equi-weighted indices approximates the growth optimal portfolio, which is also the numéraire portfolio.

Finally, by conjecturing for the normalized world stock index the dynamics of a time transformed square root process, and by establishing a list of stylized empirical facts, a two-component index model has been proposed. This model is very parsimonious and driven only by the non-diversifiable risk of the market. Via almost exact simulation this model is shown to reflect well all listed empirical stylized facts and is difficult to falsify.

Introduction

0.1 Brief Survey of Results

This section focuses on the main results presented in this thesis. The subject of the thesis is the modeling of diversified equity indices. The presentation of the conducted research is divided into four chapters, each linked to some published article or articles. Chapter 1 surveys and develops exact simulation methods for solutions of several classes of multi-dimensional stochastic differential equations, aiming to avoid numerical error propagation that may arise otherwise. Chapter 2 deals with an empirical study of world stock indices in different currency denominations. The objective of Chapter 3 is to approximate the numéraire portfolio by naive diversification. The final chapter models the typical dynamics of equity indices by conjecturing the affine nature of aggregate wealth dynamics and establishing respective empirical evidence.

Some of the new results of this thesis have already appeared in six refereed publications. There is also one extended, recent working paper from which at least one publication will follow and another working paper that is expected to generate another article. Below, most of these papers are put into the context of the thesis.

A main result of the first part of the thesis in Chapter 1 is the development and application of exact simulation methods for multi-dimensional stochastic differential equations, see Platen & Rendek (2009, 2010). The main classes of matrix stochastic processes that we consider are: the matrix Ornstein-Uhlenbeck processes, the Wishart processes and the matrix affine processes. Further examples for the simulation of matrix stochastic processes, e.g. matrix Lévy processes, are considered in the corresponding publication Platen & Rendek (2009). As an application of the proposed exact simulation technique we consider the Wonham filter problem and the resulting Zakai equation for hidden Markov chain filters. This application

shows that the issue of numerical stability can be circumvented in filtering when it is possible to simulate exact or almost-exact solutions of the Zakai equation at the observation points. This study is extracted from Platen & Rendek (2010), which illustrates the usefulness of almost-exact solutions from the practical point of view.

Chapter 2 presents other important result of this thesis published in Platen & Rendek (2008) and Ignatieva, Platen & Rendek (2011). These concern the identification of the uni- and multi-variate distributions of log-returns of world stock indices in different currency denominations. The estimated univariate distribution of log-returns of such indices turn out to be very close to the Student-t distribution with approximately four degrees of freedom. This result emerges when searching within the large family of generalized hyperbolic distributions and analyzing differently constructed world stock indices, see Platen & Rendek (2008). Since this result is established at a very high level of significance, it can be regarded as a stylized empirical fact. Moreover, a copula approach is applied to analyze the log-returns of one of the constructed indices, the EWI104s, in different currency denominations, see Ignatieva, Platen & Rendek (2011). Based on the findings corresponding to distributional properties of the marginals, the paper models the dependency of log-returns of currency denominations of the EWI104s using time-varying copulas. It has been shown that when compared to the Gaussian copula, the Student-t copula captures better the dependence structure, observed in the return data.

The Naive Diversification Theorem (NDT) is the major theoretical result of Chapter 3 and has been published in Platen & Rendek (2012b). It states that the equi-weighted index (EWI) approximates the numéraire portfolio (NP) of a given set of stocks when the number of constituents is large and the given investment universe is well securitized. The latter essentially means that the risk factors driving the underlying risky securities are sufficiently different. The practical contribution of this chapter is the construction of the EWI, which is the best performing index in terms of growth among all indices included in this thesis. Additionally, this part of the thesis raises the practically important question, whether a significantly better performance can be detected for the EWI when compared to the market capitalization weighted index (MCI) for market models where the constituents are strict supermartingales? A corresponding simulation study demonstrates that the

convergence of the EWI towards the NP appears to be remarkably robust, see also Platen & Rendek (2012c). For the interesting case of the minimal market model, where the primary security accounts when expressed in units of the NP are strict supermartingales, the EWI outperforms in the long run the MCI significantly. This can be explained by the fact that the benchmarked MCI, as the sum of strict supermartingales, yields a strict supermartingale, whereas, the benchmarked NP is simply the constant one.

The final part of this thesis considers in Chapter 4 the modeling of diversified equity indices and follows Platen & Rendek (2012a). It is conjectured that typical dynamics of aggregate wealth can be recovered from the limiting dynamics of the sum of the values of many independent economic activities and "projects" over small time periods. The main contribution of this chapter is, therefore, the identification of a parsimonious index model involving the power of a time transformed affine diffusion, a square root process. A very realistic, parsimonious one-factor, two-component model emerges by modeling via the transformed time the human behavior, which exaggerates the reactions in volatility to ups and downs of the index. These exaggerations are modeled through the transformed time via another (fast moving) square root process, which is driven by the same non-diversifiable uncertainty that drives already the fluctuations of the index. Due to its almost exact simulation, we can verify a number of stylized empirical facts pertaining to diversified equity indices. The proposed model is parsimonious with only six parameters. To demonstrate its applicability it has been applied also for the S&P500 and its volatility index VIX.

0.2 Research Motivation and Strategy

The main motivation for the need of better modeling of diversified equity indices are the discrepancies between the observed stylized empirical facts in financial data and typically available models, see e.g. Platen & Rendek (2008), Ignatieva, Platen & Rendek (2011) and Platen & Rendek (2012a). The thesis notices and explains the fact that currency denominations of a world stock index (the total wealth in the global equity market) have more clearly identifiable empirical properties than, say, an exchange rate. For instance, it is well observed that the average volatility

of a currency denomination of a world stock index is significantly smaller than the average volatility of an exchange rate. Clear statistical properties of log-returns of a world stock index in different currency denominations, which we will identify, suggest that such financial quantities are better objects for modeling than e.g. exchange rates or stock prices. Moreover, a world stock index, denominated in a given currency, is primarily driven by the nondiversifiable uncertainty of the market with respect to that currency as denominator. This fact makes it possible to potentially identify a parsimonious model, driven only by one source of uncertainty, as will be demonstrated in this thesis.

The main reason for the proposed modeling of equity indices is the, in Platen & Rendek (2012a) conjectured, theoretical affine nature of aggregate wealth dynamics. The apparent deadlock in the modeling of financial quantities seemed to be the absence of a theoretical argument that explains the potential nature of the dynamics of aggregate wealth. We conjecture the typical dynamics of aggregate wealth from studying the limiting dynamics of the sum of the values of many independent economic activities and "projects" over small time periods. In a first approximation, the variance of the increments of aggregate wealth turns theoretically out to be proportional to the number of these activities and "projects". This means that the variance is proportional to the aggregate wealth itself. Therefore, the normalized aggregate wealth itself follows approximately a time transformed square root process. With respect to transformed time the inverse of this square root process is then the resulting squared volatility. A realistic, parsimonious, one-factor, two-component model emerges by modeling via the transformed time human behavior, which exaggerates the reactions in volatility to ups and downs of the index. These exaggerations are modeled through the derivative of the transformed time via another (fast moving) square root process. This process is driven by the same nondiversifiable uncertainty of the market since there seems to be no justifiable reason to include any other source of uncertainty into the dynamics of the world stock index.

In order to be able to verify the stylized empirical facts and understand some of the effects of the proposed estimation techniques for equity indices in different currency denominations, see Platen & Rendek (2012a), the need became overwhelming for reliable exact or almost exact simulation techniques. Many simulation tech-

niques have been developed over the years, in particular those that use discrete time approximations. However, some stochastic differential equations (SDEs) can be problematic in terms of simulation. Their extreme events may not be well captured by discrete time approximations and also their long term behavior, including their transition density may be badly approximated. Therefore, it is necessary to understand and avoid the problems that may arise during the simulation of solutions of such SDEs, see Platen & Rendek (2009). During the research it became of crucial importance to identify extremely accurate simulation methods for various nonlinear types of SDEs and also for multi-dimensional SDEs. This thesis emphasizes the fact that the problem of non-Lipschitz continuous coefficients can be circumvented for some important SDEs, where we can simulate exact or almost exact solutions, see Platen & Rendek (2010). Fortunately, the class of affine diffusions that the conjectured wealth dynamics describes allows almost exact simulations.

Finally, the Naive Diversification Theorem, see Platen & Rendek (2012b), provides a way of interpreting a well diversified equity index as proxy for the numéraire portfolio of the given investment universe. This motivates us to construct and analyze various diversified world stock indices within this thesis, see Platen and Rendek (2008, 2012b, 2012c). By interpreting the analyzed indices as numéraire portfolio and fitting the proposed parsimonious model it followed clearly that the classical no-arbitrage financial modeling approach is too narrow for capturing reality, in particular, the long term dynamics of indices. The benchmark approach, see Platen & Heath (2010) and Platen (2011) generalizes classical no-arbitrage pricing toward a much richer modeling world and pricing under the real world probability measure with the numéraire portfolio as numéraire. The central building block of this pricing environment is the benchmark, the numéraire portfolio, which is also the growth optimal portfolio, see Long (1990) and Kelly (1956). This portfolio is employed as the fundamental unit of value in the analysis. It is in reality a well diversified index with clear statistical properties that can be parsimoniously modeled.

The numéraire portfolio is the strictly positive portfolio which, when used as benchmark, turns all benchmarked nonnegative portfolios into supermartingales. The pricing of derivatives under the benchmark approach applies the real world pricing formula, which yields the minimal possible price, that is, the benchmarked

derivative price is observed as the real world conditional expectation of the corresponding benchmarked payoff. Consequently, real world pricing leads to the minimal possible price and, thus, often to lower prices then suggested by classical pricing arguments. Moreover, since the existence of the equivalent risk neutral probability measure is not required, this financial framework allows for a wider and, thus, more realistic range of models in contrast to the classical approaches. The proposed model demonstrates that this modeling freedom is essential to obtain index models that remain realistic over long periods of time and fit the observed stylized empirical facts.

0.3 Related Literature

This section provides links to literature related to the main results of this thesis, which are in the areas: simulation of stochastic differential equations, analysis of the distributions of log-returns of equity indices, approximation of the numéraire portfolio and modeling the dynamics of diversified equity indices.

Simulation of Stochastic Differential Equations

Accurate scenario simulation of solutions of stochastic differential equations (SDEs) is widely applicable in stochastic analysis and its areas of application, in particular, in finance and filtering, see Kloeden & Platen (1999) and Kallianpur (1980). Monographs in the direction of stochastic numerical methods include Glasserman (2004), Jäckel (2002), Kloeden & Platen (1999), Kloeden, Platen & Schurz (2003), Milstein (1995) and Platen & Bruti-Liberati (2010).

There is a wide range of literature which deals with the problem of numerical stability. In particular, implicit and predictor-corrector methods are used to control the propagation of numerical errors. We refer here to papers e.g. by Alcock & Burrage (2006), Bruti-Liberati & Platen (2008), Hernandez & Spigler (1993), Higham (2000), Higham, Mao & Yuan (2007), Klauder & Petersen (1985), Kloeden & Platen (1992), Milstein (1988), Milstein, Platen & Schurz (1998), Platen & Shi (2008), Saito & Mitsui (1993) and Talay (1982). The issue of numerical stability can be circumvented when it is possible to simulate exact or almost exact solutions.

Beyond the Wiener process and its direct transformations, including the geomet-

ric Brownian motion and the Ornstein-Uhlenbeck process, the family of square root and squared Bessel processes is probably the most frequently used diffusion model in applications. In general, it is a challenging task to obtain efficiently a reasonably accurate trajectory of a square root process using simulation, as is documented in an increasing literature on this topic. This literature includes the use of the balanced implicit method introduced in Milstein, Platen & Schurz (1998), the adaptive Milstein scheme of Kahl (2004), the balanced Milstein methods of Alcock & Burrage (2006) and Kahl & Schurz (2005). Additionally, various other methods have been designed to approximate the square root process. Here we refer to Alfonsi (2005), Andersen (2008), Berkaoui, Bossy & Diop (2005), Bossy & Diop (2004), Broadie & Kaya (2006), Deelstra & Delbaen (1998), Diop (2003), Lord, Koekkoek & van Dijk (2006) and Smith (2007).

In finance and beyond finance the systematic construction and investigation of filters for hidden Markov chains has became an important area of research, which goes back to Fujisaki, Kallianpur & Kunita (1972), Wonham (1965) and Zakai (1969). The stochastic filtering theory is the subject of a seminal monograph by Kallianpur (1980). Later the question of finding discrete-time approximations for optimal filters was considered by Clark & Cameron (1980), Kloeden, Platen & Schurz (1993), Newton (1986), and Newton (1991). Moreover, the application of the balanced implicit method to SDEs in filtering has been considered in Fischer & Platen (1999). The paper Platen & Rendek (2010) contributes to the line of research on filtering, whereas the paper Platen & Rendek (2009) provides very accurate long term simulation tools in finance that became crucial for the study of the proposed model and can be interpreted as a contribution to the area of simulation of SDEs even beyond finance.

Distribution of Log-Returns of Equity Indices

Some of the most promising financial instruments, for which one may identify clearly a particular type of log-return density, turn out to be diversified stock indices. This observation is made in the thesis. Studies on log-returns for indices that do not emphasize this view but perform inference in this direction include the two papers by Markowitz & Usmen (1996a, 1996b), analyzing S&P500 log-returns in a Bayesian framework. These authors considered the rich family of Pearson distributions and

identified the Student-t distribution with about 4.5 degrees of freedom as the best fit to daily log-return data of the S&P500. Independently, Hurst & Platen (1997) reached a similar conclusion by studying daily log-returns of the S&P500 and other regional stock market indices using a maximum-likelihood approach. Their research was focused on a large family of normal-variance mixture distributions, see Clark (1973), which included the log-return distributions generated by several important models proposed in the literature. These distributions included among others the normal, see Samuelson (1957) and Black & Scholes (1973); the alpha-stable, see Mandelbrot (1963); the Student-t, see Praetz (1972) and Blattberg & Gonedes (1974); the normal-inverse Gaussian, see Barndorff-Nielsen (1995); the hyperbolic, see Eberlein & Keller (1995) and Küchler et al.(1999); the variance gamma, see Madan & Seneta (1990); and the symmetric generalized hyperbolic distribution, see Barndorff-Nielsen (1978) and McNeil, Frey & Embrechts (2005). In Hurst & Platen (1997) the Student-t distribution with 3.0-4.5 degrees of freedom was determined as the best fit to daily, regional stock market index log-returns. This complemented and generalized Markowitz's and Usmen's findings by the use of an alternative statistical methodology and a wider range of stock market indices. Fergusson & Platen (2006) employed a maximum likelihood ratio test, see Rao (1973), in a similar class of symmetric generalized hyperbolic distributions as mentioned above. They studied the log-return distribution of a world stock index, whose constituent weights were determined by market capitalization, and considered different currency denominations of such an index. These authors concluded, at a high level of significance, that the log-returns of their index exhibited a Student-t behavior with approximately four degrees of freedom.

For multivariate log-returns there exists an advanced statistical methodology for identifying particular generalized hyperbolic distributions as described in McNeil, Frey & Embrechts (2005). These authors showed in the application of their results to indices, exchange rates and stocks that Student-t type log-return distributions are often likely to fit the data. Usually, they did not quantify any level of significance and did not focus on diversified indices.

Breymann, Dias & Embrechts (2003), Dias & Embrechts (2004) and Dias & Embrechts (2008) analyzed the dependence structure within two-dimensional high

frequency data of exchange rate returns. They have shown that the data can be fitted best using the Student-t copula, which can be thought of as representing the dependence structure implicit in a multivariate Student-t distribution, see Embrechts, McNeil & Straumann (2001). Along the above lines of research, but focussed on indices, the papers Platen & Rendek (2008), Ignatieva, Platen & Rendek (2011) and Platen & Rendek (2012a) identify with high significance the Student-t distribution and the Student-t copula for log-returns of diversified indices.

Approximation of the Numéraire Portfolio

The ground breaking work of Markowitz (1952) heralded considerable advances in asset allocation by deriving a methodology for allocating wealth across risky investments when investors base their strategies on the means, variances and covariances of asset returns. The resulting Markowitz mean-variance approach to investing has been generalized in many ways; see Campbell & Viceira (2002) for a survey.

A recent paper by DeMiguel, Garlappi & Uppal (2009) evaluated in detail the out-of-sample performance of sample based mean-variance asset allocation strategies, relative to the naive equal value weighted counterpart. The outcome being that of the 14 estimation procedures evaluated across seven empirical datasets, no strategy consistently outperformed the naive equal value weighted strategy in terms of Sharpe ratio and other common performance criteria. The authors concluded that the theoretical gain from classical sample based mean-variance asset allocation is more than offset by estimation errors. They demonstrated that, for a portfolio of 50 assets, the estimation window needed to outperform the naive equal value weighted portfolio is about 500 years.

A basic assumption of the final part of the current thesis is concerned with the existence of the numéraire portfolio (NP), which has been studied in many papers including Long (1990), Bajeux-Besnainou & Portait (1997), Becherer (2001), Platen (2002), Bühlmann & Platen (2003), Platen (2006), Platen & Heath (2010), Karatzas & Kardaras (2007) and Kardaras & Platen (2008). The NP also appears in the literature as the Kelly portfolio or growth optimal portfolio, see Kelly (1956), which maximizes the expected logarithmic utility from terminal wealth. There is an extensive literature on the growth optimal or Kelly portfolio, which includes Markowitz (1976), Latané (1959), Breiman (1961), Hakansson (1971), Thorp (1972),

Merton (1973), Rubinstein (1976), Cover (1991), Luenberger (1998), Ziemba & Mulvey (1998), Browne (1999), Stutzer (2000), Platen & Heath (2010) and MacLean, Thorp & Ziemba (2011). In the long run the NP outperforms pathwise any other nonnegative portfolio; see e.g. Platen (2004). This is a model independent and fascinating property, which makes the NP an ideal candidate for long term asset allocation. Moreover, as shown in Merton (1973), the Markowitz-efficient frontier for a market, consisting of portfolios of stocks plus a risk-free security, can be generated by considering portfolio combinations of the NP and the risk-free security. According to the Intertemportal Capital Asset Pricing Model (ICAPM), see Merton (1973), the Sharpe ratios of all portfolios on the efficient frontier are theoretically the same, including that of the market portfolio. Since the NP plays such a key role in asset management and financial modeling the paper Platen & Rendek (2012b) provides a convenient, model independent way of approximating the NP via naive diversification.

Modeling of Diversified Indices

The standard continuous market model for an equity index has been the Black-Scholes model, see Black & Scholes (1973) and Merton (1973), which employs the exponential of a time transformed Brownian motion to describe the index dynamics, resulting in its standard version in constant volatility and Gaussian log-returns. Its historical popularity is due to its tractability and simplicity. Several shortcomings of this model became apparent since its introduction. Most striking is the observation that, in reality, the volatility of an equity index is stochastic and its return distribution is leptokurtic.

Streams of literature aiming for improvements on the standard market model by modeling volatility as a stochastic process, include the broad literature on autoregressive conditional heteroscedastic (ARCH) models and its generalizations, initiated by Engle (1982). The reader can find a systematic introduction into the extremely rich literature on continuous time stochastic volatility models, for instance, in Cont (2010). An important class of volatility models has been studied in the literature on local volatility function models; see for instance Dupire (1992) and Derman & Kani (1994b). This popular type of continuous stochastic volatility model generalizes the constant elasticity of variance (CEV) model, which goes back

to Cox (1975) and Cox & Ross (1976). The proposed model in this thesis has some similarity to CEV type models, and also to those that employ some random market activity time in the sense of subordination; see Clark (1973) and Bochner (1955). More recently, Lévy processes and jump diffusion processes have been used in asset price and index modeling, see e.g. Madan & Seneta (1990), Eberlein & Keller (1995), Barndorff-Nielsen & Shephard (2001) and Kou (2002). However, by modeling the wold stock index, driven by the continuous non-diversifiable uncertainty of the market, the thesis demonstrates that one can avoid the modeling of jumps in a first approximation of the long term dynamics of the world stock index in currency denomination. The working paper Platen & Rendek (2012a) proposes a model for a diversified stock index, which involves a time transformed square root process and a random market activity that is also determined by a square root process, which is fast moving. The proposed parsimonious model fits well a list of major empirical stylized facts. It seems to be difficult to falsify the model empirically, whereas many other popular models are clearly not consistent with observed data. The model dynamics are described under the real world probability measure, and its fit leads outside the classical no-arbitrage pricing theory. Still, the model is covered under the benchmark approach, see Platen & Heath (2010). In the very recent working paper Du, Platen & Rendek (2012) the modeling of a commodity, oil, instead of a currency has been performed. Similar results as for currencies emerge. However, the random market activity moves opposite to the one employed for the modeling of a diversified index in currency denomination.

Chapter 1

Exact Scenario Simulation for Selected Multi-dimensional Stochastic Processes

Accurate scenario simulation methods for solutions of multi-dimensional stochastic differential equations find applications in stochastic analysis, the statistics of stochastic processes and many other areas, in particular, in finance. They have been playing a crucial role as standard models in various areas and dominate often the communication and thinking in a particular field of application. Within this thesis we will repeatedly employ exact or almost exact simulation methods to demonstrate reliability and accuracy of our findings. In Platen & Rendek (2009) such simulation methods have been derived. We will survey in this chapter those methods that will become useful tools in this thesis.

Various discrete time simulation methods have been developed over the years. However, the simulation of solutions of some stochastic differential equations can be problematic due to systematic errors and numerical instabilities. In particular, one encounters often problems when simulating trajectories over long time periods or when analyzing rare events or tails of transition densities, as we will do in later chapters. Therefore, it is valuable to identify multi-dimensional stochastic differential equations with solutions that can be simulated exactly or almost exactly. This avoids several of the theoretical and practical problems encountered by those simulation methods that use discrete time approximations, see Kloeden & Platen (1999). This chapter provides a brief survey of known and new methods for the exact simulation of paths of some multi-dimensional solutions of stochastic differential equations including Ornstein-Uhlenbeck, square root, squared Bessel and Wishart type pro-

cesses. Moreover, an application of exact simulation in filtering is demonstrated in the last section, illustrating the wide applicability of the methods. This chapter is intended as a preliminary chapter and introduction to multi-dimensional solutions of stochastic differential equations, where exact or almost exact simulations are possible. It is based on the two papers Platen & Rendek (2009) and Platen & Rendek (2010).

1.1 Multi-dimensional Itô Formula

We will start our discussion by recalling some basic facts from stochastic analysis, including the multi-dimensional Itô formula. Given some family of explicitly solvable multi-dimensional stochastic differential equations (SDEs), one can obtain by application of the multi-dimensional Itô formula another family of explicitly solvable multi-dimensional SDEs. This results in a wide range of multi-dimensional SDEs that can be simulated exactly. In this section we illustrate this property by simulating a 2-dimensional Black-Scholes model. The Black-Scholes model, see Black & Scholes (1973) is the standard asset price model in finance.

Vector of Independent Wiener Processes

Let us consider an m-dimensional Wiener process

$$\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \ldots, W_t^m)^\top, t \in [0, \infty)\}. \tag{1.1}$$

Here $\boldsymbol{a}^T$ denotes the transpose of a vector. We assume that the components of this vector stochastic process $\boldsymbol{W}$, are independent. The increments of the Wiener processes $W_t^j - W_s^j$ for $j \in \{1, 2, \ldots, m\}$, $t \geq 0$ and $0 \leq s \leq t$ are then independent Gaussian random variables with mean zero and variance equal to $t - s$. Therefore, one obtains the vector increments of the standard m-dimensional Wiener process $\boldsymbol{W}_t - \boldsymbol{W}_s \sim \mathcal{N}_d(\boldsymbol{0}, (t-s)\boldsymbol{I})$ as a vector of zero mean independent Gaussian random variables with variance $t - s$. Here $\mathcal{N}_d(\cdot, \cdot)$ refers to the corresponding multivariate Gaussian distribution with dependence on mean vector and covariance matrix. $\boldsymbol{I}$ denotes here the identity matrix. We obtain for the values of the trajectory of the standard m-dimensional Wiener process at the discretization times $t_i = i\Delta, i \in$

$\{0, 1, 2, \dots\}$, with $\Delta > 0$ the following iterative formula

$$\begin{aligned} \boldsymbol{W}_0 &= \mathbf{0} \\ \boldsymbol{W}_{t_{i+1}} &= \boldsymbol{W}_{t_i} + \sqrt{\Delta}\boldsymbol{N}_{i+1}, \end{aligned} \tag{1.2}$$

where $\boldsymbol{N}_{i+1} \sim \mathcal{N}_m(\mathbf{0}, \boldsymbol{I})$ is an independent standard Gaussian random vector and $\mathbf{0}$ denotes the corresponding vector of zeros.

Multi-dimensional Itô Formula

Let be given the m-dimensional Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \dots, W_t^m)^\top, t \in [0, \infty)\}$, a d-dimensional drift coefficient vector function $\boldsymbol{a} : [0, T] \times \Re^d \to \Re^d$ and a $d \times m$-matrix diffusion coefficient function $\boldsymbol{b} : [0, T] \times \Re^d \to \Re^{d \times m}$. In this framework we assume that we have already a family of explicitly solvable d-dimensional SDEs given as

$$d\boldsymbol{X}_t = \boldsymbol{a}(t, \boldsymbol{X}_t)dt + \boldsymbol{b}(t, \boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{1.3}$$

for $t \in [0, \infty)$, $\boldsymbol{X}_0 \in \Re^d$. This means that the kth component of (1.3) equals

$$dX_t^k = a^k(t, \boldsymbol{X}_t)dt + \sum_{j=1}^{m} b^{k,j}(t, \boldsymbol{X}_t)dW_t^j. \tag{1.4}$$

For a sufficiently smooth vector function $\boldsymbol{U} : [0, T] \times \Re^d \to \Re^k$ of the solution $\boldsymbol{X}_t$ of (1.3) we obtain a k-dimensional process

$$\boldsymbol{Y}_t = \boldsymbol{U}(t, \boldsymbol{X}_t). \tag{1.5}$$

The expression for its pth component, resulting from the application of the Itô formula, satisfies the SDE

$$\begin{aligned} dY_t^p &= \left(\frac{\partial U^p}{\partial t} + \sum_{i=1}^{d} a^i \frac{\partial U^p}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{l=1}^{m} b^{i,l} b^{j,l} \frac{\partial^2 U^p}{\partial x_i \partial x_j} \right) dt \\ &\quad + \sum_{l=1}^{m} \sum_{i=1}^{d} b^{i,l} \frac{\partial U^p}{\partial x_i} dW_t^l, \end{aligned} \tag{1.6}$$

for $p \in \{1, 2, \dots, k\}$, where the terms on the right-hand side of (1.6) are evaluated at $(t, \boldsymbol{X}_t)$. It is a trivial but valuable observation that also the paths of the solution of the SDE (1.6) can be exactly simulated since $\boldsymbol{X}_{t_i}$ can be obtained exactly at all discretization points and, by (1.5), $\boldsymbol{Y}_{t_i}$ is simply a function of $\boldsymbol{X}_{t_i}$.

Vector of Correlated Wiener Processes

Let us now define a d-dimensional continuous process $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t = (\tilde{W}_t^1, \tilde{W}_t^2, \ldots, \tilde{W}_t^d)^\top,\ t \in [0,\infty)\}$ such that its components $\tilde{W}_t^1, \tilde{W}_t^2, \ldots, \tilde{W}_t^d$ are transformed scalar Wiener processes. In vector notation, such a d-dimensional transformed Wiener process can be expressed by the linear transform

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{a}t + \boldsymbol{B}\boldsymbol{W}_t, \tag{1.7}$$

where $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top$ is a d-dimensional vector, $\boldsymbol{B}$ is a $d \times m$-matrix and $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^m)^\top, t \in [0,\infty)\}$ is an m-dimensional standard Wiener process. By the application of the multi-dimensional Itô formula one obtains

$$d\tilde{W}_t^k = a_k dt + \sum_{i=1}^{m} b_{k,i} dW_t^i, \tag{1.8}$$

for $k \in \{1, 2, \ldots, d\}$. This means that $\tilde{W}_t^k$, $k \in \{1, 2, \ldots, d\}$, is constructed as a linear combination of components of the vector $\boldsymbol{W}_t$ plus some trend.

From the properties of Gaussian random variables, the following relation results

$$\begin{aligned} \tilde{\boldsymbol{W}}_0 &= \boldsymbol{0}, \\ \tilde{\boldsymbol{W}}_{t_{i+1}} &= \tilde{\boldsymbol{W}}_{t_i} + \boldsymbol{a}\Delta + \sqrt{\Delta}\tilde{\boldsymbol{N}}_{i+1}, \end{aligned} \tag{1.9}$$

for $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$ with $\Delta > 0$. For each $i \in \{0, 1, 2, \ldots\}$ the random vector $\tilde{\boldsymbol{N}}_{i+1} \sim \mathcal{N}_d(\boldsymbol{0}, \boldsymbol{\Sigma})$ is here a d-dimensional Gaussian vector with correlation matrix $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$.

Multi-dimensional Geometric Brownian Motions

Now, we describe multi-dimensional geometric Brownian motions that could yield a Black-Scholes model. This model emerges when taking the exponent of linearly transformed Wiener process. Denote by $\boldsymbol{S}_t$ a diagonal matrix with jth diagonal element S_t^j, $j \in \{1, 2, \ldots, d\}$, representing the jth asset price at time $t \in [0,\infty)$. Then the SDE for the jth Black-Scholes asset price S_t^j is defined by

$$dS_t^j = S_t^j \Big(a_t^j dt + \sum_{k=1}^{d} b_t^{j,k} dW_t^k \Big) \tag{1.10}$$

for $t \in [0,\infty)$ and $j \in \{1, 2, \ldots, d\}$. Here $W^k, k \in \{1, 2, \ldots, d\}$, denotes an independent standard Wiener process. Note that the Zakai equation for the Wonham filter

is of a similar form, see Kallianpur (1980) and Section 1.5, involving multiplicative noise. To represent the above SDE in matrix form we introduce the diagonal matrix $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ with

$$A_t^{i,j} = \begin{cases} a_t^j & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{1.11}$$

and diagonal matrix $\boldsymbol{B}_t^k = [B_t^{k,i,j}]_{i,j=1}^d$ with

$$B_t^{k,i,j} = \begin{cases} b_t^{j,k} & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{1.12}$$

for $k, i, j \in \{1, 2, \ldots, d\}$ and $t \in [0, \infty)$. If all these diagonal matrices commute in the sense

$$\boldsymbol{A}\boldsymbol{B}^l = \boldsymbol{B}^l\boldsymbol{A} \quad \text{and} \quad \boldsymbol{B}^l\boldsymbol{B}^k = \boldsymbol{B}^k\boldsymbol{B}^l \tag{1.13}$$

for all $k, l \in \{1, 2, \ldots, m\}$, then we can write the SDE (1.10) as matrix SDE

$$d\boldsymbol{S}_t = \boldsymbol{A}_t\boldsymbol{S}_t dt + \sum_{k=1}^d \boldsymbol{B}_t^k \boldsymbol{S}_t dW_t^k \tag{1.14}$$

for $t \in [0, \infty)$. Consequently, we obtain for the jth asset price the explicit solution

$$S_t^j = S_0^j \exp\left\{\int_0^t \left(a_s^j - \frac{1}{2}\sum_{k=1}^d (b_t^{j,k})^2\right) ds + \sum_{k=1}^d \int_0^t b_s^{j,k} dW_s^k\right\} \tag{1.15}$$

for $t \in [0, \infty)$ and $j \in \{1, 2, \ldots, d\}$. When taking the above exponential elementwise, the explicit solution of (1.10) can be expressed as the exponential

$$\boldsymbol{S}_t = \boldsymbol{S}_0 \exp\left\{\int_0^t \left(\boldsymbol{A}_s - \frac{1}{2}\sum_{k=1}^d \left(\boldsymbol{B}_s^k\right)^2\right) ds + \sum_{k=1}^d \int_0^t \boldsymbol{B}_s^k dW_s^k\right\} \tag{1.16}$$

for $t \geq 0$. Additionally, if the appreciation rates and volatilities are piecewise constant, then we can simulate exact solutions of the SDE (1.10). The main advantage of the multi-dimensional Black-Scholes model, which also made it so popular, is that it is highly tractable. It provides an explicit solution for the market dynamics and allows a range of explicit formulas for functionals.

Before we consider more complicated SDEs let us give a simple example for a two-dimensional Black-Scholes model with

$$\boldsymbol{B}^1 = \begin{pmatrix} b_1 & 0 \\ 0 & b_2\varrho \end{pmatrix} \quad \text{and} \quad \boldsymbol{B}^2 = \begin{pmatrix} 0 & 0 \\ 0 & b_2\sqrt{1-\varrho^2} \end{pmatrix}. \tag{1.17}$$

Here we obtain the following exact solution

$$S_t^1 = S_0^1 \exp\left\{\left(a_1 - \frac{1}{2}b_1^2\right)t + b_1 W_t^1\right\}, \tag{1.18}$$

$$S_t^2 = S_0^2 \exp\left\{\left(a_2 - \frac{1}{2}b_2^2\right)t + b_2\left(\varrho W_t^1 + \sqrt{1-\varrho^2}W_t^2\right)\right\}, \tag{1.19}$$

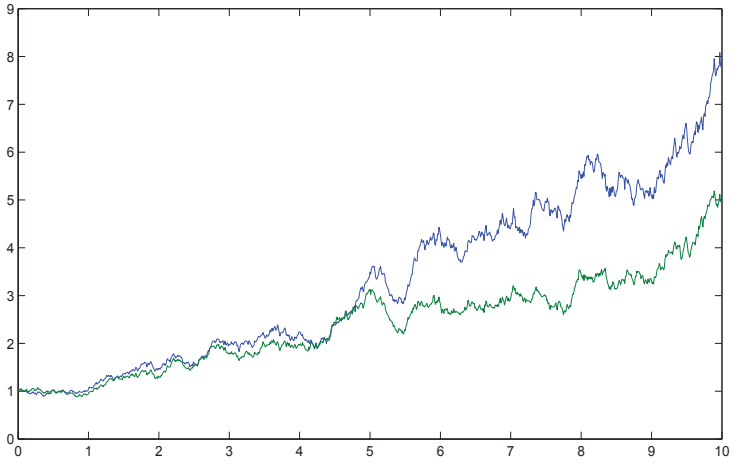

Figure 1.1: Trajectory of a two-dimensional Black-Scholes model with parameters $S_0^1 = S_0^2 = 1$, $a_1 = a_2 = 0.1$, $b_1 = b_2 = 0.2$ and $\varrho = 0.8$

for $t \in [0, \infty)$. The trajectory of this two-dimensional model is illustrated in Fig. 1.1 for the parameter choice $S_0^1 = S_0^2 = 1$, $a_1 = a_2 = 0.1$, $b_1 = b_2 = 0.2$ and $\varrho = 0.8$. One notes the high correlation between the two paths.

1.2 Matrix Ornstein-Uhlenbeck Processes

In this section we will show how to simulate matrices of Ornstein-Uhlenbeck (OU)-processes. This can be performed by using matrices of time changed Wiener processes. Therefore, we will first introduce matrix Wiener processes and show how to simulate time changed matrix Wiener processes.

Matrix Wiener Processes

Let us define a $d \times m$ standard matrix Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_t = [W_t^{i,j}]_{i,j=1}^{d,m}, t \in [0, \infty)\}$. This matrix stochastic process can be obtained by the following construction

$$\begin{aligned} \boldsymbol{W}_0 &= \boldsymbol{0} \\ \boldsymbol{W}_{t_{i+1}} &= \boldsymbol{W}_{t_i} + \sqrt{\Delta}\boldsymbol{N}_{i+1}, \end{aligned} \tag{1.20}$$

for the times $t_i = i\Delta$, $i = \{0, 1, \dots\}$, with $\Delta > 0$ and $d \times m$-matrix $\boldsymbol{0}$ of zero elements. Here $\boldsymbol{N}_{i+1} \sim \mathcal{N}_{d\times m}(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{I}_d)$ is a matrix of zero mean Gaussian distributed

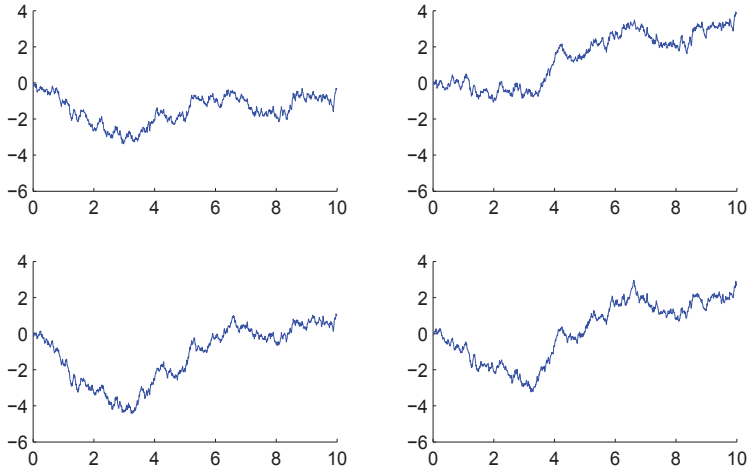

Figure 1.2: 2×2 matrix Wiener process with both correlated rows and columns

random variables. The covariance matrix $\boldsymbol{I}_m \otimes \boldsymbol{I}_d$ is an $m \times m$ block matrix with $d \times d$ block matrices as its elements, that is,

$$\boldsymbol{I}_m \otimes \boldsymbol{I}_d = \begin{pmatrix} \boldsymbol{I}_d & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_d & \dots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{I}_d \end{pmatrix}, \tag{1.21}$$

were $\boldsymbol{I}_d$ denotes the $d \times d$ identity matrix. Moreover, similar to the vector case, we are able to define a transformed matrix Wiener process $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t = [\tilde{W}_t^{i,j}]_{i,j=1}^{d,m}, t \in [0, \infty)\}$ using the above matrix Wiener process $\boldsymbol{W}$ as follows:

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{M}t + \boldsymbol{\Sigma}_1 \boldsymbol{W}_t \boldsymbol{\Sigma}_2^\top, \tag{1.22}$$

where $\boldsymbol{M}$ is a $d \times m$ matrix and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are nonsingular $d \times d$ and $m \times m$ matrices, respectively. Values of such a matrix stochastic process can be obtained at the discrete times $t_i = i\Delta$ by the following recursive computation:

$$\begin{aligned} \tilde{\boldsymbol{W}}_0 &= \boldsymbol{0} \\ \tilde{\boldsymbol{W}}_{t_{i+1}} &= \tilde{\boldsymbol{W}}_{t_i} + \boldsymbol{M}\Delta + \sqrt{\Delta}\tilde{\boldsymbol{N}}_{i+1}, \end{aligned} \tag{1.23}$$

for $i \in \{0, 1, \dots\}$ and independent $\tilde{\boldsymbol{N}}_{i+1} \sim \mathcal{N}_{d \times m}(\boldsymbol{0}, \boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)$. Here, the covariance matrix $\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1$ is an $m \times m$ block matrix of the form

$$\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_{1,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{1,2}^2 \boldsymbol{\Sigma}_1 & \dots & \sigma_{1,m}^2 \boldsymbol{\Sigma}_1 \\ \sigma_{2,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{2,2}^2 \boldsymbol{\Sigma}_1 & \dots & \sigma_{2,m}^2 \boldsymbol{\Sigma}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{m,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{m,2}^2 \boldsymbol{\Sigma}_1 & \dots & \sigma_{m,m}^2 \boldsymbol{\Sigma}_1 \end{pmatrix}, \tag{1.24}$$

where $\mathbf{\Sigma}_1 = [\sigma^1_{i,j}]^d_{i,j}$ and $\mathbf{\Sigma}_2 = [\sigma^2_{i,j}]^m_{i,j}$.

In Fig. 1.2 we illustrate a 2×2 matrix transformed Wiener process $\tilde{\boldsymbol{W}}$ for $\varrho = 0.8$, which was obtained from the standard 2×2 matrix Wiener process $\boldsymbol{W}$ by the following transformation:

$$\tilde{\boldsymbol{W}}_t = \mathbf{\Sigma}_1 \boldsymbol{W}_t \mathbf{\Sigma}_2^\top, \tag{1.25}$$

where

$$\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \begin{pmatrix} 1 & 0 \\ \varrho & \sqrt{1-\varrho^2} \end{pmatrix}. \tag{1.26}$$

We note in Fig. 1.2 the correlation effect on the trajectories on both the elements of the columns and the rows of such a 2×2 matrix valued transformed Wiener process.

Time Changed Wiener Processes

Instead of multiplying the time by some constant to scale the fluctuations of the Wiener paths, one can introduce time dependent scaling by a, so called, time change process $\varphi = \{\varphi(t), t \geq 0\}$. Let us now consider a vector of time changed standard independent Wiener processes $\boldsymbol{W} = \{\boldsymbol{W}_{\varphi(t)} = (W^1_{\varphi(t)}, \ldots, W^m_{\varphi(t)})^\top, t \in [0,\infty)\}$. Given the time discretization $t_i = i\Delta$, $i \in \{0, 1, 2, \ldots\}$, with time step size $\Delta > 0$ we obtain the corresponding time changed Wiener process at discretization times by the following iterative formula:

$$\begin{aligned} \boldsymbol{W}_{\varphi(0)} &= \mathbf{0} \\ \boldsymbol{W}_{\varphi(t_{i+1})} &= \boldsymbol{W}_{\varphi(t_i)} + \sqrt{\varphi(t_{i+1}) - \varphi(t_i)}\boldsymbol{N}_{i+1}, \end{aligned} \tag{1.27}$$

where the vector $\boldsymbol{N}_{i+1} \sim \mathcal{N}_m(\mathbf{0}, \boldsymbol{I})$ is formed by independent standard Gaussian vector random variables. Here $\boldsymbol{I}$ is the $m \times m$ identity matrix. Obviously, it is possible to apply different time changes to different elements of the vector $\boldsymbol{W}$. For instance, let us define

$$\varphi_j(t) = \frac{b_j^2}{2c_j}(e^{2c_j t} - 1) \tag{1.28}$$

for $t \in [0,\infty)$, $b_j > 0, c_j > 0$ and $j \in \{1, 2, \ldots, m\}$.

In order to obtain a time changed vector Wiener process, whose elements are correlated time changed Wiener processes, it is sufficient to define a new vector $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_{\varphi(t)} = (\tilde{W}^1_{\varphi(t)}, \ldots, \tilde{W}^d_{\varphi(t)})^\top, t \in [0,\infty)\}$ by the following transformation

$$\tilde{\boldsymbol{W}}_{\varphi(t)} = \boldsymbol{B}\boldsymbol{W}_{\varphi(t)}, \tag{1.29}$$

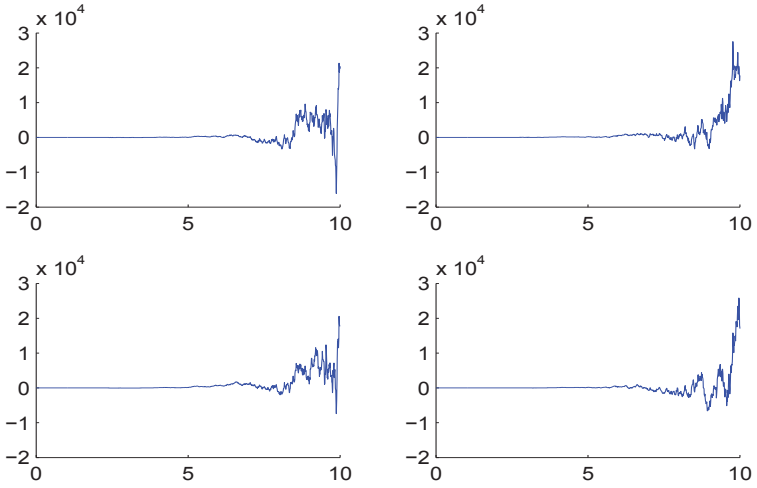

Figure 1.3: Matrix valued time changed Wiener process

where $\boldsymbol{B}$ is a $d \times m$-matrix of coefficients and $\boldsymbol{W} = \{\boldsymbol{W}_{\varphi(t)} = (W^1_{\varphi(t)}, \dots, W^m_{\varphi(t)})^\top, t \in [0, \infty)\}$ is an m-dimensional time changed Wiener process with independent components.

Additionally, let us define a $d \times m$ standard time changed matrix Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_{\varphi(t)} = [W^{j,k}_{\varphi(t)}]^{d,m}_{j,k=1}, t \in [0, \infty)\}$. Here, the independent elements of the matrix $\boldsymbol{W}_{\varphi(t)}$ are such that

$$W^{j,k}_{\varphi_{j,k}(t_{i+1})} - W^{j,k}_{\varphi_{j,k}(t_i)} \sim \mathcal{N}\left(0, \varphi_{j,k}(t_{i+1}) - \varphi_{j,k}(t_i)\right), \tag{1.30}$$

where $W^{k,j}_{\varphi_{k,j}(0)} = 0$, $t_i = i\Delta$, $i \in \{0, 1, \dots\}$ and $j \in \{1, 2, \dots, d\}$, $k \in \{1, 2, \dots, m\}$. For instance, we may define the (j,k)th time transformation by

$$\varphi_{j,k}(t) = \frac{b^2_{j,k}}{2c_{j,k}} \left(e^{2c_{j,k}t} - 1\right) \tag{1.31}$$

for $t \in [0, \infty)$, $b_{j,k} > 0, c_{j,k} > 0$, and $j \in \{1, 2, \dots, d\}$, $k \in \{1, 2, \dots, m\}$. In order to obtain a time changed matrix Wiener process with correlated elements we can use the formula (1.22).

In Fig. 1.3 we display a time changed matrix Wiener process for $d = m = 2$ with the covariance matrix $\boldsymbol{I} \otimes \boldsymbol{\Sigma}_1$, where $\boldsymbol{\Sigma}_1$ is as in (1.26), $\varrho = 0.8$ and the parameters in the time change equal $b_{j,k} = \sqrt{2}$ and $c_{j,k} = 1$ for $j, k \in \{1, 2\}$. That is, the same time change is applied to each of the elements of this matrix Wiener process. More precisely, we construct $\tilde{\boldsymbol{W}}$ by the relation

$$\tilde{\boldsymbol{W}}_{\varphi(t)} = \boldsymbol{\Sigma}_1 \boldsymbol{W}_{\varphi(t)} \boldsymbol{I}. \tag{1.32}$$

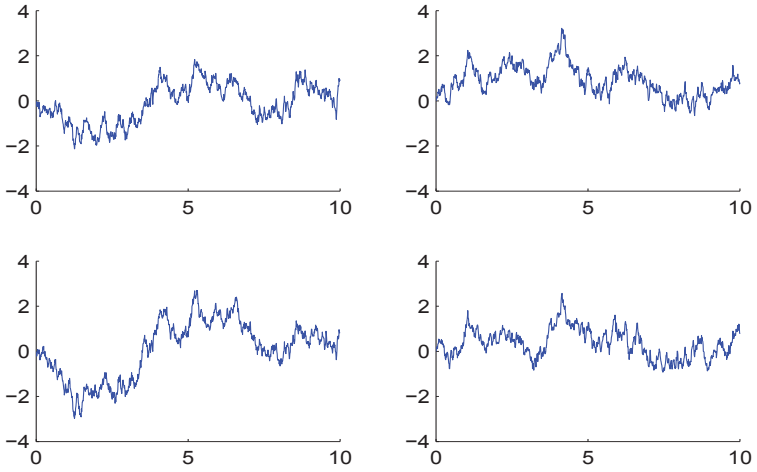

Figure 1.4: Matrix valued Ornstein-Uhlenbeck process

In this case we obtain a time changed matrix Wiener process $\tilde{\boldsymbol{W}}$ whose rows have independent elements, while its columns have dependent elements.

Multi-dimensional Ornstein-Uhlenbeck (OU)-Processes

Let us now consider vector and matrix valued OU-processes. We will here construct multi-dimensional OU-processes as time changed and scaled multi-dimensional Wiener processes. Note that given the following two functions

$$s_t = \exp\{-ct\} \quad \text{and} \quad \varphi(t) = \frac{b^2}{2c}(e^{2ct} - 1) \tag{1.33}$$

for $t \in [0, \infty)$, $b, c > 0$, a scalar OU-process $Y = \{Y_t, t \in [0, \infty)\}$ can be represented in terms of a time changed and scaled scalar Wiener process, that is

$$Y_t = s_t W_{\varphi(t)}, \tag{1.34}$$

where $W = \{W_\varphi, \varphi \geq 0\}$ is a standard Wiener process in φ-time. By Itô's formula we obtain

$$\begin{aligned} dY_t &= W_{\varphi(t)} ds_t + s_t dW_{\varphi(t)} = -\frac{Y_t}{s_t} c s_t dt + s_t \frac{b}{s_t} d\tilde{W}_t \\ &= -cY_t dt + b d\tilde{W}_t, \end{aligned} \tag{1.35}$$

where $dW_{\varphi(t)} = \frac{b}{s_t} d\tilde{W}_t$, with $\tilde{W}$ denoting a standard Wiener process in t-time.

It is straightforward to obtain a vector OU-process by

$$\boldsymbol{Y}_t = s_t \boldsymbol{W}_{\varphi(t)}, \tag{1.36}$$

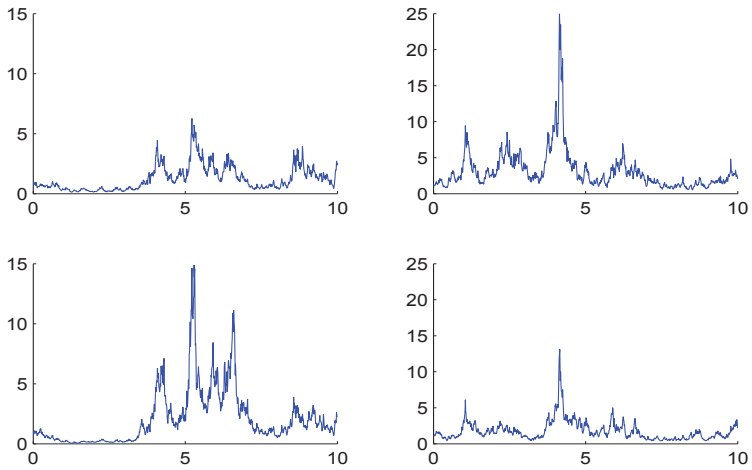

Figure 1.5: Matrix valued geometric OU-process

that is, $Y_t^j = s_t^j W^j_{\varphi_j(t)}$ for $j \in \{1, 2, \ldots, d\}$ and $t \geq 0$. The generalization to a matrix OU-process is obvious. The construction of this process starts by forming a time changed $d \times m$ matrix Wiener process. It is then scaling each element of this matrix by a function $s_t^{j,k}$ for $j \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$. Hence, the elements of such a matrix can be expressed by the relation $Y_t^{j,k} = s_t^{j,k} W^{j,k}_{\varphi_{j,k}(t)}$ for $j \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$.

We illustrate in Fig. 1.4 the matrix OU-process, obtained from the time changed matrix Wiener process in Fig. 1.3. Since, the time changed matrix Wiener process has correlated rows and independent columns, the resulting OU-process, shown in Fig. 1.4, shares this feature.

Multi-dimensional Geometric Ornstein-Uhlenbeck Processes

The Itô formula provides a general tool to generate a world of exact solutions of SDEs based on functions of the solutions of those SDEs with exact solutions we have already considered. As an example, let us generate explicit solutions for a geometric OU-process. Here each element of a matrix valued OU-process is simply exponentiated. More precisely, when denoting by $\boldsymbol{X}_t = [X_t^{j,k}]_{j,k=1}^{d,m}$ a $d \times m$ matrix OU-process value and by $\boldsymbol{Y}_t = [Y_t^{j,k}]_{j,k=1}^{d,m}$ the corresponding $d \times m$ matrix geometric OU-process value at time t, then we obtain the elements of the matrix $\boldsymbol{Y}_t$ by

$$Y_t^{j,k} = \exp\{X_t^{j,k}\} \tag{1.37}$$

for $t \in [0, \infty)$.

In Fig. 1.5 we illustrate a 2×2 matrix geometric OU-process obtained from the matrix OU-process in Fig. 1.4 by application of (1.37) to each of its elements. More complex applications of the Itô formula for generating exact solutions will be considered in the next section.

1.3 Wishart Processes

In this section we will discuss the exact simulation of Wishart processes, see Bru (1991). These are matrix valued stochastic processes where their one-dimensional version is generating squared Bessel processes. Therefore, we will start by describing the exact simulation of a squared Bessel process, which later will be generalized to its matrix equivalent, the Wishart process.

Squared Bessel Processes

A squared Bessel process (BESQ_x^δ) $X = \{X_\varphi,\ \varphi \in [\varphi_0, \infty)\}$, $\varphi_0 \geq 0$, of dimension $\delta \geq 0$ and with initial value $x > 0$, see Revuz & Yor (1999), is a fundamental stochastic process which appears in various ways in theory and applications, for instance, in financial modeling. This process can be described by the SDE

$$dX_\varphi = \delta \, d\varphi + 2 \sqrt{|X_\varphi|} \, dW_\varphi \tag{1.38}$$

for $\varphi \in [\varphi_0, \infty)$ with $X_{\varphi_0} = x \geq 0$, where $W = \{W_\varphi,\ \varphi \in [\varphi_0, \infty)\}$ is a standard Wiener process starting at the initial φ-time, $\varphi = \varphi_0$, $\delta > 0$. This means, for $\varphi \in [\varphi_0, \infty)$ one has as increment of the quadratic variation of W the difference

$$[W]_\varphi - [W]_{\varphi_0} = \varphi - \varphi_0$$

for all $\varphi \in [\varphi_0, \infty)$. Furthermore, if we fix the behavior of X_φ at the boundary zero as reflection, then the absolute sign under the square root in (1.38) can be removed, and X_φ remains nonnegative and has a unique strong solution, see Revuz & Yor (1999).

The solution of the above SDE can be simulated exactly for the case when the dimension of this process is an integer, that is $\delta \in \{1, 2, \dots\}$. More precisely, for $\delta \in \{1, 2, \dots\}$ and $x \geq 0$ the BESQ_x^δ process X can be expressed as the sum of the

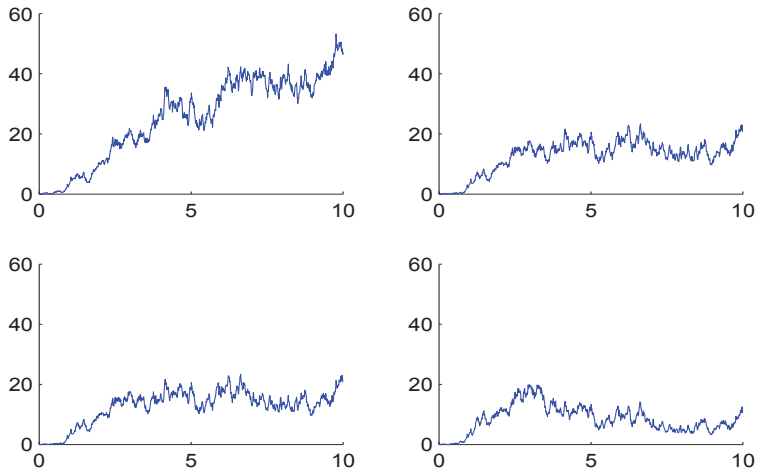

Figure 1.6: Wishart process

squares of δ independent Wiener processes $W^1, W^2, \ldots, W^\delta$ in φ-time, which start at time $\varphi = \varphi_0$ in $w^1 \in \Re, w^2 \in \Re, \ldots w^\delta \in \Re$, respectively, such that $x = \sum_{k=1}^{\delta}(w^k)^2$. We can now construct the solution of (1.38) as

$$X_\varphi = \sum_{k=1}^{\delta}(w^k + W_\varphi^k)^2 \tag{1.39}$$

for $\varphi \in [\varphi_0, \infty)$. Applying the Itô formula we obtain

$$dX_\varphi = \delta d\varphi + 2\sum_{k=1}^{\delta}(w^k + W_\varphi^k)dW_\varphi^k \tag{1.40}$$

for $\varphi \in [\varphi_0, \infty)$ with $X_0 = \sum_{k=1}^{\delta}(w^k)^2 = x$. Furthermore, by setting

$$dW_\varphi = |X_\varphi|^{-\frac{1}{2}}\sum_{k=1}^{\delta}(w^k + W_\varphi^k)dW_\varphi^k \tag{1.41}$$

we obtain the SDE (1.38). Note that we have for W_φ the quadratic variation

$$[W]_\varphi = \int_{\varphi_0}^{\varphi}\frac{1}{X_s}\sum_{k=1}^{\delta}(w^k + W_s^k)^2 ds = \varphi - \varphi_0. \tag{1.42}$$

Hence, by the Lévy theorem the process W_φ is a Wiener process in φ-time.

Wishart Process

The matrix generalization of a squared Bessel process is a Wishart process originally introduced in Bru (1991). The $m \times m$ matrix valued Wishart process with dimension

$\delta \in \{1, 2, \dots\}$ is the matrix process $\boldsymbol{S} = \{\boldsymbol{S}_t, t \geq 0\}$ with

$$\boldsymbol{S}_t = \boldsymbol{W}_t^\top \boldsymbol{W}_t \tag{1.43}$$

for $t \in \Re^+$ and initial matrix $\boldsymbol{s}_0 = \boldsymbol{W}_0^\top \boldsymbol{W}_0$. Here $\boldsymbol{W}_t$ is the value at time $t \geq 0$ of a $\delta \times m$ matrix Wiener process. Itô calculus applied to the relation (1.43) results in the following SDE

$$d\boldsymbol{S}_t = \delta \boldsymbol{I} dt + d\boldsymbol{W}_t^\top \boldsymbol{W}_t + \boldsymbol{W}_t^\top d\boldsymbol{W}_t, \tag{1.44}$$

where $\boldsymbol{I}$ is the $m \times m$ identity matrix. It can be shown that $\tilde{\boldsymbol{W}}_t$ expressed by

$$d\tilde{\boldsymbol{W}}_t = \left(\sqrt{\boldsymbol{S}_t}\right)^{-1} \boldsymbol{W}_t^\top d\boldsymbol{W}_t \tag{1.45}$$

is an $m \times m$ matrix Wiener process. Here $\sqrt{\boldsymbol{S}_t}$ represents the symmetric positive square root of $\boldsymbol{S}_t$, while $\left(\sqrt{\boldsymbol{S}_t}\right)^{-1}$ is the inverse of the matrix $\sqrt{\boldsymbol{S}_t}$. Note also that

$$\begin{aligned} d\tilde{\boldsymbol{W}}_t^\top &= d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\left(\sqrt{\boldsymbol{S}_t}\right)^{-1}\right)^\top = d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\left(\sqrt{\boldsymbol{S}_t}\right)^\top\right)^{-1} \\ &= d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\sqrt{\boldsymbol{S}_t^\top}\right)^{-1} = d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\sqrt{\boldsymbol{S}_t}\right)^{-1}, \end{aligned} \tag{1.46}$$

since $\boldsymbol{S}_t$ is a symmetric matrix. Therefore, (1.44) can be rewritten in the following form

$$d\boldsymbol{S}_t = \delta \boldsymbol{I} dt + \sqrt{\boldsymbol{S}_t} d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{S}_t} \tag{1.47}$$

for $t \in \Re^+$.

In Fig. 1.6 we plot a 2×2 Wishart process of dimension $\delta = 2$. The matrix Wiener process in this example was obtained by assuming the covariance matrix $\boldsymbol{I} \otimes \boldsymbol{\Sigma}_1$, where $\boldsymbol{\Sigma}_1$ is as in (1.26), with $\varrho = 0.8$.

1.4 Affine Matrix Processes

Another group of matrix valued stochastic processes that can be simulated exactly is that of matrices of affine processes, see Duffie & Kan (1994) or for the one-factor case Filipović (2001). This family of stochastic processes has as its special case a matrix of square root (SR) processes, which are directly linked to Wishart processes. They also can be obtained from matrices of OU-processes. Both methods of exact simulation are described below.

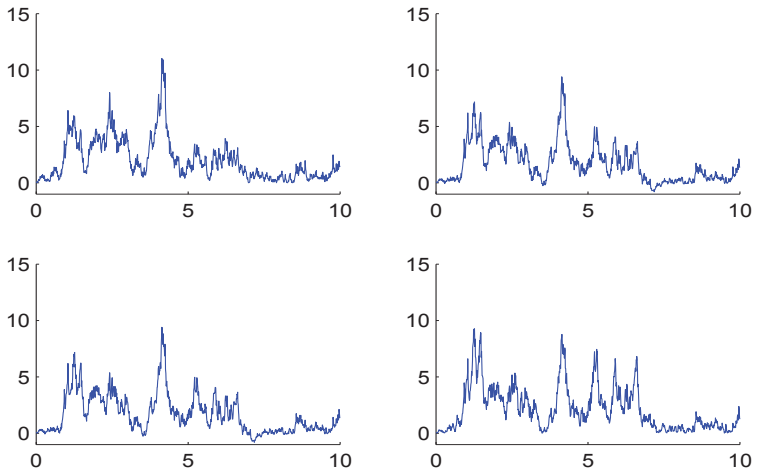

Figure 1.7: Matrix valued square root process

SR-Processes Generated via OU-Processes

Let us first consider δ standard OU-processes, that is

$$dX_t^i = -cX_t^i dt + bdW_t^i \tag{1.48}$$

for $t \in [0, \infty)$, with $X_0^i = x_0$, $c, b \in \Re$ and independent standard Wiener processes W^i for $i \in \{1, 2, \ldots, \delta\}$. The square of such an OU-process has the Itô differential

$$d(X_t^i)^2 = (b^2 - 2c(X_t^i)^2) + 2bX_t^i dW_t^i, \tag{1.49}$$

for $t \in [0, \infty)$ and $i \in \{1, 2, \ldots, \delta\}$. Furthermore, we can form the sum of the δ squared OU-processes, that is,

$$Y_t = \sum_{i=1}^{\delta} (X_t^i)^2 \tag{1.50}$$

for $t \in [0, \infty)$. The SDE for Y_t turns out to be

$$dY_t = \sum_{i=1}^{\delta} \left(b^2 - 2c(X_t^i)^2\right) dt + 2b \sum_{i=1}^{\delta} X_t^i dW_t^i \tag{1.51}$$

for $t \in [0, \infty)$. In order to simplify the above SDE we introduce another Wiener process $\bar{W} = \{\bar{W}_t, t \in [0, \infty)\}$ defined as

$$\bar{W}_t = \int_0^t d\bar{W}_s = \sum_{i=1}^{\delta} \int_0^t \frac{X_s^i}{\sqrt{Y_s}} dW_s^i \tag{1.52}$$

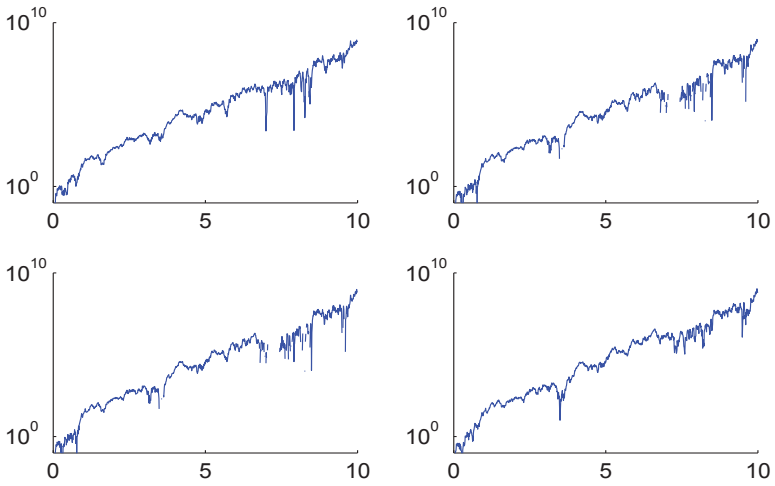

Figure 1.8: Time changed Wishart process in log-scale

for $t \in [0, \infty)$. It can be shown that the quadratic variation of $\bar{W}$ equals

$$[\bar{W}]_t = \int_0^t \sum_{i=1}^{n} \frac{(X_s^i)^2}{Y_s} ds = t. \tag{1.53}$$

Hence, by the Lévy theorem we see that $\bar{W}$ is a standard Wiener process. Therefore, we obtain an equivalent SDE for the square root process Y in the form

$$dY_t = (\delta b^2 - 2cY_t)dt + 2b\sqrt{Y_t} d\bar{W}_t \tag{1.54}$$

for $t \in [0, \infty)$ with $Y_0 = \delta(x_0)^2$. Note that this process is an SR-process of dimension $\delta \in \{1, 2, \dots\}$. It is well-known that for $\delta = 1$ the value Y_t can reach zero and is reflected at this boundary. For $\delta \in \{2, 3, \dots\}$ the process never reaches zero for $x_0 > 0$.

Matrix Valued Squares of OU-Processes

Kendall (1989) and Bru (1991) studied the matrix generalization for squares of OU-processes. Denote by $\boldsymbol{X}_t$ a $\delta \times m$ matrix solution of the SDE

$$d\boldsymbol{X}_t = -c\boldsymbol{X}_t dt + b d\boldsymbol{W}_t, \tag{1.55}$$

for $t \geq 0$, with $\boldsymbol{X}_0 = \boldsymbol{x}_0$. Here $\boldsymbol{W}_t$ is a $\delta \times m$ matrix Wiener process and $\boldsymbol{x}_0$ is a $\delta \times m$ deterministic initial value matrix; $b, c \in \Re$. By setting

$$\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t, \quad \boldsymbol{s}_0 = \boldsymbol{x}_0^\top \boldsymbol{x}_0 \tag{1.56}$$

and denoting $d\tilde{\boldsymbol{W}}_t = \sqrt{\boldsymbol{S}_t^{-1}}\boldsymbol{X}_t^\top d\boldsymbol{W}_t$ we obtain an $m \times m$ matrix SR process $\boldsymbol{S}$ of dimension $\delta = \{1, 2, \dots\}$. Note that the elements of $\tilde{\boldsymbol{W}}_t$ can be correlated. Then $\boldsymbol{S}_t$ solves the SDE

$$d\boldsymbol{S}_t = (\delta b^2 \boldsymbol{I} - 2c\boldsymbol{S}_t)dt + b(\sqrt{\boldsymbol{S}_t}d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{S}_t}) \tag{1.57}$$

for $t \geq 0$, $\boldsymbol{S}_0 = \boldsymbol{s}_0$. Here $\boldsymbol{S}_t$ corresponds to a continuous-time process of stochastic, symmetric, positive definite matrices, while $\sqrt{\boldsymbol{S}_t}$ is the positive symmetric square root of the matrix $\boldsymbol{S}_t$, see Gouriéroux & Sufana (2004). Furthermore, $\boldsymbol{S}_t^{-1}$ is the inverse of the symmetric positive definite $m \times m$ matrix $\boldsymbol{S}_t$ and $\sqrt{\boldsymbol{S}_t^{-1}}$ its square root.

Note that for $m = 1$ the transform (1.56) simplifies to equation (1.50). We illustrate in Fig. 1.7 the matrix SR-process obtained from the matrix OU-process. Note that not all non-diagonal elements of such a matrix process remain always positive. In particular, the elements $S^{1,2}$ and $S^{2,1}$ are identical and, in general, not positive. However, the diagonal elements $S^{1,1}$ and $S^{2,2}$ are correlated SR-processes, which are always positive.

SR-Processes Generated via Squared Bessel Processes

Using squared Bessel processes one can derive SR-processes by certain transformations. For this reason let $c : [0, \infty) \to \Re$ and $b : [0, \infty) \to \Re$ be given deterministic functions of time. We introduce the exponential

$$s_t = s_0 \exp\left\{\int_0^t c_u \, du\right\} \tag{1.58}$$

and the φ-time

$$\varphi(t) = \varphi(0) + \frac{1}{4}\int_0^t \frac{b_u^2}{s_u} \, du \tag{1.59}$$

for $t \in [0, \infty)$ and $s_0 > 0$. Note that we have an explicit representation for the function $\varphi(t)$ in the case of constant parameters $b_t = \bar{b} \neq 0$ and $c_t = \bar{c} \neq 0$, where

$$\varphi(t) = \varphi(0) + \frac{\bar{b}^2}{4\bar{c}s_0}(1 - \exp\{-\bar{c}t\}) \tag{1.60}$$

for $t \in [0, \infty)$ and $s_0 > 0$. Furthermore, if $\varphi(0) = -\frac{\bar{b}^2}{4\bar{c}s_0}$, then this function simply equals

$$\varphi(t) = -\frac{\bar{b}^2}{4\bar{c}s_0} \exp\{-\bar{c}t\} \tag{1.61}$$

for $t \in [0, \infty)$, $s_0 > 0$, $\bar{b} \neq 0$ and $\bar{c} \neq 0$.

Given a squared Bessel process X of dimension $\delta > 0$, using our previous notation, we introduce the SR-process $Y = \{Y_t, t \geq 0\}$ of dimension $\delta > 0$ via the relation

$$Y_t = s_t \, X_{\varphi(t)} \tag{1.62}$$

indexed by time $t \geq 0$, see also Delbaen & Shirakawa (1997).

Furthermore, by (1.38), (1.58), (1.59) and (1.62) and the Itô formula we can express (1.62) in terms of the SDE

$$dY_t = \Big(\frac{\delta}{4} \, b_t^2 + c_t \, Y_t\Big) dt + b_t \, \sqrt{Y_t} \, dU_t \tag{1.63}$$

for $t \in [0, \infty)$, $Y_0 = s_0 X_{\varphi(0)}$ and

$$dU_t = \sqrt{\frac{4 s_t}{b_t^2}} dW_{\varphi(t)}.$$

Note that U_t forms by the Lévy theorem a Wiener process, since

$$[U]_t = \int_0^t \frac{4 s_z}{b_z^2} d\varphi(z) = t. \tag{1.64}$$

The same time-change formula can be applied in the more general matrix case. Given the Wishart process $\boldsymbol{X}$ it can be shown that the matrix square root process can be obtained from the Wishart process by the following transformation

$$\boldsymbol{Y}_t = s_t \boldsymbol{X}_{\varphi(t)}, \tag{1.65}$$

where s_t and $\varphi(t)$ are as in (1.58) and (1.59), respectively. By (1.47), (1.58), (1.59) and (1.65) and the Itô formula we can express (1.65) in terms of the matrix SDE

$$d\boldsymbol{Y}_t = \left(\frac{\delta}{4} b_t^2 \boldsymbol{I} + c_t \boldsymbol{Y}_t\right) dt + \frac{b_t}{2} \left(\sqrt{\boldsymbol{Y}_t} d\boldsymbol{U}_t + d\boldsymbol{U}_t^\top \sqrt{\boldsymbol{Y}_t}\right) \tag{1.66}$$

for $t \in [0, \infty)$, $\boldsymbol{Y}_0 = s_0 \boldsymbol{X}_{\varphi(0)}$ and where $d\boldsymbol{U}_t = \sqrt{\frac{4 s_t}{b_t^2}} d\boldsymbol{W}_{\varphi(t)}$ is the stochastic differential of a matrix Wiener process.

In Fig. 1.8 we display the trajectory of the elements of a 2×2 matrix time changed Wishart process $\boldsymbol{X}_{\varphi(t)}$ in log-scale. Here the off-diagonal elements do not show any value for the time periods when the argument of the logarithm becomes negative. Such periods do not arise for the diagonal elements which are of main interest.

We now construct a trajectory of a 2×2 matrix SR-process obtained as time changed Wishart process by the use of formula (1.65). Note that this matrix SR-process is identical to the matrix SR-process in Fig. 1.7 obtained via squares of OU-processes. In Fig. 1.7 we see that the off-diagonal elements have near the time $t = 7$ indeed negative values.

Multi-dimensional Affine Processes

Let us now transform further the above obtained multi-dimensional SR-process in order to obtain multi-dimensional affine processes, see Duffie & Kan (1994). These processes have affine, that is linear drift and linear squared diffusion coefficients. In order to obtain members of this class of multi-dimensional processes we can simply shift the multi-dimensional SR-process by a nonnegative, differentiable function of time $a : [0, \infty) \to [0, \infty)$, defined through its derivative

$$a'_t = \frac{da_t}{dt} \tag{1.67}$$

for $t \in [0, \infty)$ with $a_0 \in [0, \infty)$. More precisely, we define the process $\boldsymbol{R} = \{\boldsymbol{R}_t, t \in [0, \infty)\}$ such that

$$\boldsymbol{R}_t = \boldsymbol{Y}_t + a_t \boldsymbol{I} \tag{1.68}$$

for $t \in [0, \infty)$. It is also possible to obtain more general affine processes by shifting the matrix valued SR-process by a matrix $\boldsymbol{A}_t$ of nonnegative differentiable functions of the type (1.67), that is,

$$\boldsymbol{R}_t = \boldsymbol{Y}_t + \boldsymbol{A}_t \tag{1.69}$$

for $t \in [0, \infty)$. In this case $\boldsymbol{R}_t$ solves the following matrix SDE

$$\begin{aligned} d\boldsymbol{R}_t &= \left(\frac{\delta}{4} b_t^2 \boldsymbol{I} + \boldsymbol{A}'_t - c_t \boldsymbol{A}_t + c_t \boldsymbol{R}_t\right) dt \\ &\quad + \frac{b_t}{2}\left(\sqrt{\boldsymbol{R}_t - \boldsymbol{A}_t} d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{R}_t - \boldsymbol{A}_t}\right), \end{aligned} \tag{1.70}$$

for $t \in [0, \infty)$. Here $\boldsymbol{A}'_t$ denotes the matrix of the derivatives of the type (1.67) for the shifts of each element. Obviously, we applied here the Itô formula to the equation (1.69).

We considered above the exact simulation of solutions of multi-dimensional SDEs driven by vector or matrix Wiener processes. The simulation methods described can be adapted also to multi-dimensional SDEs when these are driven by more general

vector or matrix valued Lévy processes. In principle, one can substitute the Wiener processes by some Lévy processes. For details of such simulation method we refer to Platen & Rendek (2009).

It remains to emphasize that advanced software packages, as Matlab and Mathematica, provide routines that generate a range of random variables with various distributions that are needed in the above described simulations. For some multi-dimensional distribution functions it is possible by using such software to sample corresponding vector random variables. These are directly available for computation and should have exactly the requested multivariate transition distribution function.

The above presented exact and almost exact simulation methods for multi-dimensional SDEs lead to accurate scenario simulations that are reliable over long periods of time. This is important for various applications, for instance, the pricing of insurance and pension contracts. Typically arising numerical stability problems are simply avoided by exact simulation. Finally, we remark that there is no major problem introducing further jump effects into the considered type of dynamics via a jump-adapted time discretization, see Platen (1982). This enlarges significantly the class of processes that allow exact simulation.

1.5 Quasi-exact Approximation of Hidden Markov Chain Filters

This section studies the application of exact simulation methods to SDEs with multi-dimensional multiplicative noise as appear in filtering. Filtering problems arise, for instance, when fitting adaptively regime switching models in finance. Parameters in the model may change according to a hidden Markov chain. Filtering provides an adaptive method to estimate the hidden parameter and evaluate on this basis derivatives and other functionals. The methods described in this section can be employed for randomizing the parameters in the model we will derive for diversified indices in Chapter 4. SDEs with multiplicative noise naturally occur as Zakai equation in hidden Markov chain filtering. This section proposes a quasi-exact or almost exact approximation method for hidden Markov chain filters, which can be applied when discrete time approximations, such as the Euler scheme, may not be adequate in practice.

Hidden Markov Chain Filters

First we introduce filters for hidden, continuous time, finite state Markov chains. Let $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$ with $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$ and $T \in [0, \infty)$ be the underlying filtered probability space and suppose that the hidden state process $\xi = \{\xi_t,\, t \in [0, T]\}$ is a continuous time, homogeneous Markov chain on the finite state space $\mathcal{X} = \{a_1, a_2, \ldots, a_d\}$, $d \in \{0, 1, \ldots\}$. Its d-dimensional probability vector $\boldsymbol{p}(t) = (p_1(t), \ldots, p_d(t))^\top$ at time t, with components

$$p_i(t) = P(\xi_t = a_i) \tag{1.71}$$

for each $a_i \in \mathcal{X}$, satisfies then the vector ordinary differential equation (ODE)

$$\frac{d\boldsymbol{p}(t)}{dt} = \boldsymbol{A}\,\boldsymbol{p}(t), \tag{1.72}$$

where $\boldsymbol{A}$ is the intensity matrix. The initial probability vector equals $\boldsymbol{p}(0) = \boldsymbol{p}_0$. The solution of the vector ODE (1.72) is then of the form

$$\boldsymbol{p}(t) = \exp\{\boldsymbol{A}t\}\boldsymbol{p}_0. \tag{1.73}$$

Here $\exp\{\cdot\}$ denotes the matrix exponential defined as

$$\exp\{\boldsymbol{A}\} = \sum_{k=0}^{\infty} \boldsymbol{A}^k \frac{1}{k!}. \tag{1.74}$$

In addition, suppose that the m-dimensional observation process $\boldsymbol{W} = \{\boldsymbol{W}_t,\, t \in [0, T]\}$ is the solution of the SDE

$$d\boldsymbol{W}_t = \boldsymbol{h}(\xi_t)\, dt + d\boldsymbol{W}_t^* \tag{1.75}$$

for $t \in [0, T]$ with $\boldsymbol{W}_0 = \boldsymbol{W}_0^* \in \Re^m$, $m \in \{1, 2, \ldots\}$. This type of disturbance of a signal by a Wiener process is called a Wonham filter problem, see Wonham (1965).

In the SDE (1.75) the noise process $\boldsymbol{W}^* = \{\boldsymbol{W}_t^*,\, t \in [0, T]\}$ with $\boldsymbol{W}_0^* = 0$ is an m-dimensional standard Wiener process with respect to the real world probability measure P. The Wiener process $\boldsymbol{W}^*$ is assumed to be independent of the hidden state process ξ. Finally, let

$$\mathcal{Y}_t = \sigma\,\{\boldsymbol{W}_s,\, s \in [0, t]\}$$

denote the observation sigma-algebra generated by the observations $\boldsymbol{W}_s$ for $s \in [0, t]$. This means, $\underline{\mathcal{Y}} = (\mathcal{Y}_t)_{t\in[0,T]}$ is the filtration that represents the release of observed

information, whereas $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$ with $\mathcal{A}_t = \sigma\{\xi_s, \boldsymbol{W}_s : s \in [0,t]\}$ expresses the evolution over time of the total information.

Our task is to filter as much information about the hidden state process ξ as we can from the observation process $\boldsymbol{W}$. More precisely, we shall evaluate for a given function $g : \mathcal{X} \to \Re$ the Wonham filter, which is the conditional expectation

$$\pi_T(g) = E\left(g(\xi_T) \,\middle|\, \mathcal{Y}_T\right)$$

with respect to the real world probability measure P. The function $g(\cdot)$ could be chosen, for instance, as an indicator function $\mathbf{1}_{\{\xi_T=a_i\}}$, which yields $\pi_T(g)$ as a probability. It could also represent a power function $(\xi_T)^q$, which leads to $\pi_T(g)$ describing the qth moment of the hidden Markov chain.

By application of the Girsanov transformation one obtains a probability measure $\dot{P}$, where

$$d\dot{P} = L_T^{-1}\, dP \tag{1.76}$$

with

$$L_T = \exp\left\{-\frac{1}{2}\int_0^T |\boldsymbol{h}(\xi_s)|^2\, ds + \int_0^T \boldsymbol{h}(\xi_s)^\top\, d\boldsymbol{W}_s\right\} \tag{1.77}$$

such that $\boldsymbol{W}$ is a Wiener process with respect to $\dot{P}$, while L is assumed to be a martingale. Here $L_T = \frac{dP}{d\dot{P}}$ is the corresponding Radon-Nikodym derivative. Note that we express in this situation the real world probability measure P in terms of the new probability measure $\dot{P}$ for which $\boldsymbol{W}$ is a standard vector Wiener process given by (1.75).

Let us introduce the unnormalized conditional probability $\sigma(\xi_t)^i$ for the state $a_i \in \mathcal{X}$ at time t by the conditional expectation

$$\sigma(\xi_t)^i = \dot{E}\left(\mathbf{1}_{\{\xi_t=a_i\}} L_t \,\middle|\, \mathcal{Y}_t\right) \tag{1.78}$$

with respect to the new probability measure $\dot{P}$ for $i \in \{1, 2, \ldots, d\}$ and $t \in [0,T]$. It follows from the Kallianpur-Striebel formula, see Fujisaki, Kallianpur & Kunita (1972), that the conditional probabilities of ξ_t given in $\mathcal{Y}_t$ are

$$P\left(\xi_t = a_i \,\middle|\, \mathcal{Y}_t\right) = E\left(\mathbf{1}_{\{\xi_t=a_i\}} \,\middle|\, \mathcal{Y}_t\right) = \frac{\sigma(\xi_t)^i}{\sum_{k=1}^d \sigma(\xi_t)^k} \tag{1.79}$$

for $a_i \in \mathcal{X}$ and $t \in [0,T]$. Here the d-dimensional process $\boldsymbol{\sigma}(\xi) = \{\boldsymbol{\sigma}(\xi_t) = (\sigma(\xi_t)^1, \ldots, \sigma(\xi_t)^d)^\top,\ t \in [0,T]\}$ of unnormalized conditional probabilities satisfies

the Zakai equation

$$\boldsymbol{\sigma}(\xi_t) = \boldsymbol{p}(0) + \int_0^t \boldsymbol{A}\,\boldsymbol{\sigma}(\xi_s)\,ds + \sum_{k=1}^{m} \int_0^t \boldsymbol{D}^k\,\boldsymbol{\sigma}(\xi_s)\,dW_s^k \tag{1.80}$$

for $t \in [0, T]$. This is a homogeneous linear Itô SDE. In (1.80) $\boldsymbol{D}^k$ is the $d \times d$ diagonal matrix with ith component $h_k(a_i)$ for $i \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$.

The least-squares estimate at time t for $g(\xi_t)$ with respect to the given observations at time t, that is with respect to the sigma-algebra $\mathcal{Y}_t$, is then the Wonham filter, which is given by the conditional expectation

$$\pi_t(g) = E\left(g(\xi_t) \,\middle|\, \mathcal{Y}_t\right) = \frac{\sum_{k=1}^{d} g(a_k)\,\sigma(\xi_t)^k}{\sum_{k=1}^{d} \sigma(\xi_t)^k} \tag{1.81}$$

for $t \in [0, T]$.

Quasi-exact Filters

Let us consider the following d-dimensional multiplicative noise SDE

$$d\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_t dt + \sum_{k=1}^{m} \boldsymbol{D}^k \boldsymbol{X}_t dW_t^k, \tag{1.82}$$

with a solution that is representing a vector geometric Brownian motion, where $\boldsymbol{X} = \{\boldsymbol{X}_t = (X_t^1, X_t^2, \ldots, X_t^d)^\top, t \in [0, \infty)\}$, $\boldsymbol{A} = [a^{i,j}]_{i,j=1}^d$ and $\boldsymbol{D}^k = [d^{k,i,j}]_{i,j=1}^d$, $k \in \{1, 2, \ldots, m\}$. Here, $W^k, k \in \{1, 2, \ldots, m\}$, are the elements of the vector SDE (1.75), that describes the observation process.

It turns out that if the matrices $\boldsymbol{A}, \boldsymbol{D}^1, \boldsymbol{D}^2, \ldots, \boldsymbol{D}^m$ are constant and commute, that is, if

$$\boldsymbol{A}\boldsymbol{D}^k = \boldsymbol{D}^k\boldsymbol{A} \quad \text{and} \quad \boldsymbol{D}^k\boldsymbol{D}^n = \boldsymbol{D}^n\boldsymbol{D}^k \tag{1.83}$$

for all $k, n \in \{1, 2, \ldots, m\}$, then an explicit solution of the SDE (1.82) can be expressed by the relation

$$\boldsymbol{X}_t = \boldsymbol{\Psi}_t \boldsymbol{X}_0, \tag{1.84}$$

for $t \in [0, \infty)$. Here, $\boldsymbol{\Psi}_t$ is the matrix exponential

$$\boldsymbol{\Psi}_t = \exp\left\{\boldsymbol{A}t - \frac{1}{2}\sum_{l=1}^{m} \left(\boldsymbol{D}^l\right)^2 t + \sum_{r=1}^{m} \boldsymbol{D}^r W_t^r\right\}, \tag{1.85}$$

for $t \in [0, \infty)$.

The proof follows from the Itô formula applied to (1.84). Note that

$$\begin{aligned} d\boldsymbol{X}_t = d\left(\boldsymbol{\Psi}_t \boldsymbol{X}_0\right) &= d\left(\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t + \sum_{r=1}^{m}\boldsymbol{D}^r W_t^r\right\}\right)\boldsymbol{X}_0 \\ &= d\left(\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\}\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\right)\boldsymbol{X}_0, \end{aligned} \tag{1.86}$$

since the matrices

$$\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t \quad \text{and} \quad \sum_{r=1}^{m}\boldsymbol{D}^r W_t^r \tag{1.87}$$

commute for all $t \in [0, \infty)$. Therefore,

$$\begin{aligned} d\boldsymbol{X}_t &= d\left(\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\}\right)\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\boldsymbol{X}_0 \\ &\quad + \exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\} d\left(\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\right)\boldsymbol{X}_0 \\ &= \left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\} \\ &\quad \times\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\boldsymbol{X}_0 dt + \end{aligned} \tag{1.88}$$

$$\begin{aligned} &+\sum_{k=1}^{m}\boldsymbol{D}^k\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\}\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\boldsymbol{X}_0 dW_t^k \\ &+\frac{1}{2}\sum_{k=1}^{m}\left(\boldsymbol{D}^k\right)^2\exp\left\{\left(\boldsymbol{A} - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\right)t\right\} \\ &\times\prod_{r=1}^{m}\exp\left\{\boldsymbol{D}^r W_t^r\right\}\boldsymbol{X}_0 d[W^k]_t. \end{aligned}$$

This equation simplifies to

$$d\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_t dt + \sum_{k=1}^{m}\boldsymbol{D}^k\boldsymbol{X}_t dW_t^k, \tag{1.89}$$

since the quadratic variation of the kth observation process equals

$$[W^k]_t = \int_0^t ds = t, \tag{1.90}$$

for $k \in \{1, 2, \ldots, m\}$.

The above derivation shows that an SDE of the type (1.80) has an explicit solution if the matrices $\boldsymbol{A}, \boldsymbol{D}^1, \ldots, \boldsymbol{D}^m$ commute. Note that $\boldsymbol{D}^1, \ldots, \boldsymbol{D}^m$ in (1.80) are diagonal matrices, and, thus, commute with each other. However, the matrix $\boldsymbol{A}$ is not commuting with the other matrices. Therefore, we do not have an exact explicit solution of the Zakai equation (1.80). Nevertheless, as will be illustrated later, if we formally take the matrix exponential (1.85) in the product (1.84), then one obtains a proxy of the solution of the corresponding Zakai equation. It turns out that this quasi-exact solution provides in many cases an excellent approximation of the exact solution, as we will confirm numerically. This is a practically valuable observation. The solution will be exploited to solve approximately and efficiently the Wonham filter problem. What, of course, needs to be done is to show for given matrices $\boldsymbol{A}, \boldsymbol{D}^1, \ldots, \boldsymbol{D}^m$ and initial vector $\boldsymbol{X}_0$ that the quasi-exact solution is close to the exact solution. This can be achieved by comparing the proposed approximation via discrete time simulation with a very accurately obtained numerical approximation using an extremely small time step size, see Kloeden & Platen (1999).

To prepare this type of comparison, let us now introduce the following equidistant time discretization $0 = \tau_0 < \tau_1 < \cdots < \tau_n = T$, such that $\tau_i = i\Delta$, for $i \in \{0, 1, \ldots, \frac{T}{\Delta}\}$. Denote by $\boldsymbol{Y}^{\Delta}_{\tau_i}$ at time τ_i the quasi-exact approximation of the solution $\boldsymbol{\sigma}(\xi_{\tau_i})$ of the Zakai equation (1.80), expressed by the recursive equation

$$\boldsymbol{Y}^{\Delta}_{\tau_{i+1}} = \exp\left\{\boldsymbol{A}\Delta - \frac{1}{2}\sum_{l=1}^{m}\left(\boldsymbol{D}^l\right)^2\Delta + \sum_{r=1}^{m}\boldsymbol{D}^r\Delta W^r_{\tau_{i+1}}\right\}\boldsymbol{Y}^{\Delta}_{\tau_i}, \tag{1.91}$$

at the equidistant discretization points, where the $\Delta W^r_{\tau_{i+1}}$ are increments of the rth element of the vector observation process $\boldsymbol{W}$. That is, we have

$$\Delta W^r_{\tau_{i+1}} = W^r_{\tau_{i+1}} - W^r_{\tau_i}, \tag{1.92}$$

for $r \in \{1, 2, \ldots, d\}$ and $i \in \{0, 1, \ldots, \frac{T}{\Delta}\}$.

We will consider later an example of the observation processes given by the SDE

$$dW^r_t = \xi_t dt + dW^{r*}_t, \tag{1.93}$$

for $t \in [0, T]$. We then add for such a scenario simulation all jump times of the continuous time Markov chain to the equidistant time discretization to obtain a jump adapted time discretization $0 = t_0 < t_1 < \cdots < t_{n_T} = T$ with maximum time

step size $\Delta > 0$, see Platen (1982). Note that n_T is now a random integer. The increments of this observation process at jump adapted discretization points can be obtained by exact simulation given the value of the hidden Markov chain ξ_{t_i} at time t_i. That is,

$$W^r_{t_{i+1}} - W^r_{t_i} = \xi_{t_i}(t_{i+1} - t_i) + \sqrt{t_{i+1} - t_i} Z^r_{i+1} \tag{1.94}$$

for $r \in \{1, 2, \ldots, d\}$ and $i \in \{0, 1, \ldots, n_T - 1\}$. Here $Z^r_{i+1} \sim \mathcal{N}(0, 1)$ is a standard Gaussian random variable, for $r \in \{1, 2, \ldots, d\}$ and $i \in \{0, 1, \ldots, n_T - 1\}$. Note that this process is a drifted Wiener process with piecewise constant random drift. The simulation of the increments of W^r is straightforward, since we have via simulation at our disposal the values of W^r at the times t_i for $i \in \{0, 1, \ldots, n_T\}$.

Approximate Filters

In practice, it is impossible to detect $\boldsymbol{W}$ continuously on $[0, T]$. One may, however, approximate increments of observations of $\boldsymbol{W}$ in integral form of the type

$$\int_{\tau_0}^{\tau_1} dW^j_s, \ldots, \int_{\tau_n}^{\tau_{n+1}} dW^j_s, \ldots, \int_{\tau_0}^{\tau_1} \int_{\tau_0}^{s_2} dW^j_{s_1} \, dW^k_{s_2}, \ldots$$

for each $j, k \in \{1, 2, \ldots, m\}$, $\tau_n = n\Delta$ and $n \in \{0, 1, \ldots, \frac{T}{\Delta}\}$. We shall see later on that with such integral observations it is possible to construct strong discrete-time approximations $\boldsymbol{Y}^\Delta$ with time step size Δ of the solution $\boldsymbol{\sigma}(\xi)$ of the Zakai equation (1.80). For the given function g this allows then to form the approximate Wonham filter

$$\pi^\Delta_t(g) = \frac{\sum_{k=1}^d g(a_k) \, Y^{\Delta,k}_t}{\sum_{k=1}^d Y^{\Delta,k}_t} \tag{1.95}$$

for $t \in [0, T]$.

We shall say that a discrete-time approximation $\boldsymbol{Y}^\Delta$ with time step size Δ converges on the time interval $[0, T]$ with strong order $\gamma > 0$ to the solution $\boldsymbol{X}$ of the corresponding SDE if there exists a finite constant K, not depending on Δ, and a $\delta_0 \in (0, 1)$ such that

$$\dot{E}\left(\left|\boldsymbol{\sigma}(\xi_{\tau_n}) - \boldsymbol{Y}^\Delta_{\tau_n}\right|\right) \leq K \, \Delta^\gamma \tag{1.96}$$

for all $\Delta \in (0, \delta_0)$ and $\tau_n \in [0, T]$. Note that the expectation in (1.96) is taken with respect to the probability measure $\dot{P}$ under which the observation process $\boldsymbol{W}$ is a Wiener process.

Analogously, we say that an approximate Markov chain filter $\pi^{\Delta}_{\tau_{n_t}}(g)$ with time step size Δ converges on the time interval $[0,T]$ with strong order $\gamma > 0$ to the optimal filter $\pi_{\tau_{n_t}}(g)$ for a given test function g if there exists a finite constant K, not depending on Δ, and a $\delta_0 \in (0,1)$ such that

$$E\left(\left|\pi_{\tau_{n_t}}(g) - \pi^{\Delta}_{\tau_{n_t}}(g)\right|\right) \leq K\,\Delta^{\gamma} \tag{1.97}$$

for all $\Delta \in (0,\delta_0)$ and $t \in [0,T]$. In contrast with (1.96), the expectation in (1.97) is taken with respect to the original probability measure P. In Kloeden, Platen & Schurz (1993) the following convergence result was derived.

Theorem 1.5.1 (Kloeden-Platen-Schurz)**.** *An approximate Markov chain filter $\pi^{\Delta}(g)$ with time step size Δ converges for $t \in [0,T]$ with strong order $\gamma > 0$ to the optimal filter $\pi(g)$ for a given bounded function g if the discrete-time approximation $\boldsymbol{Y}^{\Delta}$ used converges on $[0,T]$ to the solution $\boldsymbol{\sigma}(\xi)$ of the Zakai equation* (1.80) *with the same strong order γ.*

Now, we derive discrete-time strong approximations $\boldsymbol{Y}^{\Delta}$ that are converging with a given strong order $\gamma > 0$ to the solution $\boldsymbol{\sigma}(\xi)$ of the Zakai equation (1.80), which can be used to build a corresponding approximate filter.

Given an equidistant time discretization of the interval $[0,T]$ with step size $\Delta = \frac{T}{N}$ for some $N \in \{1,2,\ldots\}$, we define the partition sigma-algebra

$$\mathcal{P}^1_N = \sigma\{\Delta W^j_{i-1} : i \in \{1,2,\ldots,N\}, j \in \{1,2,\ldots,m\}\} \tag{1.98}$$

as the sigma-algebra generated by the increments

$$\Delta W^j_0 = \int_0^{\Delta} dW^j_s, \ldots, \Delta W^j_{N-1} = \int_{(N-1)\Delta}^{N\Delta} dW^j_s \tag{1.99}$$

for all $j \in \{1,2,\ldots,m\}$. Thus, $\mathcal{P}^1_N$ contains the information about the increments of $\boldsymbol{W}$ for the given time discretization.

The simplest discrete-time approximation is obtained from the Euler scheme. It has for the Zakai equation (1.80) the form

$$\boldsymbol{Y}^{\Delta}_{\tau_{n+1}} = [\boldsymbol{I} + \boldsymbol{A}\,\Delta + \boldsymbol{G}_n]\,\boldsymbol{Y}^{\Delta}_{\tau_n} \tag{1.100}$$

with

$$\boldsymbol{G}_n = \sum_{k=1}^{m} \boldsymbol{D}^k\,\Delta\,W^k_n \tag{1.101}$$

and initial value $Y_0 = \sigma(\xi_0)$, where $\boldsymbol{I}$ is the $d \times d$ unit matrix. The scheme (1.100) converges with strong order $\gamma = 0.5$ under the given assumptions. For a general SDE this is the maximum order of strong convergence that can be achieved under the partition sigma-algebra $\mathcal{P}_N^1$, as was shown in Clark & Cameron (1980). However, some commutativity property of the Zakai equation (1.80) follows from the diagonal structure of its volatility matrices, see Kloeden & Platen (1999). This allows the strong order $\gamma = 1.0$ to be attained with the information given by $\mathcal{P}_N^1$.

The Milstein scheme, which is of strong order $\gamma = 1.0$ has for the Zakai equation (1.80) the form

$$\boldsymbol{Y}^{\Delta}_{\tau_{n+1}} = \left(\boldsymbol{I} + \boldsymbol{B}\,\Delta + \boldsymbol{G}_n \left(\boldsymbol{I} + \frac{1}{2}\,\boldsymbol{D}_n \right) \right) \boldsymbol{Y}^{\Delta}_{\tau_n}, \tag{1.102}$$

where

$$\boldsymbol{B} = \boldsymbol{A} - \frac{1}{2} \sum_{k=1}^{m} (\boldsymbol{D}^k)^2. \tag{1.103}$$

Newton (1986) searched for a scheme which is asymptotically the "best" in the class of strong order $\gamma = 1.0$ schemes in the sense that it has the smallest leading error coefficient in an error estimate similar to (1.96), see Kloeden & Platen (1999). In practice, the numerical stability of discrete-time approximations for filters is highly important. Fischer & Platen (1999) studied the application of the balanced implicit method in hidden Markov chain filtering when the time between observations is rather large. They obtained quite reliable results for the balanced implicit scheme even for large step sizes. In the following we consider an example of a Wonham filter considered also by Fischer & Platen (1999). We show that our quasi-exact approximation of the Zakai equation performs remarkably well, even though the matrix $\boldsymbol{A}$ does not fully commute with the matrices $\boldsymbol{D}^1, \ldots, \boldsymbol{D}^m$ in the Zakai equation.

Wonham Filter Example

The unobservable signal process $\xi = \{\xi_t,\, t \in [0, T]\}$ in our example is a time homogeneous, continuous time, real valued Markov chain with the set $\mathcal{X} = \{a_1, a_2, \ldots, a_d\}$ of states. The scalar observation process $W = \{W_t,\, t \in [0, T]\}$ is given by the relation

$$W_t = \int_0^t h(\xi_s) ds + W_t^*, \tag{1.104}$$

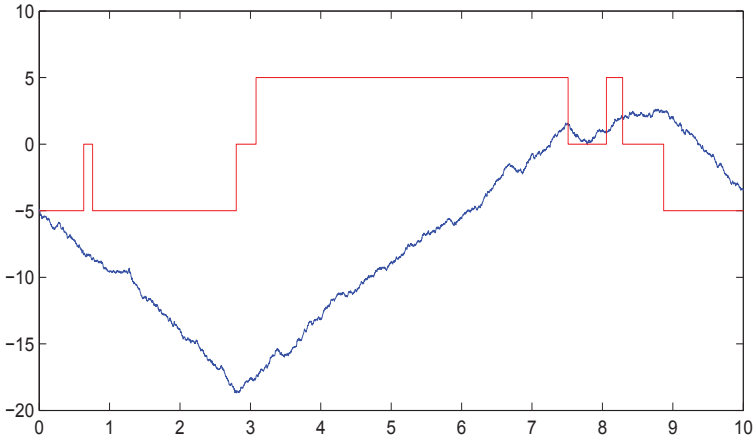

Figure 1.9: Simulation of the signal and observation processes for $\Delta = \frac{1}{500}$

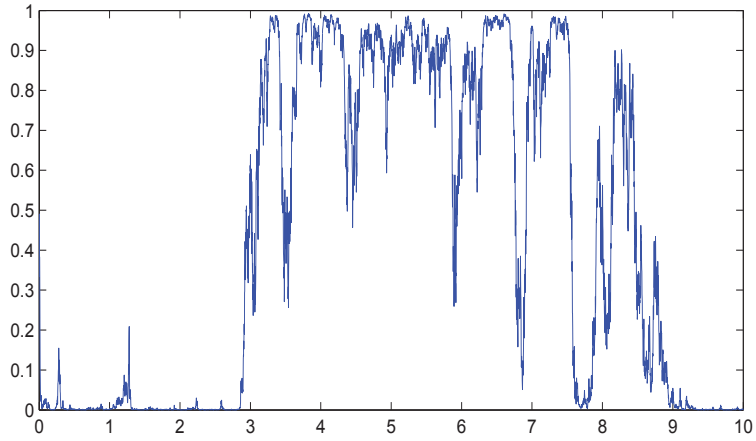

Figure 1.10: q_t^1 - obtained by the quasi-exact approximation for $\Delta = \frac{1}{500}$

for $t \in [0, T]$. One can say that W represents the time integral over the signal process that is corrupted by a Wiener process W^*. We estimate the hidden state of the Markov chain ξ by observing only the values of the process W.

It is convenient to consider a d state continuous time Markov chain $\boldsymbol{X} = \{\boldsymbol{X}_t,\, t \in [0, T]\}$ that is identical to ξ under a transformation of the state space, see Elliott, Aggoun & Moore (1995). We choose as the state space for $\boldsymbol{X}$ the set $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$ of unit vectors in $\Re^d$, with $\boldsymbol{e}_1 = (1, 0, 0, \ldots, 0)^\top$, $\boldsymbol{e}_2 = (0, 1, 0, \ldots, 0)^\top$ and so on. Then we write,

$$\xi_t = \boldsymbol{X}_t^\top \boldsymbol{a} \tag{1.105}$$

with $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top \in \Re^d$. Let $\boldsymbol{A} = [a^{i,j}]_{i,j=1}^d$ be the constant intensity matrix associated with the homogeneous, continuous time Markov chain $\boldsymbol{X}$, so that

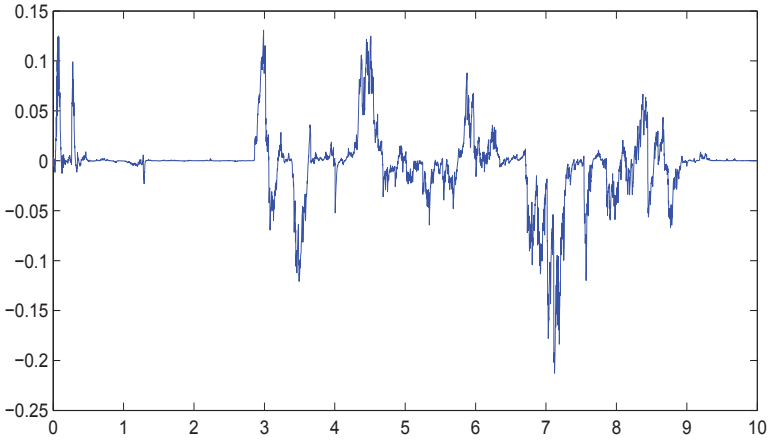

Figure 1.11: Difference between q_t^1 obtained by the quasi-exact approximation and the Euler scheme for $\Delta = \frac{1}{500}$

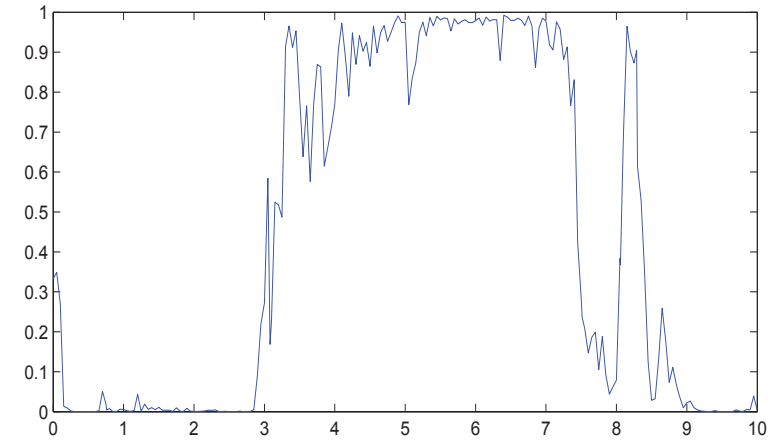

Figure 1.12: q_t^1 - obtained by the quasi-exact method with $\Delta = \frac{1}{20}$

$\boldsymbol{p}(t) = E(\boldsymbol{X}_t)$ satisfies the vector ODE

$$\frac{d\boldsymbol{p}(t)}{dt} = \boldsymbol{A}\,\boldsymbol{p}(t), \tag{1.106}$$

for $t \in [0, T]$, with given initial probability vector $\boldsymbol{p}(0)$. $\boldsymbol{H}$ is the diagonal $d \times d$ matrix that has the elements of the vector $\boldsymbol{a}$ as diagonal elements and is zero elsewhere.

Denote by $\mathcal{Y}_t$ the observation sigma-algebra generated by W up to time t. The Wonham filter for $\boldsymbol{X}$ at time t is then given as

$$\hat{\boldsymbol{X}}_t = E(\boldsymbol{X}_t \,|\, \mathcal{Y}_t).$$

As we have seen previously, the theoretical solution to the problem of calculating $\hat{\boldsymbol{X}}_t$ involves the unnormalized filter for the conditional distribution of $\boldsymbol{X}$, which is

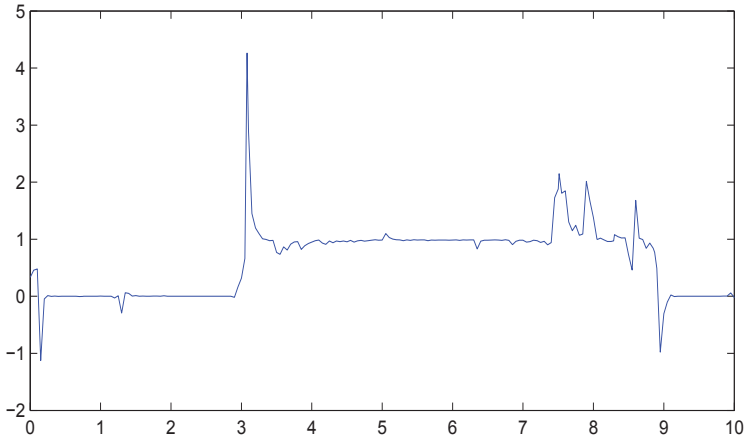

Figure 1.13: q_t^1 - obtained by the Euler method with $\Delta = \frac{1}{20}$

denoted by $\sigma(\boldsymbol{X}) = \{\sigma(\boldsymbol{X}_t),\, t \in [0,T]\}$ and satisfies the Zakai equation

$$d\sigma(\boldsymbol{X}_t) = \boldsymbol{A}\,\sigma(\boldsymbol{X}_t)\,dt + \boldsymbol{D}\,\sigma(\boldsymbol{X}_t)\,dW_t \tag{1.107}$$

for $t \in [0,T]$. The Wonham filter for $\boldsymbol{X}$ is then computed as

$$\hat{\boldsymbol{X}}_t = E(\boldsymbol{X}_t \,|\, \mathcal{Y}_t) = \frac{\sigma(\boldsymbol{X}_t)}{\sigma(\boldsymbol{X}_t)^\top \mathbf{1}} \tag{1.108}$$

for $t \in [0,T]$, see (1.81).

Let us now assume that our observation process W is a one-dimensional stochastic process of the form

$$dW_t = \xi_t dt + dW_t^*, \tag{1.109}$$

for $t \in [0,T]$. The exact simulation of this process was already discussed above.

Moreover, the parameters of the Markov chain X, or equivalently ξ, are chosen to give a realistic multiplicative noise term in the Zakai equation (1.107). The simulated hidden Markov chain X is chosen to have three states, with the vector $\boldsymbol{a}$ taken to be $\boldsymbol{a} = (5,0,-5)^\top$. Numerical experiments have shown that using more states does not change the nature of the results that we obtain. The intensity matrix $\boldsymbol{A}$ of the hidden Markov chain is chosen to be of the simple form

$$\boldsymbol{A} = \begin{bmatrix} -1.0 & 1.0 & 0 \\ 0.5 & -1.0 & 0.5 \\ 0 & 1.0 & -1.0 \end{bmatrix}. \tag{1.110}$$

This describes how the Markov chain jumps with prescribed intensities to neighboring states. For instance, the intensity to jump from level 0 to level 5 is 0.5 per unit of time.

Now, let us investigate the approximate calculation of the Wonham filter. We simulate the scenario for the signal and observation processes over the time interval $[0, T]$ with $T = 10$ using the exact simulation method (1.94). The simulated output can be seen in Fig. 1.9.

Let us now consider the probability

$$q_t^i = \hat{\boldsymbol{X}}_t^\top \boldsymbol{e}_i = E\left(\boldsymbol{X}_t^\top \boldsymbol{e}_i \,\middle|\, \mathcal{Y}_t\right) \tag{1.111}$$

which, for $i \in \{1, 2, 3\}$, denotes the filtered probability that the hidden Markov chain $\boldsymbol{X}$ is in the ith state. For illustration let us focus on q_t^1, which corresponds to the level $\xi = 5$, that is $i = 1$.

To obtain the quantity q_t^1 we have to solve the SDE (1.107). In order to do so we use our quasi-exact approximation described in this section. Applying it to the equation (1.107) and denoting by $\boldsymbol{Y}_{\tau_i}^\Delta$ the approximate solution of the Zakai equation at time τ_i, we obtain the following quasi-exact approximation

$$\boldsymbol{Y}_{\tau_{i+1}}^\Delta = \exp\left\{\boldsymbol{A}\Delta - \frac{1}{2}\boldsymbol{D}\Delta + \boldsymbol{D}\Delta W_{\tau_{i+1}}\right\}\boldsymbol{Y}_{\tau_i}^\Delta, \tag{1.112}$$

where $\Delta W_{\tau_{i+1}}$ is obtained using (1.94). Additionally, $\boldsymbol{D}$ is a diagonal matrix with elements

$$\boldsymbol{D} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -5 \end{bmatrix}. \tag{1.113}$$

In order to calculate the matrix exponential in (1.112) we use an implementation in Matlab. If a given matrix has a full set of eigenvectors $\boldsymbol{V}$ with corresponding eigenvalues $\boldsymbol{E}$, then the matrix exponential of this matrix is

$$\boldsymbol{V}\exp\{\mathrm{diag}(\boldsymbol{E})\}\boldsymbol{V}^{-1}. \tag{1.114}$$

Note also that the matrix exponential of the diagonal matrix is a diagonal matrix, whose diagonal elements are exponents of diagonal elements of the underlying matrix. If this calculation is not possible to apply, Matlab's implementation uses the Padé approximation with scaling and squaring, see Higham (2005).

Additionally, we compare our quasi-exact approximation to the results obtained via two numerical schemes: the Euler scheme and the Milstein scheme, as described previously. When the step size Δ was chosen extremely small with about $\frac{1}{500}$, then all three approximations produced virtually identical results. Fig. 1.10 displays a

plot of q_t^1 as calculated by the quasi-exact approximation with $\Delta = \frac{1}{500}$. This is compared to the result obtained by the Euler scheme for the same time step size in Fig. 1.11. In this figure we show the difference between q_t^1 obtained by the quasi-exact approximation and the Euler scheme for $\Delta = \frac{1}{500}$. We remark that small errors occur, most likely, due to the approximate nature of the Euler scheme as well as the Padé approximation of the matrix exponential used in the quasi-exact method. When using the Milstein scheme the results were practically the same. Differences between the schemes become apparent when the time step size Δ is chosen to be larger.

We emphasize that such a fine time discretization for the observation process as employed in our test, however, is often not available in real world filtering problems. Using the same realizations of the observation and signal processes that were given in Fig. 1.9, we display in Figs. 1.12 and 1.13 the plots of the filtered probability q_t^1, see (1.111), calculated by the proposed quasi-exact approximation and the Euler method when using the larger step size $\Delta = \frac{1}{20}$. For this step size we can see that the only acceptable approximation appears to be the proposed quasi-exact approximation of the hidden Markov chain filter. In the other case we even obtain negative "probabilities" and other unrealistic estimates as filter values. This undesirable effect is due to numerical instabilities of the Euler scheme. In the given example also the Milstein scheme does not yield a useful filter.

We have proposed a quasi-exact approximation of hidden Markov chain filters. It turns out that, even though the drift matrix in the Zakai equation does not perfectly commute, the quasi-exact method is rather successful in approximating hidden Markov chain filters also for large observation time steps.

The results of our experiments demonstrate that discrete time approximations, such as the Euler scheme, cannot be used as reliably as the proposed quasi-exact method. When the available observations are rare the quasi-exact method can provide useful approximations where other standard methods fail.

Chapter 2

Empirical Study of a World Stock Index in Different Currency Denominations

Before modeling theoretically the dynamics of indices it is important to develop a reasonable view on their real world dynamics. Therefore, the aim of this chapter is to document some empirical facts related to log-returns of diversified world stock indices when these are denominated in different currencies. The chapter estimates univariate distributions of log-returns for a range of world stock indices over long observation periods. The Student-t distribution with about four degrees of freedom is identified as the typical estimated log-return distribution of such indices. Owing to the observed high levels of significance, this result can be interpreted as a stylized empirical fact. The Student-t distribution is also found as the bivariate joint distribution, which suitably fits daily log-returns of a world stock index when denominated in various currencies. In the class of bivariate symmetric generalized hyperbolic distributions the maximum likelihood estimates for the tail parameters cluster in the neighborhood of those values that characterize the Student-t distribution. This empirical feature is very strong for relatively separated economies and is less pronounced for pairs of currencies involving one leading currency, such as the US dollar, or a neighboring currency such as for pairs of Scandinavian currencies. If one takes into account that the central banks of many countries influence through deliberate trading the exchange rates with respect to the US dollar or the currencies of their main trading partners, then these observed effects may be explained. Motivated by these findings, this chapter will also model the dependency in log-returns of denominations of the world stock index using time-varying copula, aiming to iden-

tify the best fitting copula family. We find that the Student-t copula is generally superior to e.g. the Gaussian copula, where the dependence structure relates to the multivariate normal distribution. We show that merely changing the distributional assumption for the marginals from normal to Student-t leads to significantly better performance. Furthermore, the Student-t copula with Student-t marginals is able to better capture dependent extreme values, which can be observed in index log-returns of different currency denominations. This chapter follows and extends the results in the following two papers: Platen & Rendek (2008) and Ignatieva, Platen & Rendek (2011).

2.1 Index Construction

This section focuses on a methodology for the construction of diversified world stock indices. Such indices are usually formed for asset allocation purposes or in order to measure the general market performance and general market risk, see Basle (1996). Market capitalization weighted indices are widely used as benchmarks in investment management. Some of the following indices also have a more theoretical motivation under the benchmark approach; see Platen & Heath (2010) and Platen & Rendek (2012b). We construct in this chapter portfolios that are all strictly positive, self-financing portfolios.

The data selected for the $d \in \mathcal{N} = \{1, 2, \dots\}$ constituents of the indices consist of daily data for the period from 1973 to 2006. We construct world stock indices from regional stock market indices and from world sector market indices. The regional stock market indices represent market capitalization weighted stock indices as constructed and provided by Thomson Reuters Datastream (Datastream). World sector indices are also constructed and provided by Datastream, and reflect the worldwide evolution of respective industries. The regional or the world sector indices are used in our study as constituents of the newly constructed indices.

Portfolio Generating Functions

We emphasize in this study four main types of indices, market capitalization weighted indices (MCIs), diversity weighted indices (DWIs) as described by Fernholz (2002), equally weighted indices (EWIs), and some type of world stock indices (WSIs) in-

troduced in Le & Platen (2006).

We assume that all the constituents of our constructed indices are capable of unbounded positive jumps and negative jumps leaving constituents arbitrarily close to zero. This reflects the fact that, in principle, any of the constituents S_t^j, $j \in \{1, 2, \ldots, d\}$, can almost default at any time. Therefore, since we consider only strictly positive portfolios in our study, their fractions of wealth always need to remain nonnegative.

In order to eliminate short sales, and to have a systematic way of generating nonnegative fractions using a wide range of methods; we introduce the notion of a portfolio generating function (PGF), which has been inspired by a similar construct described in Fernholz (2002). More precisely, a PGF $A : \Re^d \to [0, 1]^d$ maps a given vector of fractions $\boldsymbol{\pi}_{\delta,t} = (\pi_{\delta,t}^1, \pi_{\delta,t}^2, \ldots, \pi_{\delta,t}^d)^\top \in \Re^d$, into a vector of nonnegative fractions

$$\tilde{\boldsymbol{\pi}}_{\delta,t} = (\tilde{\pi}_{\delta,t}^1, \tilde{\pi}_{\delta,t}^2, \ldots, \tilde{\pi}_{\delta,t}^d)^\top = A(\boldsymbol{\pi}_{\delta,t}) \in [0, 1]^d \tag{2.1}$$

such that $\sum_{j=1}^d \tilde{\pi}_{\delta,t}^j = 1$ for all $t \in \Re^+$. Note that the given vector of fractions $\boldsymbol{\pi}_{\delta,t}$ may contain negative components. These components can be obtained by any kind of method, including the use of statistical estimates of optimal fractions. Estimates provided by experts or economically based theoretical predictions can also be used. These fractions are then translated into nonnegative fractions by a PGF. Note that we do not include the savings account in a PGF. This is also typical for most commercial indices.

Market Capitalization Weighted Indices

For an MCI we define the fraction of wealth held in the j-th constituent at time t as follows:

$$\pi_{\delta_{MCI},t}^j = \frac{\delta_t^j S_t^j}{\sum_{i=1}^d \delta_t^i S_t^i}, \tag{2.2}$$

$j \in \{1, 2, \ldots, d\}$. Here δ_t^j is the number of units of the jth constituent of the portfolio S_t^δ at time t, which is typically held constant over certain periods of time.

Diversity Weighted Indices

The so called diversity weighted indices (DWIs) are theoretically and practically interesting indices, which were proposed in Fernholz (2002). Here the PGF is a

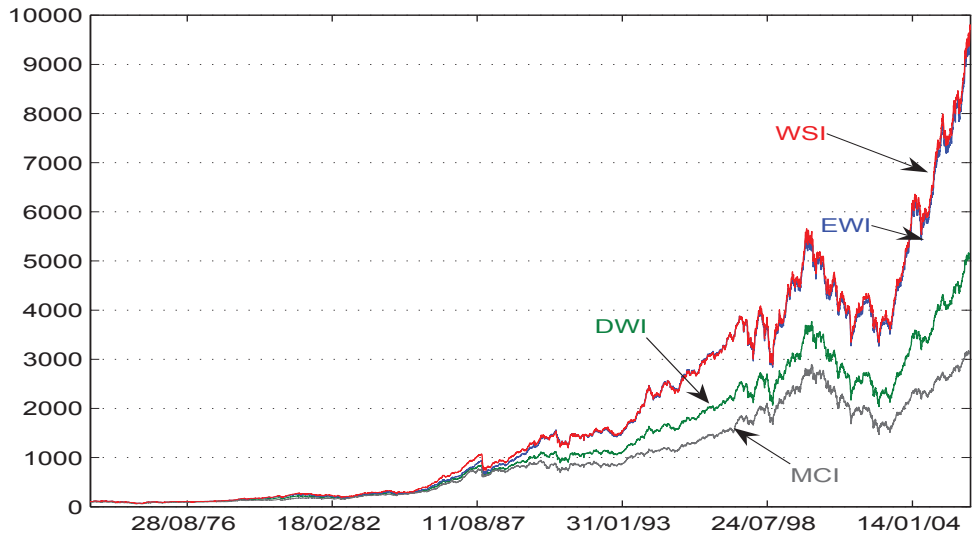

Figure 2.1: Indices constructed from regional stock market indices

function of the MCI fractions $\boldsymbol{\pi}_{\delta_{MCI},t}$, given in (2.2), and has the form

$$\tilde{\pi}^j_{\delta,t} = \frac{(\pi^j_{\delta_{MCI},t})^p}{\sum_{l=1}^d (\pi^l_{\delta_{MCI},t})^p} \tag{2.3}$$

for some choice of a real number $p \in [0,1]$, $j \in \{1,2,\ldots,d\}$, $t \in \Re^+$. The DWI has been designed to outperform the market portfolio, that is the MCI, see Fernholz (2002).

Equally Weighted Indices

An almost ideally diversified index is obtained by setting all fractions equal. The jth fraction of the equally weighted index or equi-weighted index (EWI) is then simply given by the constant

$$\pi^j_{\delta_{EWI},t} = \frac{1}{d} \tag{2.4}$$

for all $j \in \{1,2,\ldots,d\}$, where d is the number of constituents. The main advantage of this index is that it forms, in some sence, the "best" diversified portfolio and does not need the calculation of its fractions from data or other sources. We will show in this chapter that EWIs do not only exhibit excellent long term performance but also have very clear distributional properties. These distributional features, as well as their excellent long term growth rate, make EWIs important tools for theoretical investigations and practical applications. We will see in Chapter 3 that this type of

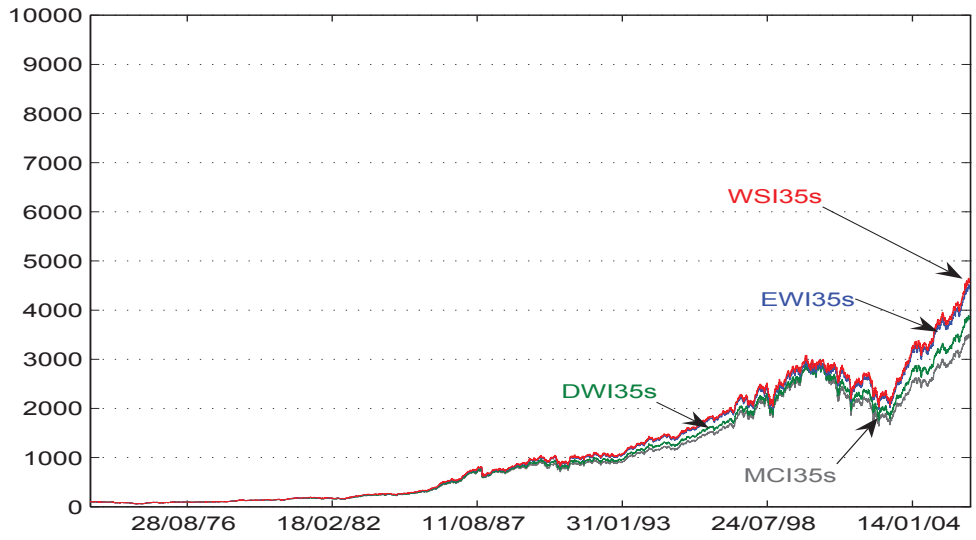

Figure 2.2: Indices constructed from sector indices based on 35 industries

portfolios approximates asymptotically the, so called, numéraire portfolio, see Long (1990).

A Family of World Stock Indices

There have been many attempts in the literature and in practice to construct investment portfolios with outstanding performance, e.g. expected growth rate or expected returns. It is evident from estimation theory, see for instance DeMiguel, Garlappi & Uppal (2009), that, in principle, hundreds of years of data are necessary to estimate risk premia with any reasonable level of significance. Such long data sets are not available. Additionally, the risk premia cannot be expected to remain constant over sufficient long periods of time. This makes it very difficult to use any statistical method successfully in investment portfolio construction. Despite the mentioned empirical difficulties, various stock indices were studied e.g. by Le & Platen (2006) with the aim of approximating the growth optimal portfolio (GOP), which is also the numéraire portfolio (NP), by using standard statistical estimates of risk premia and volatilities. No significant advantage in performance was reported for these constructed indices, as will be confirmed in this chapter.

We will study world stock indices (WSIs) as special cases in a family of indices which also include the MCI, DWI and EWI. The PGF used for the construction of

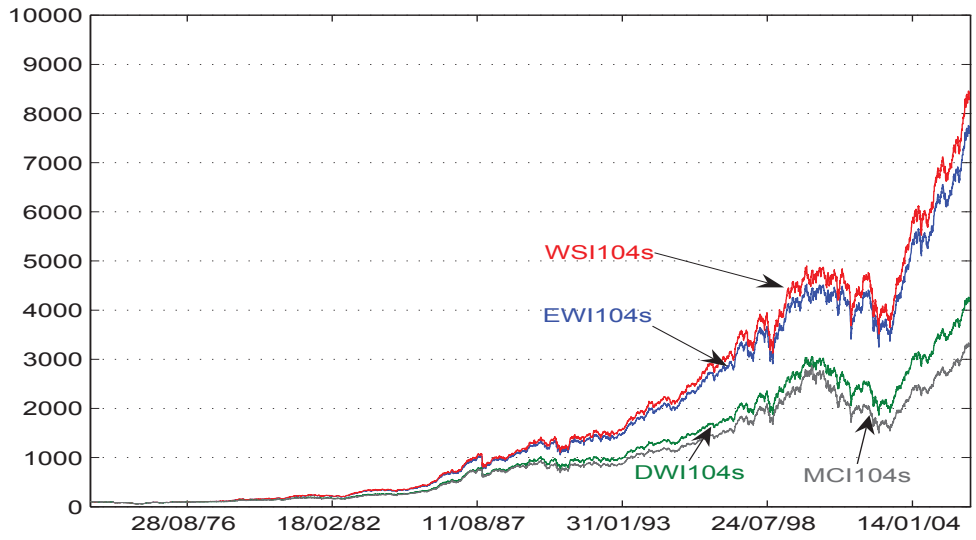

Figure 2.3: Indices constructed from sector indices based on 104 industries

this general family of indices is given by

$$\tilde{\pi}^j_{\delta,t} = \frac{(\pi^j_{\delta,t} + \mu_t)^p}{\sum_{l=1}^d (\pi^l_{\delta,t} + \mu_t)^p}, \tag{2.5}$$

for all $j \in \{1, 2, \ldots, d\}$ and $t \in \Re^+$, where $p \in [0, 1]$ is some real number. This construction is slightly more general than what has been suggested in Fernholz (2002); see (2.3). Essentially, the above PGF keeps the ranking of the fractions $\pi^j_{\delta,t}$ intact and transforms the original fractions into other positive fractions. We obtain the fractions of a DWI if $\pi^j_{\delta,t} = \pi^j_{MCI,t}$ and $\mu_t = 0$ for all $t \in \Re^+$. The GOP is the portfolio that maximizes expected logarithmic utility, see Kelly (1956). Theoretically one obtains the fractions of the GOP of the stock market in the form

$$\boldsymbol{\pi}_{\delta_*,t} = \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{a}_t - r_t \mathbf{1}) \tag{2.6}$$

for $t \in \Re^+$, see Merton (1973), Platen & Heath (2010) or Filipović & Platen (2009). Here $\boldsymbol{\Sigma}_t$ denotes the covariance matrix of returns and $\boldsymbol{a}_t$ the vector of expected returns of the constituents, while r_t is the short rate.

It is very common to estimate the covariance matrix $\boldsymbol{\Sigma}_t$ from the observation of daily returns, say, of the most recent one year period. Despite our reservations about the low significance of any estimates for the expected return vector $\boldsymbol{a}_t$, we may nevertheless try to estimate it in a standard manner from daily returns over the

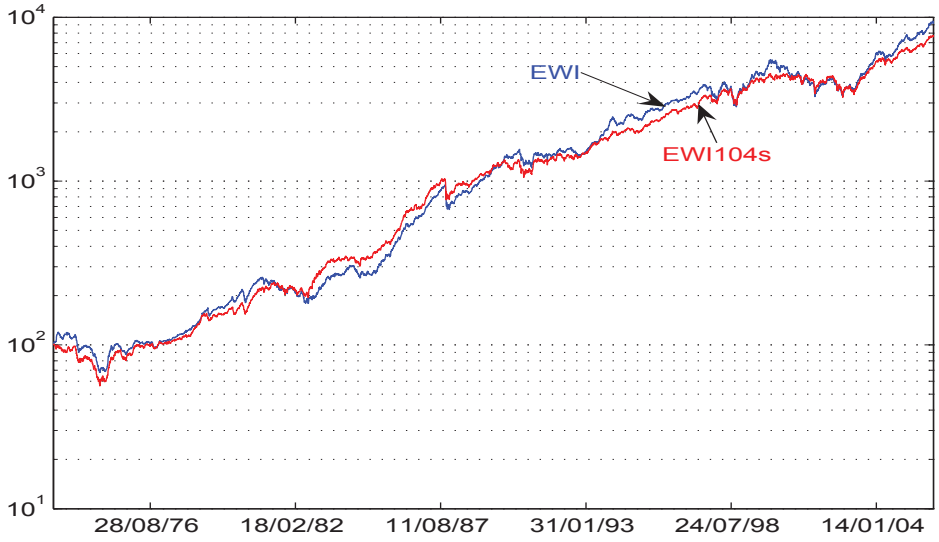

Figure 2.4: The regional EWI and sector EWI104s indices in log-scale

same one year period. This is what, in principle, is often performed in active fund management when following the sample based Markowitz approach. The resulting estimated suggested fractions of the GOP vary in an extreme manner and can be largely negative. To make the fractions positive via a PGF we set

$$\mu_t = \left| \inf_j \pi^j_{\delta,t} \right| + \mu, \tag{2.7}$$

for some choice $\mu \geq 0$, $t \in \Re^+$. The WSI used in this thesis was obtained as in Le & Platen (2006).

Comparison of Constructed Indices

Figs. 2.1-2.3 display indices constructed along the lines as described above. In Fig. 2.1 they are constructed from 38 regional stock market indices. In Fig. 2.2 they use 35 industry sector indices as constituents. Finally, in Fig. 2.3 they employ 104 industry sector indices. For all three types of constructed indices, MCIs, DWIs, EWIs and WSIs, we observe that MCIs always perform worse than DWIs, which again perform worse than EWIs and WSIs. This is a common feature of both regional and industry sector based indices. In particular, the EWI and the EWI104s perform extremely well. It can be noticed that the performance of well diversified indices is generally much better than that of less diversified indices. The diversification theorems in Platen (2005a) and Platen & Rendek (2012b), see Chapter 3, provide an explanation

for this phenomenon by stating that well diversified portfolios are likely to be better proxies for the GOP, that is, the NP. Since the GOP is known almost surely to have the best long term growth rates, see Platen & Heath (2010), this appears to be supported by the displayed constructed indices.

Furthermore, we note that the EWI and the WSI are almost identical when setting $p = 0.5$. This is due to the fact that the resulting WSI fractions are very close to those of the EWI. On the other hand, the fractions of the WSI are still flexible and make it possible for the WSI to outperform the EWI slightly. This, however, comes with a significant computational cost and a decreasing level of diversification. The real question is whether the suggested fractions for the GOP have some statistical information that is readily exploitable. Unless this is the case, it is unlikely that a WSI outperforms an EWI significantly in the long run.

In Fig. 2.4, we now plot on a logarithmic scale the best performing indices, which are the EWI, based on 38 regional stock market indices, and the EWI104s, with 104 world industry sectors as constituents. This graph shows that historically sometimes the EWI and on other occasions the EWI104s performs slightly better. They are both well diversified indices and, therefore by the above mentioned diversification theorems, can be expected to be good proxies of the GOP. This raises the question as to whether there are clear empirical features that make one of these indices better suitable as proxy of the NP. We clarify this question by studying the distribution of their log-returns.

2.2 A Class of Multivariate Log-return Distributions

This section introduces a rich class of multivariate symmetric generalized hyperbolic (SGH) densities, which covers a wide range of possible tail shapes. Several authors have proposed widely researched models that generate log-returns within the class of SGH densities, originally introduced in Barndorff-Nielsen (1978). The SGH density results as a normal-mixture density when the mixing density of the variance is a generalized inverse Gaussian (GIG) density.

More precisely, the random vector $\boldsymbol{X}$ is said to have a d-dimensional normal

variance mixture distribution if

$$\boldsymbol{X} \sim \sqrt{W}\boldsymbol{A}\boldsymbol{Z}, \tag{2.8}$$

where the k-dimensional vector $\boldsymbol{Z} \sim \mathcal{N}_k(\boldsymbol{0}, \boldsymbol{I}_k)$ has a k-variate standard normal distribution. $W \geq 0$ is a non-negative, scalar valued random variable, independent of $\boldsymbol{Z}$; and $\boldsymbol{A} \in \Re^{d\times k}$ is a matrix. It follows that $E(\boldsymbol{X}) = \boldsymbol{0}$ and $Cov(\boldsymbol{X}) = E(W)\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^\top$ when the mixing variable W has finite variance. Conditioned on the random covariance matrix $W\boldsymbol{\Sigma}$, $\boldsymbol{X}$ can be interpreted as a d-dimensional conditionally Gaussian random vector.

The GIG distribution used for the mixing variance $W \sim GIG(\lambda, \chi, \psi)$ is a three parameter distribution. The density of W equals

$$f(x) = \frac{\chi^{-\lambda}\left(\sqrt{\chi\psi}\right)^{\lambda}}{2K_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\chi x^{-1} + \psi x\right)\right\} \tag{2.9}$$

for $x > 0$, where $K_\lambda(\cdot)$ denotes a modified Bessel function of the third kind with index λ, see Abramowitz & Stegun (1972). Here, the parameters satisfy $\chi > 0$, $\psi \geq 0$ if $\lambda < 0$; $\chi > 0, \psi > 0$ if $\lambda = 0$ and $\chi \geq 0, \psi > 0$ if $\lambda > 0$.

Consequently, the multivariate SGH density has the following representation

$$f(\boldsymbol{x}) = c\, \frac{K_{\lambda-(d/2)}\left(\sqrt{(\chi + \boldsymbol{x}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x})\psi}\right)}{\left(\sqrt{(\chi + \boldsymbol{x}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x})\psi}\right)^{(d/2)-\lambda}}, \tag{2.10}$$

where the normalization constant is

$$c = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda}\psi^{d/2}}{(2\pi)^{d/2}\det(\boldsymbol{\Sigma})^{1/2}K_\lambda(\sqrt{\chi\psi})}. \tag{2.11}$$

To simplify our analysis we consider the bivariate SGH density with zero mean and variance equal to one as model for the log-returns. These properties arise from centralizing and scaling the log-return data for each currency. To express the dependence structure of pairs of log-returns we introduce the matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}, \tag{2.12}$$

which can be interpreted as the correlation matrix of a pair of conditionally Gaussian random variables, conditioned on some independent generalized inverse Gaussian random variance. Here we use in the conditionally Gaussian interpretation the

parameter $\varrho \in (-1,1)$ to measure the dependence. However, we need to emphasize that in our general context ϱ is no longer the correlation. It measures, more generally, some dependence and will be called dependence parameter.

It turns out that the following $(\bar{\alpha}, \lambda)$-parametrization, which focuses on the two tail parameters of the SGH distribution, is rather convenient and illustrative. Here we set $\alpha = \sqrt{\psi}$, $\delta = \sqrt{\chi}$ and $\bar{\alpha} = \alpha\delta$. In this parametrization the bivariate SGH density function of the vector of log-returns $\boldsymbol{X} = (X_1, X_2)^\top$ is of the form

$$\begin{aligned} f(x_1,x_2) &= \frac{1}{\delta K_\lambda(\alpha\delta)\sqrt{1-\varrho^2}} \frac{\alpha}{2\pi} \left(1 + \frac{x_1^2 - 2\varrho x_1 x_2 + x_2^2}{(1-\varrho^2)\delta^2}\right)^{\frac{1}{2}(\lambda-1)} \\ &\quad \times K_{\lambda-1}\left(\alpha\delta\sqrt{1 + \frac{x_1^2 - 2\varrho x_1 x_2 + x_2^2}{(1-\varrho^2)\delta^2}}\right), \end{aligned} \tag{2.13}$$

where $\lambda \in \Re$, $\alpha, \delta \geq 0$ and $\varrho \in (-1,1)$ is a dependence parameter, see McNeil, Frey & Embrechts (2005).. We set $\alpha > 0$ if $\lambda > 0$ and $\delta > 0$ if $\lambda \leq 0$. $K_\lambda(\cdot)$ is the modified Bessel function of the third kind with index λ. Here, the variances of X_1 and X_2 equal one, that is,

$$1 = \begin{cases} \frac{\delta^2}{-2(\lambda+1)} & \text{if} \quad \alpha = 0 \quad \text{for} \quad \lambda < 0, \\ \frac{2\lambda}{\alpha^2} & \text{if} \quad \delta = 0 \quad \text{for} \quad \lambda > 0, \\ \frac{\delta^2 K_{\lambda+1}(\bar{\alpha})}{\bar{\alpha} K_\lambda(\bar{\alpha})} & \text{otherwise.} \end{cases} \tag{2.14}$$

The above bivariate SGH density is, therefore, a three parameter density. The two shape parameters for the tails were chosen as λ and $\bar{\alpha} = \alpha\delta$. The third parameter $\varrho \in (-1,1)$ determines the degree of dependence between the log-returns.

Within the class of bivariate SGH densities we will describe in the following four special cases of SGH densities that are related to the log-return densities of important asset price models:

In Praetz (1972) and Blattberg & Gonedes (1974), models were proposed with log-returns following a Student-t density with degrees of freedom $\nu > 2$. This kind of log-return density can also be estimated for the minimal market model, proposed in Platen (2001). The Student-t density follows from the above SGH density for the parameters $\lambda = -\frac{1}{2}\nu < 0$, $\alpha = 0$ and $\delta = \sqrt{\nu - 2}$, that is, for the shape parameters $\lambda = -\frac{1}{2}\nu$ and $\bar{\alpha} = 0$. Using these parameter values the Student-t density function

for the log-return vector $\boldsymbol{X} = (X_1, X_2)^\top$ has then the form

$$f(x_1, x_2) = \frac{\Gamma(\frac{1}{2}\nu + 1)}{\sqrt{(\nu-2)(1-\varrho^2)}\pi\Gamma(\frac{1}{2}\nu)} \left(1 + \frac{x_1^2 - 2\varrho x_1 x_2 + x_2^2}{(1-\varrho^2)(\nu-2)}\right)^{-\frac{1}{2}\nu - 1}, \qquad (2.15)$$

where $\Gamma(\cdot)$ is the gamma function.

Equation (2.15) expresses the bivariate probability density of a Student-t distributed random variable with ν degrees of freedom. The above bivariate Student-t density is a two parameter density. The dependence between the two log-returns is governed by the dependence parameter $\varrho \in (-1, 1)$. Note that ϱ is here not simply the correlation of the random variables X_1 and X_2 because these are not Gaussian. The degree of freedom $\nu = -2\lambda$ is the shape parameter, with smaller ν indicating larger tail heaviness for the Student-t density. Furthermore, when the degree of freedom increases, that is $\nu \to \infty$, then the marginal Student-t density asymptotically approaches the standard Gaussian density.

In Madan & Seneta (1990) and Geman, Madan & Yor (2001) it has been proposed that log-returns may be distributed according to a normal-variance gamma mixture distribution. This case of the variance gamma (VG) model is obtained when the parameters are set to $\lambda > 0$, $\delta = 0$ and $\alpha = \sqrt{2\lambda}$. This means that one has the shape parameters $\lambda > 0$ and $\bar{\alpha} = 0$. Using these parameter values the bivariate VG density function for the log-return vector $\boldsymbol{X} = (X_1, X_2)^\top$ has then the form

$$\begin{aligned} f(x_1, x_2) &= \frac{\lambda}{\pi 2^{\lambda-1}\Gamma(\lambda)} \left(\frac{2\lambda\,(x_1^2 + 2x_1x_2\varrho + x_2^2)}{1-\varrho^2}\right)^{\frac{1}{2}(\lambda-1)} \\ &\quad \times K_{\lambda-1}\left(\sqrt{\frac{2\lambda\,(x_1^2 - 2x_1x_2\varrho + x_2^2)}{1-\varrho^2}}\right), \end{aligned} \qquad (2.16)$$

with

$$f(0,0) = \frac{\lambda}{\pi 2^{\lambda-1}\Gamma(\lambda)} 2^{\lambda-2}\Gamma(\lambda - 1). \qquad (2.17)$$

The resulting bivariate VG density is a two parameter density. The dependence parameter ϱ acts here differently than in the case of the bivariate Student-t density (2.15). The parameter λ is the shape parameter with smaller λ implying larger tail heaviness. Furthermore, when $\lambda \to \infty$, the marginal VG density asymptotically approaches the standard Gaussian density.

In Barndorff-Nielsen (1995) log-returns were proposed to follow a normal-inverse Gaussian (NIG) mixture distribution. The corresponding density arises from the

SGH density when the shape parameter $\lambda = -\frac{1}{2}$ is chosen. The resulting bivariate NIG density is a two parameter density. The parameter $\bar{\alpha}$ is the shape parameter with smaller $\bar{\alpha}$ implying larger tail heaviness. Furthermore, when $\bar{\alpha} \to \infty$ the marginal NIG densities can be shown to asymptotically approach the standard Gaussian density.

In Eberlein & Keller (1995) and Küchler, Neumann, Sørensen & Streller (1999) models are proposed, where log-returns appear to be hyperbolically distributed. This occurs for the choice of the shape parameter $\lambda = \frac{d+1}{2}$ in the d-variate SGH density. The bivariate hyperbolic density that results is a two parameter density with $\lambda = 1.5$. The parameter $\bar{\alpha}$ is the shape parameter with smaller $\bar{\alpha}$ implying larger tail heaviness. Furthermore, when $\bar{\alpha} \to \infty$ the marginal hyperbolic densities asymptotically approach the standard Gaussian density.

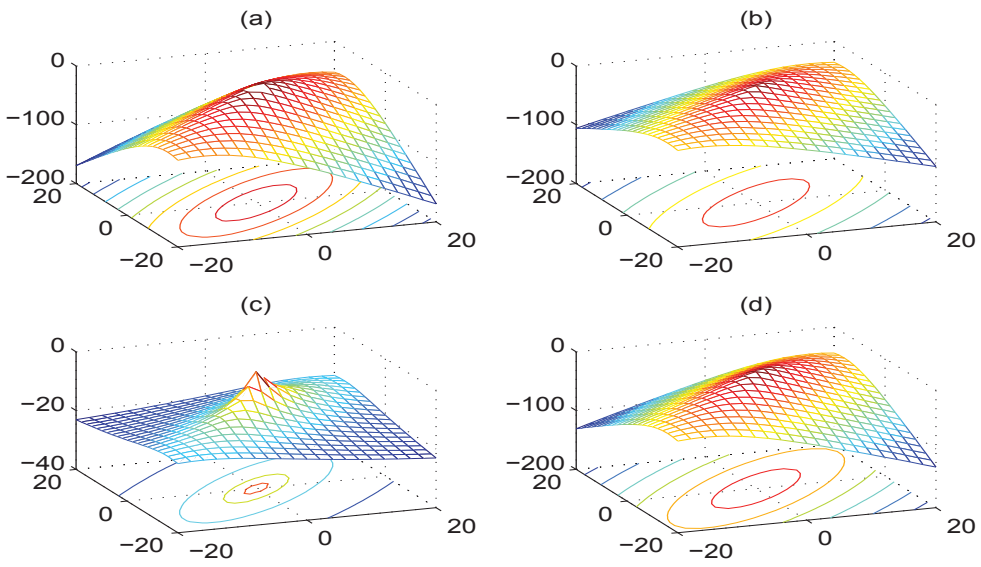

Figure 2.5: Log-scale plots of: (a) VG density; (b) hyperbolic density; (c) Student-t; (c) NIG

By choosing the parameters according to the estimates, as they result from the maximum likelihood estimation below, we illustrate in Fig. 2.5 four special cases of the bivariate SGH density, where the dependence parameter ϱ is chosen to be equal to 0.8: (a) VG density for $\lambda = 4$; (b) hyperbolic density for $\lambda = 1.5$, $\bar{\alpha} = 1$; (c) Student-t density for $\lambda = -2$ and (d) NIG density for $\lambda = -1/2$, $\bar{\alpha} = 4$. In this illustration we choose the parameters $\bar{\alpha}$, λ so that (2.14) is satisfied and the marginal

variances equal one. One notes the heavy tails in each of the bivariate distributions for the extreme values, in particular, for the Student-t distribution.

Likelihood Ratio Test

In order to test the hypothesis that a candidate log-return distribution is acceptable or not, we follow the classical maximum likelihood ratio test; see Rao (1973). The likelihood ratio is defined by the expression

$$\Lambda = \frac{\mathcal{L}^*{}_{model}}{\mathcal{L}^*{}_{nesting\ model}}, \tag{2.18}$$

here $\mathcal{L}^*{}_{model}$ represents the maximized likelihood function of the specific nested density, while $\mathcal{L}^*{}_{nesting\ model}$ represents the maximized likelihood function of the nesting density. Here the likelihood is defined as the product of the marginal densities. For example, we will later choose in some cases the SGH density as nesting model and the symmetric Student-t density as one of the nested models. Note that in the process of maximizing the likelihood, we optimize with respect to the parameters of the given parameterized distribution. Hence we obtain both the optimal parameters and the optimal value of the likelihood function. It can be shown that for increasing number of observations $n \to \infty$ the test statistic

$$L_n = -2\ln(\Lambda) \tag{2.19}$$

is asymptotically distributed as a chi-square distribution; see Rao (1973). Additionally, the degrees of freedom of this chi-square distribution are determined by the difference between the number of parameters of the nesting and the nested models. Specifically, the bivariate SGH density is a three-parameter density, while the four special cases we consider: the symmetric variance gamma, Student-t, hyperbolic and generalized inverse Gaussian densities are two-parameter densities, which implies that their test statistic is chi-square distributed with one degree of freedom.

Note that asymptotically it can be shown that

$$P(L_n < \chi^2_{1-\alpha,1}) \approx F_{\chi^2(1)}(\chi^2_{1-\alpha,1}) = 1-\alpha, \tag{2.20}$$

where $F_{\chi^2(1)}$ denotes the chi-square distribution with one degree of freedom and $\chi^2_{1-\alpha,1}$ is its $100(1-\alpha)\%$ quantile. We cannot reject on a 90% level of significance the hypothesis that the suggested density is the true underlying density, if the relation

$$L_n < \chi^2_{0.1,1} \approx 0.015791 \tag{2.21}$$

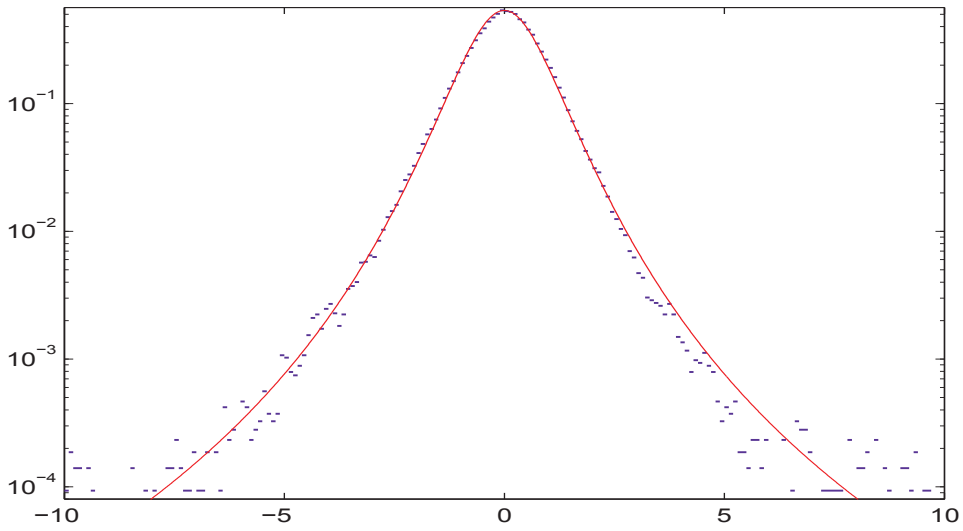

Figure 2.6: Log-histogram of the EWI104s log-returns and Student-t density with four degrees of freedom

is satisfied. We then reject the hypothesis that the suggested density is the true underlying density at the 99% level of significance if the following relation is not satisfied

$$L_n < \chi^2_{0.01,1} \approx 0.000157. \tag{2.22}$$

If we require greater precision, then the hypothesis is not rejected at the 99.9% level of significance if

$$L_n < \chi^2_{0.001,1} \approx 0.000002. \tag{2.23}$$

To conclude the above discussion, we call the density with the smallest test statistic L_n the best fit in the given family of distributions.

2.3 Fitted Univariate Log-return Distributions

This section is devoted to the analysis of the univariate log-returns of twelve world stock indices. We distinguish between the world stock indices constructed from the regional stock market indices, the world sector indices based on the 35 industry indices, and the world sector indices based on the 104 industry indices mentioned previously. The region based indices consist of the: MCI, DWI, EWI and WSI, while the sector based indices are represented by the: MCI35s, DWI35s, EWI35s and WSI35s, as well as, the MCI104s, DWI104s EWI104s and WSI104s. We use daily

Table 2.1: Empirical moments for log-returns of the EWI104s in various currency denominations

Country	$\hat{\mu}_y$	$\hat{\sigma}_y$	$\hat{\beta}_y$	$\underline{\hat{\kappa}}_y$
Australia	0.000565	0.008573	1.255285	38.042243
Austria	0.000423	0.008932	-0.560596	9.818562
Belgium	0.000473	0.008616	-0.572222	8.617690
Brazil	0.000774	0.010628	0.470175	15.449772
Canada	0.000519	0.007150	-0.655434	11.930284
Denmark	0.000491	0.008637	-0.493104	10.819425
Finland	0.000524	0.008646	-0.305966	9.590003
France	0.000512	0.008546	-0.522862	8.826783
Germany	0.000426	0.008627	-0.611260	8.901389
Greece	0.000867	0.009325	0.530904	27.355595
Hong Kong	0.000600	0.007379	-0.709914	16.946884
India	0.000698	0.008086	0.260450	20.710839
Ireland	0.000554	0.008863	-0.359713	35.363948
Italy	0.000622	0.008481	-0.504651	9.270888
Japan	0.000394	0.008151	-0.742019	9.827505
Korea S.	0.000593	0.009045	0.453250	36.812322
Malaysia	0.000533	0.007845	-0.664893	15.829299
Netherlands	0.000439	0.008558	-0.598923	8.979067
Norway	0.000502	0.008365	-0.431793	10.298253
Portugal	0.000714	0.009343	0.168636	13.568042
Singapore	0.000499	0.007228	-1.116746	17.150500
Spain	0.000594	0.008756	0.204594	16.586917
Sweden	0.000560	0.008372	0.256363	17.623515
Taiwan	0.000502	0.007456	-0.956096	16.691994
Thailand	0.000634	0.009012	1.861305	62.413554
UK	0.000536	0.008165	-0.593624	9.567338
USA	0.000501	0.007004	-0.819822	14.237813

Table 2.2: Results for log-returns of the EWI104s

	SGH	Student-t	NIG	Hyperbolic	VG
σ	0.9807068	0.7191163	0.9697258	0.9584118	0.9593693
$\bar{\alpha}$	0.0000000		0.9694605	0.7171357	
λ	-2.1629649				1.4912414
ν		4.3259646			
$\ln(\mathcal{L}^*)$	-285796.3865295	-285796.3865297	-286448.9371892	-287152.0787956	-287499.8259143
L_n		**0.0000004**	1305.1013194	2711.3845322	3406.8787696

data from 1973 to 2006, provided by Datastream, for all the components underlying our indices. In the following, we mainly report the results for the log-returns of the EWI104s when denominated in 27 currencies.

A summary of the main empirical moments of the log-returns of the EWI104s when denominated in different currencies, is presented in Table 2.1. Note that we obtain here the average empirical mean $\hat{\mu}_y = 0.000557$, the average empirical standard deviation $\hat{\sigma}_y = 0.008437$, the average sample skewness $\hat{\beta}_y = -0.213288$ and the average sample excess kurtosis $\underline{\hat{\kappa}}_y = 17.823402$. We do not remove any extreme values as potential outliers from our data set, hence market crashes and other sudden market corrections are not discarded. Removing outliers would also not be appropriate, as the proper modeling of extreme log-returns is of great importance in risk management.

First, to get a visual impression of the shape of the log-return density of the

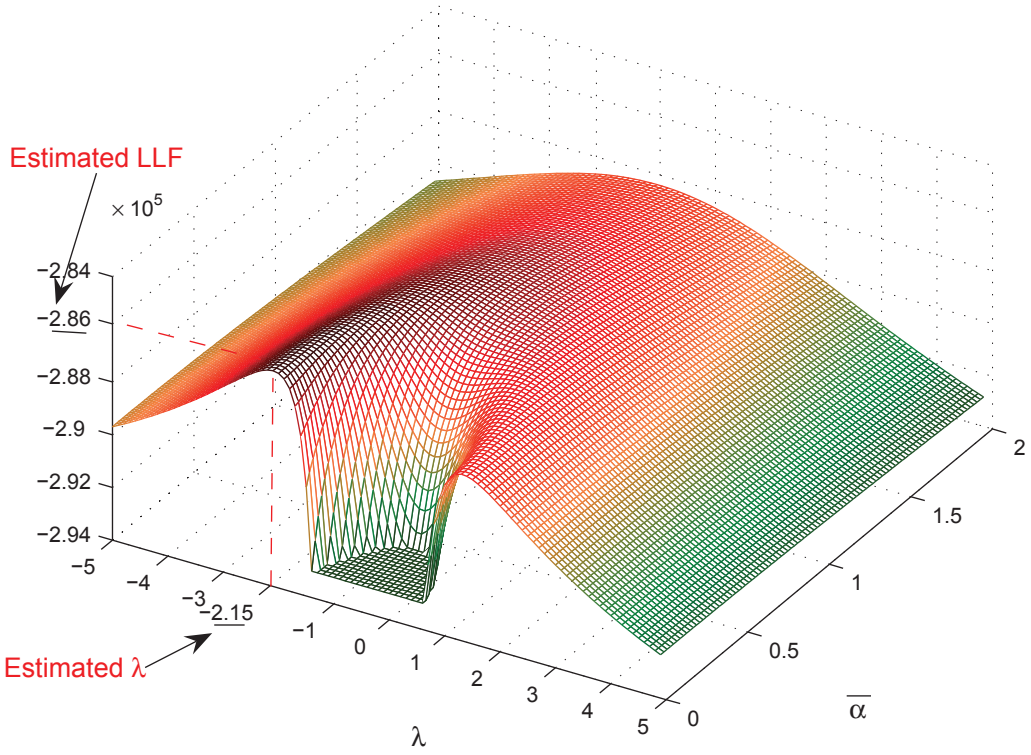

Figure 2.7: Log-likelihood function based on the EWI104s

EWI104s, we exploit all log-returns that we observe from this index in all 27 currency denominations. For this purpose we combine appropriately shifted and scaled log-returns of all currency denominations. More precisely, for each currency denomination of the index we shift all the obtained log-return values so that their sample mean becomes zero, and scale them in order to obtain a sample variance of one. In Fig. 2.6 we present the resulting histogram of the total cohort of shifted and scaled log-returns on a logarithmic scale. Note that this histogram is based on $214,658$ observations, which makes it very reliable. Additionally, in Fig. 2.6 we show in log-scale the theoretical Student-t density for $\nu = 4$ degrees of freedom. We observe visually an excellent fit of the log-returns of the EWI104s to the Student-t density. For the other constructed indices, a similar visual impression is obtained, with the EWI104s seeming to fit best.

The maximum likelihood methodology is then employed to estimate the parameters of the SGH density. Note that in our analysis we do not take the autocorrelation of the log-returns into account. In fact, we are more interested in the identification

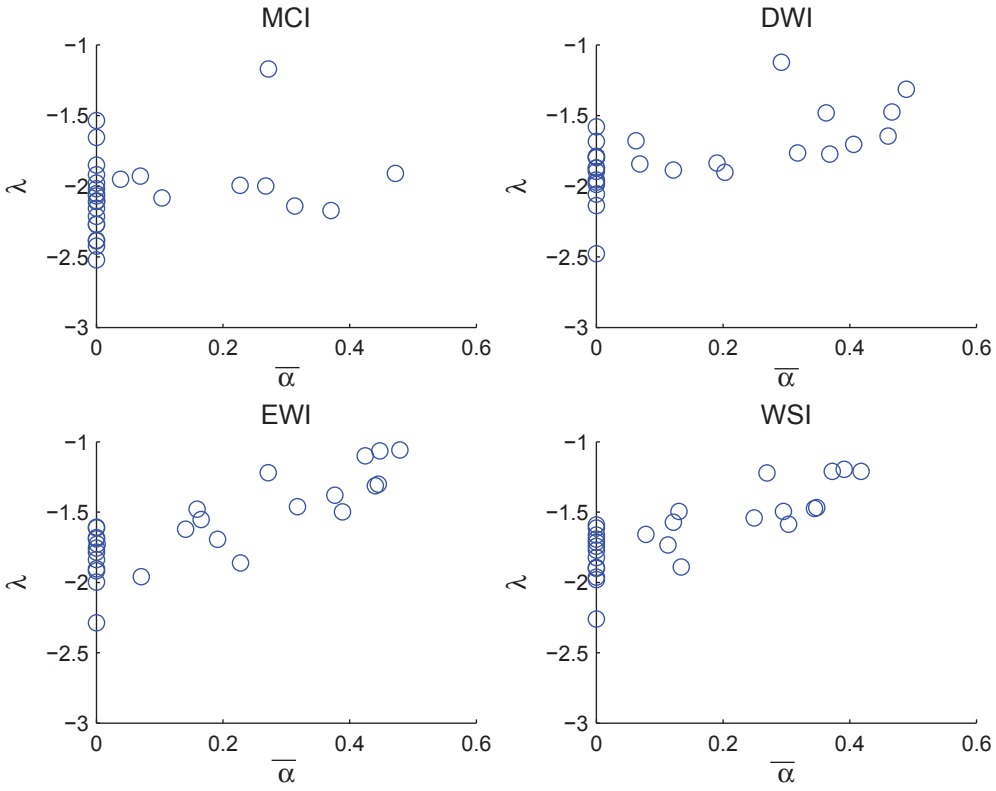

Figure 2.8: $(\bar{\alpha}, \lambda)$-plot for log-returns of indices in 27 different currencies constructed from regional stock market indices as constituents

of the stationary density resulting for the underlying model of the normalized equity index. This will be further clarified in Chapter 4 of this thesis. In Chapter 4 we will perform similar likelihood analysis for the simulated trajectory of the equity index and find that the simulated log-returns exhibit similar distributional behavior when obtained from the in this thesis developed equity index model.

For the same sample used to produce the histogram in Fig. 2.6, we exhibit the log-likelihood function for the SGH in Fig. 2.7 as a function of the parameters λ and $\bar{\alpha}$. One notes a clear, flat global maximum around the point $\bar{\alpha} = 0$ and $\lambda = -2$, which refers to a Student-t distribution with four degrees of freedom. We then apply the maximum likelihood method to the log-returns of the EWI104s for four special cases of the SGH distribution. These cases concern the Student-t density, the NIG density, the hyperbolic density and the VG density. We maximize the corresponding log-likelihood functions with respect to the parameters σ, $\bar{\alpha}$, λ and ν of the SGH, as shown in Table 2.2, and display their estimates in its second column. Additionally,

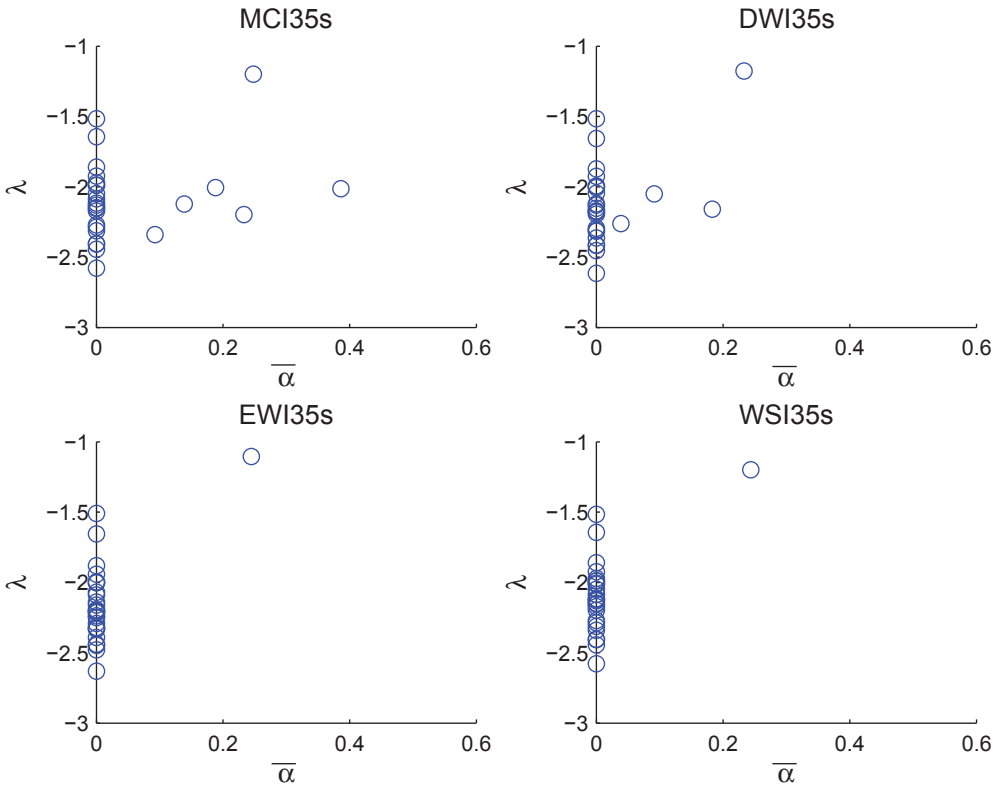

Figure 2.9: $(\bar{\alpha}, \lambda)$-plot for log-returns of indices in 27 different currencies constructed from 35 sector indices as constituents

row six of Table 2.2 contains the estimated parameters and maximized values of the log-likelihood functions for the SGH density and each of the four considered special cases. In the last row of Table 2.2 we calculated the value of the test statistic L_n. The almost zero value obtained for L_n suggests a very good fit of the Student-t distribution with $\nu \approx 4.3$ degrees of freedom to our set of data in the class of SGH distributions. The extremely small value of the test statistic, $L_n = 0.0000004$, allows us to conclude that the H_0 hypothesis of a Student-t fit to the EWI104s cannot be rejected at least on a 99.9% level of significance, since $\chi^2_{0.001,1} \approx 0.000002$. We emphasize that this is an extremely high level of significance. Note also that the estimated parameter value $\bar{\alpha} = 0$ for the SGH density suggests that a Student-t density already represents the best fit to the log-returns of the EWI104s when searching among the family of SGH densities.

For all the components underlying our indices we now again use daily data from 1973 to 2006, provided by Datastream. Furthermore, we denominate all twelve

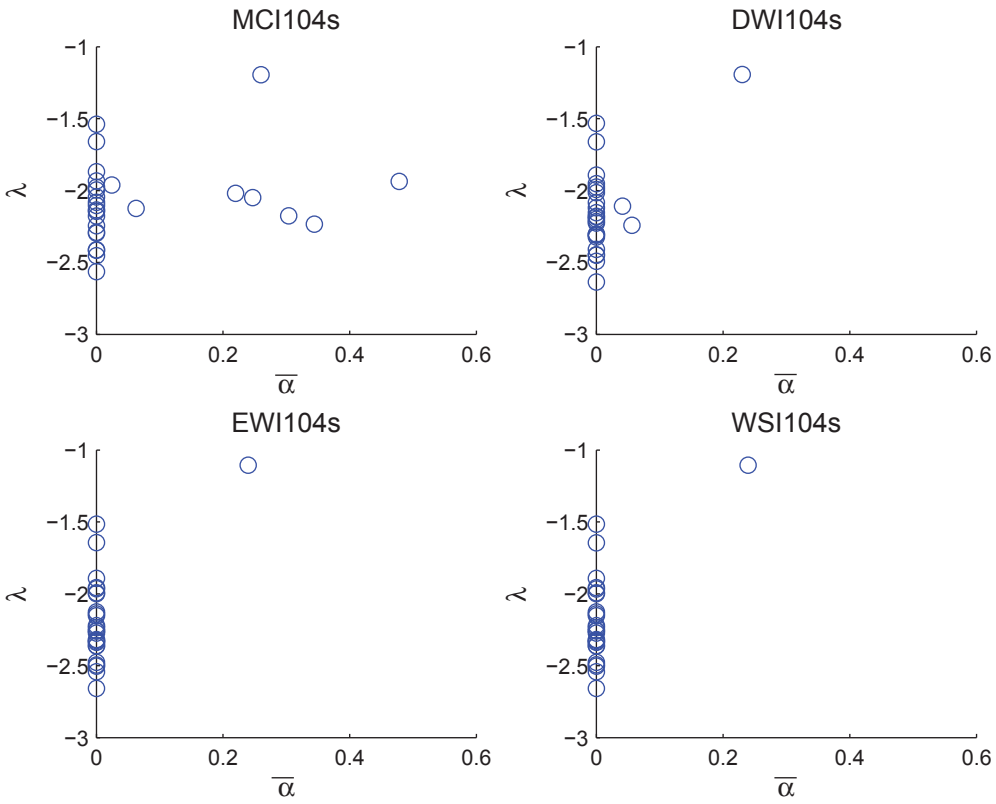

Figure 2.10: $(\bar{\alpha}, \lambda)$-plot for log-returns of indices in 27 different currencies constructed from 104 sector indices as constituents

constructed indices in 27 different currencies and study their log-returns for each of the denominations separately. Note that the denomination of a diversified world stock index in a given currency reflects in its fluctuations the general market risk with respect to this currency, see Platen & Stahl (2003a). We mention that the data for all exchange rates were not available from 1973. For instance, the time series for the Brazilian real starts only in 1995.

For convenience, we shift and scale the log-returns in order to obtain sample means equal to zero and sample standard deviations equal to one for all currency denominations. This does not change the generality of our analysis, but standardizes the testing procedure.

We first apply the maximum likelihood estimation methodology for both the generalized hyperbolic (GH) and the SGH distributions for the log-returns of each index and for each currency denomination. The GH distribution contains the extra parameter $\gamma \in \Re$, which represents the level of skewness of this distribution.

Table 2.3: L_n test statistic of the EWI104s for different currency denominations

Country	Student-t	NIG	Hyperbolic	VG	ν
Australia	0.000000	76.770817	150.202282	181.632971	4.281222
Austria	0.000000	39.289103	77.505683	102.979330	4.725907
Belgium	0.000000	31.581622	60.867570	83.648470	4.989912
Brazil	2.617693	5.687078	63.800349	60.078395	2.713036
Canada	0.000000	47.506215	79.917741	104.297607	5.316154
Denmark	0.000000	41.509921	87.199686	114.853658	4.512101
Finland	0.000000	28.852844	68.677271	88.553080	4.305638
France	0.000000	26.303544	57.639325	80.567283	4.722787
Germany	0.000000	27.290205	52.667918	71.120798	5.005856
Greece	0.000000	60.432172	104.789463	125.601499	4.674626
Hong.Kong	0.000000	42.066531	100.834255	122.965326	3.930473
India	0.000000	74.773701	163.594078	198.002956	3.998713
Ireland	0.000000	77.727856	136.505582	170.013644	4.761519
Italy	0.000000	25.196598	55.185625	75.481897	4.668983
Japan	0.000000	37.630363	77.163656	102.967380	4.649745
Korea.S.	0.000000	120.904983	304.829431	329.854620	3.289204
Malaysia	0.000000	79.714054	186.013963	221.061290	3.785195
Netherlands	0.000000	26.832761	51.625813	71.541627	5.084056
Norway	0.000000	42.243851	89.012090	115.059003	4.472349
Portugal	0.000000	61.177624	137.681039	165.689683	3.984860
Singapore	0.000000	36.379685	77.600590	98.124375	4.251472
Spain	0.000000	56.694545	109.533768	138.259224	4.517153
Sweden	0.000000	77.618384	143.420049	178.983373	4.546640
Taiwan	0.000000	41.162560	96.283628	115.186585	3.914719
Thailand	0.000000	78.250621	254.590254	267.508143	3.032038
UK	0.000000	26.693076	55.937248	80.678494	4.952843
USA	0.000000	40.678242	79.617362	100.901197	4.636661

Our study, however, reveals that the estimated parameter values for γ are of the order 10^{-6} or less. Since γ and μ turn out to be extremely small we do not report the statistical findings for the GH distribution and just concentrate on the SGH distribution, which appears to be almost identical to what we reported previously.

In order to visualize the fitted log-return distributions of our twelve constructed indices, we plot the estimated parameter λ versus the estimated parameter $\bar{\alpha}$ in Figs. 2.8, 2.9 and 2.10 for each constructed index, respectively, in all 27 currency denominations. Fig. 2.8 presents the results for each of the MCI, DWI, EWI and WSI for 27 currency denominations. Note that the estimated values of $\bar{\alpha}$ are in most cases close to zero, which already confirms the Student-t property. Furthermore, the estimated values of the parameter λ range from around -2.5 to -1.0. This indicates that the best fit can be expected for Student-t distributions with degrees of freedom of about $\nu = -2\lambda$ ranging from around 2 to 5. This is emphasized by the cluster of points in Fig. 2.8 located on the negative λ axis near -2.

Note that in the group of indices shown in Fig. 2.8 the log-returns of the MCI seem to fit visually the Student-t distribution best. There are just nine points which do not sit on the negative λ axis. It can be also noticed that the range of the estimated values of λ in the case of the EWI and WSI is narrower for $\bar{\alpha} = 0$.

Fig. 2.9 illustrates the estimated parameters of the SGH distribution for the indices constructed on the basis of 35 world industry sectors as constituents. Here we used the observed log-returns of the MCI35s, DWI35s, EWI35s and WSI35s for the estimation of the log-return distribution. Similarly, as for the case of regional stock index based indices as constituents, we obtain estimates of $\bar{\alpha}$ that are close to zero and values for λ, which range from approximately -1.0 to -2.5. In the case of these sector based indices in 27 currencies we observe an even better fit of the Student-t distribution. This is, in particular, visible in Fig. 2.9 for the EWI35s and the WSI35s. In these two cases, only the log-returns denominated in the Brazilian real do not exhibit a proper Student-t fit, as the estimate of $\bar{\alpha}$ is not zero in this case but is approximately equal to 0.25. This is probably a consequence of the short length of our data series on the Brazilian exchange rate. It is obvious that data sets of sufficient length are necessary in order to obtain a proper fit to the underlying true distribution.

In Fig. 2.10 we analyzed log-returns of the MCI104, DWI104s, EWI104s and WSI104s, based on 104 world industry sector indices as constituents. We again obtain estimates for the parameters $\bar{\alpha}$ and λ, which indicate a good Student-t fit to the observed log-returns of all four indices considered. We note that the improved diversification of the indices in Fig. 2.10 did not greatly improve the Student-t fit when compared with Fig. 2.9. The best fits are here again obtained for the EWI104s and the WSI104s.

In conclusion, in all three figures the estimated $(\bar{\alpha}, \lambda)$ points are localized near the negative λ-axis, which indicates an approximate Student-t density. Moreover, one can notice that the Student-t density represents a better fit for two types of the world industry sector based indices: It fits very well in the case of the EWI35s, as well as the EWI104s. These fits are remarkable and can be interpreted as a stylized empirical fact.

In the second column of Table 2.3 we show only six digits for the Student-t test statistics, which is sufficient to decide whether these values are less than the 99.9% quantile 0.000002 of the chi-square distribution with one degree of freedom. We emphasize that the estimated degrees of freedom of the Student-t density obtained, are in the range of around 3 to 5, with a concentration near 4, as can be concluded

from the last column of Table 2.3.

One possible explanation of the above documented empirical facts is that the above diversified indices are driven by the nondiversifiable uncertainty of the market, which generates Student-t log-returns with four degrees of freedom for a well diversified index. There are many continuous time models thinkable that would generate such log-returns. For instance, the minimal market model (MMM), introduced in Platen (2001) and further described in Platen & Heath (2010), is one of such models. The MMM models the discounted GOP by a time transformed squared Bessel process of dimension four. The squared volatility of this process is the inverse of a square root process, which has as its stationary density a gamma density with four degrees of freedom. Therefore, the mixing density for the variance of the returns of the GOP is that of the inverse of a gamma distributed random variable. Consequently, log-returns generated by the MMM, when estimated over a sufficiently long time period, would appear to be Student-t distributed with four degrees of freedom. Chapter 4 will propose even more realistic model that exhibits this property.

2.4 Fitted Bivariate Log-return Distributions

To clarify further in which direction a realistic model for a diversified world index should go, we focus in this section on the bivariate distribution of log-returns of the EWI104s when denominated in different currencies. In the previous section and Platen & Rendek (2008) the univariate Student-t distribution was identified for all these currency denominations. In fact, we could not reject the Student-t hypothesis for the log-returns of the EWI104s at the high 99.9% level of significance in all cases considered above.

We now apply the maximum log-likelihood methodology to pairs of log-returns of the EWI104s in different currency denominations. At first, let us consider the bivariate SGH densities that are estimated when considering Australian dollar log-returns of the EWI104s in conjunction with log-returns of the EWI104s denominated in other currencies. Table 2.4 summarizes for the Student-t, the NIG, the hyperbolic and the VG model the resulting test statistics. For the Student-t model the second

Table 2.4: Test statistics for Australian dollar denomination of the EWI104s paired with other currency denominations

AU vs	$L_n^{Student-t}$	L_n^{VG}	$L_n^{hyperbolic}$	L_n^{NIG}	ν	ϱ
Austria	**0.00274025**	236.55845406	173.09166361	61.14818639	3.90382049	0.51958549
Belgium	**0.00268553**	189.22777798	135.47978607	42.43451789	3.90270597	0.52975278
Canada	**0.00945525**	383.63742806	302.15085711	138.52025651	3.99853063	0.78240493
Denmark	**0.01071392**	317.88771606	242.54225166	94.98967385	3.83420592	0.64844959
Finland	6.10051720	141.10617451	109.77042154	13.91296972	3.41369825	0.58007350
France	**0.00201669**	192.18917100	136.03978111	39.29902652	3.82551267	0.53854522
Germany	**0.00343527**	179.94846900	129.68118003	44.50772697	4.00381705	0.52056947
India	**0.00174492**	474.64922765	407.03338053	139.81107398	3.32763559	0.71795534
Ireland	**0.00059705**	281.18124814	224.12702161	87.33135209	3.76171049	0.55436625
Italy	0.12204592	158.81412998	115.04959011	24.11820486	3.65327244	0.56227292
Japan	0.06537189	277.21608865	212.11667266	77.10560911	3.81720949	0.62409861
Malaysia	**0.01542705**	555.36165251	477.97086928	145.05643860	3.14328556	0.75017320
Netherl.	**0.00361761**	178.40304644	124.10893727	37.29233197	3.94087465	0.52899993
Norway	**0.00235773**	285.96573474	218.37037709	78.84324303	3.77162766	0.66697989
Portugal	**0.00021500**	300.31208740	239.27207897	69.18348116	3.47527269	0.54116762
Spain	**0.00106233**	204.61053444	161.12945242	41.20480843	3.48860083	0.57467597
Sweden	**0.00089575**	361.00317250	294.34436968	123.46493060	3.78422621	0.67659006
UK	**0.00064994**	254.75967167	193.25178409	81.62402590	4.06922065	0.66731532
USA	0.03393279	266.78659148	227.97996642	63.55741960	3.33472650	0.79242332

Table 2.5: Test statistics for US dollar denomination of the EWI104s paired with other currency denominations

USA vs	$L_n^{Student-t}$	L_n^{VG}	$L_n^{hyperbolic}$	L_n^{NIG}	ν	ϱ
Australia	0.03393297	266.78659099	227.97996643	63.55741957	3.33446659	0.79242332
Austria	**0.00256224**	134.73059219	97.71377792	35.73898997	4.32153080	0.67294050
Belgium	3.27984294	77.85490511	48.95878079	11.14186455	4.36981718	0.68569372
Canada	7.40430763	120.01666437	74.73407431	15.18540666	4.22211760	0.91527388
Denmark	**0.00323445**	149.30811831	108.10201571	41.11709989	4.28846201	0.67495392
Finland	2.95289908	107.01498821	85.36523117	21.10238174	3.87650221	0.73612270
France	3.64652056	86.96479253	56.22828540	13.16659792	4.24625684	0.69863871
Germany	2.77322928	75.27644297	46.50740906	10.95284501	4.47611255	0.67756222
India	40.40298515	705.21429964	970.99037909	20.04038067	1.88004348	0.95479770
Ireland	**0.00225433**	178.45506759	149.01226074	68.77685881	4.15355644	0.70489287
Italy	6.76679376	62.97646549	42.03537752	6.75574398	4.10152450	0.72286305
Japan	3.96736350	112.53857499	78.32095430	16.36172747	3.97418294	0.67690443
Malaysia	1.02180400	1327.00601731	1784.37475954	228.41900838	1.94684216	0.96582832
Netherl.	3.28833251	75.82687729	45.23534025	9.22019399	4.45911761	0.68774633
Norway	**0.00319992**	152.67795003	104.21581327	35.19242207	4.27513262	0.71233960
Portugal	**0.00631108**	294.32335669	255.16166144	100.62776920	3.73382193	0.71223257
Spain	**0.00297697**	109.70420788	95.09407429	33.04763827	3.88099965	0.73704326
Sweden	**0.00158380**	233.42505334	178.66537682	82.95266849	4.28756905	0.71095526
UK	5.55822674	68.41485288	40.24352898	6.09768583	4.36293689	0.71860250

last column provides the estimate for the degrees of freedom and the last column the estimated dependence parameter ϱ. A similar study is summarized in Table 2.5 for the log-returns of the EWI104s denominated in US dollar paired with other currencies.

We highlight those pairs of log-return denominations for which the Student-t

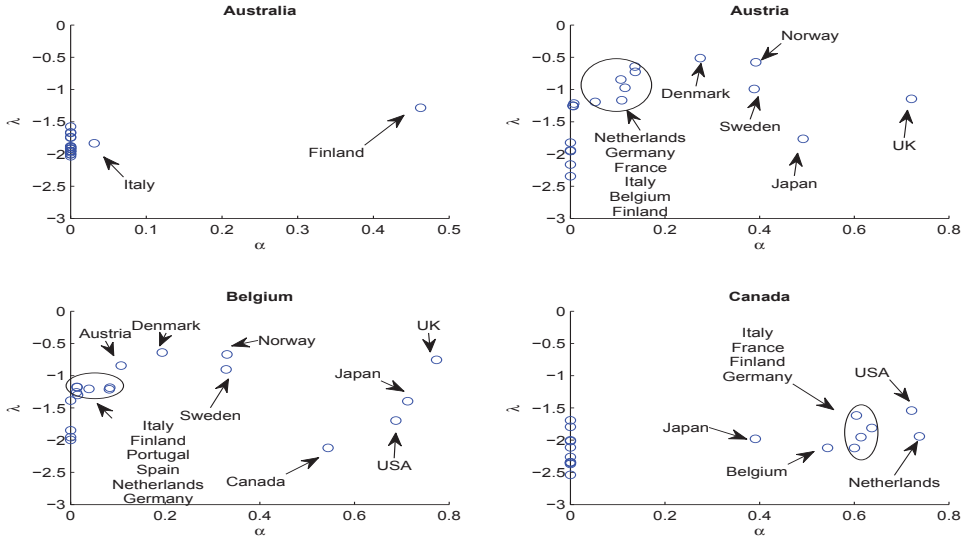

Figure 2.11: $(\bar{\alpha}, \lambda)$-plot for pairs of log-returns with reference to Australia, Austria, Belgium and Canada

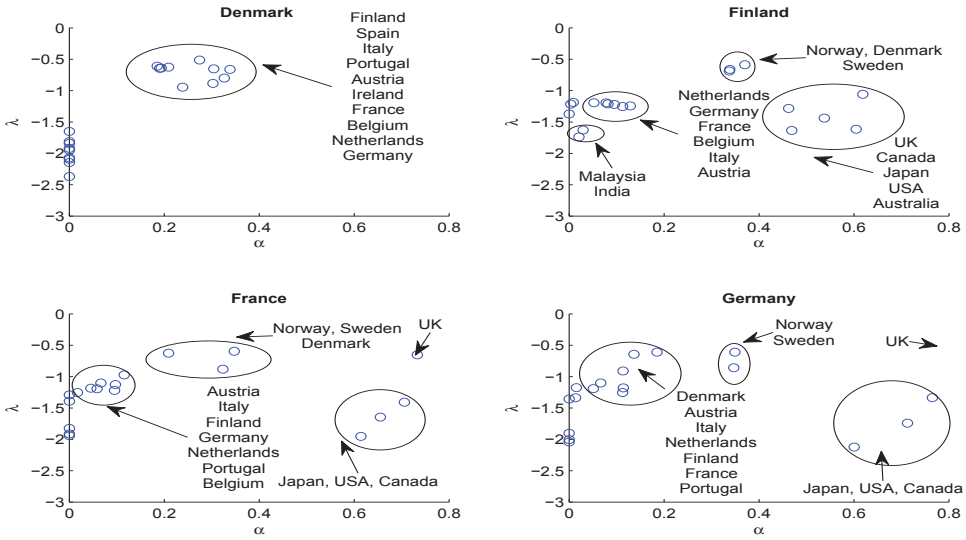

Figure 2.12: $(\bar{\alpha}, \lambda)$-plot for pairs of log-returns with reference to Denmark, Finland, France and Germany

hypothesis cannot be rejected at the 90% level of significance by writing the corresponding test statistic in bold numbers. For the Australian dollar denomination paired with other currency denominations we cannot reject the Student-t hypothesis in fifteen out of nineteen cases at the 90% level of significance. While, for the US dollar denomination paired with other currency denominations we cannot reject the

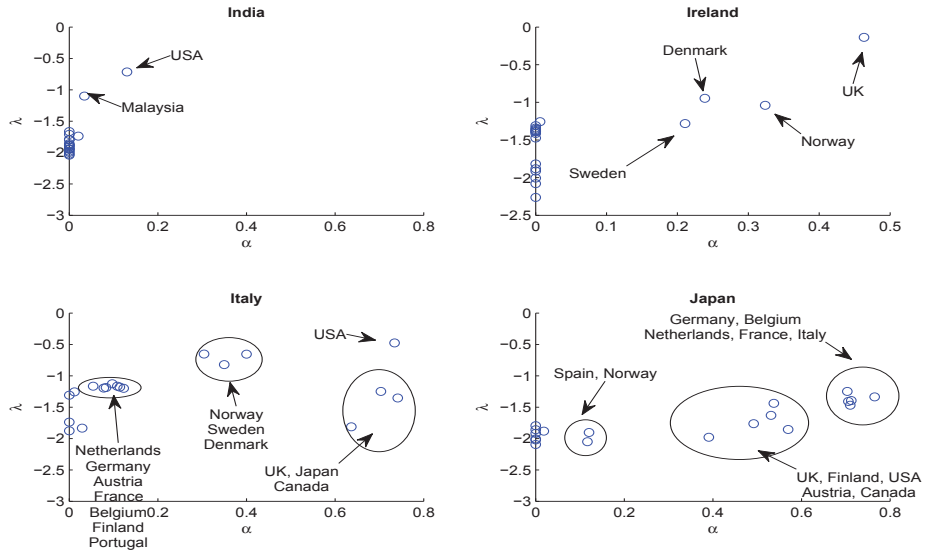

Figure 2.13: $(\bar{\alpha}, \lambda)$-plot for pairs of log-returns with reference to India, Ireland, Italy and Japan

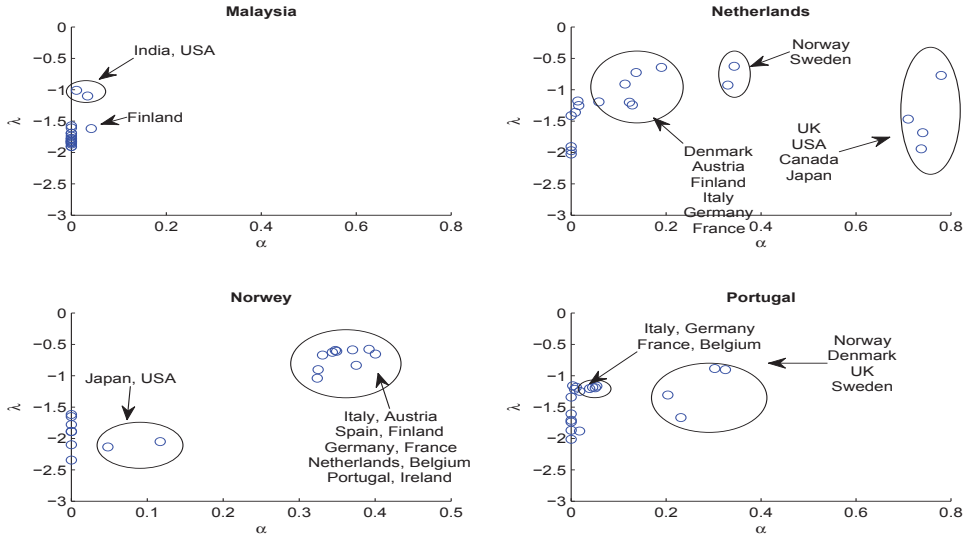

Figure 2.14: $(\bar{\alpha}, \lambda)$-plot for pairs of log-returns with reference to Malaysia, Netherlands, Norway and Portugal

Student-t hypothesis in seven out of nineteen cases at the 90% level of significance. Additionally, as shown in the second last column in Tables 2.4 and 2.5, the bivariate Student-t density fitted to pairs of log-returns has approximately three to four degrees of freedom. For the US dollar pairings a dependence parameter ϱ of about 0.7 to 0.9 resulted. For the Australian dollar pairings a lower dependence parameter

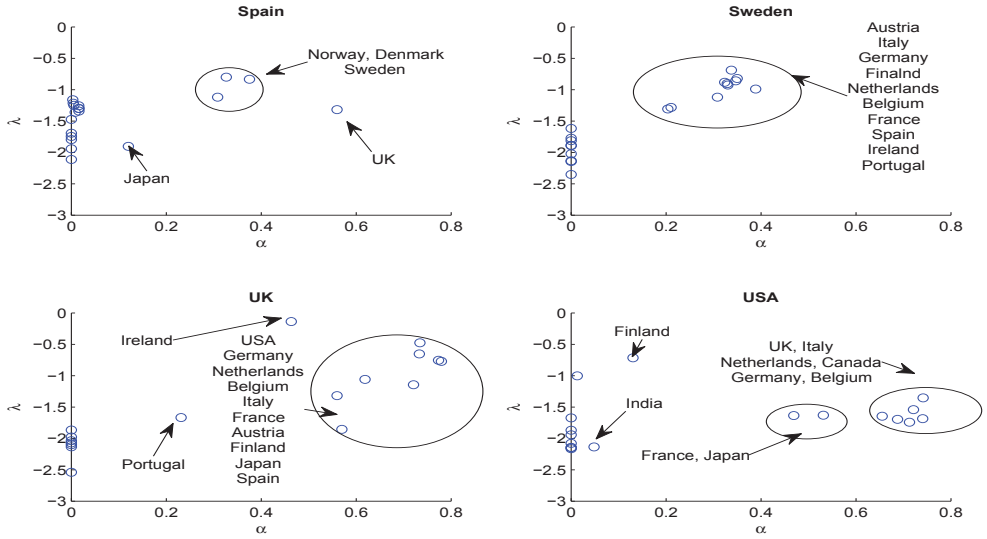

Figure 2.15: $(\bar{\alpha}, \lambda)$-plot for pairs of log-returns with reference to Spain, Sweden, UK and USA

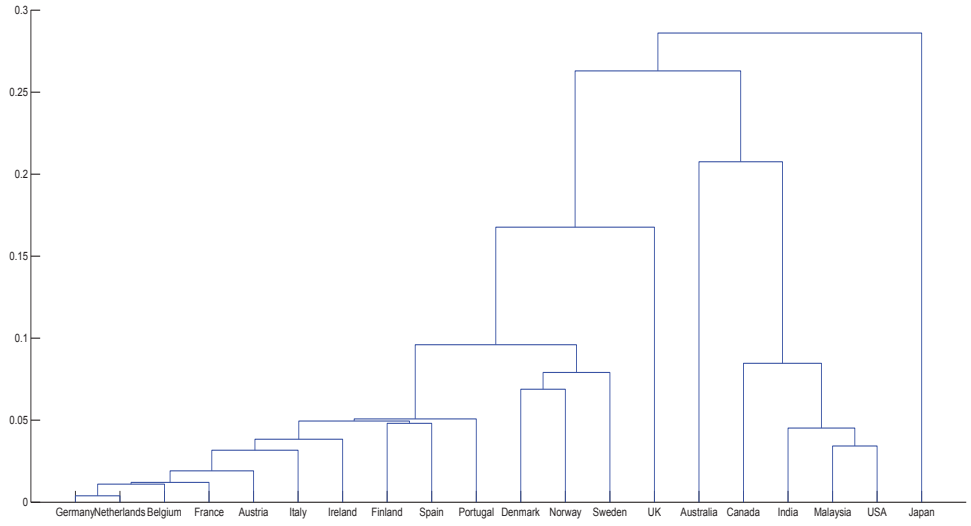

Figure 2.16: Cluster analysis plot for the log-returns of the EWI104s in different currency denominations based on the estimated dependence parameter ϱ

value of approximately 0.5 to 0.7 was typically estimated. This reflects the lead role of the US currency by being closely linked to many currencies. On the other hand, the Australian currency is not so closely linked to most other currencies.

It may be important to note that when assuming an unperturbed currency dynamics yileds a Student-t joint distribution for log-returns of diversified indices,

then a more freely floating currency, as the Australian dollar, would show more test statistics that support the Student-t hypothesis than the US dollar, which is heavily targeted by many central banks as reference currency. The central bank interventions could, for instance, reduce the occurance of extreme log-returns in exchange rates.

Now, it is of interest to see the results for other pairs of log-returns. For pairs of EWI104s log-returns related to different currencies, the estimated values of the shape parameters $\bar{\alpha}$ and λ from the bivariate SGH density are displayed in Figs. 2.11-2.15 in $(\bar{\alpha}, \lambda)$-scatter plots, generating a point for each of the analyzed currency pairs.

Interestingly, the estimated $(\bar{\alpha}, \lambda)$-points are approximately located near the negative λ-axis. It is the Student-t density that arises for the $(\bar{\alpha}, \lambda)$-parameter points that can be found near the negative λ-axis. Most clusters for the estimated parameter points are located rather close to the coordinates $\bar{\alpha} = 0$ and $\lambda = -2$. As previously, these parameters refer to the Student-t distribution with four degrees of freedom. For comparison, a VG density is reflected by $(\bar{\alpha}, \lambda)$-parameter points near the positive λ-axis. None of such points have been estimated. The NIG density refers in Figs. 2.11-2.15 to the horizontal line with $\lambda = -0.5$. There were several pairs of related currencies that generated such points. The hyperbolic density gives estimated parameters at the horizontal line $\lambda = 1.5$. For this case no currency pair showed any reasonable likelihood.

Our results visually suggest that the observed log-returns appear to be pairwise bivariate Student-t distributed with only a few exceptions. The clustering near $\lambda = -2$ and $\bar{\alpha} = 0$ is strongest for relatively independent economies, as can be seen in Fig. 2.11 for Australia, but also in Fig. 2.13 for India and also for Malaysia in Fig. 2.14. All these countries have rather independent economies and other central banks are unlikely to target the above three currencies in their trading activities to stabilize their own currencies via the exchange rate. In the case of European countries, as well as the USA, Canada and Japan, we observe some distortion from an ideal bivariate Student-t fit. This phenomenon could be due to the trading activities of the central banks of the respective countries, when they aim to remove extreme moves of the respective exchange rates. This trading activity then makes the joint distribution less leptokurtic, which means that estimates can be expected

in the $(\bar{\alpha}, \lambda)$-plot more to the right from the negative λ-axis and slightly above -2 for the $\bar{\alpha}$ parameter. This is indeed the case for most of the currencies of close trading partners. If one wishes, then can even identify in the Figs. 2.11-2.15 groupings of currencies, as those of the European countries, the Scandinavian countries without Finland and the North American countries.

Let us now study in a broader manner the dependence structure of the considered currencies. Note that, as the log-returns are not Gaussian correlation may not be ideal to measure dependence. For instance, a zero correlation does, in general, not imply independence. We perform now a cluster analysis based on the estimated values of the dependence parameters. This is not the same as if one performs a standard correlation analysis. The resulting cluster graph is displayed in Fig. 2.16. In this figure we can see that all the currencies are linked to some degree. They all depend on each other. The dependence is clearly noticeable for the European countries including: Germany, Netherlands, Belgium, France, Austria, Italy, Ireland, Finland, Spain and Portugal. This cluster is seen to be linked in Fig. 2.16 with another cluster which is the one of the Scandinavian countries including: Denmark, Norway and Sweden. Interestingly, Finland seems to be closer linked to the already mentioned European countries. Furthermore, these two clusters are linked with the UK. Additionally, we can distinguish the cluster of Canada, India, Malaysia and the USA. Interestingly, Australia and Japan seem to form their own two additional clusters.

Note that these findings are consistent with the statistical results obtained previously for pairings of currencies. For instance, in the case of India and Malaysia in Figs. 2.13 and 2.14, we note that most of the pairings result in the bivariate Student-t fit except for the USA and Malaysia when paired with India, as well as the USA and India when paired with Malaysia. Similarly, we observe that the Student-t fit is distorted for pairs of the European currencies, which form a large cluster in Fig. 2.16. Note, when performing a standard cluster analysis based on correlation the dependencies are not as clearly identified as shown in Fig. 2.16. The reason is not likely that the Student-t bivariate density fits better the historical dependence structure than the Gaussian bivariate density.

We have examined the properties of pairs of log-returns of the EWI104s when

denominated in different currencies. As a result of this study we identify the bivariate Student-t distribution with about four degrees of freedom for the pairs of log-returns. This feature of log-returns is, however, distorted for related economies. The inconsistencies with the bivariate Student-t distribution may be caused by the trading of central banks with the aim to influencing exchange rates.

2.5 Copulas and Dependence

We aim now to further clarify empirical properties of log-returns of diversified indices in different denominations. This section refers to results in Ignatieva, Platen & Rendek (2011). When modeling a joint distribution of a random vector of risk factors, one has to take into account both, the marginal behavior of individual risk factors, as well as their dependence structure. The copula approach provides a way to measure the dependence structure by separating it from modeling the marginal distribution. Copulae define the multivariate distribution functions on the unit cube $[0,1]^d$, which allows to connect their one-dimensional uniform-[0,1] marginals to the joint cumulative distribution function; see Nelsen (1998) for the formal definition.

According to Sklar's theorem, every distribution function can be decomposed into its marginal distribution and a copula. On the other hand, every distribution function can be obtained by coupling marginal distributions with the dependence structure given by a copula. More precisely, Sklar's theorem states that if F is a d-dimensional distribution function with marginals $F_1 \ldots, F_d$, then there exists a copula C with

$$F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\} \tag{2.24}$$

for every $x_1, \ldots, x_d \in [-\infty, \infty]$. If $F_1, \ldots, F_d$ are continuous, then C is unique. On the other hand, if C is a copula and $F_1, \ldots, F_d$ are distribution functions, then the function F defined in (2.24) is a joint distribution function with marginals $F_1, \ldots, F_d$. A proof of Sklar's theorem can be found e.g. in Joe (1997).

Thus, for a random vector $X = (X_1, \ldots, X_d)^\top$ with a joint distribution $X \sim F_X$ and continuous marginals $X_j \sim F_{X_j}$, $j \in \{1, \ldots d\}$, one defines the multivariate copula using Sklar's theorem, as the distribution function C_X of $u = (u_1, \ldots, u_d)^\top \in [0,1]^d$, where $u_j = F_{X_j}(x_j)$:

$$C_X(u_1, \ldots, u_d) = F_X\{F_{X_1}^{-1}(u_1), \ldots, F_{X_d}^{-1}(u_d)\}. \tag{2.25}$$

For an absolutely continuous copula C one can define the copula density as

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d}. \tag{2.26}$$

If $X = (X_1, \ldots, X_d)^\top$ has an absolute continuous distribution function F and a copula C_X, then the density c_X is obtained by differentiating C_X in (2.25). That is,

$$c_X(u_1, \ldots, u_d) = \frac{f\{F_{X_1}^{-1}(u_1), \ldots, F_{X_d}^{-1}(u_d)\}}{\prod_{j=1}^d f_j\{F_{X_j}^{-1}(u_j)\}}, \tag{2.27}$$

where f is the joint density of F_X and f_j is the density of F_{X_j}. The density function of X is then given by

$$f(x_1, \ldots, x_d) = c_X(u_1, \ldots, u_d) \cdot \prod_{j=1}^d f_j(x_j)$$

with $x_j = F_{X_j}^{-1}(u_j)$.

Sklar's identity (2.24) can also be applied to a, so-called, survival copula. If F is the distribution function of the random vector $X = (X_1, \ldots, X_d)^\top$ with marginals $F_1 \ldots, F_d$, then there exists a copula C^* with

$$\overline{F}(x_1, \ldots, x_d) = C^*\{\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)\}, \tag{2.28}$$

where $\overline{F}(x_1, \ldots, x_d) = P(X_1 > x_1, \ldots, X_d > x_d)$ and $\overline{F}_i(x_i) = P(X_i > x_i)$, $i \in \{1, \ldots d\}$. C^* is the survival copula related to C. In particular, for the bivariate case the survival copula can be defined as

$$C^*(u_1, u_2) = 1 - u_1 - u_2 + C(1 - u_1, 1 - u_2), \tag{2.29}$$

see Nelsen (1998).

Dependence and Tail Dependence

Common approaches to model dependency among random variables include the Pearson correlation coefficient r, Spearmann's ρ and Kendall's τ, see e.g. McNeil, Frey & Embrechts (2005). While Pearson's linear correlation depends on the distribution of the univariate marginals (i.e., keeping the dependence structure constant, different marginals might lead to different values for the joint distribution; see Dias (2004)), the other two rank correlations are independent of the univariate marginal distributions. For properties of dependence measures one can refer to Embrechts, McNeil & Straumann (2001).

While the above coefficients consider only linear dependence among random variables, tail dependence coefficients allow to measure the extreme dependence in the tails of the multivariate distribution. These appear to be particularly useful in insurance and risk management when modeling the joint (dependent) risk; see e.g. Wang (1997) and Embrechts, McNeil & Straumann (2001). The concepts of lower and upper tail dependence refer to the study of the dependence between extreme values in the lower and in the upper tails. The notion of tail dependence in relation to copula first appeared in Joe (1997). For the bivariate case, the upper and the lower tail dependence coefficients can be defined as follows: If (U_1, U_2) is a pair of uniform variables on the unit square $[0,1]^2$, then the upper tail dependence coefficient $\lambda_u \in [0,1]$ is defined as

$$\lambda_u = \lim_{u \to 1-} P(U_1 > u | U_2 > u) = \lim_{u \to 1-} \frac{C^*(u,u)}{1-u}. \tag{2.30}$$

Similarly, the lower tail dependence coefficient $\lambda_l \in [0,1]$ is defined as

$$\lambda_l = \lim_{u \to 0+} P(U_1 \leq u | U_2 \leq u) = \lim_{u \to 0+} \frac{C(u,u)}{u}. \tag{2.31}$$

If the upper tail dependence coefficient λ_u falls into the interval $(0,1]$, then U_1 and U_2 are said to be asymptotically dependent in the upper tail, and if $\lambda_u = 0$, then U_1 and U_2 are said to be asymptotically independent in the upper tail. Similarly, if $\lambda_l \in (0,1]$ or $\lambda_l = 0$, then U_1 and U_2 are said to be asymptotically dependent or independent, respectively, in the lower tail. For properties of the lower and the upper tail dependence coefficients the reader is referred to Embrechts, McNeil & Straumann (2001) and McNeil, Frey & Embrechts (2005). Hu (2006) reviews dependence and tail dependence measures for mixture copula models. The following result for the Student-t copula, as well as the derivations for other copula models can be found in Embrechts, McNeil & Straumann (2001) and McNeil, Frey & Embrechts (2005).

Copula Examples

This section will concentrate mostly on two popular copula families: the elliptical copula family and the Archimedean copula family. Some d-dimensional copula from these parametric copula families, with copula parameter controlling the degree of dependence, are presented below. Further copula models, in particular, Ali-Mikhail-Haq and Plakett copula, can be found e.g. in Joe (1993) and Nelsen (1998).

Elliptical copula families, see Lindskog, McNeil & Schmock (2001), have a dependence structure generated by elliptical distributions, which include among others the normal and Student-t distributions, as well as the stable distribution class discussed in e.g. Rachev & Mittnik (2000) and Rachev & Han (2000). For the modeling of dependencies using elliptical distributions one can refer to Hult & Lindskog (2001), Fang, Fang & Kotz (2002) and Frahm, Junker & Szimayer (2003). Its applications in finance and risk management are discussed, for instance, in Breymann, Dias & Embrechts (2003), McNeil, Frey & Embrechts (2005) and Dias & Embrechts (2008). In the following we consider the Gaussian copula and the Student-t copula.

The Archimedean copula family includes, for instance, the Gumbel, Clayton and Frank copula, which have simple closed forms and have been studied e.g. in relation to the modeling of portfolio credit risk in McNeil, Frey & Embrechts (2005), Dias (2004) and Wu, Valdez & Sherris (2007).

In addition to the one-parametric copula, the current section will consider some mixture copula models of the Archimedean copula as introduced in Joe (1993). Here the distribution function has the form of a convex combination of two or more copula. Denoting by C^A and C^B copula with dependence parameters θ_1 and θ_2, respectively, the mixture copula model has the form:

$$C_X(u_1, \ldots, u_d, \theta) = \theta_3 C_X^A(u_1, \ldots, u_d, \theta_1) + (1 - \theta_3) C_X^B(u_1, \ldots, u_d, \theta_2), \qquad (2.32)$$

where $\theta_3 \in [0, 1]$. The following will consider four mixture models studied in Dias (2004). These include Clayton & survival Clayton, Clayton & Gumbel, survival Clayton & survival Gumbel and Gumbel & survival Gumbel copula. Dias (2004), Angel Canela & Pedreira Collazo (2012) and Hu (2006) study mixture models for modeling dependencies across international financial markets.

Gaussian Copula

The Gaussian copula expresses the dependence structure of the multivariate normal distribution, i.e. normal marginal distributions are combined with a Gaussian copula to form a multivariate normal distribution. If $Y_j \sim N(0, 1)$ and $Y = (Y_1, \ldots, Y_d)^\top \sim N_d(0, \Psi)$, where Ψ denotes a correlation matrix, then an explicit expression for the Gaussian copula is given by

$$C_\Psi^{Ga}(u_1, \ldots, u_d) = F_Y\{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\} \qquad (2.33)$$

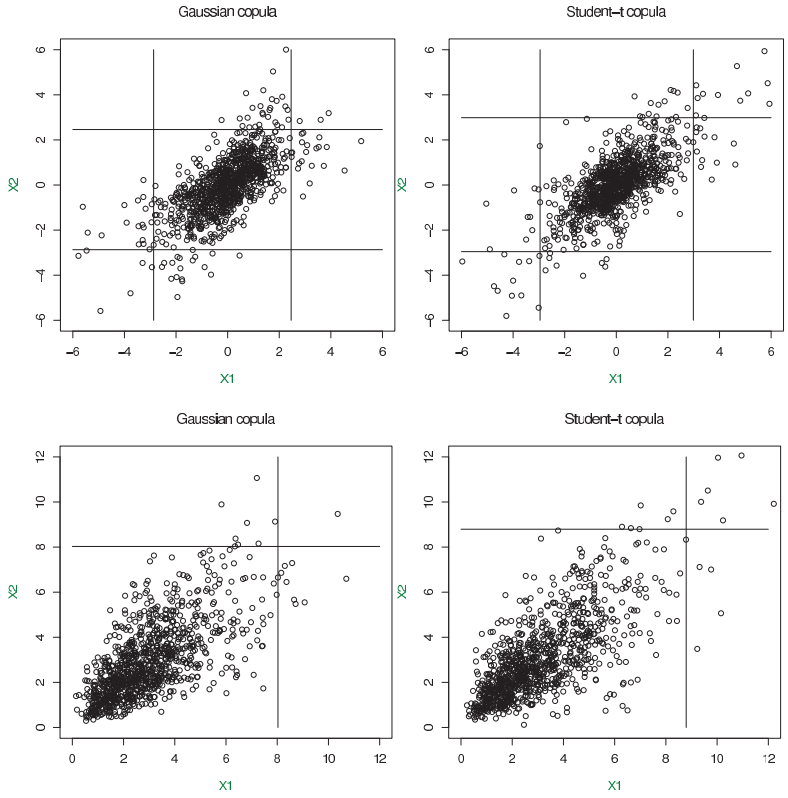

Figure 2.17: 1000 bivariate realizations simulated from the Gaussian copula (left panel) and Student-t copula (right panel) with identical Student-t marginal distributions with four degrees of freedom (upper panel) and Gamma(3,1) marginal distributions (lower panel), and identical correlation $\theta = 0.7$

$$= \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_d)} 2\pi^{-\frac{d}{2}} \mid \Psi \mid^{-\frac{1}{2}} \exp\left(-\frac{1}{2} r^\top \Psi^{-1} r\right) dr_1 \dots dr_d.$$

Defining $\zeta_j = \Phi^{-1}(u_j)$, $\zeta = (\zeta_1, \dots, \zeta_d)^\top$, the density of the Gaussian copula can be written as

$$c_\Psi^{Ga}(u_1, \dots, u_d) = \mid \Psi \mid^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\zeta^\top(\Psi^{-1} - \mathcal{I}_d)\zeta\right\}. \tag{2.34}$$

Student-t Copula

The Student-t copula inherits the dependence structure from the multivariate Student-t distribution. Let $X = (X_1, \dots, X_d)^\top \sim t_d(\nu, \mu, \Sigma)$ have a multivariate Student-t distribution with ν degrees of freedom, mean vector μ and positive-definite dispersion or scatter matrix Σ. The copula remains invariant under a standardization of the marginal distributions[1]. This means that the copula of a $t_d(\nu, \mu, \Sigma)$ distribution

[1]In fact, it remains invariant under any series of strictly increasing transformations of the components of the random vector X; see Nelsen (1998).

is identical to that of a $t_d(\nu, 0, \Psi)$ distribution, where Ψ is the correlation matrix associated with Σ. The unique Student-t copula is the copula $C_X = C^t_{\nu,\Psi}$. For $u = (u_1, \ldots, u_d)^\top \in [0,1]^d$, it is given by

$$C^t_{\nu,\Psi}(u_1, \ldots, u_d) = t_{\nu,\Psi}\{t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)\}, \tag{2.35}$$

where t_ν^{-1} is the quantile function of the univariate t-distribution with ν degrees of freedom. The density of the t-copula is given as

$$c^t_{\nu,\Psi}(u_1, \ldots, u_d) = \frac{t_{\nu,\Psi}\{t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)\}}{\prod_{j=1}^d t_{\nu,\Psi}\{t_\nu^{-1}(u_j)\}}. \tag{2.36}$$

With $\zeta_j = t_\nu^{-1}(u_j)$ the density of the t-copula can be expressed as

$$c^t_{\nu,\Psi}(u_1, \ldots, u_d) = \mid \Psi \mid^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+d}{2}) \left\{\Gamma(\frac{\nu}{2})\right\}^{d-1} \left(1 + \frac{1}{\nu}\zeta^\top \Psi^{-1} \zeta\right)^{-\frac{\nu+d}{2}}}{\left\{\Gamma(\frac{\nu+1}{2})\right\}^d \prod_{j=1}^d \left(1 + \frac{1}{\nu}\zeta_j^2\right)^{-\frac{\nu+1}{2}}}. \tag{2.37}$$

The Student-t copula generates symmetric tail dependence. Its tail dependence coefficients, as defined in equations (2.30) and (2.31), take the following relatively simple form:

$$\lambda_u = \lambda_l = 2t_{\nu+1}\sqrt{(\nu+1)(1-\rho)/(1+\rho)}. \tag{2.38}$$

Here t_ν denotes the Student-t distribution function, ν is the number of degrees of freedom, and ρ is the correlation coefficient. Figure 2.17 shows 1000 bivariate realizations simulated using two different dependence structures. Both models use identical marginal distributions, the Gamma(3,1) distribution in the upper panel and the Student-t distribution with four degrees of freedom in the lower panel, where both have identical dependence coefficient $\theta = 0.7$. However, the Gaussian copula is applied to model the dependency in the left panel and the Student-t copula is used to model the dependency in the right panel. We observe that the second model favors the simultaneous occurrence of extreme values, which is from the point of view of risk management the more dangerous case. Since extreme losses have a tendency to occur simultaneously in log-returns, it appears also to be more realistic just from this observation alone. However, the standard RiskMetrics methodology, which assumes a Gaussian copula, assumes only very few extreme losses to appear jointly, similar as shown in the upper left panel of Figure 2.17.

Clayton Copula

The Clayton copula with the dependence parameter $\theta \in (0, \infty)$ is defined as

$$C_\theta(u_1, \ldots, u_d) = \left\{ \left(\sum_{j=1}^{d} u_j^{-\theta} \right) - d + 1 \right\}^{-1/\theta} \tag{2.39}$$

with density:

$$c_\theta(u_1, \ldots, u_d) = \prod_{j=1}^{d} \{1 + (j-1)\theta\} u_j^{-(\theta+1)} \left(\sum_{j=1}^{d} u_j^{-\theta} - d + 1 \right)^{-(1/\theta + d)} . \tag{2.40}$$

As the copula parameter θ tends to infinity, the dependence becomes maximal and as θ tends to zero, one has independence. The Clayton copula can mimic lower tail dependence but no upper tail dependence.

Gumbel Copula

The Gumbel copula with the dependence parameter $\theta \in [1, \infty)$ is of the form:

$$C_\theta(u_1, \ldots, u_d) = \exp\left[- \left\{ \sum_{j=1}^{d} (-\log u_j)^\theta \right\}^{1/\theta} \right]. \tag{2.41}$$

For $\theta > 1$ this copula generates an upper tail dependence, while for $\theta = 1$ it reduces to the product copula (i.e. independence): $C_\theta(u_1, \ldots, u_d) = \prod_{j=1}^{d} u_j$. Maximal dependence is achieved when θ tends to infinity.

2.6 Copula Estimation

The following section focuses on the estimation of parametric copula, including the estimation of the marginal parameters, as well as the dependence structure given by the dependence parameter. Consider a vector of random variables $X = (X_1, ..., X_d)^\top$ with parametric univariate marginal distributions $F_{X_j}(x_j, \delta_j)$, $j \in \{1, ..., d\}$. Furthermore, let a copula belong to a parametric family $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$. From Sklar's Theorem the distribution of X can be expressed as

$$F_X(x_1, \ldots, x_d) = C\{F_{X_1}(x_1; \delta_1), \ldots, F_{X_d}(x_d; \delta_d); \theta\} \tag{2.42}$$

and its density as

$$f(x_1, \ldots, x_d; \delta_1, \ldots, \delta_d, \theta) = c\{F_{X_1}(x_1; \delta_1), \ldots, F_{X_d}(x_d; \delta_d); \theta\} \prod_{j=1}^{d} f_j(x_j; \delta_j), \tag{2.43}$$

where

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d} \tag{2.44}$$

is a copula density. For a sample of observations $\{x_t\}_{t=1}^T$, where $x_t = (x_{1,t}, \ldots, x_{d,t})^\top$, and a vector of parameters $\alpha = (\delta_1, \ldots, \delta_d, \theta)^\top \in \mathbb{R}^{d+1}$ the likelihood function is given by the product

$$L(\alpha; x_1, \ldots, x_T) = \prod_{t=1}^{T} f(x_{1,t}, \ldots, x_{d,t}; \delta_1, \ldots, \delta_d, \theta) \tag{2.45}$$

and the corresponding log-likelihood function is the sum

$$\begin{aligned} \ell(\alpha; x_1, \ldots, x_T) &= \sum_{t=1}^{T} \ln \left[c\{F_{X_1}(x_{1,t}; \delta_1), \ldots, F_{X_d}(x_{d,t}; \delta_d); \theta\} \right] \\ &\quad + \sum_{t=1}^{T} \sum_{j=1}^{d} \ln \left[f_j(x_{j,t}; \delta_j) \right]. \end{aligned} \tag{2.46}$$

The objective is to maximize this log-likelihood function numerically. The estimation can be performed, for instance, in the following three different ways, employing the exact maximum likelihood (EML), the inference for marginals (IFM) and the canonical maximum likelihood (CML) method.

Exact Maximum Likelihood (EML)

The exact maximum likelihood (EML) method is well-known and straightforward. It estimates the parameter α in one step through maximizing the log-likelihood function $\ell(\alpha)$. The estimates $\tilde{\alpha}_{EML} = (\tilde{\delta}_1, \ldots, \tilde{\delta}_d, \tilde{\theta})^\top$ solve then the first order condition

$$(\partial \ell / \partial \delta_1, \ldots, \partial \ell / \partial \delta_d, \partial \ell / \partial \theta) = 0. \tag{2.47}$$

The drawback of the EML method is that with an increasing scale of the problem the algorithm becomes computationally challenging.

Inference for Marginals (IFM)

In the inference for marginals (IFM) method parameters for marginals and copula are estimated separately, which represents a sequential two-step maximum likelihood method; see e.g. McLeish & Small (1988) and Joe (1997). The parameters δ_j from the marginal distributions are estimated in the first step and the dependence

parameter θ is estimated in the second step after the estimated marginal distributions have been substituted into the copula. For $j \in \{1, \ldots, d\}$ the log-likelihood function for each of the marginal distributions is given by

$$\ell_j(\delta_j) = \sum_{t=1}^{T} \ln\left(f_j\left(x_{j,t}; \delta_j\right)\right) \tag{2.48}$$

and the estimated marginal parameter obtained as $\ell_j(\delta_j)$. The pseudo log-likelihood function

$$\ell(\theta, \hat{\delta}_1, \ldots, \hat{\delta}_d) = \sum_{t=1}^{T} \ln\Big(c\{F_{X_1}(x_{1,t}; \hat{\delta}_1), \ldots, F_{X_d}(x_{d,t}; \hat{\delta}_d); \theta\}\Big) \tag{2.49}$$

is maximized over θ to obtain the estimator $\hat{\theta}$ for the dependence parameter θ. The estimates $\hat{\alpha}_{IFM} = (\hat{\delta}_1, \ldots, \hat{\delta}_d, \hat{\theta})^\top$ solve the first order condition

$$(\partial\ell_1/\partial\delta_1, \ldots, \partial\ell_d/\partial\delta_d, \partial\ell/\partial\theta) = 0. \tag{2.50}$$

Canonical Maximum Likelihood (CML)

In contrast to the EML and IFM methods, where one has to make an assumption about the parametric form of the marginal distributions, the canonical maximum likelihood (CML) method maximizes the pseudo log-likelihood function with empirical marginal distributions in the form:

$$\ell(\theta) = \sum_{t=1}^{T} \ln\Big(c\{\hat{F}_{X_1}(x_{1,t}), \ldots, \hat{F}_{X_d}(x_{d,t}); \theta\}\Big). \tag{2.51}$$

Here the empirical marginal cumulative distribution function is given by

$$\hat{F}_{X_j}(x) = \frac{1}{T+1} \sum_{t=1}^{T} \mathbf{1}_{\{X_{j,t} \leq x\}}, \tag{2.52}$$

see Genest & Rivest (2002). Using this method, the parameter can be estimated in one step by using the estimate $\ell(\theta)$.

Modeling Dependency by Dynamic Copulae

As argued in Platen & Stahl (2003b), a diversified world stock index is very suitable for measuring the general market performance and general market risk; see Basel (1996) as well as more recent documents on Besel II and Solvency II. Such indices are widely used as benchmarks when evaluating investment strategies. The

study of diversified world stock indices is motivated by the benchmark approach; see Platen & Heath (2010), which puts a well diversified portfolio into the center of financial modeling. By diversification theorems in Platen (2005b) and Platen & Rendek (2012b), a diversified portfolio can be interpreted as a good proxy for the growth optimal portfolio (GOP), which is in many ways the best performing portfolio, see also Chapter 3. As mentioned earlier, it can be characterized as the portfolio maximizing the expected log-utility from terminal wealth. It is the portfolio with almost surely maximal long-term growth rate. It outperforms pathwise in the long run almost surely the trajectory of any other strictly positive portfolio. The equally weighted index of a given investment universe can be regarded, under minor regularity assumptions, as a good proxy for the GOP, as will be shown in the next chapter.

Data

An equally weighted index is an almost ideally diversified index, all fractions are equal, that is, $\pi_{\delta_{EWI},t} = \frac{1}{d}$ with $\sum_{j=1}^{d} \pi_{\delta_{EWI},t} = 1$, $j \in \{1, 2, ..., d\}$. The EWI104s used previously, is also employed in the following analysis. It is based on 104 world industry sector indices as constituents that are provided by Datastream; see Le & Platen (2006) and Platen & Rendek (2008). Figure 2.3 plots the EWI104s denominated in USD for the time period from 02 January, 1973 to 10 March, 2006. The analysis of this section uses 20 currency denominations, the respective currencies are determined by the availability of currency data for the entire period. These denominations include the currencies of the following countries: Australia(Aus), Austria(Au), Belgium(Be), Canada(Ca), Denmark(Dk), Finland(Fi), France(Fr), Germany(Ge), India(Ind), Ireland(Ire), Italy(It), Japan(J), Malaysia(Ma), Netherlands(Nth), Norway(Nor), Portugal(Por), Spain(Sp), Sweden(Swe), UK, USA.

Fitting Static and Time-varying Copulae

The following study aims to fit a parametric copula, that is, it estimates the copula dependence parameter, assuming that the unknown marginal distributions are nuisance parameters[2]. Following Section 2.3, we assume that the marginals have a

[2]For the EML and IFM method one needs to make an assumption about the parametric form of the marginal distributions, whereas CML uses empirical marginal distributions.

Copula model	mean($\hat{\theta_1}$)($\hat{s.e.}$)	med($\hat{\theta_1}$)	min($\hat{\theta_1}$)	max($\hat{\theta_1}$)	mean($\hat{\theta_2}$)($\hat{s.e.}$)	med($\hat{\theta_2}$)	min($\hat{\theta_2}$)	max($\hat{\theta_2}$)	mean($\hat{\theta_3}$)($\hat{s.e.}$)	med($\hat{\theta_3}$)	min($\hat{\theta_3}$)	max($\hat{\theta_3}$)	AIC
Clayton	3.278(2.437)	2.196	1.265	19.283	-	-	-	-	-	-	-	-	-7767.55(10)
Frank	10.96(6.824)	8.007	5.040	56.636	-	-	-	-	-	-	-	-	-8925.03(7)
Gaussian	0.810(0.082)	0.784	0.647	0.9899	-	-	-	-	-	-	-	-	-7980.49(9)
Gumbel	3.243(1.743)	2.477	1.833	15.340	-	-	-	-	-	-	-	-	-9063.73(6)
Plakett	814.9(1541)	23.61	9.984	4376.4	-	-	-	-	-	-	-	-	-8054.06(8)
Student-t	0.831(0.092)	0.809	0.631	0.9930	-	-	-	-	-	-	-	-	**-10713.2(1)**
Clayton & surv.Clayton	8.630(19.59)	3.296	0.698	182.5	6.720(14.46)	3.560	0.634	102.43	0.534(0.075)	0.529	0.253	0.732	-10130.44(5)
Clayton & Gumbel	1.306(1.191)	0.955	0.264	8.263	7.085(9.693)	4.118	1.010	94.970	0.320(0.100)	0.314	0.046	0.728	-10384.80(4)
surv. Clayton & surv. Gumbel	1.168(2.003)	0.608	0.010	15.36	6.344(6.278)	3.803	2.640	50.000	0.277(0.108)	0.279	0.021	0.990	-10406.78(3)
Gumbel & surv. Gumbel	9.204(11.31)	4.689	1.415	50.00	6.720(14.46)	3.560	0.634	102.43	0.534(0.075)	0.529	0.253	0.732	**-10673.96(2)**

Table 2.6: An average over 190 pairs of currency denominations estimating the copula dependence parameter with standard errors

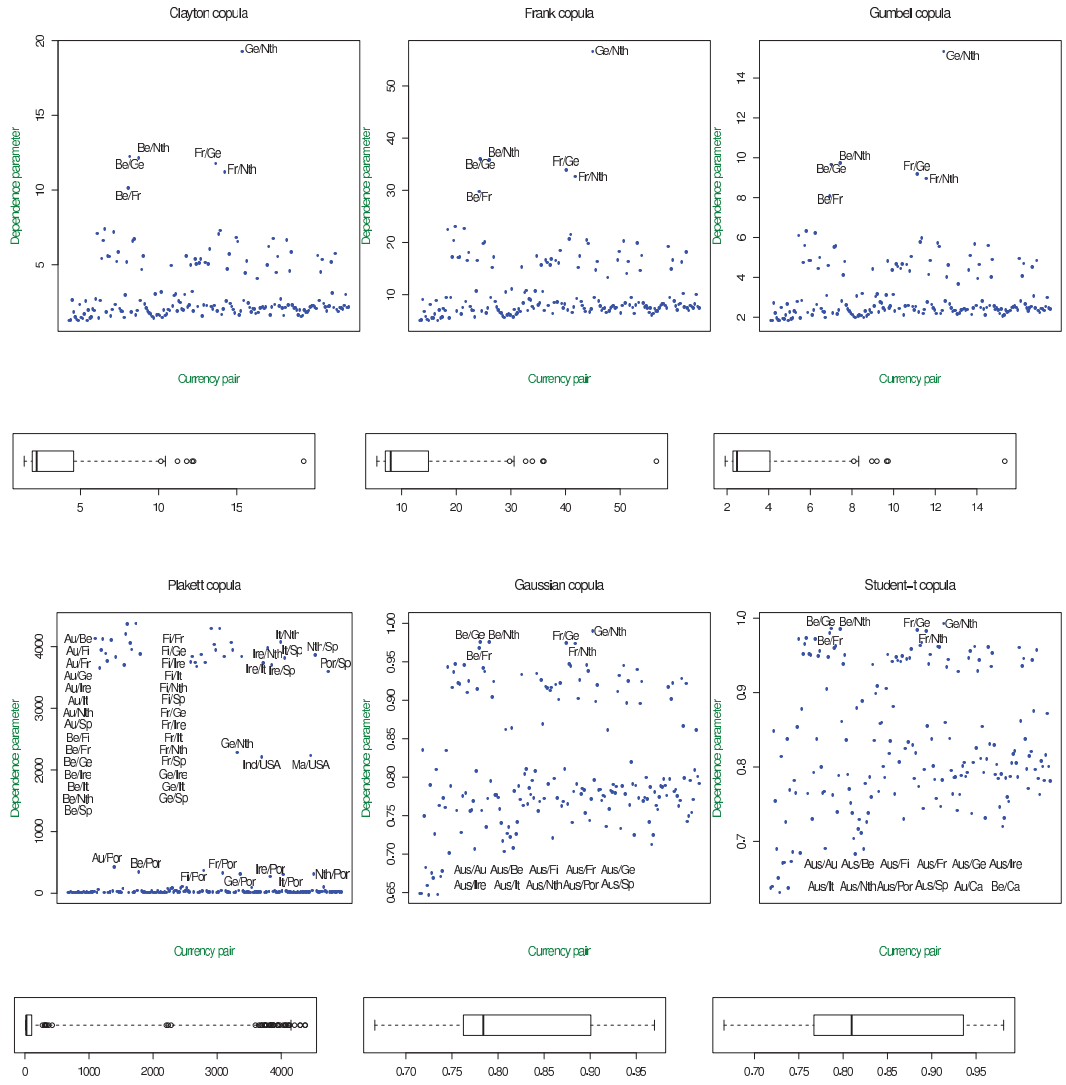

Figure 2.18: Estimated copula dependence parameter for 190 pairs of currency denominations and different copula models, together with box-plots

Student-t distribution. The estimation of the dependence parameter of the Student-t copula proceeds as follows: First, we transform the original data to the "copula scale" by applying a probability integral transform to obtain uniformly $[0, 1]$-distributed values. Then we apply the IFM method to estimate different copula assuming that the marginals are Student-t. The estimated number of degrees of freedom for the marginals is taken from Table 2.3 for each of the currency denominations of the EWI104s index.

A static copula is assumed to estimate the average dependence parameter using log-return data from the time interval covering the entire time period from 02 January, 1973 to 10 March, 2006. This section fixes different one parametric copula

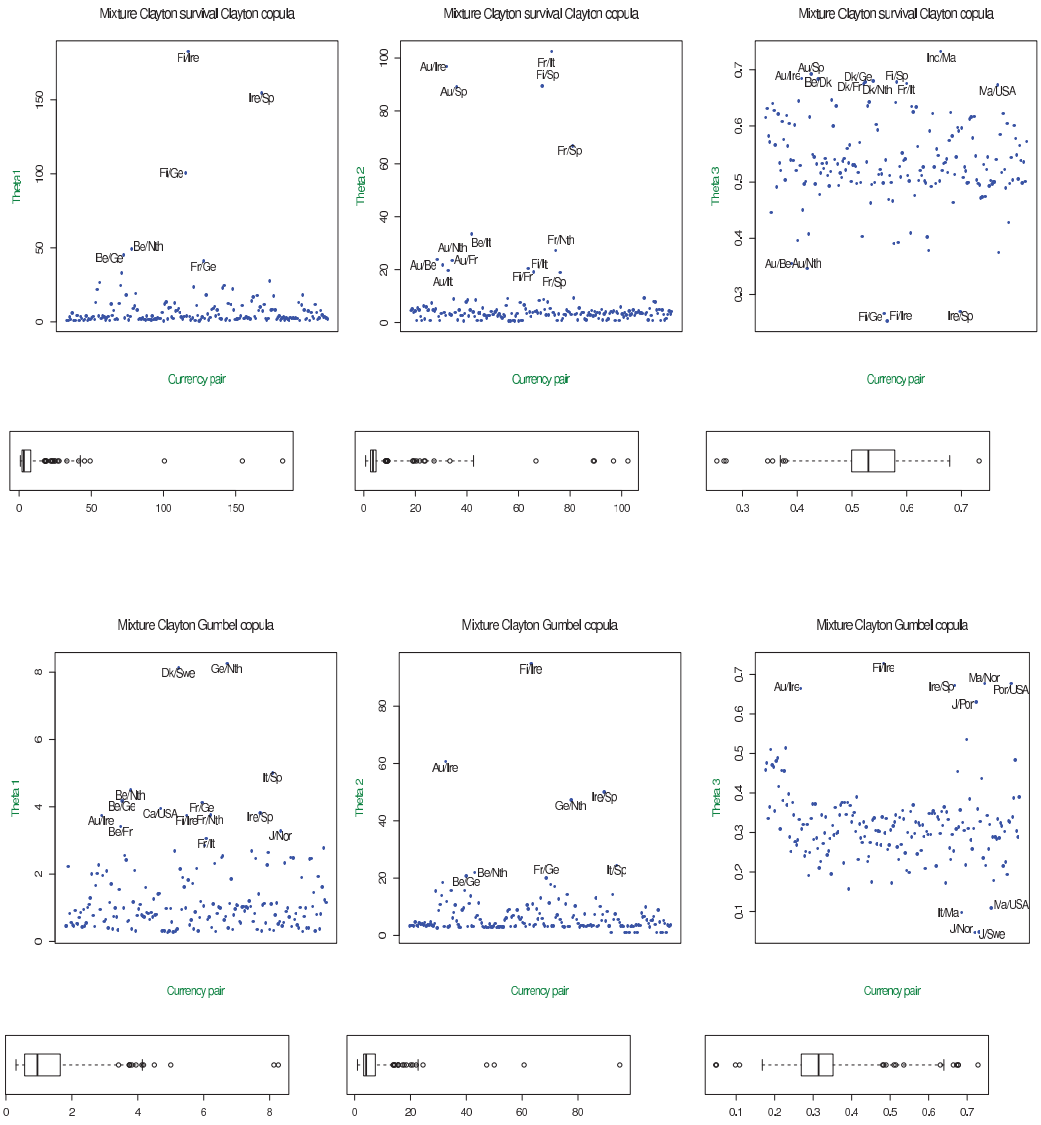

Figure 2.19: Estimated copula dependence parameters for 190 pairs of currency denominations and two mixture copula models: mixture Clayton & survival Clayton and mixture Clayton & Gumbel

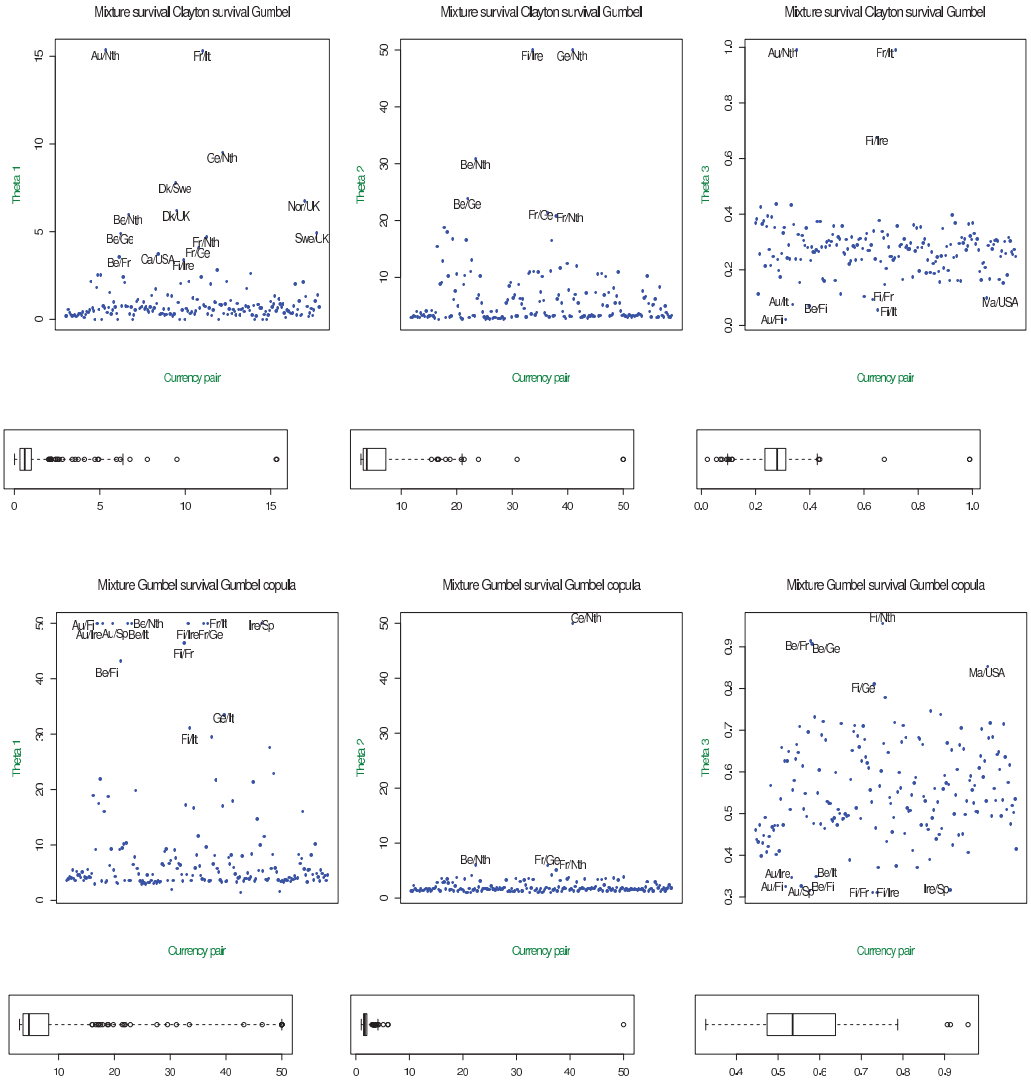

Figure 2.20: Estimated copula dependence parameters for 190 pairs of currency denominations and two mixture copula models: mixture survival Clayton & survival Gumbel and mixture Gumbel & survival Gumbel

models, as well as mixture copula models summarized in Table 2.6, and uses the 20 different currency denominations listed above. It obtains 190 estimated copula of pairs for each of the following copula models: Clayton, Frank, Gaussian, Gumbel, Plakett, Student-t, and mixture models Clayton & survival Clayton, Clayton & Gumbel, survival Clayton & survival Gumbel, Gumbel & survival Gumbel. Figure 2.18 shows the estimated dependence parameter for each pair of currency denominations and one-parametric copula models. Figures 2.19 and 2.20 show the estimated copula dependence parameters for 190 pairs of currency denominations and different mixture copula models. Parameters θ_1 and θ_2 plotted in the first and the second panels are the dependence parameters for the first and second terms of the mixture, respectively, and the mixture parameter θ_3 plotted in the third panel gives the proportion of the first term. Box-plots exhibited on the bottom of each plot show the extreme of the lower whisker (2.5%-quantile), low quartile, median, upper quartile and the extreme of the upper whisker (97.5%-quantile) of the estimated parameters. Not surprising, one observes high dependence between European countries and lower dependence between countries on different continents. Visually, the Student-t copula in Figure 2.18 shows rather evenly distributed estimated dependence parameters. Also the mixture Gumbel survival Gumbel copula in Figure 2.20 shows such behavior. The average estimated copula dependence parameters taken over 190 currency pairs together with their standard errors, minimum and maximum values are given in the first three columns of Table 2.6. In the case of mixture models, the parameters θ_1 and θ_2 are the dependence parameters for the first and second terms of the mixture, respectively, and θ_3 is the mixture parameter which gives the proportion of the first term.

To judge the performance of each model fitted, the Akaike information criterion (AIC) is provided, see Akaike (1974):

$$AIC = -2L(\alpha; x_1, \ldots, x_T) + 2q,$$

where q is the number of parameters of the family of distributions fitted. AIC combines two components: the log-likelihood L, which measures the goodness-of-fit, and a penalty term $2q$, accounting for model complexity. The better the model fits the data, the larger is the likelihood. That means, smaller values of AIC indicate a better fit. In the last column of Table 2.6 all models are ranked by their AIC

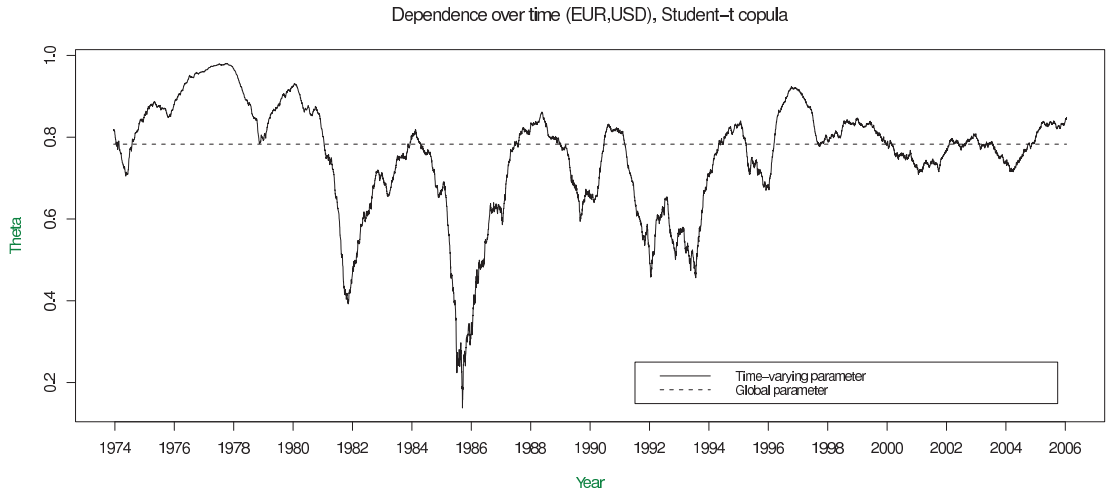

Figure 2.21: Copula dependence parameter for denominations of the EWI104 in Euro and USD; estimated using a Student-t copula with Student-t marginals for the log-return data from 02 January, 1973 to 10 March, 2006

(model ranking is given in parentheses). One observes that the Student-t copula provides the best fit across different copula models, followed by the mixture Gumbel & survival Gumbel copula. This coincides also with our previous visual impression that these copula can well identify the specifics of the joint distributions observed.

Additionally, the time-varying dependence parameter is estimated using log-returns corresponding to a moving window with a fixed size of n=250, corresponding to one year of observations. A sub-sample of log-returns $\{\widehat{X}_t\}_{t=s-n+1}^{s}$ is scrolling in time for $s \in \{n, ..., T\}$ generating a time-series for the dependence parameter $\{\widehat{\theta}_t\}_{t=n}^{T}$. The static case, on the contrary, estimates the dependence parameter based on the whole series of observations. Figure 2.21 displays the estimated time-varying dependence parameter between the denominations of the EWI104s in Euro and US Dollar, estimated using the IFM method for a Student-t copula with Student-t marginals. The dashed line corresponds to the static case. One observes that the dependence structure between denominations of the EWI104s in different currencies is not constant but varies over time. In addition, Figure 2.22 plots the 250 observations used for the estimation from the moving window corresponding to maximal dependence on 11 October, 1977 (left panel) and minimal dependence on 30 September, 1985 (right panel). One notes that just before the 1987 crash the dependence was rather low but increased dramatically with the crash. These type of empirical findings are important for modeling diversified indices in continuous time. Advanced models

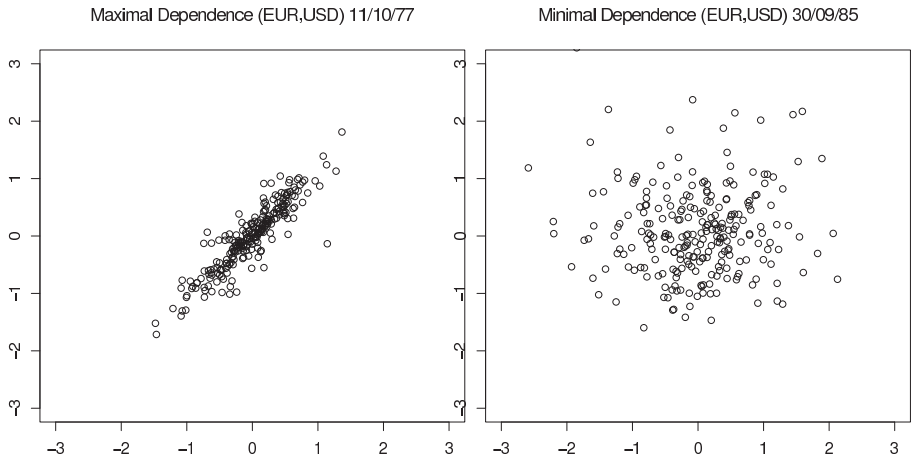

Figure 2.22: Log-returns of denominations of the EWI104s in Euro and USD at maximal dependence on 11 October, 1977 (left panel) and minimal dependence on 30 September, 1985 (right panel)

should reflect this type of dependency in a dynamic manner.

Chapter 3

Approximating the Numéraire Portfolio by Naive Diversification

As we will demonstrate in Chapter 4, the classical no-arbitrage modeling framework is too narrow to capture realistically the long term dynamics of a well diversified index in currency denomination. The more general benchmark approach, with its central building block, the numéraire portfolio (NP) or growth optimal portfolio (GOP), provides an adequate modeling framework also over long periods of time. Within this section we approximate the growth optimal portfolio (GOP), which is also the numéraire portfolio (NP); see Platen & Heath (2010). We significantly improve the performance of the resulting proxies of the NP when compared with results in Le & Platen (2006) and Platen & Rendek (2008) presented in Section 2.1. To identify the exact NP one would need to have a perfect model and needs to perfectly estimate rates of return and covariance of returns for thousands of assets. However, estimation theory has shown, due to the limited estimation window available for real asset data, the sample based Markowitz mean-variance approach produces unreliable weights which fluctuate substantially over time. This chapter proposes an alternate approach to portfolio optimization, being the use of naive diversification to approximate the NP. The NP is the strictly positive portfolio that, when used as benchmark, makes all benchmarked nonnegative portfolios supermartingales. Furthermore, it is the GOP and, therefore, maximizes expected logarithmic utility and outperforms any other strictly positive portfolio in the long run. This chapter proves theoretically and demonstrates empirically for the well-securitized global equity market that the naive equal value weighted portfolio converges to the numéraire portfolio when the number of constituents tends to infinity. This result is model independent and,

therefore, very robust. The systematic construction of diversified stock indices by naive diversification from real data is demonstrated. Even when taking transaction costs into account, these indices significantly outperform the corresponding market capitalization weighted indices in the long run, indicating empirically their asymptotic proximity to the numéraire portfolio. Finally, in time of financial crisis, a large equi-weighted fund, carrying the investments of major pension funds and insurance companies, would provide important liquidity to the market. It would not only dampen the drawdown of the market capitalization weighted index in a crisis but would also moderate the excesses of an asset price bubble when the market is overheated. This chapter is based on the two papers: Platen & Rendek (2012b) and Platen & Rendek (2012c).

3.1 Diversified Indices

From an economic point of view it is intuitively appealing that only the nondiversifiable risk should attract a risk premium. This view has been supported by variants of the ICAPM, see e.g. Merton (1973), where the market capitalization weighted index (MCI) carries the nondiversifiable risk. This chapter systematically employs naive diversification, and captures in this way via the equi-weighted index (EWI) nondiversifiable risk. It demonstrates empirically that EWIs perform extremely well, and appear to approach the numéraire portfolio (NP) asymptotically, as the number of constituents increases. The Naive Diversification Theorem, presented in Section 3.3, will explain theoretically the approximation of the NP by EWIs.

Industry Classification Benchmark

To be systematic in capturing the nondiversifiable risk of the global equity market, the diversifiable risk is systematically removed according to the economic structure of the market. The index constructions performed are based on common types of economic activity, and use accordingly an industry classification, more precisely, the Industry Classification Benchmark (ICB); see ICB (2008). This classification provides a comprehensive structure for the analysis of companies and industries, and has been jointly created by the FTSE and Dow Jones. Four layers of classification for each company are provided. More precisely, the ICB distinguished around

2010/2011 10 industries, 19 supersectors, 41 sectors and 114 subsectors; covering all major industries represented in the world stock market. For instance, industries in the ICB structure include: Oil & Gas, Basic Materials, Industrials, Consumer Goods, Health Care, Consumer Services, Telecommunications, Utilities, Financials and Technology. The total number of companies covered by the ICB classification system exceeds $60,000$. Altogether, companies from 72 countries are currently classified. The ICB classifies a company as belonging to a subsector by using the subsector definition that most accurately describes the nature of its business as determined by its largest source of revenue. This chapter argues that the ICB classification system is well suited for identifying the nondiversifiable risk generated by the main types of economic activity in the world economy. Note that other similar industry classification systems are likely to provide analogous results. We employ the ICB structure as it was used by Thomson Reuters Datastream (Datastream), when evaluating their own industry, supersector, sector and subsector indices.

Datastream Indices

A representative sample of stocks is chosen by Datastream in each country, for each of the industry subsectors applicable to that country, covering 75 to 80 per cent of the total market capitalization. Moreover, the number of stocks in each subsector is determined by the size of that market sector. The inclusion of a stock by Datastream depends on its total market value and the availability of data. The Datastream subsector and industry indices are updated regularly. This ensures that these indices include the top stocks by market capitalization and reflect new industries and subsectors appropriately. Delisted stocks are removed from an index when notification of a delisting is received. Temporarily suspended stocks remain in an index unless it is believed that the suspension is going to be long term, in which case they are removed from an index. Datastream has a well-defined set of rules that determines the inclusion of stocks in their indices.

The Datastream database is hierarchical and self-contained. For instance, the world Oil & Gas subsector index was constructed from 52 country Oil & Gas subsector indices as at the end of April 2011. Furthermore, each country Oil & Gas index is available together with its constituents. Moreover, the database includes fixed index datatypes and recalculated index datatypes. Fixed index datatypes are

not recalculated historically when constituents change, allowing the effects of dead stocks to be observed. The recalculated index datatypes reflect the performance of current constituents, therefore, they avoid any distortion due to stocks entering or leaving an index. Datastream does not provide constituent histories for their sector indices. Consequently, to construct indices that avoid survival bias, this chapter uses the fixed datatype, country industry subsector indices as constituents. There are 1969 country industry subsector indices offered by Datastream, which are total return MCIs. When forming comparable MCIs and EWIs, the same set of stocks is used, so that the effect of naive diversification can be gauged realistically.

Index Construction

The data are observed daily. Denote by t_n, $n \in \{1, 2, \dots\}$, the time of the nth observation. Let $S_{t_n}^{j,k}$ denote the value of the jth industry, supersector, sector or subsector index for the kth country at time t_n, as given by Datastream. The number of constituents at the dth classification level, $d \in \{1, 10, 19, 41, 114\}$, available in the database equals $\sum_{j=1}^{d} \ell_{d,j}$, where $\ell_{d,j}$ denotes the number of countries for which the jth industry, supersector, sector or subsector was available.

This chapter constructs various diversified portfolios, where it applies naive diversification in a systematic way. In a first study it forms five indices $S_{t_n}^{\delta_{EWId}}$, with equally weighted constituents. The value of the equi-weighted index, EWId, having d constituents, $d \in \{1, 10, 19, 41, 114\}$, is obtained by the following recursive formula:

$$S_{t_n}^{\delta_{EWId}} = S_{t_{n-1}}^{\delta_{EWId}} \left(1 + \sum_{j=1}^{d} \sum_{k=1}^{\ell_{d,j}} \pi_{\delta_{EWId}, t_{n-1}}^{j,k} \frac{S_{t_n}^{j,k} - S_{t_{n-1}}^{j,k}}{S_{t_{n-1}}^{j,k}} \right), \tag{3.1}$$

with initial value $S_0^{\delta_{EWId}} = 100$ and portfolio weights $\pi_{\delta_{EWId}, t_{n-1}}^{j,k} = \frac{1}{d\ell_{d,j}}$, $j \in \{1, 2, \dots, d\}$, $k \in \{1, 2, \dots, \ell_{d,j}\}$, chosen at time t_{n-1}, for those countries which have constituents of this type. Recall that $d \in \{1, 10, 19, 41, 114\}$ refers to the number of constituents available at the corresponding classification level according to the ICB structure, and $\ell_{d,j}$ is the corresponding number of countries. Note that the number of constituents in the above indices can change over time. This has been accommodated in the index construction by the use of time dependent fractions.

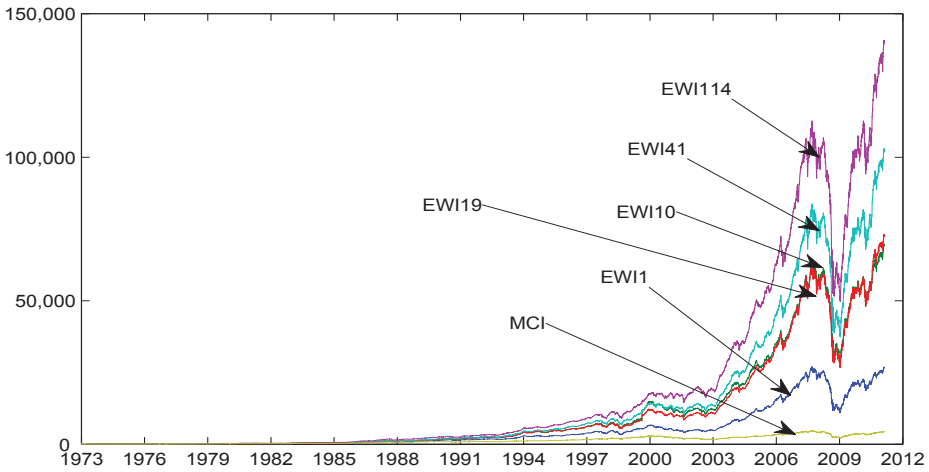

Figure 3.1: The MCI and five equi-weighted indices: EWI1 (market), EWI10 (industry), EWI19 (supersector), EWI41 (sector), EWI114 (subsector).

Index Comparison

The resulting family of indices, denominated in US dollars, is shown in Fig. 3.1 for the period from January 1973 until April 2011, and is explained further below. For $d = 1$ the EWI1 weights equally the 54 country market capitalization-weighted indices, as provided by Datastream.

For $d = 10$, the ICB distinguishes between 10 industries. Here one starts with the market capitalization-weighted country industry indices, as provided by Datastream. For each industry, each country industry index is then equally weighted to yield a corresponding world industry index. Then, by equally weighting the resulting 10 world industry indices, the EWI10 with value $S_{t_n}^{\delta_{EWI10}}$ is obtained.

Analogously, the case $d = 19$ considers 19 ICB supersectors and starts with the corresponding market capitalization-weighted country supersector indices. By equally weighting the country supersectors, one obtains world supersector indices, which are then equally weighted to yield the EWI19. Furthermore, the case $d = 41$ uses 41 sectors and provides the EWI41. Finally, the case $d = 114$ builds on all 114 subsectors, first weighting equally all country subsector indices. It then equally weights the resulting world subsectors to yield the EWI114 with value $S_{t_n}^{\delta_{EWI114}}$. Fig. 3.1 also displays the US dollar value of the MCI, which represents here the Datastream fixed datatype market capitalization weighted total return world stock index.

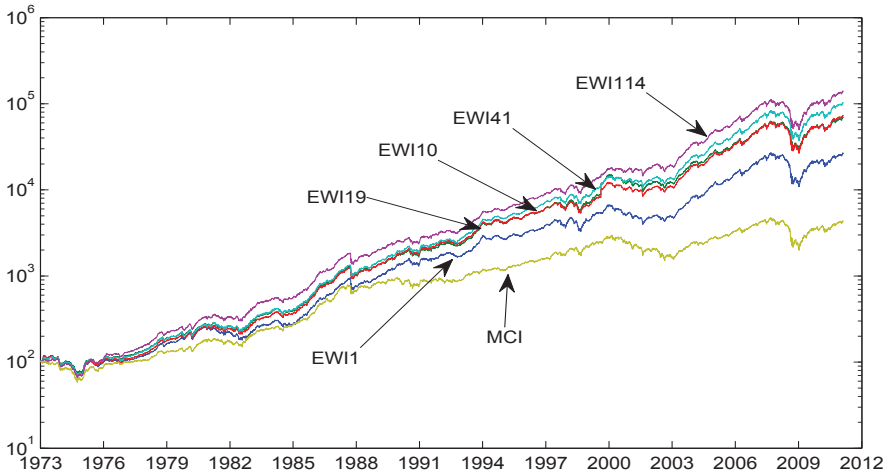

Figure 3.2: The MCI and five equi-weighted indices in log-scale: EWI1 (market), EWI10 (industry), EWI19 (supersector), EWI41 (sector), EWI114 (subsector).

In the construction of the MCI the specific risks of large industries and large economies enjoy substantial weight. The MCI performs worst among the indices exhibited in Fig. 3.1. A better display of earlier periods and a visualization of the average long term growth is given by the logarithms of the above described indices, exhibited in Fig. 3.2. One may conjecture that the MCI is exposed to significant diversifiable risks and, therefore, may not yield such a strong average long term growth as some better diversified portfolios. The Naive Diversification Theorem will later provide theoretical support for such intuition. Among the EWIs displayed in Fig. 3.1, the best performing one is the EWI114, and the worst performing one the EWI1. The EWI10, EWI19 and EWI41 are relatively close to each other. One observes that diversifying with more market capitalization weighted constituents over the same investment universe, systematically enhances the average long term growth rate. This is suggesting that smaller fractions achieve visually not only a better diversification effect, but also a higher average long term growth rate. Additionally, one notes from Fig. 3.2 that all EWIs are driven by almost exactly the same continuous uncertainty. This uncertainty can be interpreted as the nondiversifiable risk of the equity market. The MCI is, of course, strongly exposed to this nondiversifiable risk. However, it carries to some extent also the uncertainties of some economies and industries that are heavily weighted by market capitalization. This leads to slightly different fluctuations of the MCI when compared to those of

the EWIs, and, as will be explained by the NDT, to lower long term growth.

3.2 Law of Large Numbers and Diversification

After having observed that naive diversification works well in practice when focusing on the different types of economic activity, it seems now appropriate to aim for a theoretical understanding of the observed diversification effect. Consider a continuous time market which trades nonnegative primary security accounts. These could represent stocks of companies with all dividends reinvested. In this chapter we use total return indicies of industry sectors and other market segments as primary security accounts. As formulated in the introduction, the key assumption of the chapter is that there exists a numéraire portfolio (NP). The following arguments employ this strictly positive portfolio as numéraire or benchmark and denominate all primary security accounts in its units. Denote the value of the jth benchmarked primary security account by $\hat{S}_t^j$, $j \in \{1, 2, \dots\}$. Since the real market has tens of thousands of stocks, this chapter considers an infinite number of benchmarked primary security accounts. This allows the convenient study of the asymptotic behavior of sequences of portfolios with increasing number of constituents.

For a portfolio strategy $\delta = \{\delta_t = (\delta_t^1, \delta_t^2, \dots)^\top, t \geq 0\}$ the quantity δ_t^j describes the number of units of the jth benchmarked primary security account held at time t. The benchmarked value of this portfolio is given by

$$\hat{S}_t^\delta = \sum_{j=1}^{\infty} \delta_t^j \hat{S}_t^j. \tag{3.2}$$

This chapter considers only strategies that are observable under the available information and are such that the corresponding benchmarked portfolio remains finite in finite time. Furthermore, the portfolios are assumed to be self-financing, which means that changes in the portfolio value are due only to changes in the values of the primary security accounts.

To introduce for the following discussion the notion of a martingale, denote by $E_t(X)$ the conditional expectation of a random variable X under the real-world probability measure P, given the information available at time t. If the stochastic process $Y = \{Y_t, t \geq 0\}$ satisfies the equation

$$Y_t = E_t(Y_s) \tag{3.3}$$

for all $s \geq t$, then it is called a martingale, see e.g. Shiryaev (1984). Intuitively, Y_t is the best forecast at time t of all future values of the process Y.

More generally, a process $Z = \{Z_t, t \geq 0\}$ is called a supermartingale if

$$Z_t \geq E_t(Z_s) \tag{3.4}$$

for all $s \geq t$; see e.g. Shiryaev (1984) or Platen & Heath (2010) for details and examples. Intuitively, in the long term the forecast of future values of a supermartingale is downward trending or has no trend. In the latter case, it is a martingale.

This chapter uses the NP as benchmark. A portfolio when denominated in units of the NP is called a benchmarked portfolio. The defining property of the NP is that the values of all benchmarked nonnegative portfolios form supermartingales. This means that the NP performs so well that over any time period the current benchmarked value of any nonnegative portfolio is greater than or equal to any of its expected future benchmarked values. In this simple and specific sense the NP is "best" performing. As in Platen & Heath (2010) one can argue that the above supermartingale property is the central property of a benchmarked portfolio, and several fundamental statements can be directly derived. For example, one can deduce that the path of the NP exceeds asymptotically over time that of any other strictly positive portfolio that starts with the same initial capital; see e.g. Platen & Heath (2010) for a proof of this model independent result.

It should be emphasized that the assumption on the existence of the NP is satisfied for all arbitrage free financial market models. In particular, the condition that a market admits an equivalent risk-neutral probability measure is a much stronger assumption.

Given the fact that the NP is in many ways the "best" performing portfolio, it is of great interest from a practical point of view to have direct access to such a tradeable instrument in the global stock market. As indicated in the introduction, the work by DeMiguel, Garlappi & Uppal (2009) has shown that it is unrealistic to hope to identify parameters that estimate trends in asset prices with any useful accuracy using sample-based estimation methods. On the other hand, Section 3.1 has demonstrated empirically that naively diversified portfolios perform extremely well in the long run. This chapter will now prove under very weak technical assumptions that the equi-weighted index (EWI) approximates the NP if the number

of constituents tends to infinity. This means that there is no need to estimate the parameters of a market model in order to construct a proxy for the NP. Instead, one needs only to construct an EWI with many sufficiently different constituents.

To illustrate the nature of the diversification effect, consider for fixed $t \in [0, \infty)$ and $h > 0$ the simple one period case, where the returns $R^j_{t,h} = \frac{\hat{S}^j_{t+h} - \hat{S}^j_t}{\hat{S}^j_t}$, $j \in \{1, 2, \dots\}$, of the benchmarked primary security accounts $\hat{S}^1, \hat{S}^2, \dots$ are independent with zero conditional mean and finite conditional variance. More precisely, assume that

$$E_t(R^j_{t,h}) = 0 \tag{3.5}$$

and

$$E_t\left(R^j_{t,h} R^i_{t,h}\right) = \begin{cases} (\sigma^j_{t,h})^2 & \text{for} \quad j = i \\ 0 & \text{otherwise,} \end{cases} \tag{3.6}$$

where

$$(\sigma^j_{t,h})^2 \leq \bar{\sigma}^2_{t,h} < \infty \tag{3.7}$$

can be random. In a classical arbitrage-free continuous market, the zero mean property (3.5) reflects the well-known fact, see Long (1990), that benchmarked primary security accounts are martingales. If the benchmarked primary security accounts have independent uncertainties, due to the different economic activities that they model, then the conditions (3.6) and (3.7) are very reasonable.

The return of the benchmarked equi-weighted index $\hat{S}^{\delta_{EWI\ell}}_t$, generated by the first ℓ benchmarked primary security accounts $\hat{S}^1_t, \hat{S}^2_t, \dots, \hat{S}^\ell_t$, is

$$R^{\delta_{EWI\ell}}_{t,h} = \frac{1}{\ell} \sum_{j=1}^{\ell} R^j_{t,h} \tag{3.8}$$

over the given period. Obviously, by (3.5), one has

$$E_t\left(R^{\delta_{EWI\ell}}_{t,h}\right) = 0. \tag{3.9}$$

Given the independence of the returns of the benchmarked primary security accounts, it follows by (3.6) and (3.7) that

$$E_t\left(\left(R^{\delta_{EWI\ell}}_{t,h}\right)^2\right) = \frac{1}{\ell^2} \sum_{j=1}^{\ell} \left(\sigma^j_{t,h}\right)^2 \leq \frac{\bar{\sigma}^2_{t,h}}{\ell}. \tag{3.10}$$

Note that for an increasing number ℓ of constituents, the conditional variance of $R^{\delta_{EWI\ell}}_{t,h}$ vanishes asymptotically as $\ell \to \infty$. Consequently, one obtains zero return

for the limiting benchmarked index. Since the benchmarked returns of the NP are trivially zero, one notes that the limiting portfolio can only be the NP itself. Below it will become clear that this statement is a particular version of the Naive Diversification Theorem that we will derive in the next section.

The above arguments rely on the key idea of using the NP as benchmark. Beyond that they are, in principle, an application of the classical Law of Large Numbers. It has been intuitively clear for almost a century that the Law of Large Numbers plays a key role in explaining the effect of diversification. This chapter makes this intuition precise by using the conceptual framework of the benchmark approach; see Platen & Heath (2010). The next section removes the above restrictive assumption on the independence of the returns of benchmarked primary security accounts. It is already clear from the above arguments that a wide range of sequences of diversified portfolios approximates the NP. As will be shown in the next section, it is the sequence of EWIs, which requires minimal assumptions to ensure convergence towards the NP.

3.3 Naive Diversification Theorem

The Law of Large Numbers is largely model independent and, therefore, very robust. This applies also to the Naive Diversification Theorem (NDT), which will be derived below in general terms. To make the following statements rigorous, stochastic calculus will be applied, which has been widely used in continuous time finance since the work in Merton (1973). In this context the notion of the quadratic variation of a continuous stochastic process $X = \{X_t, t \geq 0\}$ is important. It can be defined as the limit in probability of the sum of squared increments of X when choosing finer and finer equidistant time-discretizations. For instance, the quadratic variation $[X]_t$ of a standard Brownian motion X equals the time t, that is, $[X]_t = t$. For ease of presentation, this chapter focuses on continuous market models. Without any major loss of generality the driving sources of traded uncertainty for benchmarked primary security accounts will be modeled by the independent Brownian motions $W^1, W^2, \ldots$.

The jth benchmarked primary security account value $\hat{S}_t^j$ at time t satisfies the

driftless Itô stochastic differential equation (SDE)

$$\frac{d\hat{S}_t^j}{\hat{S}_t^j} = \sum_{k=1}^{\ell} \sigma_t^{j,k} dW_t^k. \tag{3.11}$$

One can refer to Merton (1973), Long (1990), Platen & Heath (2010) and Filipović & Platen (2009) for the derivation of such driftless SDEs. Recall from Platen & Heath (2010) that the NP is the growth optimal portfolio which maximizes expected logarithmic utility. It yields in a continuous market the SDE (3.11) when used as numéraire for the jth primary security account. Assume that the volatility processes $\sigma^{j,k} = \{\sigma_t^{j,k}, t \geq 0\}$ are such that the corresponding stochastic integrals in (3.11) exist. Note that the driftless benchmarked primary security accounts are supermartingales. This does not mean that they have to be martingales, see Platen & Heath (2010) for details on this issue. The benchmarked self-financing portfolio process $\hat{S}^\delta = \{\hat{S}_t^\delta, t \geq 0\}$, with strategy $\delta = \{\delta_t = (\delta_t^1, \delta_t^2, \dots)^\top, t \geq 0\}$ is characterized by the SDE

$$d\hat{S}_t^\delta = \sum_{j=1}^{\infty} \delta_t^j d\hat{S}_t^j, \tag{3.12}$$

which is driftless. With the introduction of the fraction of wealth invested in the jth primary security account,

$$\pi_{\delta,t}^j = \frac{\delta_t^j \hat{S}_t^j}{\hat{S}_t^\delta}, \tag{3.13}$$

where $\sum_{j=1}^{\infty} \pi_{\delta,t}^j = 1$, one can rewrite the SDE (3.12) in the form

$$\frac{d\hat{S}_t^\delta}{\hat{S}_t^\delta} = \sum_{j=1}^{\infty} \pi_{\delta,t}^j \frac{d\hat{S}_t^j}{\hat{S}_t^j} = \sum_{j=1}^{\infty} \pi_{\delta,t}^j \sum_{k=1}^{\ell} \sigma_t^{j,k} dW_t^k. \tag{3.14}$$

The ℓth equi-weighted index (EWIℓ) invests the fractions

$$\pi_{\delta_{EWI\ell},t}^j = \begin{cases} \frac{1}{\ell} & \text{for} \quad j \in \{1, 2, \dots, \ell\} \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

Since the benchmarked NP equals the constant one, it follows that there exists a strategy $\delta_* = \{\delta_{*,t} = (\delta_{*,t}^1, \delta_{*,t}^2, \dots)^\top, t \geq 0\}$ such that

$$\frac{d\hat{S}_t^{\delta_*}}{\hat{S}_t^{\delta_*}} = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \pi_{\delta_*,t}^j \sigma_t^{j,k} dW_t^k = 0. \tag{3.16}$$

This means that the return process of the benchmarked NP equals zero.

It is now the aim to construct sequences of portfolios that approximate the NP in a mathematically precise and practically useful sense. The limits of the

return processes of such sequences of benchmarked portfolios should be zero. More precisely, the return process $\hat{Q}^\delta = \{\hat{Q}^\delta_t, t \geq 0\}$ of a benchmarked portfolio $\hat{S}^\delta$, given by the SDE

$$d\hat{Q}^\delta_t = \frac{1}{\hat{S}^\delta_t} d\hat{S}^\delta_t \tag{3.17}$$

for $t \geq 0$ with $\hat{Q}^\delta_0 = 0$, has to have small fluctuations to be a good proxy of the NP.

Definition 1 *A sequence* $(\hat{S}^{\delta_\ell})_{\ell \in \{1,2,\dots\}}$ *of strictly positive benchmarked portfolios, with initial values equal to one, is called a sequence of benchmarked approximate numéraire portfolios if for each* $\varepsilon > 0$ *and* $t \geq 0$ *one has*

$$\lim_{\ell \to \infty} P\left(\frac{d}{dt} \left[\hat{Q}^{\delta_\ell} \right]_t > \varepsilon \right) = 0. \tag{3.18}$$

The intuition is that, if one can construct a sequence of benchmarked portfolios where the quadratic variation of the return process vanishes asymptotically, then the limit can only be the constant one, that is, the benchmarked NP itself.

It seems reasonable to say that the returns of a benchmarked primary security account capture its specific or idiosyncratic traded uncertainty against the market as a whole; see Platen & Stahl (2003a). Due to the given structure of the market with different types of economic activity in different industry sectors of the economy, it is reasonable to assume that a particular specific uncertainty drives only the returns of a restricted number of benchmarked primary security accounts. If this were the case, then one could say that the securitization of the market is sufficiently developed and a diversification effect can be expected. To capture this property of a market in a mathematically precise manner, one can introduce the following notion:

Definition 2 *A financial market is well-securitized if there exists a real number* $q > 0$ *and a stochastic process* $\underline{\sigma}^2 = \{\underline{\sigma}^2_t, t \geq 0\}$ *with finite mean such that for all* $\ell, k \in \{1, 2, \dots\}$ *and* $t \geq 0$ *one has*

$$\frac{1}{\ell} \left| \sum_{j=1}^{\ell} \sigma^{j,k}_t \right|^2 \leq \frac{1}{\ell^q} \underline{\sigma}^2_t \tag{3.19}$$

P-almost surely.

Note that for independent benchmarked primary security accounts, as assumed in the previous section, condition (3.19) is easily verified in the presence of finite second moments for the volatility processes. If only a bounded number of benchmarked primary security accounts is driven by the same source of uncertainty, then

finite second moments for individual volatilities secure condition (3.19). Even if the number of related returns of benchmarked primary security accounts grows at a rate slower then $\sqrt{\ell}$, then the market still remains well-securitized as long as it has finite second moments for individual volatilities.

We will show below that the proof of the main result of this chapter, the NDT, can be based on an even weaker but slightly more technical assumption than formulated in condition (3.19).

Theorem (Naive Diversification Theorem) *In a well-securitized financial market the sequence of benchmarked equi-weighted indices, with fractions given by* (3.15)*, is a sequence of benchmarked approximate numéraire portfolios.*

We prove this result by using a lemma, which we prepare first. Note that the return process of the ℓth benchmarked EWI has at time t the value

$$\hat{Q}_t^{\delta_{EWI\ell}} = \sum_{j=1}^{\ell} \frac{1}{\ell} \sum_{k=1}^{\ell} \int_0^t \sigma_s^{j,k} dW_s^k. \tag{3.20}$$

The quadratic variation of this return process is of the form

$$\left[\hat{Q}^{\delta_{EWI\ell}}\right]_t = \frac{1}{\ell} \int_0^t \sum_{k=1}^{\ell} \left| \frac{1}{\sqrt{\ell}} \sum_{j=1}^{\ell} \sigma_s^{j,k} \right|^2 ds. \tag{3.21}$$

As can be seen below, this particular structure allows one to use a weaker, slightly more technical assumption, than imposed in (3.19) to guarantee a sequence of benchmarked approximate numéraire portfolios.

Lemma *Assume for all $\varepsilon > 0$ and $t \geq 0$ that one has the limit in probability*

$$\lim_{\ell \to \infty} P\left(\frac{1}{\ell} \sum_{k=1}^{\ell} \left| \frac{1}{\sqrt{\ell}} \sum_{j=1}^{\ell} \sigma_t^{j,k} \right|^2 > \varepsilon \right) = 0, \tag{3.22}$$

then the sequence of equi-weighted indices is a sequence of benchmarked approximate numéraire portfolios.

Proof of Lemma:

The statement of the above lemma is obtained via (3.18) in Definition 1, together with (3.21) and (3.22) since for $\varepsilon > 0$ and $t > 0$ one obtains directly that

$$\lim_{\ell \to \infty} P\left(\frac{d}{dt} \left[\hat{Q}^{\delta_{EWI\ell}}\right]_t > \varepsilon \right) = \lim_{\ell \to \infty} P\left(\frac{1}{\ell} \sum_{k=1}^{\ell} \left| \frac{1}{\sqrt{\ell}} \sum_{j=1}^{\ell} \sigma_t^{j,k} \right|^2 > \epsilon \right) = 0. \tag{3.23}$$

□

Proof of the Naive Diversification Theorem:

It is straightforward to see by the well-known Markov inequality that in a well-securitized market the assumption (3.22) is satisfied due to condition (3.19) since

$$\lim_{\ell\to\infty} P\left(\frac{1}{\ell}\sum_{k=1}^{\ell}\left|\frac{1}{\sqrt{\ell}}\sum_{j=1}^{\ell}\sigma_t^{j,k}\right|^2 > \varepsilon\right) \leq \lim_{\ell\to\infty} P\left(\frac{1}{\ell^q}\underline{\sigma}_t^2 > \varepsilon\right) \leq \lim_{\ell\to\infty}\frac{1}{\ell^q}\frac{1}{\varepsilon}E\left(\underline{\sigma}_t^2\right) = 0, \tag{3.24}$$

which proves the Naive Diversification Theorem. □

Note from the above derivation that if condition (3.22) does not hold, then also the statement (3.23) cannot be true. This shows that in this case naive diversification does not lead to convergence in probability towards the NP.

The statement of the NDT is very robust. Already, under condition (3.22), which is weaker than (3.19), it covers a wide range of models. Essentially, one can say that the NDT is model independent since no particular assumptions about the underlying market model need to be made. We emphasize, the NDT, and more generally the above lemma, are based on very few probabilistic arguments, and no major economic assumptions have to be made. The presence of a large number of reasonably different economic activities, securitized via stocks in the market, is sufficient to let the sequence of EWIs automatically approximate the NP. This provides an alternative view to economic modeling. Simply the laws of statistics are sufficient in the naturally structured world of securitized economic activity to provide the fact that a naively diversified portfolio should be a good proxy of the NP. The latter portfolio forms the central building block of Platen's benchmark approach. This approach represents a generalization of the classical arbitrage pricing theory and modern portfolio theory, see Platen & Heath (2010)

3.4 Transaction Costs and Reallocation Frequency

This section demonstrates that the impact of transaction costs is almost negligible from the perspective of large funds. The observed long term growth of the EWI114 is, in principle, obtainable by an investable portfolio. Additionally, the influence of the length of the time between capital reallocation is studied in this section. Only minor changes in performance are observed for the cases with time periods between

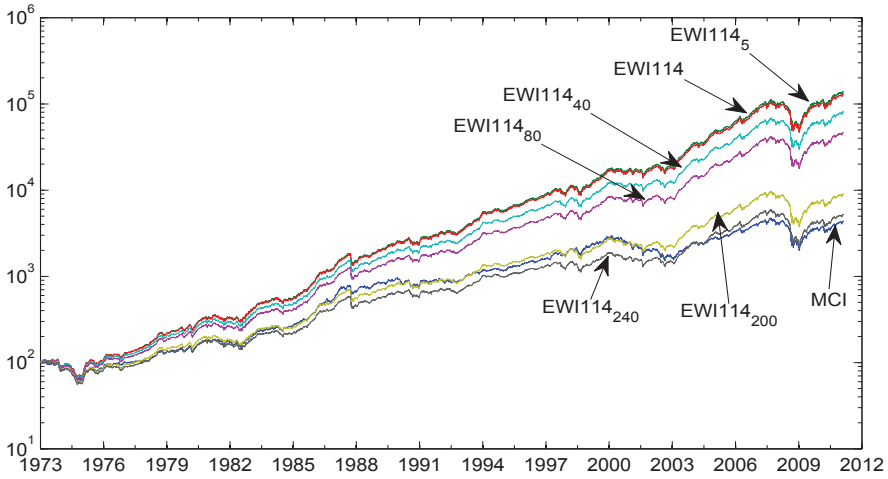

Figure 3.3: Logarithms of MCI, EWI114 without transaction cost and EWI114$_\xi$ with transaction costs of 5,40,80,200 and 240 basis points.

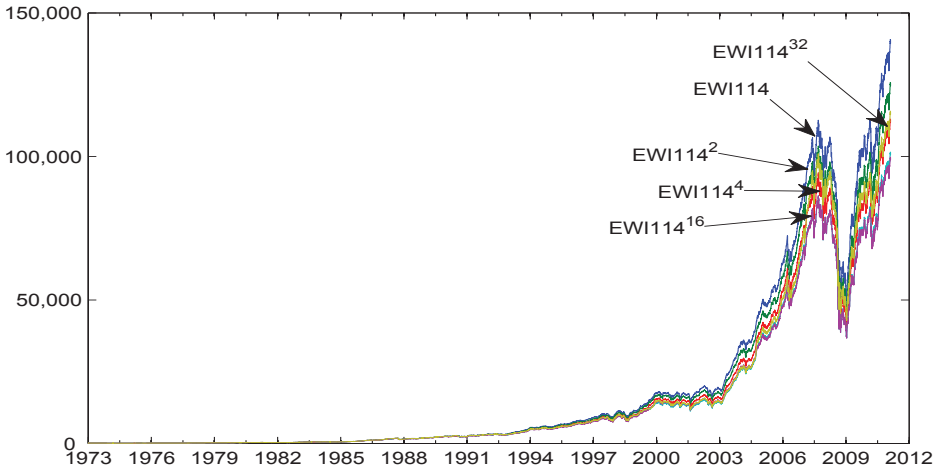

Figure 3.4: EWI114^m reallocated daily and every 2, 4, 8, 16 and 32 days.

capital reallocations of up to one month.

Transaction Cost

In the case when constructing equi-weighted indices it is necessary to reallocate regularly capital according to the naive diversification strategy. Naive diversification constitutes a form of hedging. To compare realistically the long term performance of portfolios one needs to account for transaction costs incurred by reallocations. That is, the value of an EWI at each reallocation has to be reduced by the respective transaction cost. When forming the equi-weighted index $S^{\delta_{EWI_\xi}}$ in US dollars, the

Table 3.1: Summary for various EWI114 versions with transaction cost equal to $\xi \in \{0, 5, 40, 80, 200, 240\}$ basis points

Transaction cost	0	5	40	80	200	240
Reallocation terms	1					
Final value	139338.64	130111.93	80543.07	46555.04	8988.23	5194.46
Annualised average return	0.1979	0.1961	0.1834	0.1689	0.1254	0.1109
Annualised volatility	0.1135	0.1135	0.1135	0.1135	0.1134	0.1134
Sharpe ratio	1.4205	1.4046	1.2930	1.1654	0.7822	0.6544
Reallocation terms	2					
Final value	124542.04	119369.00	88697.63	63166.73	22808.64	16240.40
Annualised average return	0.1949	0.1938	0.1859	0.1770	0.1500	0.1411
Annualised volatility	0.1134	0.1134	0.1134	0.1134	0.1135	0.1136
Sharpe ratio	1.3955	1.3856	1.3163	1.2369	0.9987	0.9193
Reallocation terms	4					
Final value	111899.82	108230.16	85698.25	65628.82	29467.48	22562.42
Annualised average return	0.1921	0.1912	0.1850	0.1780	0.1568	0.1497
Annualised volatility	0.1135	0.1135	0.1134	0.1134	0.1134	0.1135
Sharpe ratio	1.3699	1.3622	1.3080	1.2459	1.0591	0.9967
Reallocation terms	8					
Final value	100505.37	97963.66	81881.57	66705.62	36055.10	29367.02
Annualised average return	0.1892	0.1885	0.1837	0.1783	0.1621	0.1566
Annualised volatility	0.1127	0.1127	0.1127	0.1127	0.1128	0.1128
Sharpe ratio	1.3531	1.3471	1.3051	1.2569	1.1119	1.0634
Reallocation terms	16					
Final value	98775.24	96892.29	84677.43	72588.14	45711.97	39177.91
Annualised average return	0.1887	0.1882	0.1847	0.1806	0.1684	0.1643
Annualised volatility	0.1130	0.1130	0.1130	0.1130	0.1131	0.1131
Sharpe ratio	1.3463	1.3418	1.3102	1.2740	1.1647	1.1281
Reallocation terms	32					
Final value	114592.50	112929.09	101939.59	90678.85	63804.84	56744.28
Annualised average return	0.1927	0.1923	0.1896	0.1865	0.1772	0.1741
Annualised volatility	0.1131	0.1131	0.1131	0.1131	0.1133	0.1133
Sharpe ratio	1.3797	1.3763	1.3522	1.3245	1.2408	1.2127

study assumes proportional transaction costs ξ to be charged in the following way:

$$S_{t_n}^{\delta_{EWI_\xi}} = S_{t_n-}^{\delta_{EWI_\xi}} - \sum_{j=1}^{d} \frac{\xi}{10000} S_{t_n}^j \left| \delta_{EWI_\xi,t_n}^j - \delta_{EWI_\xi,t_{n-1}}^j \right|, \tag{3.25}$$

where the number d of constituents is partly suppressed in the notation, and ξ denotes the parameter for the proportional transaction cost in basis points (bp). Note that one has $\delta_{EWI_\xi,t_n}^j = \frac{1}{d}$ in (3.25). In the above formula $S_{t_n}^j$ denotes the value of the jth primary security account in US dollars at time t_n. Furthermore, $S_{t_n-}^{\delta_{EWI_\xi}}$ denotes the value of the EWI in US dollars just before reallocation takes place at time t_n, whereas $S_{t_n}^{\delta_{EWI_\xi}}$ equals its value after reallocation and after accounting for transaction costs. Note that the results are, in principle, the same if one uses another base currency.

This section studies only indices constructed similarly to the EWI114 but with different values for the transaction cost parameter ξ. Later different time intervals between capital reallocations will be allowed. In general, one expects larger transac-

tion costs to yield lower average long term growth. Indeed, this can be observed in Fig. 3.3, which displays the logarithms of the MCI, the EWI114 and the EWI114_{ξ}, constructed for the transaction cost parameter values $\xi \in \{5, 40, 80, 200, 240\}$ under daily capital reallocation. The EWI114_5 has only marginally lower values than the EWI114 and both are almost indistinguishable. In the long run the MCI is located approximately between the EWI114_{200} and the EWI114_{240}, suggesting that it takes roughly between 200 and 240 bp in proportional transaction costs to bring the long term growth of the EWI114_{ξ} down to that of the MCI when reallocating daily. Transaction costs of 200 bp are significantly beyond those charged to large funds. Note in Fig. 3.3 that all EWIs are driven by almost exactly the same nondiversifiable risk factor. The MCI is exposed to some additional risk factors since it does not always move perfectly parallel to the EWIs.

Impact of Reallocation Frequency

Let us now study, first without transaction costs, the resulting EWIs when reallocating every 2,4,8,16 and 32 days. Fig. 3.4 displays these indices together with the EWI114. Denote by EWI114^m the index which reallocates every m days without transaction costs. The EWI114^m does not outperform the daily reallocated EWI114, for $m \in \{2, 4, 8, 16, 32\}$. However, if the reallocations are less frequent, then by chance occasional outperformance of the EWI114 by such constructed indices can be expected. The Naive Diversification Theorem predicts that in the long-run the EWIs with sufficiently frequent capital reallocations should approximate well the NP and, thus, outperform other indices.

Further below, this chapter studies the combined impact of different transaction cost parameters and different time intervals between reallocations. Table 3.1 displays the final value of the respective indices, the estimated annualized expected return $\hat{\mu}$, the estimated annualized volatility $\hat{\sigma}$ and the resulting Sharpe ratio $\hat{s} = \frac{\hat{\mu}-\hat{r}}{\hat{\sigma}}$. The Sharpe ratio is here evaluated as the ratio of the difference between the annualized sample mean of the returns minus an assumed average US short rate $\hat{r}$ of 0.05, over the annualized sample standard deviation $\hat{\sigma}$. Recall that the MCI has the estimated volatility of 0.134 and the estimated expected return of 0.108, which yields a Sharpe ratio of 0.537.

The results for the indices constructed with different days $m \in \{1, 2, 4, 8, 16, 32\}$

between capital reallocations and with transaction cost parameter $\xi \in \{0, 5, 40, 80, 200, 240\}$ are displayed in Table 3.1. In this table the estimated volatilities remain very close to each other up to quarterly reallocations. With an annualized value of approximately 0.11 they are clearly smaller than the estimated volatility of 0.13 for the MCI. The largest Sharpe ratio reported in Table 3.1 with value 1.42 is that of the EWI114 with daily reallocations and without transaction costs. The Sharpe ratio of the EWI114^m seems to decline for increasing $m \in \{1, 2, 4, 8, 16\}$. Even with very realistic transaction costs of 80 bp the Sharpe ratios of the different equi-weighted indices remain in this table above twice the Sharpe ratio of the MCI. In summary, we see in Table 3.1 that less frequent reallocations and higher transaction costs diminish, in general, the Sharpe ratio.

This chapter aims to give a better understanding of the basic principle underpinning the well-observed phenomenon of diversification, fundamental to fund and risk management. It approximates the numéraire portfolio, which is in many ways the "best" performing portfolio, by a sequence of naive (equi-weighted) portfolios with increasing number of constituents. The Naive Diversification Theorem has been established, which states that a naively diversified portfolio with many constituents is a good proxy for the numéraire portfolio of a well-securitized market. By applying naive diversification on real equity data, this chapter identifies an equi-weighted index, formed by equi-weighted country subsector indices, as a good proxy of the numéraire portfolio of the global stock market. This proxy turns out to be the best performing index in the family of diversified portfolios constructed in this chapter. Its outstanding long term performance has been predicted by the Naive Diversification Theorem. An equi-weighted index should show an excellent average long term growth rate due to its proximity to the numéraire portfolio.

This chapter demonstrates that naive diversification resolves practical difficulties encountered with the sample-based Markowitz mean-variance approach to portfolio optimization. There is no need for estimating drift parameters. Even if one accounts for transaction costs in the construction of an equi-weighted index, it still shows a significantly higher Sharpe ratio when compared to the one of the corresponding market capitalization weighted index. The proposed approach of approximating the numéraire portfolio by naive diversification is robust and also very general. With the

possible construction of an investable proxy for the numéraire portfolio, this chapter opens new lines of research. The numéraire portfolio is not only in several ways the "best" performing portfolio, it is also the natural numéraire for derivative pricing under the real-world probability measure, see Platen & Heath (2010). Furthermore, as a consequence of the Naive Diversification Theorem it captures the nondiversiable risk of the market. This general market risk is highly relevant for risk measurement and risk management. Moreover, when equi-weighted world indices of the proposed type would become major investment vehicles in the financial market architecture used by large pension funds and insurance companies, they would have a stabilizing effect. In time of financial crisis they would provide liquidity to those asset classes under stress. On the other hand, in the case of an asset bubble they would moderate the excesses by selling highly priced assets. This goes along with the intuitive reasoning behind naive diversification, which always sells high and buys low.

3.5 Simulation of Diversified Portfolios in Continuous Financial Markets

By using techniques and results that we prepared in Chapter 1 on scenario simulation, this section analyzes the simulated long-term behavior of well diversified portfolios in continuous financial markets. For simulated markets we know exactly the numéraire portfolio and can compare it with other portfolios. We focus on the equi-weighted index and the market portfolio, and rely on results in Platen & Rendek (2012c). This section will illustrate that the equai-weighted index constitutes a good proxy of the numéraire portfolio. The multi-asset market models considered include the Black-Scholes model, the Heston model, the ARCH diffusion model and a multi-asset version of the minimal market model. All these models are simulated exactly or almost exactly over an extremely long period of time to analyze the long term growth of the respective portfolios. This section illustrates the robustness of the diversification phenomenon captured by the Naive Diversification Theorem when approximating the numéraire portfolio by the equi-weighted index. Significant outperformance in the long run of the market capitalization weighted portfolio by the equi-weighted index is documented for different market models.

We will see that under the multi-asset minimal market model the equi-weighted

index outperforms remarkably well the market portfolio. In this case the benchmarked market portfolio is a strict supermartingale, whereas the benchmarked equiweighted index is a martingale. Equal value weighting overcomes the strict supermartingale property that the benchmarked market portfolio inherits from its strict supermartingale constituents under this model.

For given market dynamics the study simulates for the long period of $T = 150$ years the benchmarked trajectories of $d = 1000$ primary security accounts, sampling 100 times per year. For simplicity, the interest rate is set to zero, therefore, the inverse of the benchmarked savings account provides the numéraire portfolio (NP) when denominated in units of the domestic currency, that is,

$$S_t^{\delta_*} = (\hat{S}_t^0)^{-1}. \tag{3.26}$$

The product of the NP with the jth benchmarked primary security account yields the value of the jth primary security account denominated in domestic currency, that is,

$$S_t^j = \hat{S}_t^j \, S_t^{\delta_*}, \tag{3.27}$$

for $j \in \{0, 1, \ldots, d\}$.

In reality, the market capitalization of stocks is very different from each other. Statistical analysis of market data suggests that the size of companies and similarly the market capitalization of their stocks seem to be Pareto distributed; see e.g. Simon & Bonini (1958). To generate realistic market capitalization weighted indices (MCIs) in the following simulations, the initial values S_0^j, $j \in \{0, 1, \ldots, d\}$, of primary security accounts follow here a Pareto distribution

$$F_j(x) = 1 - \left(\frac{x_0}{x}\right)^{\lambda} \tag{3.28}$$

with the parameters $\lambda = 1.1$ and $x_0 = \frac{\lambda - 1}{\lambda}$; see Simon & Bonini (1958).

The simulations below will be performed by first simulating the benchmarked primary security accounts including the benchmarked savings account according to a given market model. Then the benchmarked EWI and the benchmarked MCI are obtained from these benchmarked constituents by choosing appropriate fractions, see Section 2.1. Since the NP is the inverse of the benchmarked savings account, the EWI (MCI) is simply equal to the NP times the benchmarked EWI (benchmarked MCI).

Black-Scholes Model

When using discrete time numerical schemes for the simulation of solutions of SDEs, as studied in Kloeden & Platen (1999), there may be issues arising when dealing with non-Lipschitz continuous drift or diffusion coefficients. Furthermore, problems concerning numerical stability may emerge for long-term simulations or even negative values could be simulated by standard discrete time approximations for strictly positive price processes; see Platen & Bruti-Liberati (2010). In the following sections exact and almost exact simulations of various market models are described. These avoid the above indicated numerical problems in scenario simulations when using standard discrete time approximations. For the following simulation studies market models have been selected where exact or almost exact simulations are possible. A contribution of this section is, therefore, also the description of highly accurate scenario simulation methods for several frequently used market model classes in finance.

Recall first the simulation under the standard market model, which is the multi-asset version of the Black-Scholes model; see Black & Scholes (1973). Under this model the benchmarked primary security accounts satisfy the following vector SDE

$$d\hat{\boldsymbol{S}}_t = \sum_{k=0}^{d} \boldsymbol{B}^k \hat{\boldsymbol{S}}_t dW_t^k, \tag{3.29}$$

for $t \in [0, \infty)$. Here $\hat{\boldsymbol{S}} = \{\hat{\boldsymbol{S}}_t = (\hat{S}_t^0, \hat{S}_t^1, \ldots, \hat{S}_t^d)^\top, t \in [0, \infty)\}$ is a vector of benchmarked primary security accounts, and $\boldsymbol{B}^k = [B^{k,i,j}]_{i,j=0}^d$ is a $(d+1)\times(d+1)$ diagonal volatility matrix, with elements

$$B^{k,i,j} = \begin{cases} b^{j,k} & \text{for} \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{3.30}$$

for $k, i, j \in \{0, 1, 2, \ldots, d\}$. Note that the benchmarked primary security accounts $\hat{S}^j$, $j \in \{0, 1, 2, \ldots, d\}$, represent under this model martingales.

The multi-asset Black-Scholes model can be simulated exactly. The vector SDE (3.29) has an explicit solution. The jth benchmarked primary security account is represented by the exponential

$$\hat{S}_t^j = \hat{S}_0^j \exp\left\{ -\frac{1}{2} \sum_{k=1}^{d} \left(b^{j,k}\right)^2 t + \sum_{k=1}^{d} b^{j,k} W_t^k \right\}. \tag{3.31}$$

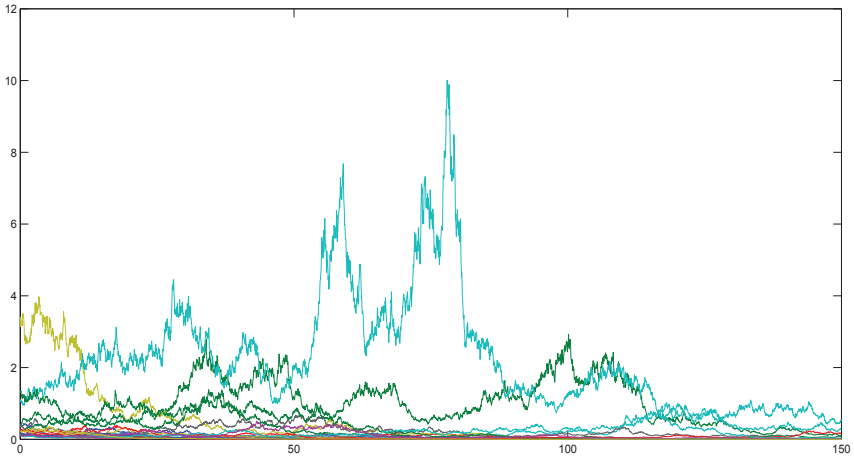

Figure 3.5: Simulated benchmarked primary security accounts under the Black-Scholes model

Throughout all simulation studies this section employs the equi-distant time discretization $0 = t_0 < t_1 < \ldots$, where $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, $\Delta > 0$. For simplicity, the simulation of benchmarked primary security accounts is performed for the case of independent benchmarked primary security accounts. However, for most of the market models the case of dependent benchmarked primary security accounts can be handled analogously, as shown in Platen & Rendek (2009) and Section 1.1. For the jth independent benchmarked primary security account one obtains at time t_{i+1} under the Black-Scholes model the exponential

$$\hat{S}^j_{t_{i+1}} = \hat{S}^j_0 \exp\left\{-\frac{1}{2}\left(b^{j,j}\right)^2 t_{i+1} + b^{j,j} W^j_{t_{i+1}}\right\}, \tag{3.32}$$

for $j \in \{0, 1, \ldots, d\}$.

The phenomenon of diversification is now illustrated by simulating in a Black-Scholes market the NP, the EWI and the MCI. This simulation can be exactly performed without generating any error by using the explicit expression described above. The study simulates $d = 1000$ benchmarked primary security accounts over 150 years, each with volatility $b^{j,j} = 0.2$ for $j \in \{1, \ldots, 1000\}$. The first 20 simulated trajectories are displayed in Fig. 3.5. Note that these benchmarked primary security accounts are modeled by martingales. As mentioned earlier, the independent initial values $\hat{S}^j_0$ are generated using a Pareto distribution. Consequently, there is great variety in the market capitalization of stocks.

Fig 3.6 shows the resulting simulated trajectories of the NP, the EWI and the MCI denominated in domestic currency. In this case the EWI approximates rather

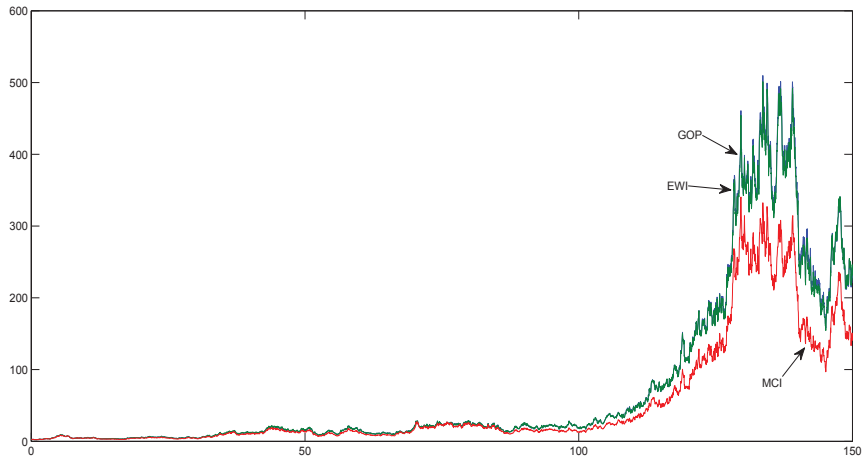

Figure 3.6: Simulated NP, EWI and MCI under the Black-Scholes model

well the NP, which is also the GOP, and the differences between both portfolio values are even difficult to see, in particular, in the earlier parts of the trajectories. We see in Fig. 3.7 that the MCI seems to be initially a reasonable proxy of the NP, however, after some time it diverges from the NP due to its lower long term performance. As can be seen in Fig. 3.5, large total values for some stocks are observed in the simulated market. The resulting large fractions of these stocks in the MCI are likely to distort its long run performance because the market portfolio appears to be not well diversified in the sense of this chapter. The emerging fractions of the corresponding primary security accounts are probably too large to be acceptable as those of a diversified portfolio. This phenomenon becomes even clearer in Fig. 3.7, which displays the constant benchmarked NP, $\hat{S}_t^{\delta_*} = 1$, as well as, the benchmarked EWI and the benchmarked MCI. Note that initially, sometimes the benchmarked EWI and sometimes the benchmarked MCI performed better. However, in the long run the benchmarked EWI fluctuates around the benchmarked NP, while the benchmarked MCI diverges downwards. Additionally, one notes in this figure that the benchmarked MCI has much larger volatility than the benchmarked EWI. This is typical also for the other market models considered and for the real market as well. The described simulation has been repeated many times for other scenarios. The better performance in the long run of the EWI over the MCI was similarly evident.

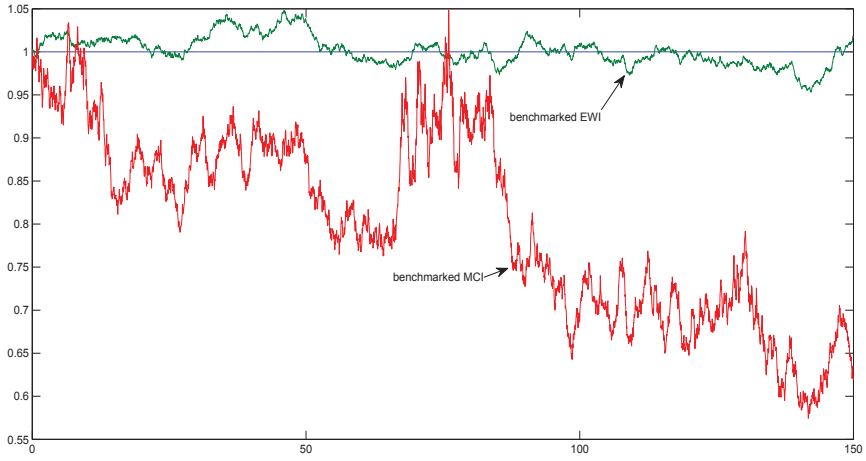

Figure 3.7: Simulated benchmarked NP, EWI and MCI under the Black-Scholes model

Heston Model

The multi-asset version of the Heston model, see Heston (1993), can be described by a set of two vector SDEs in the form

$$d\hat{\boldsymbol{S}}_t = \operatorname{diag}\left(\sqrt{\boldsymbol{V}_t}\right)\operatorname{diag}\left(\hat{\boldsymbol{S}}_t\right)\left(\boldsymbol{A}d\tilde{\boldsymbol{W}}_t^1 + \boldsymbol{B}d\tilde{\boldsymbol{W}}_t^2\right), \tag{3.33}$$

$$d\boldsymbol{V}_t = (\boldsymbol{a} - \boldsymbol{E}\boldsymbol{V}_t)\,dt + \boldsymbol{F}\operatorname{diag}\left(\sqrt{\boldsymbol{V}_t}\right)d\tilde{\boldsymbol{W}}_t^1, \tag{3.34}$$

for $t \in [0,\infty)$. Here $\hat{\boldsymbol{S}} = \{\hat{\boldsymbol{S}}_t = (\hat{S}_t^0, \hat{S}_t^1, \ldots, \hat{S}_t^d)^\top, t \in [0,\infty)\}$ is a vector of benchmarked primary security accounts, which are local martingales but not necessarily martingales; see Platen & Heath (2010). Additionally, $\tilde{\boldsymbol{W}}^1 = \{\tilde{\boldsymbol{W}}_t^1 = (\tilde{W}_t^{1,0}, \tilde{W}_t^{1,1}, \ldots, \tilde{W}_t^{1,d})^\top, t \in [0,\infty)\}$ and $\tilde{\boldsymbol{W}}^2 = \{\tilde{\boldsymbol{W}}_t^2 = (\tilde{W}_t^{2,0}, \tilde{W}_t^{2,1}, \ldots, \tilde{W}_t^{2,d})^\top, t \in [0,\infty)\}$ are independent vectors of correlated Wiener processes. That is, one has

$$\tilde{\boldsymbol{W}}_t^k = \boldsymbol{C}^k \boldsymbol{W}_t^k, \tag{3.35}$$

where $\boldsymbol{C}^k = [C^{k,i,j}]_{i,j=0}^d$, and with $\boldsymbol{W}^k = \{\boldsymbol{W}_t^k = (W_t^{k,0}, W_t^{k,1}, \ldots, W_t^{k,d})^\top, t \in [0,\infty)\}$ for $k \in \{1,2\}$, denoting again a vector of independent Wiener processes.

Additionally, $\boldsymbol{A} = [A^{i,j}]_{i,j=0}^d$ is a diagonal matrix with elements

$$A^{i,j} = \begin{cases} \varrho_i & \text{for} \quad i = j \\ 0 & \text{otherwise,} \end{cases} \tag{3.36}$$

and $\boldsymbol{B} = [B^{i,j}]_{i,j=0}^d$ is a diagonal matrix with elements

$$B^{i,j} = \begin{cases} \sqrt{1-\varrho_i^2} & \text{for} \quad i = j \\ 0 & \text{otherwise.} \end{cases} \tag{3.37}$$

Moreover, $\boldsymbol{V} = \{\boldsymbol{V}_t = (V_t^0, V_t^1, \ldots, V_t^d)^\top, t \in [0, \infty)\}$ is a vector of squared volatilities, $\boldsymbol{a} = (a_0, a_1, \ldots, a_d)^\top$. $\boldsymbol{E} = [E^{i,j}]_{i,j=0}^d$ is a diagonal matrix with elements

$$E^{i,j} = \begin{cases} \kappa_i & \text{for} \quad i = j \\ 0 & \text{otherwise,} \end{cases} \tag{3.38}$$

and $\boldsymbol{F} = [F^{i,j}]_{i,j=0}^d$ is a diagonal matrix with elements

$$F^{i,j} = \begin{cases} \gamma_i & \text{for} \quad i = j \\ 0 & \text{otherwise.} \end{cases} \tag{3.39}$$

One method for the exact simulation of the Heston model has been suggested in Broadie & Kaya (2006). The study in the current section uses a slightly different, possibly more convenient, almost exact simulation for benchmarked primary security accounts under the Heston model. The method involves exact simulation of the squared volatility processes and some almost exact simulation of the independent benchmarked primary security accounts. The latter simulation is conditional on the exactly simulated trajectories of the squared volatilities; see Platen & Rendek (2009) and Section 1.4.

One obtains the value of the jth squared volatility $V_{t_{i+1}}^j$ at time t_{i+1}, $i \in \{0, 1, \ldots\}$, by sampling directly from the noncentral chi-square distribution $\chi'^2_{\nu_j}(\lambda_j)$ with ν_j degrees of freedom and noncentrality parameter λ_j, that is,

$$V_{t_{i+1}}^j = \frac{\gamma_j^2 \left(1 - \exp\{-\kappa_j \Delta\}\right)}{4\kappa_j} \chi'^2_{\nu_j} \left(\frac{4\kappa_j e^{-\kappa_j \Delta}}{\gamma_j^2 \left(1 - e^{-\kappa_j \Delta}\right)} V_{t_i}^j \right), \tag{3.40}$$

where $\nu_j = \frac{4a_j}{\gamma_j^2}$. Details on sampling from noncentral chi-square distributions can be found, for instance, in Glasserman (2004). The resulting simulation method for the jth squared volatility V^j is exact.

It remains to describe the almost exact simulation of the vector of the logarithms of benchmarked assets. By following first Broadie & Kaya (2006), the jth value $X_{t_{i+1}}^j = \ln\left(\hat{S}_{t_{i+1}}^j\right)$ at time t_{i+1} can be represented in the form

$$\begin{aligned} X_{t_{i+1}}^j &= X_{t_i}^j + \frac{\varrho_j}{\gamma_j} \left(V_{t_{i+1}}^j - V_{t_i}^j - a_j \Delta\right) + \left(\frac{\varrho_j \kappa_j}{\gamma_j} - \frac{1}{2}\right) \int_{t_i}^{t_{i+1}} V_u^j du \\ &\quad + \sqrt{1 - \varrho_j^2} \int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j}. \end{aligned} \tag{3.41}$$

Furthermore, the distribution of the integral increment

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j}, \tag{3.42}$$

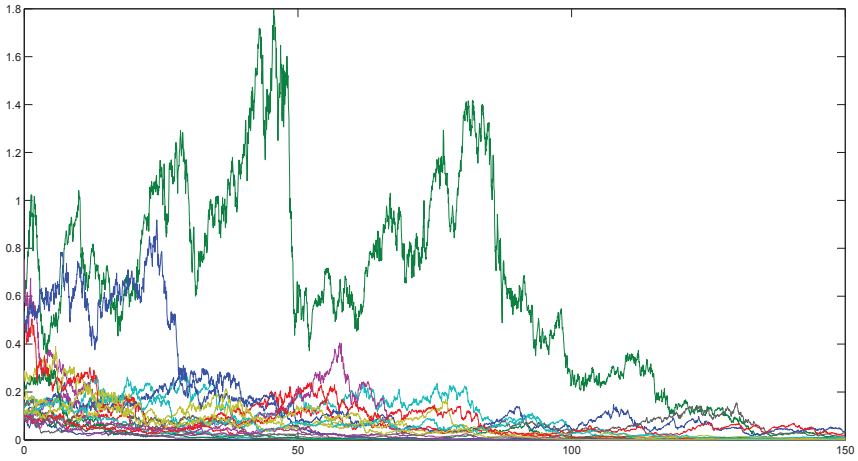

Figure 3.8: Simulated benchmarked primary security accounts under the Heston model

given the path of V^j, is conditionally Gaussian with mean zero and variance $\int_{t_i}^{t_{i+1}} V_u^j du$, since V^j is independent of the Brownian motion $W^{2,j}$ for all $j \in \{0, 1, 2, \ldots, d\}$. Moreover, one needs to evaluate the variance $\int_{t_i}^{t_{i+1}} V_u^j du$ conditioned on the path of the process V^j. The proposed almost exact simulation method uses as an approximation via the trapezoidal rule

$$\int_{t_i}^{t_{i+1}} V_u^j du \approx \frac{\Delta}{2} \left(V_{t_i}^j + V_{t_{i+1}}^j \right). \tag{3.43}$$

It is well-known that this quadrature rule generates by its symmetry excellent approximations; see Kloeden & Platen (1999). Consequently, one has the approximate conditionally Gaussian random variable

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j} \approx \mathcal{N}\left(0, \frac{\Delta}{2} \left(V_{t_i}^j + V_{t_{i+1}}^j \right)\right) \tag{3.44}$$

to compute. This approximation can be achieved with practically negligible error by using a sufficiently small time step size. Since the aim here is to illustrate the diversification effect, this section omits any particular error analysis. The above approximation converges in distribution when the time step size decreases, as can be deduced from (3.44). For the above multi-asset Heston model this results in an efficient, almost exact simulation technique. Alternatively, one could have used the Broadie & Kaya (2006) exact simulation method, which leads from the perspective of this section to the same conclusions.

Benchmarked primary security accounts are simulated under the multi-asset Heston model. Fig. 3.8 displays the first 20 simulated benchmarked primary secu-

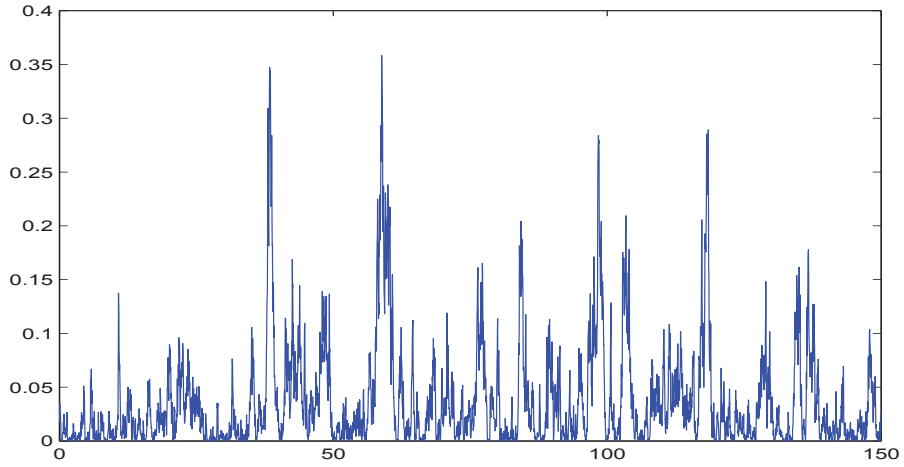

Figure 3.9: Simulated squared volatility under the Heston model

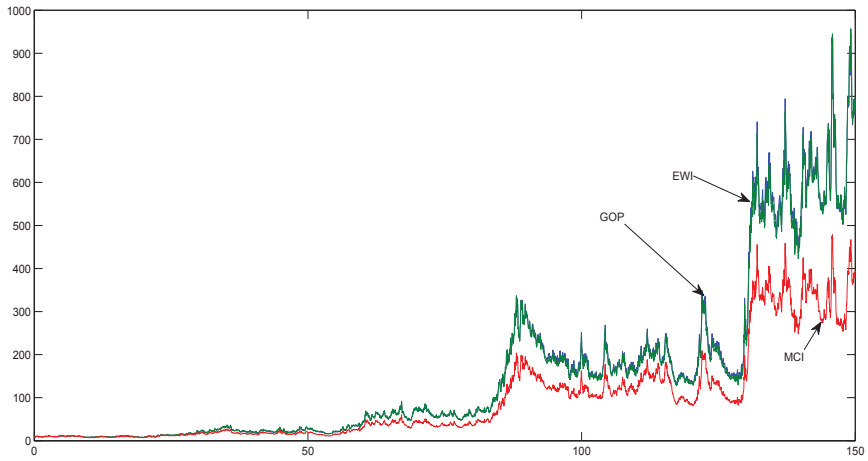

Figure 3.10: Simulated NP, EWI and MCI under the Heston model

rity accounts. These benchmarked primary security accounts are nonnegative local martingales and, thus, supermartingales. The parameters in (3.40) are chosen according to market data, see Gatheral (2006). The jth squared volatility process is simulated according to (3.40) for the initial value $V_0^j = 0.0174$, $a_j = 0.0469$, $\kappa_j = 1.3253$, $\gamma_j = 0.3877$, $j \in \{1, \ldots, 1000\}$. The correlation parameter is here set to $\varrho_j = -0.7165$ in order to reflect the typically observed leverage effect, see Black (1976). The initial values $\hat{S}_0^j$ are again generated from the Pareto distribution mentioned earlier. For illustration, Fig. 3.9 displays a typical trajectory of the squared volatility process V^j under the Heston model, simulated exactly according to the formula (3.40).

Fig 3.10 exhibits the simulated NP, EWI and MCI under the Heston model. Also

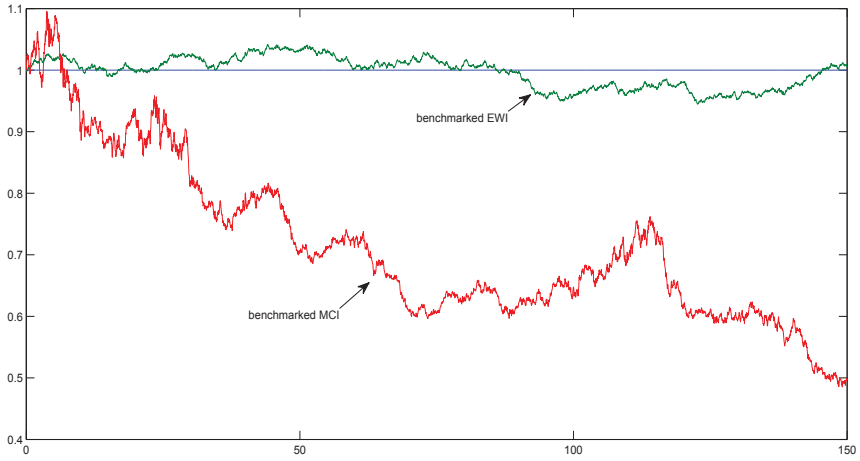

Figure 3.11: Simulated benchmarked NP, EWI and MCI under the Heston model

here the EWI provides an excellent proxy for the NP, and it is difficult to distinguish both trajectories. The MCI, however, does not perform as well as the EWI. Finally, Fig. 3.11 illustrates the constant benchmarked NP $\hat{S}_t^{\delta_*} = 1$, the benchmarked MCI and the benchmarked EWI. This plot shows the closeness of the benchmarked EWI to the benchmarked NP. Note that, the benchmarked MCI has a larger volatility than the benchmarked EWI and fluctuates in the long run significantly.

ARCH-diffusion Model

Some continuous time limits of popular time series models in finance, including several ARCH and GARCH models, can be captured by the multi-dimensional ARCH-diffusion model that is considered below. The class of ARCH and GARCH time series models was initiated in Engle (1982). The ARCH-diffusion model is obtained as a continuous time limit of the innovation process of the GARCH$(1,1)$ and NGARCH$(1,1)$ models; see Nelson (1990) and Frey (1997). The ARCH-diffusion model can be described by the following two vector SDEs

$$d\hat{\boldsymbol{S}}_t = \operatorname{diag}\left(\sqrt{\boldsymbol{V}_t}\right)\operatorname{diag}\left(\hat{\boldsymbol{S}}_t\right)\left(\boldsymbol{A}d\tilde{\boldsymbol{W}}_t^1 + \boldsymbol{B}d\tilde{\boldsymbol{W}}_t^2\right), \tag{3.45}$$

$$d\boldsymbol{V}_t = (\boldsymbol{a} - \boldsymbol{E}\boldsymbol{V}_t)\,dt + \boldsymbol{F}\operatorname{diag}(\boldsymbol{V}_t)\,d\tilde{\boldsymbol{W}}_t^1, \tag{3.46}$$

for $t \in [0,\infty)$. Here $\hat{\boldsymbol{S}} = \{\hat{\boldsymbol{S}}_t = (\hat{S}_t^0, \hat{S}_t^1, \ldots, \hat{S}_t^d)^\top, t \in [0,\infty)\}$ denotes again a vector of benchmarked primary security accounts, where one knows that these are local martingales and, thus, supermartingales. Furthermore, $\tilde{\boldsymbol{W}}^1 = \{\tilde{\boldsymbol{W}}_t^1 = (\tilde{W}_t^{1,0}, \tilde{W}_t^{1,1}, \ldots, \tilde{W}_t^{1,d})^\top, t \in [0,\infty)\}$ and $\tilde{\boldsymbol{W}}^2 = \{\tilde{\boldsymbol{W}}_t^2 = (\tilde{W}_t^{2,0}, \tilde{W}_t^{2,1}, \ldots, \tilde{W}_t^{2,d})^\top,$

$t \in [0,\infty)\}$ are independent vectors of correlated Wiener processes. Additionally, $\boldsymbol{A} = [A^{i,j}]_{i,j=0}^{d}$ is a diagonal matrix with elements as in (3.36), and $\boldsymbol{B} = [B^{i,j}]_{i,j=0}^{d}$ is a diagonal matrix with elements given in (3.37). Moreover, $\boldsymbol{V} = \{\boldsymbol{V}_t = (V_t^0, V_t^1, \ldots, V_t^d)^\top, t \in [0,\infty)\}$ is a vector of squared volatilities, $\boldsymbol{a} = (a_0, a_1, \ldots, a_d)^\top$ is a vector; $\boldsymbol{E} = [E^{i,j}]_{i,j=0}^{d}$ is a diagonal matrix with elements as in (3.38); and $\boldsymbol{F} = [F^{i,j}]_{i,j=0}^{d}$ is a diagonal matrix with elements as in (3.39).

In the given case one can simulate the jth squared volatility process V^j, $j \in \{0,1,2,\ldots\}$, almost exactly by approximating the time integral via the trapezoidal rule in the following exact representation

$$\begin{aligned} V_{t_{i+1}}^j &= \exp\left\{\left(-\kappa_j - \frac{1}{2}\gamma_j^2\right) t_{i+1} + \gamma_j W_{t_{i+1}}^{1,j}\right\} \\ &\quad \times \left(V_{t_0}^j + a_j \sum_{k=0}^{i} \int_{t_k}^{t_{k+1}} \exp\left\{\left(\kappa_j + \frac{1}{2}\gamma_j^2\right) s - \gamma_j W_s^{1,j}\right\} ds\right). \end{aligned} \tag{3.47}$$

This yields the approximation

$$\begin{aligned} V_{t_{i+1}}^{j,\Delta} &= \exp\left\{\left(-\kappa_j - \frac{1}{2}\gamma_j^2\right) t_{i+1} + \gamma_j W_{t_{i+1}}^{1,j}\right\} \\ &\quad \times \left(V_{t_0}^j + a_j \frac{\Delta}{2} \sum_{k=0}^{i} \left[\exp\left\{\left(\kappa_j + \frac{1}{2}\gamma_j^2\right) t_k - \gamma_j W_{t_k}^{1,j}\right\}\right.\right. \\ &\quad \left.\left. + \exp\left\{\left(\kappa_j + \frac{1}{2}\gamma_j^2\right) t_{k+1} - \gamma_j W_{t_{k+1}}^{1,j}\right\}\right]\right) \end{aligned} \tag{3.48}$$

for $i \in \{0,1,\ldots\}$.

The following describes the almost exact simulation of the logarithms of stocks $X_{t_{i+1}}^j = \ln\left(\hat{S}_{t_{i+1}}^j\right)$, $j \in \{0,1,2,\ldots,d\}$. One can represent the value of $X_{t_{i+1}}^j$ at time t_{i+1} as

$$\begin{aligned} X_{t_{i+1}}^j &= X_{t_i}^j - \frac{1}{2}\int_{t_i}^{t_{i+1}} V_u^j du + \frac{2\varrho_j}{\gamma_j}\left(\sqrt{V_{t_{i+1}}^j} - \sqrt{V_{t_i}^j}\right) \\ &\quad - \frac{2\varrho_j}{\gamma_j}\int_{t_i}^{t_{i+1}} \left(\frac{a_j}{2\sqrt{V_u^j}} - \left(\frac{\kappa_j}{2} + \frac{\gamma_j^2}{8}\right)\sqrt{V_u^j}\right) du \\ &\quad + \sqrt{1-\varrho_j^2}\int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j}. \end{aligned} \tag{3.49}$$

Furthermore, the distribution of

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j}, \tag{3.50}$$

conditioned on the path of V^j, is conditionally Gaussian with mean zero and variance $\int_{t_i}^{t_{i+1}} V_u^j du$ because V^j is independent of the Brownian motion $W^{2,j}$ for all $j \in$

$\{0, 1, 2, \ldots, d\}$. Moreover, it is possible to approximate $\int_{t_i}^{t_{i+1}} V_u^j du$ given the path of the process V^j. One can use here again the trapezoidal approximation

$$\int_{t_i}^{t_{i+1}} V_u^j du \approx \frac{\Delta}{2} \left(V_{t_i}^j + V_{t_{i+1}}^j \right) \tag{3.51}$$

to obtain the conditionally Gaussian integral increment

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u^j} dW_u^{2,j} \approx \mathcal{N}\left(0, \frac{\Delta}{2} \left(V_{t_i}^j + V_{t_{i+1}}^j \right)\right). \tag{3.52}$$

Similarly, it is possible to approximate the second integral on the right hand side of (3.49) given in the form

$$\int_{t_i}^{t_{i+1}} \left(\frac{a_j}{2\sqrt{V_u^j}} - \left(\frac{\kappa_j}{2} + \frac{\gamma_j^2}{8} \right) \sqrt{V_u^j} \right) du \approx \tag{3.53}$$
$$\frac{\Delta}{2} \left(\frac{a_j}{2\sqrt{V_{t_i}^j}} - \left(\frac{\kappa_j}{2} + \frac{\gamma_j^2}{8} \right) \sqrt{V_{t_i}^j} + \frac{a_j}{2\sqrt{V_{t_{i+1}}^j}} - \left(\frac{\kappa_j}{2} + \frac{\gamma_j^2}{8} \right) \sqrt{V_{t_{i+1}}^j} \right).$$

This approximation can be achieved with high accuracy when the time step size is small by using again the trapezoidal quadrature formula. In this manner one obtains an efficient almost exact simulation technique for the multi-asset ARCH-diffusion model, which converges in distribution as the time step size decreases.

Now, benchmarked primary security accounts are simulated as multi-dimensional ARCH diffusions. For simplicity, the same squared volatility process is used for all benchmarked primary security accounts, where $a = 0.0469$, $\kappa = 1.3253$, $\gamma = 1$ and $V_0 = 0.0174$, similar as for the Heston squared volatility model. Furthermore, the driving noise of each of the benchmarked asset prices is correlated with $\varrho = -0.7165$ to the noise that drives the squared volatility process. The extra Wiener processes that drive the benchmarked asset prices are independent from each other. Fig. 3.12 shows the first 20 simulated benchmarked risky primary security accounts with Pareto distributed initial values. A typical trajectory of the squared volatility under the ARCH-diffusion model is displayed in Fig. 3.13.

Fig. 3.14 shows the constant benchmarked NP, the benchmarked EWI and the benchmarked MCI. Also here the EWI appears to be an excellent proxy of the NP, while the MCI diverges from the NP in the long run. Also here one notes the smaller volatility of the benchmarked EWI compared to the benchmarked MCI.

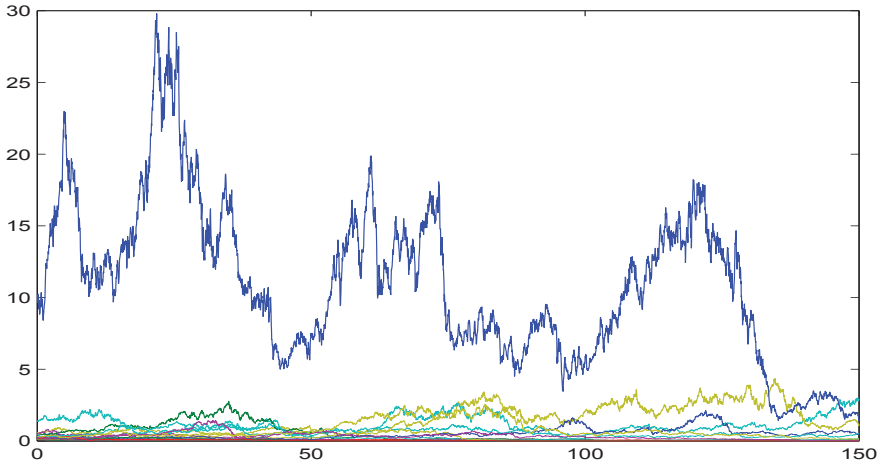

Figure 3.12: Simulated benchmarked primary security accounts under the ARCH-diffusion model

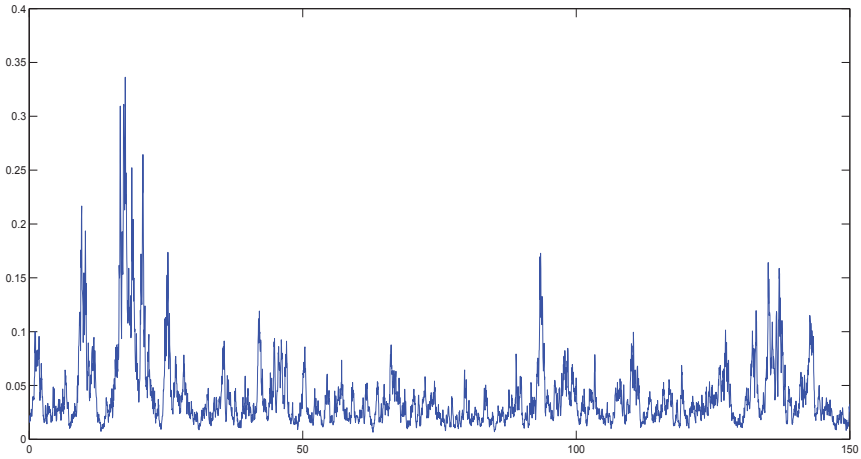

Figure 3.13: Simulated squared volatility under the ARCH-diffusion model

Stylized Minimal Market Model

The previous models can be interpreted as rather direct generalizations of the Black-Scholes model, obtained by introducing some stochastic volatility process. In most versions of these models, when applied in practice, the benchmarked primary security accounts are martingales or rather close to martingales. Consider now the stylized minimal market model (MMM), which is similar to the version of the MMM described in Platen (2001) and Platen & Heath (2010). This model generates by its nature strict supermartingales as benchmarked primary security accounts out of scalar diffusion processes. Each benchmarked primary security account is the inverse of a time transformed squared Bessel process of dimension four and, thus, a strict supermartingale; see Revuz & Yor (1999). The MMM models the jth benchmarked

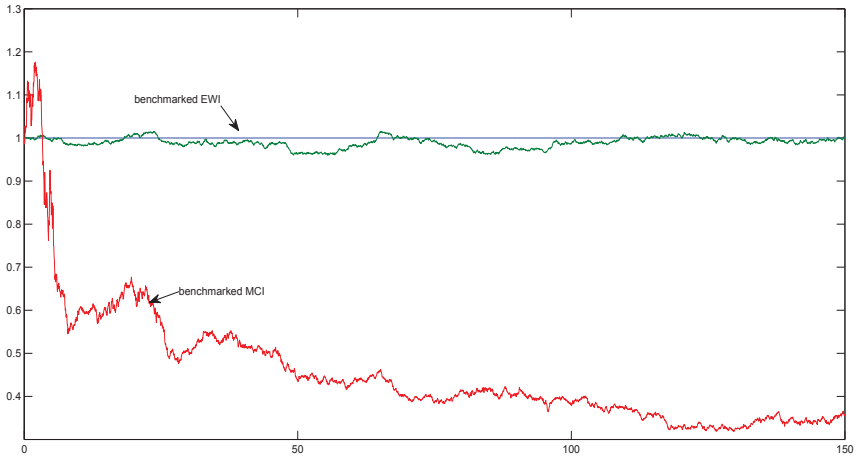

Figure 3.14: Simulated benchmarked NP, EWI and MCI under the ARCH-diffusion model

primary security account by the expression

$$\hat{S}_t^j = \frac{1}{Y_t^j \alpha_t^j}, \tag{3.54}$$

where $\alpha_t^j = \alpha_0^j \exp\{\eta^j t\}$, $j \in \{0, 1, \ldots, d\}$. Here η^j is the jth net growth rate for $j \in \{0, 1, \ldots, d\}$, and Y_t^j is the time t value of the square root process Y^j, which satisfies the SDE

$$dY_t^j = \left(1 - \eta^j Y_t^j\right) dt + \sqrt{Y_t^j} d\tilde{W}_t^j \tag{3.55}$$

for $t \in [0, \infty)$, where $Y_0^j = \frac{1}{\eta^j}$ for $j \in \{0, 1, \ldots, d\}$. Here $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t = (\tilde{W}_t^0, \tilde{W}_t^1, \ldots, \tilde{W}_t^d)^\top, t \in [0, \infty)\}$ is a vector of correlated Wiener processes. Obviously, there are similarities with the square root processes we considered in Section 1.4. Note also the similarity of the equation (3.55) and the equation (4.3) describing the affine nature of the normalized aggregate wealth dynamics in Chapter 4.

Note that under the MMM, $S^j(\varphi^j(t)) = Y_t^j \alpha_t^j$ is a squared Bessel process of dimension four in the, so called, φ^j-time. That is, one has the SDE

$$dS^j(\varphi^j(t)) = 4d\varphi^j(t) + 2\sqrt{S^j(\varphi^j(t))} d\bar{W}^j(\varphi^j(t)) \tag{3.56}$$

for $t \in [0, \infty)$, where

$$\varphi^j(t) = \frac{\alpha_0^j}{4\eta^j} \left(\exp\{\eta^j t\} - 1\right) \tag{3.57}$$

and

$$d\bar{W}^j(\varphi^j(t)) = d\tilde{W}_t^j \sqrt{\frac{d\varphi^j(t)}{dt}}. \tag{3.58}$$

The inverse $\hat{S}^j(\varphi^j(t))$ of the jth squared Bessel process of dimension four in φ^j-time satisfies the SDE

$$d\hat{S}^j(\varphi^j(t)) = -2\left(\hat{S}^j(\varphi^j(t))\right)^{\frac{3}{2}} d\bar{W}^j(\varphi^j(t)) \tag{3.59}$$

for $t \in [0, \infty)$. Note that, $\hat{S}^j$ is a local martingale. More precisely, it can be shown that $\hat{S}^j$ is a nonnegative strict local martingale and, thus, a strict supermartingale, see Revuz & Yor (1999).

It remains to explain the exact simulation of benchmarked primary security accounts under the MMM. Given the time discretization $0 < t_0 < t_1 < \ldots$, where $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, one first generates by (3.57) for $j \in \{0, 1, \ldots, d\}$ the φ^j-time at the physical time t_i.

The next step is to simulate four independent Wiener processes $\bar{W}^{k,j}$, $k \in \{1, 2, 3, 4\}$, in φ^j-time. This can be achieved by calculating

$$\bar{W}^{k,j}_{t_{i+1}} = \bar{W}^{k,j}_{t_i} + \sqrt{\varphi^j(t_{i+1}) - \varphi^j(t_i)} Z^k_{i+1}, \tag{3.60}$$

where $Z^k_{i+1} \sim \mathcal{N}(0, 1)$ is a standard Gaussian random variable. Here $k \in \{1, 2, 3, 4\}$, $j \in \{0, 1, \ldots, d\}$ and $i \in \{0, 1, \ldots\}$.

Then the jth benchmarked primary security account at time t_{i+1} is obtained by the expression

$$\hat{S}^j_{t_{i+1}} = \hat{S}^j\left(\varphi^j(t_{i+1})\right) = \frac{1}{\sum_{k=1}^4 \left(w^k + \bar{W}^{k,j}_{t_{i+1}}\right)^2}, \tag{3.61}$$

for $i \in \{0, 1, \ldots\}$, where $\hat{S}^j_0 = \sum_{k=1}^4 (w^k)^2$.

Now, one can simulate the independent benchmarked risky primary security accounts according to the MMM dynamics. The simulation uses the net growth rate $\eta^j = 0.09$ and the scaling parameter $\alpha^j_0 = 0.05$. Fig. 3.15 plots the first 20 simulated benchmarked primary security accounts. These processes are strict supermartingales and one clearly observes their systematic long run downward trend. Remarkable are the extreme values of benchmarked asset prices that typically appear from time to time under the MMM. The jth volatility equals under the MMM at time t_i the expression $\hat{S}^j_{t_i}\alpha^j_{t_i}$. Fig. 3.16 plots a typical path of the squared volatility under the MMM.

The constant benchmarked NP, benchmarked EWI and benchmarked MCI are shown in Fig. 3.18. The EWI represents a good proxy for the NP. In this case the

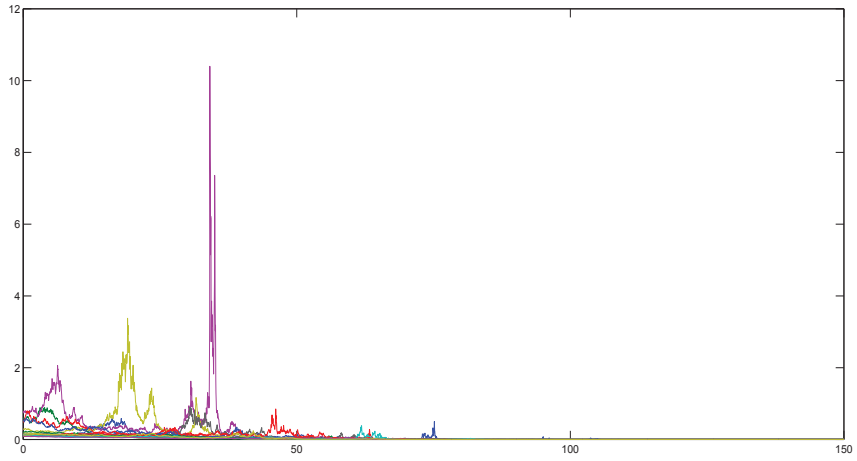

Figure 3.15: Simulated benchmarked primary security accounts under the MMM

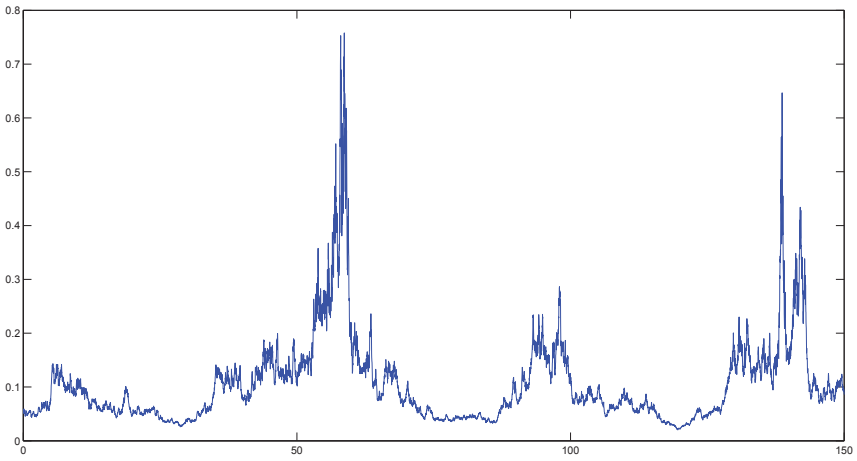

Figure 3.16: Simulated squared volatility under the MMM

benchmarked MCI does by far not reach the long run performance of the benchmarked EWI. One observes a qualitative difference between these two processes. This difference is explained by the fact that the benchmarked MCI, as the sum of strict supermartingales, is a strict supermartingale. It, therefore, exhibits a significant long run downward trending mean. On the other hand, the return process of the EWI, by keeping the fractions constant, makes the expected quadratic variation of the square integrable benchmarked EWI finite, whereas that of the benchmarked MCI is infinite. It follows from the Naive Diversification Theorem that the return process for the benchmarked EWI tends for the increasing number of constituents to zero, which makes its benchmarked value a constant. Equal value weighting can be interpreted as a form of hedging. It turns out that one needs to hedge to avoid

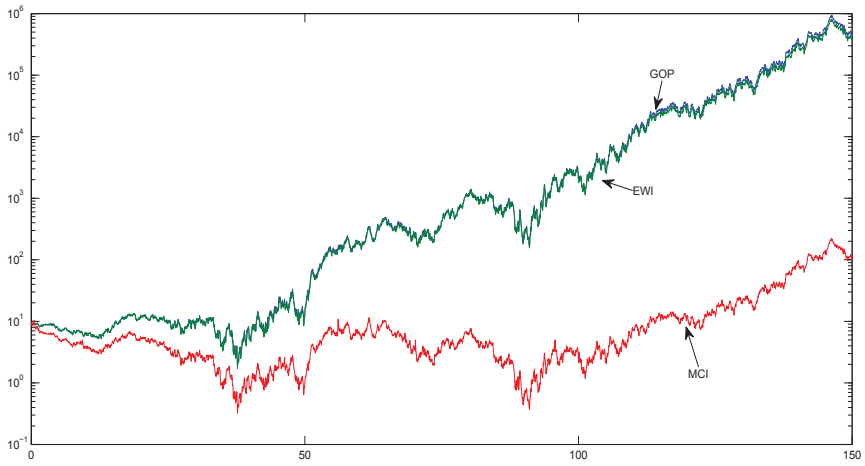

Figure 3.17: Simulated NP, EWI and MCI under the MMM in log-scale

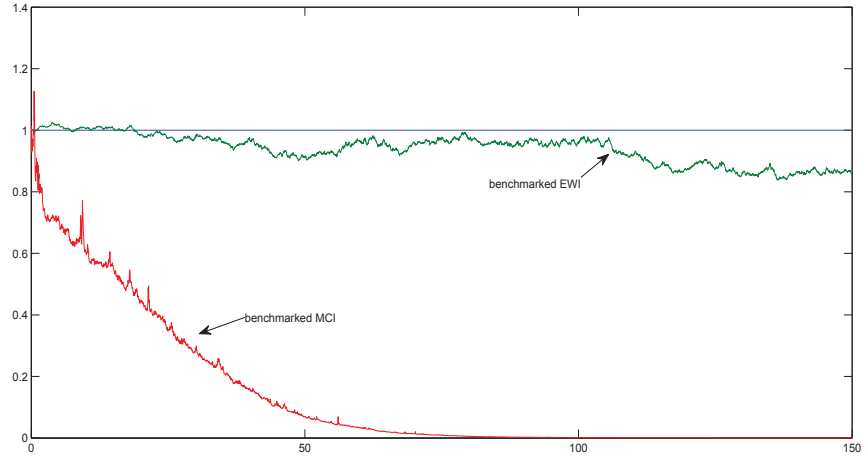

Figure 3.18: Simulated benchmarked NP, EWI and MCI under the MMM

the strict supermartingale property of the constituents to be inherited by the resulting benchmarked portfolio. It is not enough to diversify over the constituents as is achieved by the MCI. The exposure to each source of uncertainty needs to be controlled by some kind of hedging. In this section the hedging is facilitated by equal value weighting.

The important observation of this section is that if benchmarked primary security accounts are, in reality, strict local martingales, then portfolio management should construct well diversified portfolios that when benchmarked form martingales. These portfolios when appropriately diversified would perform in the long run better than the market capitalization weighted portfolio, which is then a strict supermartingale.

Chapter 4

The Affine Nature of Aggregate Wealth Dynamics

This chapter derives a parsimonious two-component affine diffusion model for a world stock index to capture the dynamics of aggregate wealth. The observable state variables of the model are the normalized index and the inverse of the stochastic market activity, both modeled as square root processes. The square root process in market activity time for the normalized aggregate wealth emerges from the affine nature of aggregate wealth dynamics, which will be derived under basic assumptions and does not contain any parameters that have to be estimated. The proposed model employs only three well interpretable structural parameters, which determine the market activity dynamics, and three initial parameters. It is driven by the continuous, nondiversifiable uncertainty of the market and no other source of uncertainty. The model, to be valid over long time periods, needs to be formulated in a general financial modeling framework beyond the classical no-arbitrage paradigm. It reproduces a list of major stylized empirical facts, including Student-t distributed log-returns, which we observed in Chapter 2, and typical volatility properties. Robust methods for fitting and simulating this model are demonstrated.

4.1 Dynamics of Aggregate Wealth

To obtain an idea what type of stochastic process may be well suited for modeling a normalized world stock index, we ask the question, what would be the type of limiting diffusion process that may likely emerge for the aggregate wealth of an economy? We consider normalized units of wealth such that the long term exponential growth of the economy is taken out on average and some kind of equilibrium may be ob-

served. For simplicity, consider a multi-period discrete time setting, and assume at time $\tau_i = i\Delta, i \in \{0, 1, \dots\}$, $0 < \Delta < 1$, that the economy has accumulated the total wealth $Y^{\Delta}_{\tau_i}$, each wealth unit worth $\sqrt{\Delta}$, $Y^{\Delta}_0 = Y_0 > 0$. This means, one has about $\frac{Y^{\Delta}_{\tau_i}}{\sqrt{\Delta}}$ wealth units at time τ_i. During the period until time $\tau_{i+1} = \tau_i + \Delta$ each wealth unit invests in a "project" or pursues some economic activity, which "consumes" on average $\eta\Delta$ units of wealth, $\eta > 0$. Furthermore, $\beta\sqrt{\Delta}$ new wealth units are generated, on average, during this time period, $\beta > 0$. This means, the mean of the increment of aggregate wealth for the time period equals $(\beta - \eta Y^{\Delta}_{\tau_i})\Delta$.

Most important for understanding the nature of aggregate wealth dynamics is the fact that, in a first approximation, it is appropriate to assume that the outcomes of the "projects" and economic activities are independent of each other. This means, if we assume for simplicity that each "project" and economic activity generates in the period $[\tau_i, \tau_{i+1})$ wealth with variance $\upsilon^2\Delta^{\frac{3}{2}}$, then the variance of the increment of the aggregate wealth amounts to $\upsilon^2 Y^{\Delta}_{\tau_i}\Delta$, since one has about $\frac{Y^{\Delta}_{\tau_i}}{\sqrt{\Delta}}$ wealth units at time τ_i. Obviously, its deviation is then $\upsilon\sqrt{Y^{\Delta}_{\tau_i}\Delta}$. Note that, very naturally, a square root of aggregate wealth appears in the deviation of aggregate wealth. Note that it does not matter much what type of distribution the individual wealth outcomes have when these are generated by different "projects" or economic activities. Intuitively, by the Central Limit Theorem the increment $Y^{\Delta}_{\tau_{i+1}} - Y^{\Delta}_{\tau_i}$ of aggregate wealth is for $\Delta \to 0$ asymptotically conditionally Gaussian distributed with the above mean and variance. Moreover, the difference equation

$$Y^{\Delta}_{\tau_{i+1}} - Y^{\Delta}_{\tau_i} = \left(\beta - \eta Y^{\Delta}_{\tau_i}\right)\Delta + \upsilon\sqrt{Y^{\Delta}_{\tau_i}\Delta}W_{\tau_i}, \tag{4.1}$$

with ΔW_{τ_i} denoting a random variable with approximately mean zero and variance Δ, resembles an Euler scheme, see Kloeden & Platen (1999), of the square root process $Y = \{Y_\tau, \tau \geq 0\}$ with stochastic differential equation (SDE)

$$dY_\tau = (\beta - \eta Y_\tau)\, d\tau + \upsilon\sqrt{Y_\tau} dW(\tau), \tag{4.2}$$

$\tau \geq 0, Y_0 > 0$. Here $W = \{W(\tau), \tau \in [0, \infty)\}$ is a standard Brownian motion. Along the lines of Alfonsi (2005) and Diop (2003) it follows under general assumptions, that for vanishing time step size $\Delta \to 0$ the aggregate wealth process Y^{Δ} converges in a weak sense, see also Kloeden & Platen (1999), to the affine diffusion process $Y = \{Y_\tau, \tau \in [0, \infty)\}$, satisfying the SDE (4.2). This type of process was introduced

in Section 1.4. Note that the solution $Y = \{Y_t, t \in [0, \infty)\}$ of the SDE (4.2) is a square root process and, thus, an affine process, see e.g. Duffie & Kan (1994). More precisely, a square root process is sometimes referred to as CIR (Cox, Ingersoll & Ross (1985)) interest rate process. This nonnegative affine process has an explicitly known transition density, see Revuz & Yor (1999).

Most important for the nature of aggregate wealth dynamics is the fact that the diffusion coefficient turns out to be proportional to the square root of the total wealth Y_τ. This is a consequence of the independence of outcomes of "projects" and economic activities. The drift is here in a first approximation, and for natural reasons, linear in Y_τ, which makes Y a highly tractable affine process.

We know a wide range of important properties of this process. For example, when assuming $\beta > \frac{v^2}{2}$, the value Y_τ never reaches zero; see Revuz & Yor (1999). Note that the volatility of the limiting aggregate wealth process emerges as $\frac{v}{\sqrt{Y_\tau}}$. This kind of volatility process models the well-known leverage effect without involving considerations on company value to debt ratios, as e.g. suggested in Black (1976). Under the above approach volatility emerges naturally as a consequence of the uncertain nature of economic activity. The market capitalization weighted index represents, in a first approximation, the aggregate wealth of the economy, which is here modeled.

Economic and market activity change over time. People react to the observed aggregate wealth evolution in their economic activity, including their trading intensity. It is well known, see Kahneman & Tversky (1979), that human behavior exaggerates for decisions under risk small probabilities, as described in prospect theory. This exaggeration applies also to the above extracted "natural" dynamics of volatility. One may assume that the human participants in the economy overreact to the observed ups and downs of aggregate wealth. When volatility should be high under the above derived normalized index dynamics in τ-time it becomes even higher in physical time and vice versa.

In Section 4.3 we will integrate the above revealed affine nature of aggregate wealth dynamics in an index model, which incorporates then also the typical overreactions in human behavior via a random market activity time. This time will be driven by the same nondiversifiable uncertainty that drives the index dynamics and

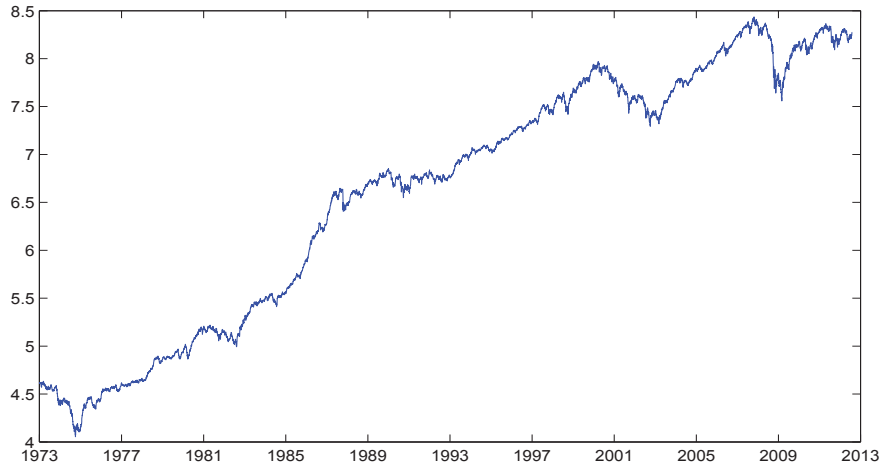

Figure 4.1: Logarithm of a well diversified world stock index MCI

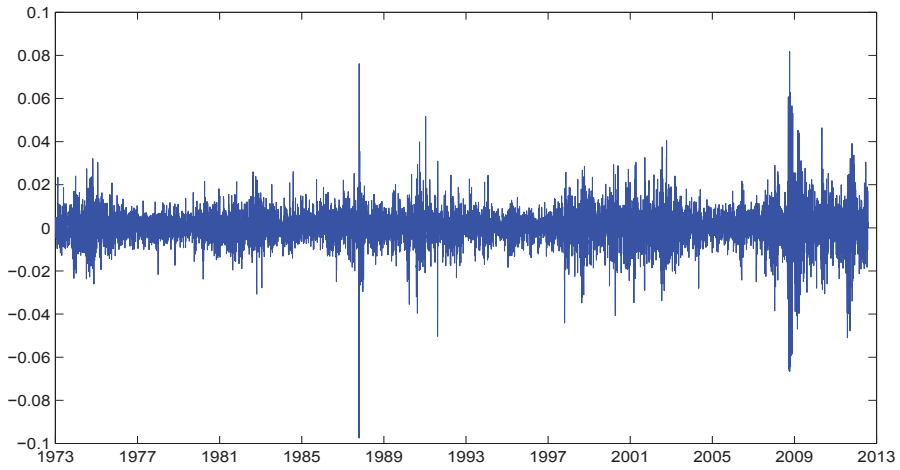

Figure 4.2: Log-returns of the MCI

will exaggerate the volatility to the extent observed in reality.

4.2 Stylized Empirical Facts

According to Popper (1934) one cannot "prove" the validity of a model. The best one can achieve is to find that one cannot falsify the candidate model but can falsify competing models based on empirical evidence. In this section we employ standard statistical and econometric techniques to establish for a well diversified equity index a list of stylized empirical facts. Deliberately, we rely on common and robust techniques and avoid any technicalities.

We use for our empirical analysis the total market equity index (with mnemon-

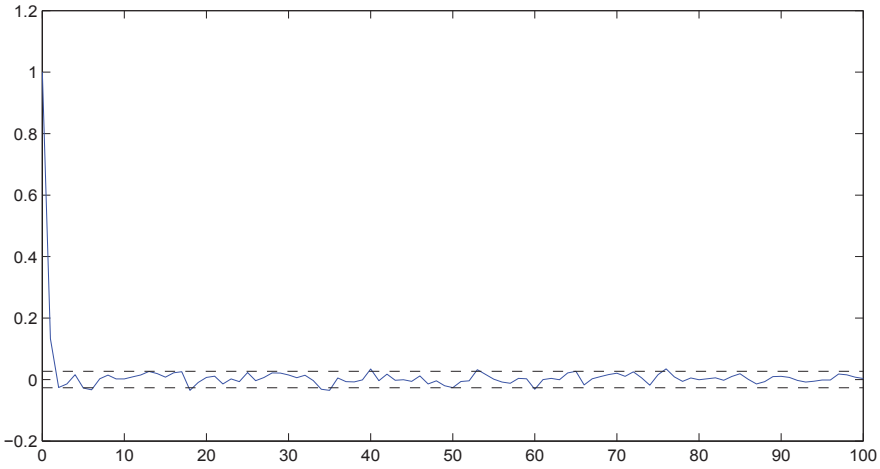

Figure 4.3: Average autocorrelation function for log-returns of the MCI in different currency denominations

ics TOTMKWD), downloaded from Thomson Reuters Datastream in August 2012. This index is a market capitalization weighted world stock index, and very similar to the MSCI world index. In this chapter, we call this equity index, the market capitalization index (MCI). Fig. 4.1 displays the logarithm of the MCI while Fig. 4.2 shows the log-returns of the MCI.

The list of stylized empirical facts, established below, will allow us to circle in a parsimonious, tractable model that cannot be easily falsified with respect to these facts:

(i) Uncorrelated Returns

First, we studied the MCI denominated in the 26 currencies of the countries listed in Table 4.2. For each currency the autocorrelation of the log-returns was calculated for each shown number of days between observations and then averaged over the currencies, which provided in Fig. 4.3 the well-known typical graph; see e.g. Ghysels, Harvey & Renault (1996). It shows practically no correlation between the different returns. Here the horizontal axis gives the number of days between the observed returns. The autocorrelation function in Fig. 4.3 is, in principle, located between the displayed 95% upper and lower confidence bounds. In the literature this is a well accepted property of index returns. Therefore, we formulate our first stylized empirical fact:

Index log-returns are not correlated over time.

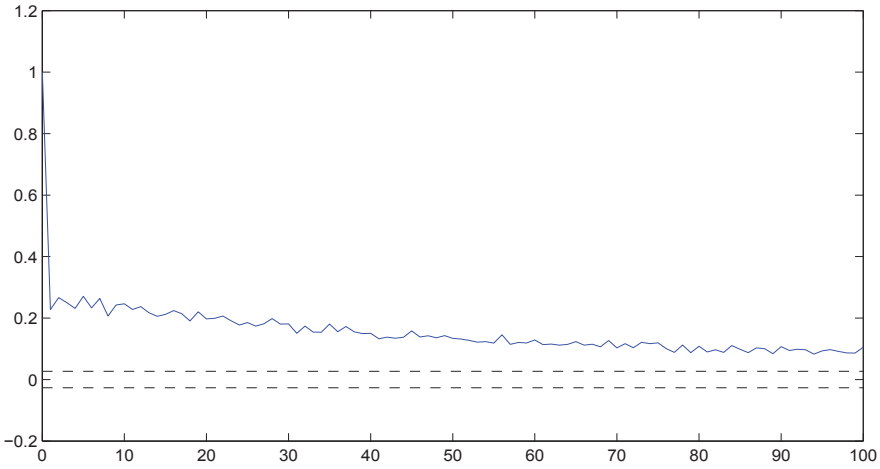

Figure 4.4: Average autocorrelation function for the absolute log-returns of the MCI in different currency denominations

(ii) Correlated Absolute Returns

Similarly as above, also the autocorrelation for the absolute log-returns of the MCI has been estimated for each currency denomination. In Fig. 4.4 the resulting average of the estimated autocorrelation functions of the absolute log-returns for the 26 currency denominations of the MCI is displayed in dependence on the time lag in days. As known from many other studies, see e.g. Ghysels, Harvey & Renault (1996), one observes that this sample autocorrelation does not die away fast and remains after 100 days still far from the 95% confidence bound for the hypothesis that there may be no autocorrelation. Furthermore, the decay of the average autocorrelation does not seem to be exponential. Even for large lags of several months there is still some not negligible autocorrelation of absolute log-returns present. Note that the Black-Scholes model and many of its extensions, including exponential Lévy models, do not exhibit this property. Also this is in the literature a widely accepted empirical property of equity indices and we formulate the second stylized empirical fact:

The sample autocorrelation of absolute log-returns does not die away exponentially, and cannot be neglected for large lags.

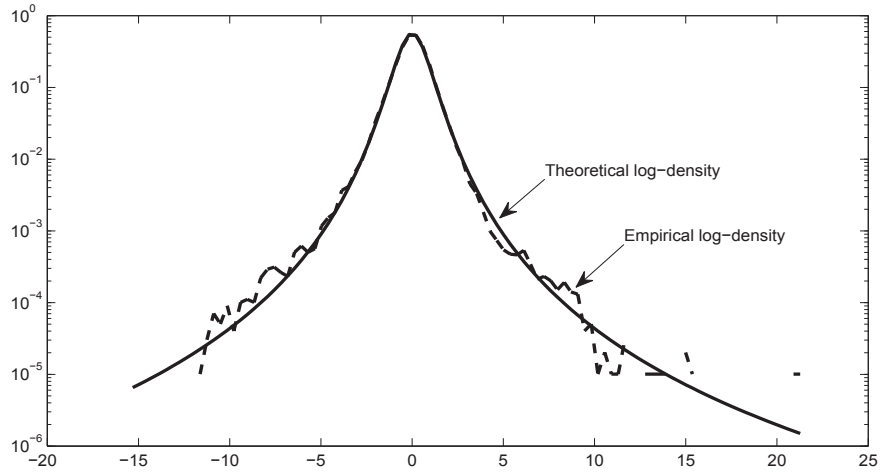

Figure 4.5: Logarithms of empirical density of normalized log-returns of the MCI and Student-t density with 3.5 degrees of freedom

Table 4.1: Log-Maximum likelihood test statistic for daily, weekly and fortnightly log-returns of the MCI

Frequency	Student-t	NIG	Hyperbolic	VG	ν
Daily	3.016652	1336.306309	4453.111995	5367.048317	3.495706
Weekly	0.000000	211.666854	519.162808	701.370317	4.307228
Fortnightly	3.338603	60.216977	204.104996	47.288206	4.216611

(iii) Student-t Distributed Returns

From the 26 currency denominations with mostly 39 years of daily observations we have a substantial dataset of log-returns available, which provide a total observation period of about 1000 years. Similarly as in Fergusson & Platen (2006) and Platen & Rendek (2008) as well as Section 2.3, for each of the 26 currency denominations we shifted the log-returns of the MCI so that their average for a given currency denomination becomes zero. For each currency denomination these shifted returns are then scaled such that a variance of one is estimated. Finally, all shifted and scaled returns are joined in one large dataset of $268,684$ daily normalized returns. Note, by choosing as object of study a global benchmark, the MCI, it became possible to form such a large dataset for log-returns where one can expect clear statistical results. The corresponding log-histogram for the entire sample of log-returns is shown in Fig. 4.5. This

Table 4.2: Log-Maximum likelihood test statistic for different currency denominations of the MCI

Country	Student-t	NIG	Hyperbolic	VG	ν
Australia	0.000000	58.129392	144.865134	180.029220	3.918181
Austria	1.000815	41.266882	157.562238	190.955439	3.435910
Belgium	0.976188	40.671955	154.030262	189.996360	3.458126
Canada	0.000000	42.354744	94.626898	124.885496	4.454300
Denmark	0.194751	35.909603	110.311722	144.178062	3.907690
Finland	3.953519	29.507246	156.290841	176.470426	3.192224
France	2.454694	33.126251	153.720450	186.278677	3.332276
Germany	1.089218	37.984695	146.623206	178.879859	3.462573
Greece	0.087889	47.616770	159.745413	185.004929	3.368001
Hong Kong	0.955483	33.163802	142.601238	168.218205	3.293006
India	0.000000	58.935308	170.125958	208.765488	3.648135
Ireland	0.007503	66.903516	218.951365	254.213268	3.361688
Italy	3.975738	26.302040	144.201210	167.455552	3.244692
Japan	1.203367	42.032471	186.745356	218.934846	3.228296
Korea S.	0.000000	100.310795	314.042722	345.682862	3.107851
Malaysia	0.040659	66.371816	210.194378	253.251366	3.438201
Netherlands	0.900068	40.155097	151.802950	185.510512	3.474006
Norway	0.089357	39.067234	112.438014	152.293627	4.001805
Portugal	2.253549	39.426104	188.698882	209.615521	3.150901
Singapore	0.315454	27.494392	99.564426	123.186925	3.592624
Spain	0.715989	47.820695	192.030088	219.961909	3.263888
Sweden	0.000000	68.297279	159.571420	200.409252	4.020877
Taiwan	0.723684	29.571185	120.256752	140.772751	3.358707
Thailand	0.972236	43.834495	225.116046	244.468484	2.924801
UK	0.949227	19.643975	68.567717	103.504763	4.253508
USA	0.138002	46.862320	155.995200	192.490293	3.566637

figure displays also the log-density of a Student-t distribution with 3.5 degrees of freedom. Remarkable is the excellent visual fit for the medium range of the log-return values and also for extreme log-returns. The latter are most important in risk management; see McNeil, Frey & Embrechts (2005). We notice also a slightly negative skew in the log-returns when compared to the shown symmetric Student-t log-density, which has been observed also in other studies; see e.g. Ghysels, Harvey & Renault (1996).

To obtain an idea about the significance of the above observed Student-t property, a standard maximum likelihood ratio test, described in Section 2.2, has been performed for the rich class of symmetric generalized hyperbolic (SGH) distributions, as introduced in detail in Section 2.2; see also Barndorff-Nielsen (1977) and McNeil, Frey & Embrechts (2005). It turns out that the log-maximum likelihood test statistics for the weekly log-returns is with $0.00000012 < \chi^2_{0.001,1} \approx 0.000002$ such that one cannot reject with 99.9% significance the hypothesis that the Student-t distribution with 4.3 degrees

of freedom is the underlying distribution, see Table 4.1. For daily and fortnightly log-returns the Student-t maximum likelihood test statistic is clearly the smallest when compared to the other special cases of the SGH distribution. A similar study was performed in Section 2.3 with an equi-weighted world index, the EWI104s, which did not include the financial crisis 2007/2008. In that study for daily log-returns, the hypothesis that the Student-t distribution with 4.3 degrees of freedom is the underlying distribution could not be rejected at the 99.9% level of significance. It can be noticed in these kind of studies on diversified world stock indices that a few extreme log-returns, e.g. those of the recent financial crisis, can slightly distort the otherwise perfect Student-t fit. We conclude that it is difficult to identify with extreme significance the distribution and the exact parameters from the available data. However, the Student-t distribution as best fit for well diversified equity index log-returns appears to be a stylized fact.

We studied also separately daily log-returns of the MCI when denominated in each of the 26 different currencies with the results listed in Table 4.2. This table reports the log-maximum likelihood test statistics for different currency denominations of the MCI. In 5 out of the 26 currency denominations the hypothesis that the Student-t density is the true density, when nested in the class of SGH densities, cannot be rejected on a significance level of 99.9%. Models generating the following respective log-returns can be rejected for all currency denominations: the normal inverse Gaussian (NIG) density, appearing e.g. in asset price models of Barndorff-Nielsen (1997); the hyperbolic density, resulting from models e.g. in Eberlein & Keller (1995); and the variance gamma (VG) density, typical for models developed in Madan & Seneta (1990) and Carr et al. (2004). Additionally, in the last column of Table 4.2 we report the estimated degrees of freedom for the Student-t density, which range between 2.9 and 4.5. Here we emphasize again that a few extreme log-returns in a more than 30 year daily dataset can bring the estimate for the degrees of freedom easily down by half a degree.

Also other authors estimated the Student-t distribution from returns of equity indices; see e.g. Markowitz & Usmen (1996a, 1996b) and Hurst & Platen

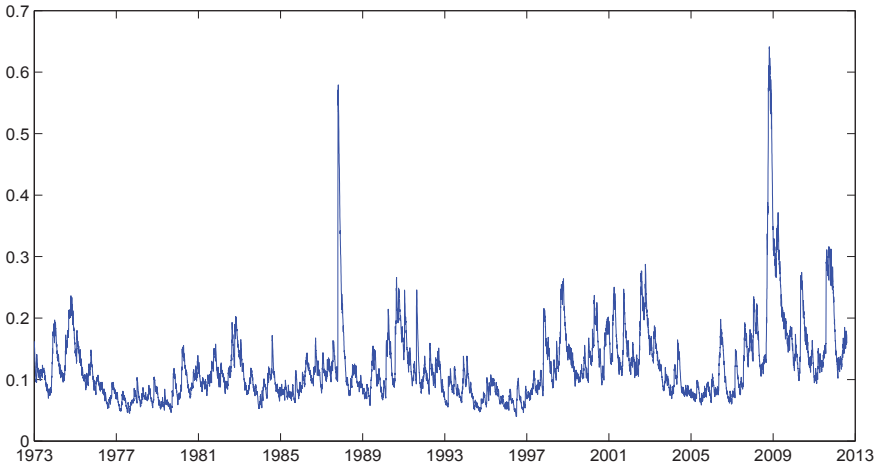

Figure 4.6: Estimated volatility from log-returns of the MCI

(1997). Recently, Alparslan, Tessitore & Usmen (2012) compared the best of the Pearson family of distributions with the best of the family of stable distributions. Their results strongly support the Student-t distribution for index log-returns when using a Bayesian approach.

We summarize the third stylized empirical fact:

Short and longer term log-returns appear, when estimated, as being Student-t distributed with about four degrees of freedom, exhibiting a slight negative skew.

(iv) Volatility Clustering

It is well-accepted that the volatility of an index is stochastic and clusters occasionally over time; see e.g. Ghysels, Harvey & Renault (1996) and the log-returns of the MCI displayed in Fig. 4.2. A standard estimation of volatility from the daily observed discounted MCI in US dollar denomination is shown in Fig. 4.6, where the squared volatility was estimated via a moving average with weight $\alpha\sqrt{\Delta} = 0.0569$ to each newly observed squared log-return, $\alpha = 0.92$, $\Delta = \frac{1}{261}$. This figure confirms for the MCI that its volatility is stochastic and shows clusters of higher values during some time periods. Furthermore, it is reasonable to conclude from the available observations of the MCI in different currency denominations that the volatility has approximately a stationary density. One has to acknowledge the fact that empirically one has so far only access to some moving average type estimate, as shown in Fig. 4.6. The hidden

accurate theoretical volatility path of an index is almost impossible to observe. A volatility index, like the VIX, shown in Fig. 4.11 for the S&P500, gives a market perspective on some quantity closely related to the hidden volatility of the index.

The already mentioned wide range of literature on volatility modeling, including the ARCH and GARCH literature, which originated with Engle (1982), agrees on stochasticity, stationarity and clustering of volatility. The observable stationarity property is not reflected by CEV type models, see e.g. Cox (1975). Also models like the popular SABR model, see Hagan et al. (2002), which uses geometric Brownian motion to model volatility, are not consistent with volatility observations over long time periods. We formulate the fourth stylized empirical fact:

Volatility appears to be stochastic, exhibiting approximately stationary dynamics, with occasional clusters of higher volatility.

(v) Long Term Exponential Growth

For long term risk management, as required for pensions and insurance contracts, the long term average growth of securities is important. In Fig. 4.1 the logarithm of the MCI, discounted by the US savings account, has been displayed. It seems to be reasonable to fit in Fig. 4.1 a trend line. Its slope measures the long term average growth rate of the discounted MCI. By removing this observed average growth from the logarithm of the discounted MCI one obtains the logarithm of the resulting normalized discounted MCI, which is displayed as upper graph in Fig. 4.7. It is reasonable to expect that the normalized discounted MCI could be modeled by a stochastic process that has approximately a stationary density. Many popular models are not consistent with the above mentioned stationarity property, e.g. generalizations of the Black-Scholes model, including popular stochastic volatility models like the Heston model, see Heston (1993). We summarize our observation in the fifth stylized empirical fact:

The discounted index exhibits in the long term on average exponential growth, and the accordingly normalized index appears to have approximately a station-

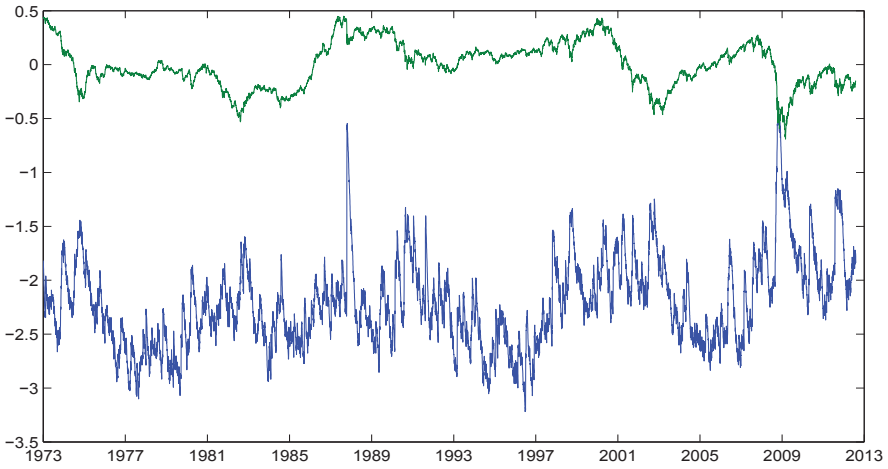

Figure 4.7: Logarithms of normalized discounted MCI (upper graph) and its volatility (lower graph)

ary density.

(vi) Leverage Effect

It has been well documented since the work of Black (1976) and Rubinstein (1976) that the volatility of a normalized equity index appears visually to be negatively correlated with its volatility. Fig. 4.7 visualizes this property by plotting as upper graph the logarithm of the normalized discounted MCI and as lower graph the logarithm of the estimated volatility, previously shown in Fig. 4.6. Visually one notes that both processes appear to fluctuate mostly in opposite directions. Note that the logarithm of the volatility fluctuates in a much wider range. Despite the visual impression of strong negative "dependence", the estimated correlation between the increments of both processes amounts only to -0.1. As has been made clear in recent work by Ait-Sahalia, Fan & Li (2012), and as will be confirmed later via simulation, the hidden theoretical volatility can be perfectly negatively correlated with the normalized index, but the estimated volatility may show only minor negative correlation with the normalized index. Given the strong visual negative "dependence" of the MCI and its volatility, this makes it difficult to identify any other source of uncertainty than the nondiversifiable uncertainty, which from an economic perspective may well drive both, the index fluctuations and the volatility fluctuations.

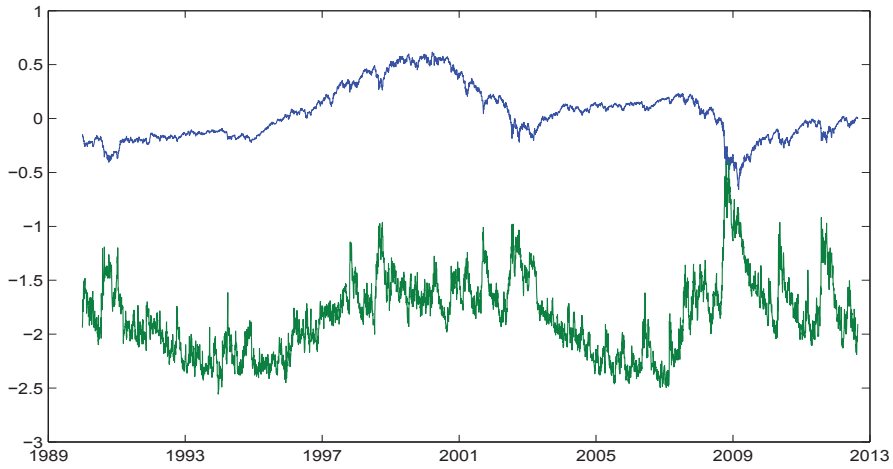

Figure 4.8: Logarithms of normalized discounted S&P500 (upper graph) and VIX (lower graph)

That there may be only one driving noise for the index and its volatility, is supported by the following fact: Strong negative correlation is detectable for equity indices and their volatility index, as long as the latter is available. Fig. 4.8 shows the logarithm of the normalized discounted S&P500 together with the logarithm of its volatility index, the VIX. The estimated correlation between their increments is -0.71. This shows that there is significant negative correlation between these two trajectories. Note that visually there is not much difference between Fig. 4.7 and Fig. 4.8. Simulations of an index with perfectly correlated volatility will later show that the estimated correlation is of similar magnitude as observed for the S&P500 and the VIX. We summarize the above observations in the sixth stylized empirical fact:

The dynamics of an index shows the leverage effect, where the volatility moves visually up when the normalized index moves down and vice versa. Both are potentially driven by the nondiversifiable uncertainty of the market.

(vii) Extreme Volatility at Major Market Downturns

In Fig. 4.7, similarly as in Fig. 4.11, one observes at major market downturns, e.g. in 1987 and 2007/2008, that the logarithm of volatility increases substantially more than the logarithm of the discounted index moves down. However, in periods of more moderate index movements, the magnitudes of volatility and index movements have been more in line. This means, in times of crises

the increase of the volatility is much stronger than could be typically explained by, say, a simple function of the normalized index as model for volatility. On the other hand, there is a significant increase in trading activity observed in times of index decline, which contribute to increased volatility see e.g Ané & Geman (2000). This leads to the seventh and final stylized empirical fact in our list:

At major index downturns volatility increases significantly; more than a functional link between volatility and normalized index could explain. Moreover, market activity increases substantially during sharp index declines, which seems to contribute to increased volatility.

Any reasonably accurate model for the dynamics of the discounted MCI needs to reflect the above described stylized empirical facts. There may not exist too many parsimonious one- or two-component continuous diffusion models that could reflect to a sufficient degree all the above mentioned stylized empirical facts. For identifying the most suitable model class it would be extremely valuable to understand the nature of the index dynamics in a stylized manner. In Section 4.1 we conjectured a particular affine diffusion process, which theoretically should capture well the normalized aggregate wealth dynamics. We propose in the following a more general and still reasonably tractable class of parsimonious models that we will confront with the observed stylized empirical facts. The data should then tell us which specification of this model seems to be most likely.

4.3 Index Model

This section considers an index model, which leans on the theoretically derived normalized aggregate wealth dynamics. It employs two Markovian components: a square root process for modeling the normalized index in some market activity time; and the inverse of a faster moving square root process for modeling the respective market activity. The normalized index and the market activity are driven by the same Brownian motion, which is modeling the nondiversifiable continuous market uncertainty.

Based on the above derived theoretical normalized aggregate wealth dynamics

we propose a continuous time two-component model for the discounted index. Denote by S_t the discounted value of a (total return) index in calendar time $t \geq 0$, denominated in units of the domestic savings account. This discounted index will be expressed by the product

$$S_t = A_{\tau_t}(Y_{\tau_t})^q \tag{4.1}$$

for $t \geq 0$. An exponential function A_{τ_t} of a given τ-time, the market activity time (to be specified below), models the long term average growth of the discounted index as

$$A_{\tau_t} = A \exp\{a\tau_t\} \tag{4.2}$$

for $t \geq 0$.

We use in (4.2) the initial parameter $A > 0$ and the long term average net growth rate $a \in \Re$ with respect to market activity time. One could interpret A_{τ_t} as the average value function, along which aggregate wealth of the economy evolves over time and to which the discounted index reverts back in the long run. We refer to Fig. 4.1, exhibiting the logarithm of the discounted MCI and its trendline, as some visual support for this assumption. It reflects part of the stylized empirical fact (v) on long term exponential growth.

Normalized Index

As a consequence of equation (4.1), the ratio $(Y_{\tau_t})^q = \frac{S_t}{A_{\tau_t}}$ denotes the normalized index at time t. This normalized index is assumed to form an ergodic diffusion process evolving according to τ_t-time. This means it represents a scalar diffusion process with stationary probability density. This property accommodates the remaining part of the stylized empirical fact (v), which requests a normalized index with stationary density. Fig. 4.2 shows the normalized MCI. It seems indeed realistic to model its dynamics by an ergodic process in the above described manner. Motivated by the arguments of Section 4.1, the dynamics of the normalized index are modeled by the qth power, $q > 0$, of a square root process $Y = \{Y_\tau, \tau \geq 0\}$ in τ-time; given by the SDE (4.3) below. We assume that the square root process has dimension $\delta > 2$ to keep it strictly positive. The square root process Y, as a

component of the model, is highly tractable. We assume that it satisfies the SDE

$$dY_\tau = \left(\frac{\delta}{4} - \frac{1}{2} \left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)} \right)^{\frac{1}{q}} Y_\tau \right) d\tau + \sqrt{Y_\tau}\, dW(\tau), \tag{4.3}$$

for $\tau \geq 0$ with $Y_0 > 0$. Only the two parameters $\delta > 2$ and $q > 0$ enter the SDE (4.3) together with its initial value $Y_0 > 0$. Note that the reference level for the square root process Y equals $\frac{\delta}{2} \left(\frac{\Gamma\left(\frac{\delta}{2}\right)}{\Gamma\left(\frac{\delta}{2}+q\right)} \right)^{\frac{1}{q}}$, which is then also its long run mean and one has asymptotically

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau Y_s ds = \frac{\delta}{2} \left(\frac{\Gamma\left(\frac{\delta}{2}\right)}{\Gamma\left(\frac{\delta}{2}+q\right)} \right)^{\frac{1}{q}} \quad \text{P-a.s..} \tag{4.4}$$

Here $\Gamma(\cdot)$ denotes the gamma function and $W = \{W(\tau), \tau \geq 0\}$ denotes a Brownian motion in τ-time, which we specify further below.

Similarly, as discussed for the SDE (4.2), the choice of a positive power of the square root process Y in equation (4.1) creates a leverage effect in a simple and robust manner. This accommodates the stylized empirical fact (vi) on the leverage effect. The power of the square root process in the model has similarities to variants of the constant elasticity of variance (CEV) model, see Ross (1976). However, the volatility of the CEV model is not an ergodic process. The proposed model is a generalization of the minimal market model, see Chapter 13 in Platen & Heath (2010). It falls also into the wider category of local volatility function models, see Dupire (1993) and Derman & Kani (1994a).

The stationary density of the quantity $(Y_\tau)^q$ is explicitly given by the formula

$$p_{Y^q}(y) = \frac{\Gamma\left(\frac{\delta}{2}+q\right)^{\frac{\delta}{2q}}}{q\Gamma\left(\frac{\delta}{2}\right)^{\frac{\delta}{2q}+1}} y^{\frac{\delta}{2q}-1} \exp\left\{ - \left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)} y \right)^{\frac{1}{q}} \right\} \tag{4.5}$$

for $y \geq 0$, see e.g. Platen & Heath (2010). The parametrization in (4.3) is chosen such that the long run mean for the process $(Y)^q$ equals one. More precisely, one obtains by the ergodic theorem

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau (Y_s)^q ds = 1 \quad \text{P-a.s.} \tag{4.6}$$

The stationary density of the square root process Y is a gamma density with δ degrees of freedom and long run mean given in (4.4). This density generates for the normalized index and, thus, for the index, long term log-returns that when estimated

appear to be Student-t distributed with δ degrees of freedom. This Student-t distribution results as a normal mixture distribution, where the inverse of the variance of the log-returns is proportional to Y_τ. The square root process has a stationary gamma density with δ degrees of freedom. In this manner, the qth power of the square root process in (4.1) accommodates the part of the empirical stylized fact (iii), which concerns the observed Student-t distribution for longer term returns. Due to the negative "dependence" between normalized index and its volatility in τ-time one models automatically the observed slight negative skewness of log-returns, expressed in the stylized empirical fact (iii).

The first component of the model, the qth power of the square root process Y describes how the normalized index reverts in τ-time back to its long term mean. This process models in τ-time the overall long term feedback mechanism of the normalized index, which brings in the long run the aggregate wealth back into the range of its long term average level. Furthermore, it ensures that the variance of increments of aggregate wealth is in τ-time proportional to total wealth.

Market Activity Time

Let us now focus on the second component of the model. According to the stylized empirical fact (iii), a Student-t distribution has to be obtained also when fitting short term index returns. Therefore, we introduce in this subsection market activity to reflect human behavior in economic activity and trading, which exaggerates the movements of volatility in reaction to moves of the normalized index.

The market activity process $M = \{M_t, t \geq 0\}$, reflects also the observation expressed in the stylized empirical fact (vii), where the market reacts to a major decrease in the level of the normalized index with significantly increased market activity. Trading activity is known to increase in relative terms when an index is falling and to decrease when it is rising, see e.g. Ané & Geman (2000). This results in speeding up or slowing down, respectively, the time scale under which the above discussed natural normalized index moves. To model conveniently the cumulative effect of speeding up and slowing down market activity, the above mentioned τ-time has been introduced, which is integrated market activity and called market activity time.

More precisely, we model the market activity time τ_t via the ordinary differential

equation

$$d\tau_t = M_t dt \tag{4.7}$$

for $t \geq 0$ with $\tau_0 \geq 0$. Here we call the derivative of τ-time with respect to calendar time t the market activity $\frac{d\tau_t}{dt} = M_t$ at time $t \geq 0$. We will later see that M_t is, in reality, a fast moving process when compared to the square root process Y_{τ_t}. The stylized empirical fact (iii) requests that not only longer term log-returns but also short term returns of an index are estimated as being Student-t distributed. As we can see in Fig. 4.2, the overall long term feedback mechanism of Y_{τ_t} moves relatively slowly. Therefore, the market activity M_t needs to evolve such that it generates, via a mixture of normals, short term log-returns that when estimated indicate a Student-t distribution. Consequently, the market activity process M needs to have as stationary probability density an inverse gamma density.

Additionally, we argued earlier that potentially only the continuous nondiversifiable uncertainty should drive both the normalized index and the market activity. The squared volatility with respect to τ-time equals $\frac{1}{Y_\tau}$, where Y_τ is the square root process in (4.3). According to the stylized empirical fact (vii) the moves of the volatility in τ-time are in t-time exaggerated. Therefore, we model market activity, similarly to squared volatility, as the inverse of a square root process. More precisely, the process $\frac{1}{M} = \{\frac{1}{M_t}, t \geq 0\}$ is assumed to be a fast moving square root process in t-time with the dynamics

$$d\left(\frac{1}{M_t}\right) = \left(\frac{\nu}{4}\gamma - \epsilon\frac{1}{M_t}\right)dt + \sqrt{\frac{\gamma}{M_t}}dW_t, \tag{4.8}$$

for $t \geq 0$ with $M_0 > 0$, where $\gamma > 0$, $\nu > 2$ and $\epsilon > 0$. The Brownian motion $W = \{W_t, t \geq 0\}$ in t-time models the continuous nondiversifiable uncertainty of the market in t-time, and will be linked below in relation (4.11) to $W(\tau_t)$, which is the respective Brownian motion in τ-time. It is straightforward to confirm that the stationary density of the resulting market activity process M is an inverse gamma density with ν degrees of freedom. The resulting stationary density is given by the formula

$$p_M(y) = \frac{\left(\frac{2\epsilon}{\gamma}\right)^{\frac{1}{2}\nu}}{\Gamma\left(\frac{1}{2}\nu\right)} y^{-\frac{1}{2}\nu-1} \exp\left\{-\frac{2\epsilon}{\gamma}\frac{1}{y}\right\}, \tag{4.9}$$

for $y \geq 0$; see e.g. Section 12.4 in Platen & Heath (2010). As the result of normal mixing, this density generates short term index log-returns that when estimated,

appear as being Student-t distributed with about ν degrees of freedom. One can show in the given parametrization that the long term average level for the inverse of market activity equals $\frac{\nu}{4}\frac{\gamma}{\epsilon}$, that is,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \frac{1}{M_s}ds = \frac{\nu}{4}\frac{\gamma}{\epsilon} \quad \text{P-a.s.} \tag{4.10}$$

The Brownian motion $W(\tau)$, which models in market activity time the long term nondiversifiable uncertainty driving Y_τ, is linked to the standard Brownian motion $W = \{W_t, t \geq 0\}$ in t-time in the following manner:

$$dW(\tau_t) = \sqrt{\frac{d\tau_t}{dt}}dW_t = \sqrt{M_t}dW_t \tag{4.11}$$

for $t \geq 0$ with $W_0 = 0$.

The proposed index model is a Markovian two-component model, driven only by one Wiener process $W = \{W_t, t \geq 0\}$. The two components are Y_{τ_t} and $\frac{1}{M_t}$ that solve the SDEs (4.3) and (4.8), respectively. When substituting in the SDE (4.3) formally τ by τ_t, to characterize together with the SDE (4.8) the dynamics of the two-dimensional state variable $(Y_{\tau_t}, \frac{1}{M_t})$, the model is only driven by the Brownian motion $W = \{W_t, t \geq 0\}$. The full specification of the dynamics of the model is, therefore, given by the two SDEs (4.3) and (4.8) together with the respective initial conditions, where $W = \{W_t, t \geq 0\}$ models nondiversifiable uncertainty in t-time.

We can now study and apply the proposed model in many ways. First, let us study the resulting expected rate of return and volatility of the discounted index.

Expected Rate of Return

By application of the Itô formula one obtains from (4.1), (4.2), (4.3), (4.7) and (4.8) for the discounted index S_t the SDE

$$dS_t = S_t\left(\mu_t dt + \sigma_t dW_t\right) \tag{4.12}$$

for $t \geq 0$, with initial value $S_0 = A_0(Y_0)^q$ and expected rate of return

$$\mu_t = \left(\frac{a}{M_t} - \frac{q}{2}\left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)}\right)^{\frac{1}{q}} + \left(\frac{\delta}{4}q + \frac{1}{2}q(q-1)\right)\frac{1}{M_t Y_{\tau_t}}\right) M_t. \tag{4.13}$$

Obviously, short term index log-returns are approximately uncorrelated, as requested by the stylized fact (i); see also Fig. 4.21.

Volatility

The volatility with respect to t-time emerges in the form

$$\sigma_t = q\sqrt{\frac{M_t}{Y_{\tau_t}}}. \tag{4.14}$$

Obviously, it is a function of the two square root processes Y_{τ_t} and $\frac{1}{M_t}$.

Autocorrelation

Due to the volatility (4.14), the absolute returns of S near time t are on average approximately proportional to $\sqrt{\frac{M_t}{Y_{\tau_t}}}$. Consequently, since Y is slow moving the correlation for two absolute log-returns near time t and $t+T$, respectively, is for small T determined by the correlation between $\sqrt{M_t}$ and $\sqrt{M_T}$, which can be shown to decline exponentially fast since M is a fast moving scalar diffusion process; see (4.8). On the other hand, when the time lag T is relatively large, then the correlation between the respective absolute returns is approximately that between $\frac{1}{\sqrt{Y_{\tau_t}}}$ and $\frac{1}{\sqrt{Y_{\tau_t+T}}}$. As we have discussed earlier, $\frac{1}{\sqrt{Y_{\tau_t}}}$ is moving rather slowly and the resulting correlation remains substantial even after several months. Similar to Fig. 4.4 the Fig. 4.22 shows the average autocorrelation of absolute returns under the proposed model obtained from simulated data, which confirms the above remarks. This reflects the stylized empirical fact (ii), where the correlation of absolute log-returns does not die out fast and is also not exponentially declining.

Obviously, the volatility is stochastic and generates under the model clusters of outbursts of volatility, as will be also confirmed when simulating paths of the index and its corresponding volatility in Section 4.5; see e.g. Fig. 4.15. This property accommodates the stylized empirical fact (iv).

3/2 Volatility Model

Of particular interest for the long term dynamics of squared volatility under the proposed model is by relation (4.14) the inverse $\frac{1}{Y_\tau}$ of the square root process with respect to τ-time. By (4.3) one obtains via the Itô formula the SDE

$$d\left(\frac{1}{Y_\tau}\right) = \left(\frac{1}{2}\left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)}\right)^{\frac{1}{q}}\frac{1}{Y_\tau} + \left(1-\frac{\delta}{4}\right)\left(\frac{1}{Y_\tau}\right)^2\right)d\tau - \left(\frac{1}{Y_\tau}\right)^{\frac{3}{2}}dW(\tau) \tag{4.15}$$

for $\tau \geq 0$. This means that the factor $\frac{1}{Y_\tau}$ in the formula for squared volatility

$$\sigma_t^2 = \frac{q^2 M_t}{Y_{\tau_t}}, \tag{4.16}$$

see (4.14), follows in τ-time a $\frac{3}{2}$-volatility model, which was suggested e.g. in Platen (1997). Some versions of $\frac{3}{2}$-volatility models appeared, for instance, in Lewis (2000) and more recently in Carr & Sun (2007). Interestingly, the latter authors provided arguments from the perspective of volatility derivative pricing and hedging which support the choice of a $\frac{3}{2}$-volatility model for a diversified index. We will see in Section 4.5 that our model is in line with empirical evidence on volatility derivatives. Note that the long term average level for $\frac{1}{Y_\tau}$ can be calculated and equals $\frac{2}{\delta-2}\left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)}\right)^{\frac{1}{q}}$. One obtains by the ergodic theorem

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau \frac{1}{Y_s}ds = \frac{2}{\delta-2}\left(\frac{\Gamma\left(\frac{\delta}{2}+q\right)}{\Gamma\left(\frac{\delta}{2}\right)}\right)^{\frac{1}{q}} \quad \text{P-a.s..} \tag{4.17}$$

Finally, we recall that due to the perfect negative "dependence" between the normalized index and its volatility the log-returns of the index exhibit under the model a slight negative skew, which has been mentioned under the stylized empirical fact (iii).

Market Activity

By the fact that M_t has an inverse gamma density with ν degrees of freedom and negative first moment $\frac{\nu}{4}\frac{\gamma}{\epsilon}$, the mean of M_t is explicitly known. By the ergodic theorem this mean amounts to

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t M_s ds = \frac{4}{\nu-2}\frac{\epsilon}{\gamma} \quad \text{P-a.s.} \tag{4.18}$$

Hence, for the proposed model one obtains the asymptotic relation

$$\lim_{h\to\infty}\frac{\tau_{t+h}-\tau_t}{h} = \frac{4}{\nu-2}\frac{\epsilon}{\gamma} \quad \text{P-a.s.} \tag{4.19}$$

for $t \geq 0$. Thus, increments of market activity time can be approximated asymptotically over long time periods h in t-time by the formula

$$\tau_{t+h} - \tau_h \approx \frac{4}{\nu-2}\frac{\epsilon}{\gamma}h \tag{4.20}$$

for h sufficiently large, $h \in (0, \infty)$. Therefore, when quantifying long-term effects, as is the case for many long dated derivatives, it may turn out to be sufficient to employ only average τ-time by using the approximation

$$\tau_t \approx \tau_0 + \frac{4}{\nu - 2}\frac{\epsilon}{\gamma}t, \tag{4.21}$$

$t \geq 0$. This is a convenient property of the model. Essentially, in the long run we have only to deal with a one component model, characterized by the square root process Y, which is highly tractable and runs then in average τ-time for long dated contracts. For short term index derivatives the market activity is relevant and cannot be neglected.

Due to the volatility formula (4.14) and the SDEs (4.3) and (4.8) the fluctuations of both processes $\frac{1}{Y_{\tau_t}}$ and M_t are similarly driven by the fluctuations of the same Brownian motion, $W = \{W_t, t \geq 0\}$. Consequently, one obtains a much stronger reaction of the volatility process to extreme moves of the nondiversifiable uncertainty W than would be typical for standard scalar diffusion models or subordinated stochastic volatility models with independent subordinator. The product of two scalar diffusions in the formula (4.14) for the volatility makes the proposed model very realistic, as will be confirmed in Section 4.4. It encapsulates the fact that there is a close relationship between the random moves of the normalized index value Y_{τ_t}, that is, the index itself and those of the market activity M_t, which is largely driven by the behavior of market participants.

The proposed model can easily be made more flexible via extensions, e.g. allowing parameters to be time dependent, including a more flexible local volatility function, introducing a second Brownian motion or adding jumps. However, in Section 4.4 we will be able to demonstrate when fitting the model that we can reduce the number of flexible parameters and extract a parsimonious stylized model, which represents the affine aggregate wealth dynamics conjectured in Section 4.1.

Now, we are going to discuss the proposed model in a financial modeling framework that goes considerably beyond the currently widely used classical no-arbitrage framework.

Benchmark Approach

The benchmark approach, see Platen (2011) and Platen & Heath (2010), is a general financial market modeling framework that goes beyond the classical no-arbitrage paradigm; see Ross (1976), Harrison & Kreps (1979) and Delbaen & Schachermayer (1994). The approach uses the numéraire portfolio (NP), see Long (1990), as central building block. It has been shown in Chapter 3 (see also Platen & Heath (2010) and Platen & Rendek (2012b)) that a diversified equity index, like the MCI, is a good proxy of the NP. Consequently, under the proposed model the SDE (4.12) could be interpreted as the SDE describing the evolution of the discounted NP of the given investment universe.

Due to the SDE (4.12) and the Itô formula, the dynamics for the benchmarked savings account $\hat{B}_t = \frac{1}{S_t}$, which is the inverse of the discounted NP, is characterized by the SDE

$$d\hat{B}_t = \hat{B}_t \left(\left(-\mu_t + \sigma_t^2 \right) dt - \sigma_t dW_t \right), \tag{4.22}$$

for $t \geq 0$, see (4.13) and (4.14). It follows for

$$\sigma_t^2 \leq \mu_t \tag{4.23}$$

for all $t \geq 0$ that the benchmarked savings account $\hat{B}_t$ forms an $(\underline{\mathcal{A}}, P)$-supermartingale. This is a stochastic process where its current value is greater or equal than its expected future values. This supermartingale property is the key property of any benchmarked nonnegative security under the benchmark approach, see Platen (2011). Since a nonnegative supermartingale that reaches zero will always remain at zero, this property eliminates any possibility for, so called, *strong arbitrage* in the sense of Platen (2011), which is equivalent to the notion of arbitrage in Loewenstein & Willard (2000).

Assumptions on the Model

To guarantee almost surely in the proposed model the absence of strong arbitrage, that is the inequality (4.23), one has by (4.13) and (4.14) to satisfy the following two conditions:

Assumption 1. *First, the dimension δ of the square root process Y needs to satisfy the equality*

$$\delta = 2(q+1). \tag{4.24}$$

Assumption 2. *The long term average net growth rate a with respect to τ-time has to satisfy the inequality*

$$\frac{q}{2}\left(\frac{\Gamma\left(2q+1\right)}{\Gamma\left(q+1\right)}\right)^{\frac{1}{q}} \leq a. \tag{4.25}$$

Assumption 1 is necessary for the supermartingale property of the benchmarked savings account. Assumption 2 guaranties then that the long term average net growth rate, denoted by a, is high enough to make the benchmarked savings account a supermartingale. These two assumptions guarantee jointly that the savings account, when denominated in units of the NP, becomes an $(\underline{\mathcal{A}}, P)$-supermartingale under the proposed model. This means, $\hat{B}_t$ is then satisfying the SDE

$$d\hat{B}_t = \hat{B}_t\left(\left(\frac{q}{2}\left(\frac{\Gamma\left(2q+1\right)}{\Gamma\left(q+1\right)}\right)^{\frac{1}{q}} - a\right)M_t dt - q\sqrt{\frac{M_t}{Y_{\tau_t}}}dW_t\right) \tag{4.26}$$

with zero or negative drift for all $t > 0$. The latter property makes $\hat{B}$ an $(\underline{\mathcal{A}}, P)$-supermartingale.

Beyond Classical No-arbitrage

If equality holds in relation (4.25), then the SDE (4.26) is driftless. Note however, that this does not mean that $\hat{B}$ forms then a true martingale. It is in this case a nonnegative strict local martingale and, thus, a strict supermartingale and not a martingale.

In a complete market the benchmarked savings account, normalized to one at the beginning, represents the Radom-Nikodym derivative for the risk neutral measure. Under the classical paradigm the benchmarked savings account would have to be a true martingale; see Delbaen & Schachermayer (1994). It is important to emphasize that we identified here a model that goes beyond classical no-arbitrage settings. It goes even beyond the benchmark approach as formulated in Platen & Heath (2010), where the benchmarked savings account is assumed to be a local martingale. For the proposed model this may be only the case when equality holds in (4.23). However, the proposed model is well covered by an extension of the benchmark approach formulated in Platen (2011). The SDE for the benchmarked savings account does not need to be driftless. It can have a negative drift. We will see in Section 4.4 the fit of the model to historical data, where a clear negative drift is evident for the analyzed index. This leads the model far away from the classical no-arbitrage

paradigm, and allows it to explain under the benchmark approach "puzzles" and "anomalies" that have been pointed out in the literature when applying the classical no-arbitrage approach. The following indicates such a "puzzle".

Risk Premium Puzzle

By applying the Assumptions 1 and 2 to the proposed model, one obtains by (4.13) the risk premium

$$\mu_t = \sigma_t^2 + M_t \left(a - \frac{q}{2} \left(\frac{\Gamma(2q+1)}{\Gamma(q+1)} \right)^{\frac{1}{q}} \right) \geq \sigma_t^2 \tag{4.27}$$

for all $t \geq 0$. Obviously, since the discounted index is the NP in our complete market, μ_t can be here higher than the classical risk premium σ_t^2 of the NP, which is the square of the volatility of the NP. Under the benchmark approach, described in Platen (2011), such a higher risk premium is permitted. We emphasize that the risk premium puzzle, see Mehra & Prescott (1985), is in our generalized benchmark framework not a "puzzle". It simply states the fact that the expected return of the NP is in reality larger than the square or its volatility. There is no economic reason to stop the expected return to reach a certain level. The classical no-arbitrage paradigm is making restrictive assumptions that do not coincide with reality. For empirical reasons we will have to go in our modeling beyond classical no-arbitrage assumptions to capture realistically the observed long term dynamics of the aggregate wealth.

Pricing of Derivatives

Under the benchmark approach one uses the NP as numéraire or benchmark. The pricing of derivatives applies the real world pricing formula, see Platen (2011), which yields the benchmarked derivative price as real world conditional expectation of the corresponding benchmarked payoff. To calculate such a price for a contingent claim that involves the discounted NP, one can employ the proposed model. The transition density of the square root process in τ-time is a non-central chi-square density with $\delta = 2(q+1)$ degrees of freedom, which makes this component of the model very tractable. For many long dated contingent claims the τ-time can be expected to be well approximated via average τ-time, according to (4.20).

For the pricing of short dated derivatives under the proposed model, numerical techniques can be employed, e.g. similar to those described in Section 13.4 in

Platen & Heath (2010). The proposed model can be expected to recover well the typically negatively skewed implied volatility surface of index options. The random fluctuations of market activity will turn out to be crucial for realistic modeling of short dated index derivatives. They generate the observed strong curvature of the implied volatility surface of index options near the strike and close to maturity.

The proposed model uses only one Brownian motion to drive the two diffusion processes. The three initial and five structural parameters are: the parameter $A \geq 0$ in (4.2) for fitting the initial value of the average exponential growth part; the initial value $M_0 > 0$ of the market activity; the initial value $Y_0 > 0$ of the square root process; the long term average net growth rate $a > 0$ (with respect to τ-time) of the discounted index; the power $q > 0$ of the square root process when forming the normalized index; the parameter q determines also the dimension $\delta = 2(q+1) > 2$, see (4.24), of the square root process; the degrees of freedom $\nu > 2$, see (4.8), of the stationary gamma density of the inverse of the market activity; the reference level parameter ϵ for the market activity, see (4.8) and (4.20); and the scaling parameter $\gamma > 0$ in the diffusion coefficient of the SDE (4.8) for the inverse of market activity. All eight parameters have a clear meaning and can be directly estimated from time discrete observations of the index, as will be demonstrated in the next section. Moreover, we will demonstrate in the next section that by reasoning, as presented in Section 4.1, and by empirical evidence one can reduce the number of structural parameters to three.

4.4 Fitting the Model

It is paramount that the proposed model can be easily fitted in a robust manner to historical index data. This section illustrates a simple and robust step by step method for fitting the proposed model to historical index data. It will turn out that we can fix some of the parameters, which will yield a stylized version of the model. Below we separate the estimation of the parameters into the following steps:

Step 1: Normalization of Index

A linearly regressed function of calendar t-time to the path of the logarithm of the daily observations of the discounted MCI, as shown in Fig. 4.1, yields the straight

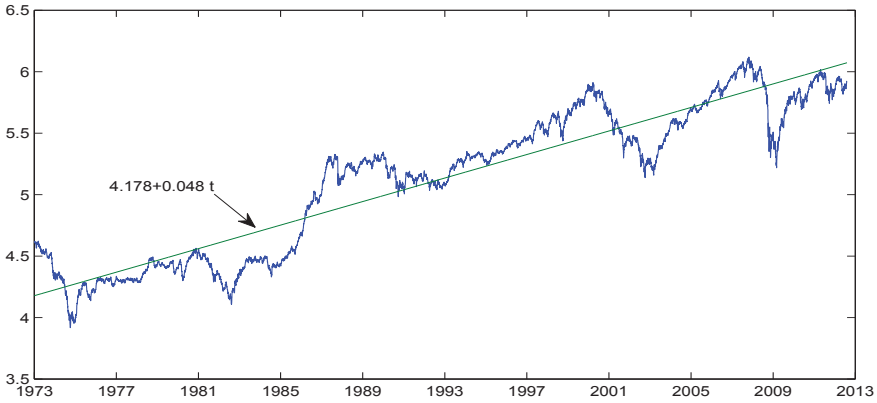

Figure 4.1: Logarithm of the discounted MCI and linear fit

line $4.178+0.048t$. The normalized MCI is then obtained by dividing the discounted index by the long term approximation for the exponential function A_{τ_t}, given on the right hand side of relation (4.28), obtained from (4.2) using (4.20). This means, we set

$$A_{\tau_t} \approx A \exp\Big\{\frac{4a\epsilon}{\gamma(\nu-2)}t\Big\}, \tag{4.28}$$

for $t \geq 0$, and read off from the above mentioned linear regression the estimate $\frac{4a\epsilon}{\gamma(\nu-2)} \approx 0.048$. The resulting discounted MCI $(Y_{\tau_t})^q$ is plotted in Fig. 4.2 with respect to physical t-time, where the initial parameter $A = 65.21$ is estimated by making the average of $(Y_{\tau_t})^q$ approximately to one, as required by (4.6).

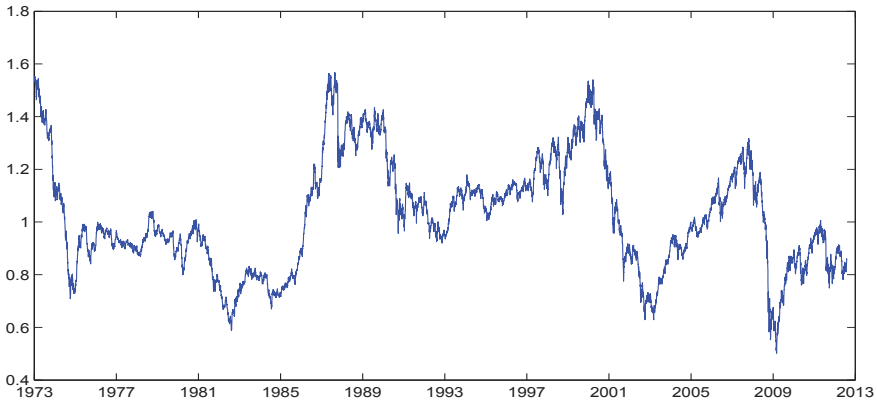

Figure 4.2: Normalized MCI

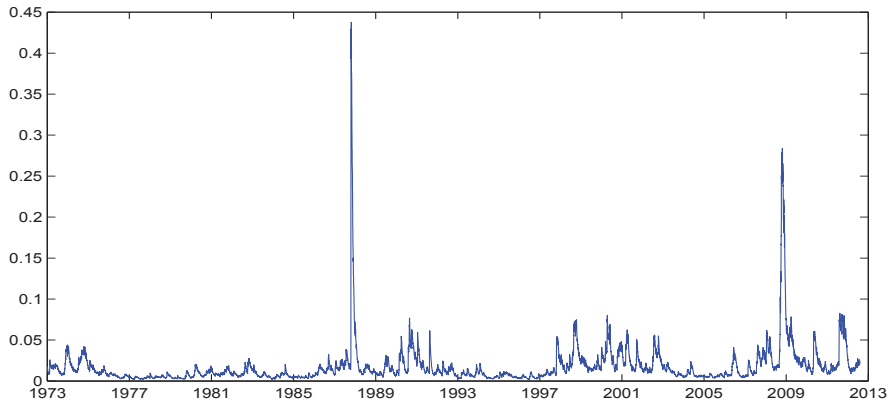

Figure 4.3: Estimated trajectory of the market activity M

Step 2: Power q

We recall our results for the normalized and scaled log-returns for the 26 currency denominations of the MCI, the logarithm of their histogram was displayed in Fig. 4.5. Maximum-likelihood Student-t fits are provided in Table 4.1 for daily, weekly and fortnightly log-returns. They estimate about four degrees of freedom. We emphasized that one needs a very large time window to estimate with high significance the degrees of freedom. One can see later in Fig. 4.23 and Table 4.3, using simulated data under the proposed model with $\delta = 4$ that for 40 years of daily observations it is only possible to estimate the degrees of freedom of log-returns with an error where the estimate could easily yield one degree more or less than four. For instance, the less than 40 years of daily data for the US denomination of the MCI are still not sufficient for the task of estimating the degrees of freedom more precisely than expecting it to be between 3 and 5. Fortunately, it will turn out in our step by step estimation procedure that there is not much influence of the choice by the degrees of freedom $\delta = 2(q+1)$ on the estimated values of the other parameters when fitting the model. Consequently, when following the economic arguments in Section 4.1, which conjectured asymptotically a time transformed square root process for the dynamics of the normalized index, there is a reason for setting $q = 1$, which provides the square root in the diffusion coefficient of the normalized index. Therefore, according to the theoretically derived aggregate wealth dynamics we specify the proposed model by fixing the power q to one. Thus, the dimension of the square root process amounts to $\delta = 4$ because of (4.24).

Step 3: Observing Market Activity

In Ané & Geman (2000) it has been shown that when subordinating intraday log-returns on observed trading activity, one obtains Gaussian distributed conditional returns. This is what our proposed model would predict for intraday log-returns. Intraday data were not available to us for sufficiently long time periods that we could use to fit our model. Instead we construct an observable proxy for market activity in the following way: By (4.3), (4.7) and an application of the Itô formula, one obtains as time derivative of the quadratic variation for $\sqrt{Y_{\tau_t}}$ the expression

$$\frac{d[\sqrt{Y}]_{\tau_t}}{dt} = \frac{1}{4}\frac{d\tau_t}{dt} = \frac{M_t}{4}, \tag{4.29}$$

which is proportional to market activity. The observed market activity time $\tau_t = \int_0^t M_s ds$ is shown in Fig. 4.6. It is a well observable process which tells us with its slope when market activity is high or low. In order to obtain an estimated proxy for the trajectory of the market activity process $M = \{M_t, t \geq 0\}$, appearing in (4.29), we perform some exponential smoothing of the empirical derivative of the quadratic variation $[\sqrt{Y}]_t$. Of course, other smoothing methods could potentially be used. However, we realized that most smoothing methods yielded very similar outcomes. We found the standard exponential smoothing method to be sufficient and rather robust with respect to the choice of the weight parameter $\alpha > 0$, which will be specified below.

The estimation of the trajectory of the market activity process M is performed using daily observations. First, the "raw" time derivative $Q_t = \frac{d[\sqrt{Y}]_{\tau_t}}{dt}$ at the ith observation time $t = t_i$ is estimated from the finite difference

$$\hat{Q}_{t_i} = \frac{[\sqrt{Y}]_{\tau_{t_{i+1}}} - [\sqrt{Y}]_{\tau_{t_i}}}{t_{i+1} - t_i} \tag{4.30}$$

for $i \in \{0, 1, \dots\}$. Second, exponential smoothing is applied to the observed finite differences according to the recursive standard moving average formula

$$\tilde{Q}_{t_{i+1}} = \alpha\sqrt{t_{i+1} - t_i}\hat{Q}_{t_i} + (1 - \alpha\sqrt{t_{i+1} - t_i})\tilde{Q}_{t_i}, \tag{4.31}$$

$i \in \{0, 1, \dots\}$, with weight parameter $\alpha > 0$. It is clear that the above smoothing depends on the observation frequency and the weight parameter with which new values enter the moving average calculation. We found that a smoothing parameter

of about $\alpha \approx 0.92$ delivered a robust estimate for the trajectory of the market activity process. This parameter and its neighboring values provided for daily but also for the two-day observation frequency a very similar trajectory that we use here as proxy for the discretely observed market activity process M. Fig. 4.3 displays the resulting trajectory of M_t for daily observations, when interpreting this value as estimate of $4\frac{d}{dt}[\sqrt{Y}]_{\tau_t}$, for $t \geq 0$. Here an initial value of $M_0 \approx 0.0175$ emerged and the time average of the trajectory of $(M_t)^{-1}$ amounted to 113.92.

The estimated trajectory of the market activity process, shown in Fig. 4.3, appears to be that of a rather "fast" moving process when compared with the trajectory of the square root process Y, representing the normalized MCI shown in Fig. 4.2. This means, we are dealing in our model with two different time scales. These are the t-time and the τ-time. The "slow" moving square root process Y moves in τ-time and is modeled according to the affine nature of aggregate wealth dynamics. The "fast" moving market activity process M evolves in t-time and models the reactions of market participants to changes in the level of the normalized index, generating some exaggerations in volatility. Changes in market activity are triggered by observed random ups and downs of the normalized index. One can see this when comparing Fig. 4.2 and Fig. 4.3. Market activity, when compared to the squared volatility $\frac{1}{Y_{\tau_t}}$ in τ-time, appears to move mostly together up and down. Therefore, it appears sufficient that we assumed that, all fluctuations in the model are driven by the nondiversifiable uncertainty of the market. It seems to be difficult to find evidence that the driving uncertainty for the normalized index and the market activity have to be different.

Step 4: Parameter γ

Now, we would like to identify the parameter γ in the diffusion coefficient in the SDE (4.8) for the inverse $\frac{1}{M_t}$ of the market activity process. Fig. 4.4 plots the quadratic variation of the square root of the estimated process $\frac{1}{M}$. Given that $\frac{1}{M}$ is assumed to be a square root process, this quadratic variation should be ideally a straight line. In a first approximation the graph in Fig. 4.4 confirms this. However, as can be seen from (4.4) and will be shown in Section 4.6 with simulated data, such discretely formed proxy for the quadratic variation, derived from the estimated proxy of $\frac{1}{M}$, is not a perfect straight line, see Fig. 4.19. The deviations result from the procedure

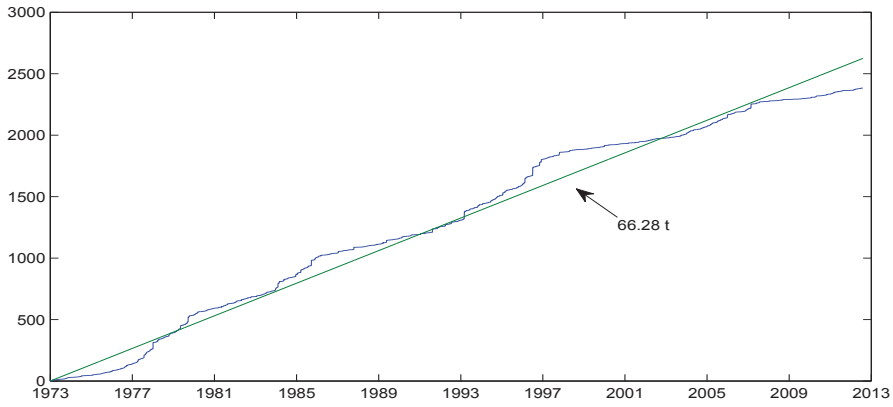

Figure 4.4: Quadratic variation of the square root of the estimated trajectory of $\frac{1}{M}$ with linear fit

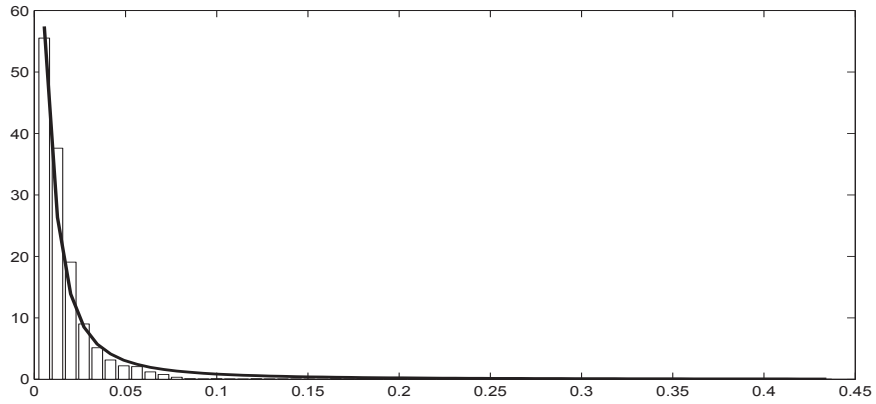

Figure 4.5: Histogram of market activity M with inverse gamma fit

of exponential smoothing when estimating the proxy for the market activity process M. We assume now that a straight line can be fitted (by linear regression) to such observed quadratic variation and the parameter γ can be estimated from the slope of this line. Our estimate for the slope equals here 66.28. Since under the proposed model we have $\frac{d}{dt}\left[\sqrt{\frac{1}{M}}\right]_t = \frac{1}{4}\gamma$, we obtain $\gamma \approx 265.12$.

Step 5: Parameters ν and ϵ

Fig. 4.5 plots the histogram for the trajectory of the proxy of the market activity M using daily observations. This fit appears to be realistic. The maximum likelihood estimation of the stationary inverse gamma density of the market activity process yields $\nu \approx 3.75$ degrees of freedom and mean $\frac{4\epsilon}{\gamma(\nu-2)} \approx 0.0175$; see (4.18). Since $\gamma \approx 265.12$ this yields $\epsilon \approx 2.18$.

The stylized fact (iii) requests for long term and short term log-returns of the discounted index a Student-t distribution with about the same degrees of freedom. Therefore, we simplify the model further. For the stylized version of the model we assume that it generates short term log-returns with the same degrees of freedom as long term log-returns. This means, we have set in the stylized version of the proposed model $\nu = \delta = 4$. This reduces the number of parameters to six that have to be estimated for the stylized version of the model.

Step 6: Long Term Average Net Growth Rate

Since we have now estimated the parameters γ, ν and ϵ, and know that $\frac{d\tau_t}{dt} \approx \frac{4\epsilon}{\gamma(\nu-2)} \approx 0.0175$, it follows from (4.29) and the slope $0.048 \approx a\frac{4\epsilon}{\gamma(\nu-2)}$ of the trend line in Fig. 4.1 that for the long term average net growth rate we have an estimate of about $a \approx 2.55$. This estimated value of a satisfies clearly the inequality (4.25), where its left hand side equals $\frac{1}{2}\frac{\Gamma(3)}{\Gamma(2)} = 1$. Most importantly we observe that when the MCI is interpreted as the numéraire portfolio (NP) under the benchmark approach, the benchmarked savings account is by (4.25) an $(\underline{\mathcal{A}}, P)$-supermartingale with negative drift under the proposed model. It seems to be far from being a local martingale, which by (4.25) would have been the case for a parameter value a near the level one. The fact that the estimate of the growth rate a is significantly greater than the level one, creates a challenge for classical pricing and hedging under the fitted model. However, one can employ the generalized version of the benchmark approach with the real world pricing formula, which has been outlined in Platen (2011).

We have now determined all parameters needed for the stylized version of the model. The resulting six parameter estimates are: $A \approx 65.21$, $Y_0 \approx 1.53$, $M_0 \approx 0.0175$, $a \approx 2.55$, $\epsilon \approx 2.18$ and $\gamma \approx 265.12$. Note that we have reduced the originally eight parameters of the proposed model to six parameters, where we fixed $q = 1.0$, $\delta = 4.0$ and $\nu = 4.0$. Since the three parameters A, Y_0 and M_0 of the model are initial parameters, the model has become rather parsimonious, with only three structural parameters remaining. All three structural parameters have a clear economic meaning. The resulting model is consistent with the theoretically derived affine nature of aggregate wealth dynamics. Note that he SDE (4.3) for the normalized index in τ-time has no parameter in the drift or diffusion coefficient that has to be estimated. This means, all three parameters in the SDE (4.2) for the normalized

aggregate wealth became identified as one. Therefore, we found that normalized aggregate wealth in market activity time follows most likely a very particular square root process, which is, fortunately, very tractable.

In case one restricts the dynamics of the benchmarked savings account to a process, that is a local martingale, the case that is assumed in Platen & Heath (2010), one is forced to set $a = 1$, and has only the two structural parameters ϵ and γ remaining. In this case the stylized model becomes a five parameter model with two structural parameters and three initial parameters.

One can apply the fitted stylized model in many ways. As one possible application let us visualize for the stylized model the trajectories of the resulting market activity time and volatility, respectively. In Fig. 4.6 we show the market activity time, the τ-time, as it emerges from our estimation. One notes periods of high and low market activity generating steeper and flatter slopes, respectively, of the τ-time. The estimated τ-time is obtained directly from the observed quadratic variation of the square root of the normalized index Y when multiplied by four. One can use the modeled state variables Y_{τ_t} and $\frac{1}{M_t}$ for pricing and hedging of derivatives and for measuring and managing risk under the benchmark approach, see Platen & Heath (2010) and Platen & Bruti-Liberati (2010).

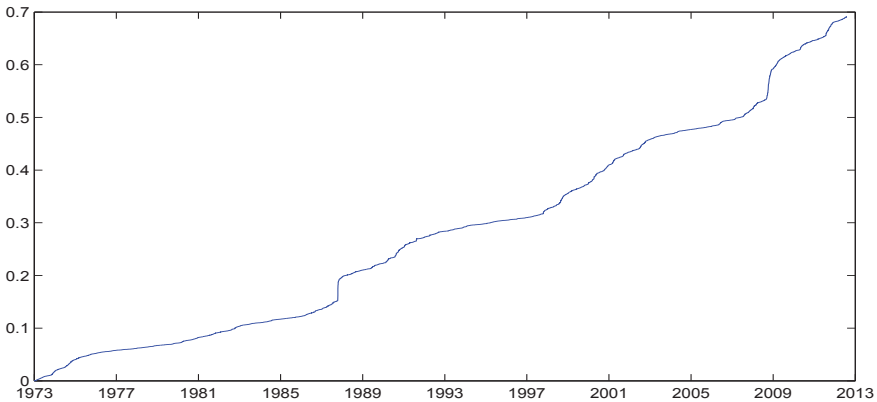

Figure 4.6: Market activity time

With the proposed model one can measure many quantities that are of interest for valuation and risk management. For instance, in recent years volatility derivatives became an important asset class. Their underlying is a volatility index of a diversified index.

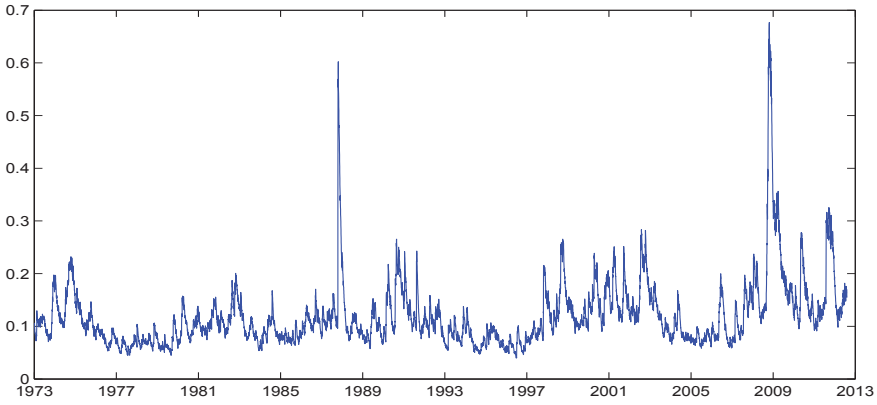

Figure 4.7: Calculated volatility of the discounted MCI for daily observations

For some indices the respective volatility index is available and expresses the market's view on where the (hidden) volatility may be. By formulas (4.14) and (4.29)-(4.31) the volatility of the discounted MCI equals approximately

$$\sigma_t \approx 2\sqrt{\frac{\tilde{Q}_t}{Y_{\tau_t}}}. \tag{4.32}$$

Fig. 4.7 plots the calculated volatility in t-time of the discounted MCI, obtained via formula (4.32). The volatility displayed in Fig. 4.7 is obtained from daily observations, for the period from 01/01/1973 until 10/08/2012. The sample mean of the calculated volatility in t-time equals 0.119 for the given period, which yields an estimated average volatility of about 11.9% for the MCI. This is also what one estimates from the observed log-returns of the MCI.

The proposed stylized model has been fitted by the same procedure to various diversified equity indices, including the EWI114_{40} constructed in Section 3.4, the EWI104s analyzed in Section 2.3-2.6 and the MSCI world index. In each case we obtained robust results, similar to those obtained above for the MCI. By assuming that the discounted index evolution follows also in future similar dynamics with the estimated constant parameters, one obtains a probabilistic description for the future dynamics of the discounted index. Additionally, one can calibrate the model to observed index derivative prices and obtains in this case a reflection of the market's view on the future evolution of the index and its volatility.

To demonstrate along these lines that the proposed stylized model and fitting procedure applies well, even to some regional index and its options, we report in

the next section results for the S&P500, where we compare the calculated volatility with the VIX, the volatility index of the S&P500.

4.5 Modeling the S&P500 and its Volatility Index VIX

Let us now study the case of an important regional market capitalization weighted equity index, the discounted S&P500. We consider the same time period from 01/01/1973 until 10/08/2012.

Fig. 4.8 plots the logarithm of the discounted S&P500 together with a linearly regressed trendline in calendar t-time, yielding the straight line $3.95 + 0.0378t$. This straight line provides the estimate for the parameter A with a value of about $A \approx 52.09$ and the estimated net growth rate in t-time at about a level of $\frac{4\epsilon a}{\gamma(\nu-2)} \approx 0.0378$, see (4.28). The normalized S&P500 is then obtained by dividing the discounted S&P500 by the exponential function $A_{\tau_t} \approx 52.09 \exp\{0.0378t\}$, for $t \geq 0$. Using the proposed stylized model, the resulting normalized S&P500, denoted by Y_{τ_t}, is plotted in Fig. 4.9 with respect to calendar t-time, where we set $Y_0 = 2.27$.

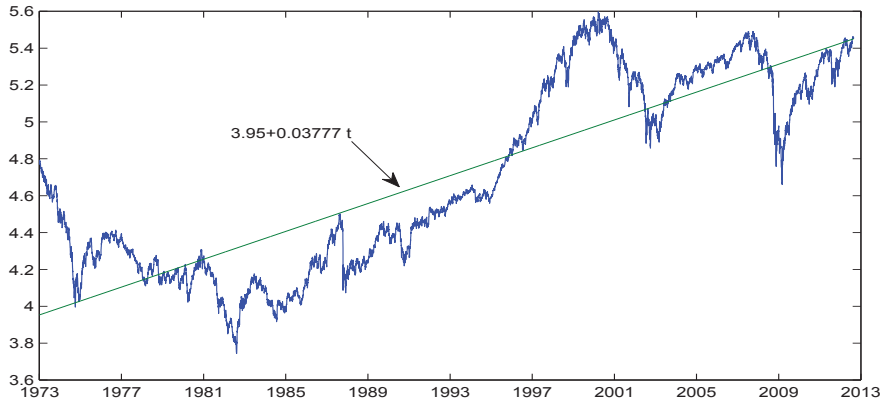

Figure 4.8: Logarithm of the discounted S&P500 and linear fit

Fig. 4.10 plots the estimated market activity process M obtained from the normalized discounted S&P500 with daily observations shown for the period from 1973 until 2012. The process M was obtained by the same steps and using the same weight parameter $\alpha = 0.92$ for exponential smoothing, as in the case of the MCI. From the trajectory of the estimated market activity process M we obtain $\epsilon \approx 2.15$, $\gamma \approx 172.3$ and consequently $a \approx 1.5$. Note that in this case the Assumption 2 is

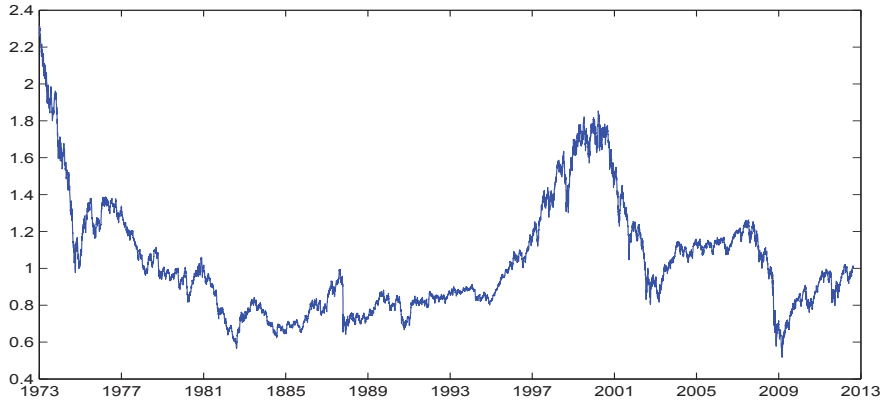

Figure 4.9: Normalized S&P500

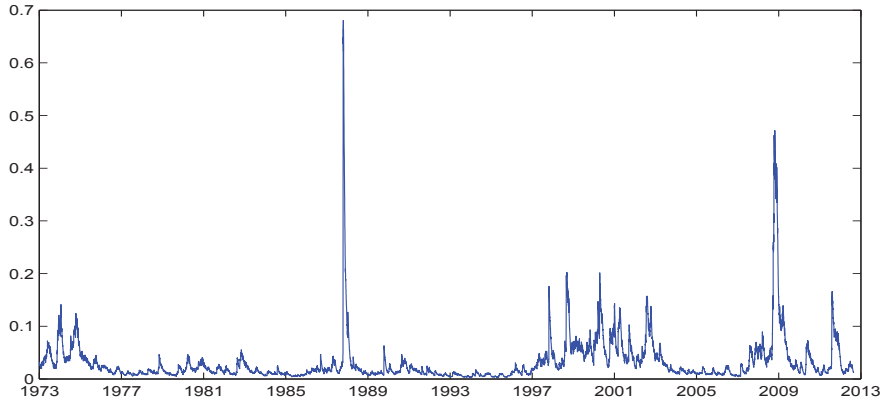

Figure 4.10: Estimated market activity process M for the S&P500 from daily data

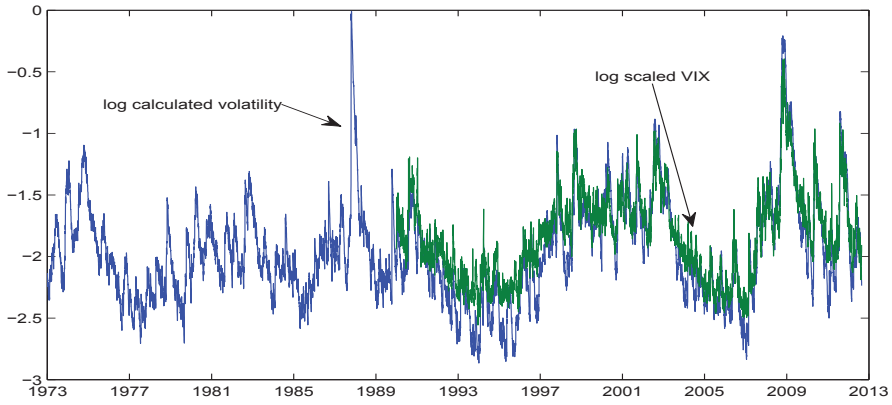

Figure 4.11: Logarithms of scaled VIX and calculated volatility of the discounted S&P500

approximately satisfied, and one may use in relation (4.25) the equality sign. This means, the benchmark approach, as described in Platen & Heath (2010), is poten-

tially applicable for the S&P500 as proxy of the NP of the US market, and the stylized model employs then only five parameters.

The volatility index VIX is provided by the Chicago Board Options Exchange. It is the volatility index for the S&P500. The literature points out that the VIX is a biased estimator of the volatility implied from index options, see e.g. Fleming, Ostdiek & Whaley (1995). Hence, it is common to match the short term at the money implied volatility by scaling the VIX. Typically, the VIX, which stems from daily observations, is known to multiply the implied volatility by a factor of about 1.2; see e.g. Blair, Poon & Taylor (2001). That is, when denoting by VIX_t the value of the quoted VIX at time t, then one obtains the scaled VIX value V_t via the formula

$$V_t = \frac{\text{VIX}_t}{100}\frac{1}{1.2}, \tag{4.33}$$

for $t \geq 0$, which provides approximately the at the money implied volatility of short term S&P500 options.

Fig. 4.11 plots the calculated volatility in calendar t-time of the discounted S&P500, calculated according to formula (4.32), and obtained from daily observations, for the period from 02/01/1990 until 10/08/2012. The sample mean of the calculated volatility in t-time equals 0.1625 for the given period. For comparison, the daily data for the logarithm of the scaled VIX, according to (4.33), are also shown in Fig. 4.11. The sample mean of the scaled VIX is with 0.1707 close to that of the calculated volatility. The logarithm of the scaled VIX behaves in Fig. 4.11 very similarly to the logarithm of the calculated volatility of the discounted S&P500, obtained under the proposed stylized model. To generate the volatility, the proposed model combines in formula (4.14) the normalized discounted S&P500, modeled by Y_{τ_t} and displayed in Fig. 4.9 with the market activity process M, displayed in Fig. 4.10. As one can see, via formula (4.14) the proposed model recovers visually with good accuracy the trajectory of the scaled VIX from the trajectories of the two observable processes Y_{τ_t} and M_t. We emphasize that the proposed stylized model is parsimonious by construction and uses only two constant structural parameters for its characterization. These parameters seem to remain the same also over long periods of time. Furthermore, the model employs only one Brownian motion, which makes it a one-factor model with two state variables.

Since calculated volatility recovers the VIX accurately it can be expected to recover also well short dated, close to at the money European option prices on the S&P500 index.

4.6 Simulation Study

This section explains the steps we propose for the simulation of the stylized version of the model. Both square root processes $\frac{1}{M}$ and Y are simulated jointly by sampling from their non-central chi-square transition density with the same sources of randomness, see also Platen & Rendek (2009) or Platen & Bruti-Liberati (2010). We will see that this simulation is performed without any error for $\frac{1}{M}$ and almost exactly for Y. The second part of this section checks for the seven stylized empirical facts, listed in Section 4.2.

The following illustration of the simulation of the stylized version of the model uses the six parameter values estimated in Section 4.4 for the MCI. The simulation is performed using the following steps:

1. Simulation of the Process $\frac{1}{M}$

The inverse $\frac{1}{M}$ of the market activity process we have to simulate first. It is described by the SDE (4.8) and is a square root process of dimension $\nu = 4$.

This process can be sampled exactly due to its non-central chi-square transition density of dimension $\nu = 4$. That is, we have

$$\frac{1}{M_{t_{i+1}}} = \frac{\gamma(1 - e^{-\epsilon(t_{i+1}-t_i)})}{4\epsilon} \left(\chi^2_{3,i} + \left(\sqrt{\frac{4\epsilon e^{-\epsilon(t_{i+1}-t_i)}}{\gamma(1 - e^{-\epsilon(t_{i+1}-t_i)})}} \frac{1}{M_{t_i}} + Z_i \right)^2 \right), \quad (4.34)$$

for $t_i = \Delta i, i \in \{0, 1, \dots\}$; see also Broadie & Kaya (2006). Here Z_i is an independent standard Gaussian distributed random variable and $\chi^2_{3,i}$ is an independent chi-square distributed random variable with three degrees of freedom. Then the right hand side of (4.34) becomes a non-central chi-square distributed random variable with the requested non-centrality and four degrees of freedom.

A simulated path of M_{t_i}, according to (4.34), is displayed in Fig. 4.12 for a period of 40 years using the previously estimated parameters for the MCI. It resembles the type of trajectory shown in Fig. 4.3 with more pronounced spikes than those observed in Fig. 4.3 that were caused by the 1987 crash and the GFC in 2007/2008.

We will see later that the estimated market activity obtained from the simulated index resembles strongly the estimated trajectory of the market activity of the MCI, displayed in Fig. 4.3, which showed less peaked spikes.

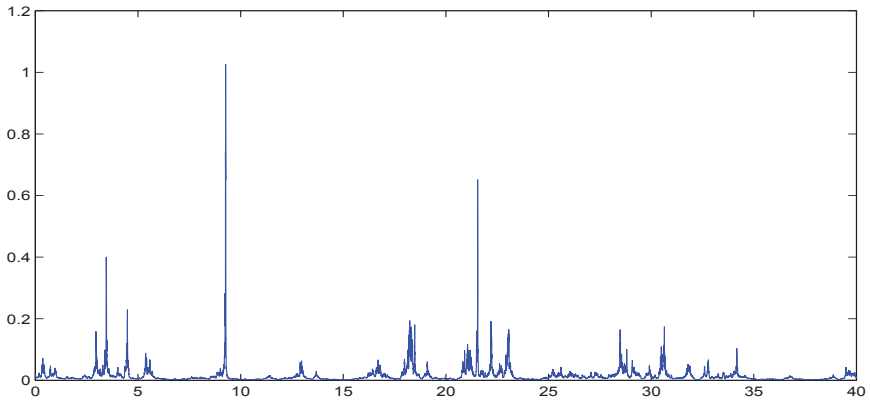

Figure 4.12: Simulated path of M

2. Calculation of τ-Time

The next step of the simulation generates the market activity time, the τ-time. By (4.7) one aims for the increment

$$\tau_{t_{i+1}} - \tau_{t_i} = \int_{t_i}^{t_{i+1}} M_s ds, \tag{4.35}$$

$i \in \{0, 1, \dots\}$. To avoid any anticipation of future uncertainty the integral on the right hand side of equation (4.35) is numerically approximated by the product

$$M_{t_i}(t_{i+1} - t_i). \tag{4.36}$$

Fig. 4.13 displays the resulting τ-time, the market activity time, which resembles well that shown in Fig. 4.6.

3. Calculation of the Y Process

The simulation of the Y process is very similar to the simulation of the square root process $\frac{1}{M}$. Both processes are square root processes of dimension four and both are driven by the same source of uncertainty. We therefore employ in each time step the same Gaussian random variable Z_i and the same chi-square distributed random variable $\chi^2_{3,i}$, as in (4.34), to obtain the new value of the Y process,

$$Y_{\tau_{t_{i+1}}} = \frac{1 - e^{-(\tau_{t_{i+1}} - \tau_{t_i})}}{4} \left(\chi^2_{3,i} + \left(\sqrt{\frac{4e^{-(\tau_{t_{i+1}} - \tau_{t_i})}}{1 - e^{-(\tau_{t_{i+1}} - \tau_{t_i})}} Y_{\tau_{t_i}}} + Z_i \right)^2 \right), \tag{4.37}$$

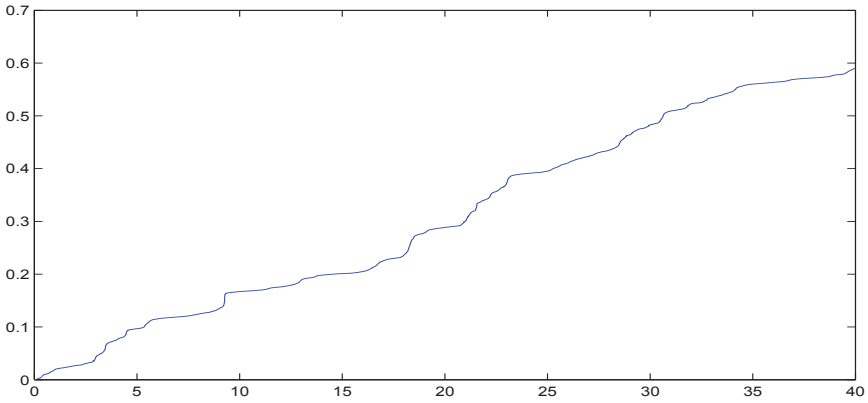

Figure 4.13: Simulated τ-time, the market activity time

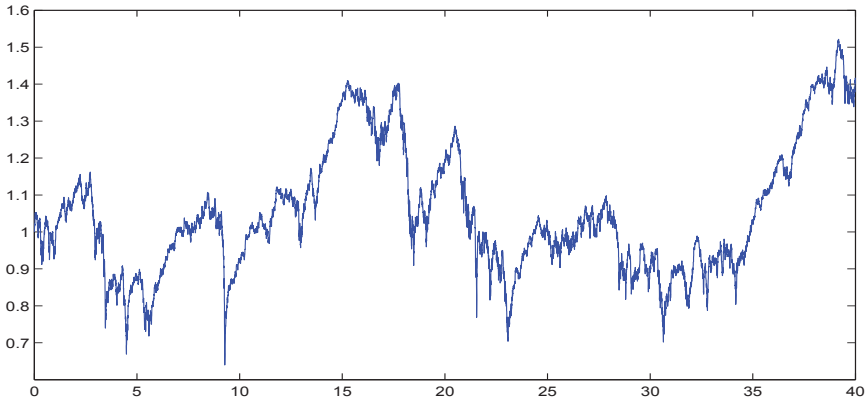

Figure 4.14: Simulated trajectory of the normalized index Y_{τ_t}

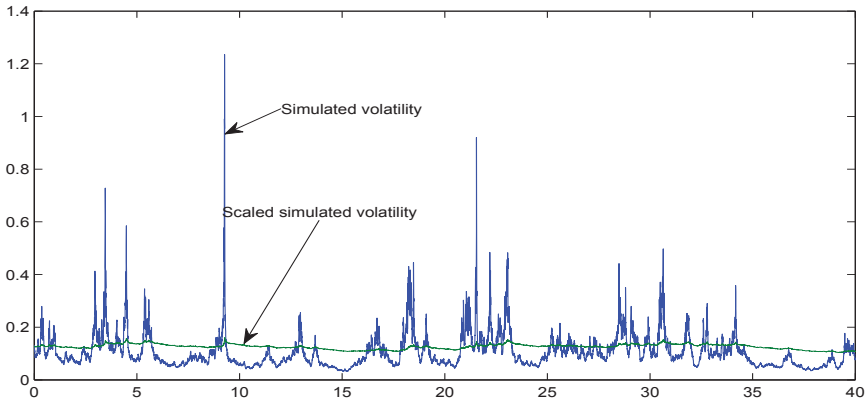

Figure 4.15: Simulated volatility of the index and simulated scaled volatility

for $t_i = \Delta i, i \in \{0, 1, \dots\}$. Note that the difference $\tau_{t_{i+1}} - \tau_{t_i}$ was approximated using in (4.36) the market activity of the previous step. For small step size the approximation of (4.35) by (4.36) is rather accurate and we say that Y_{τ_t} is here

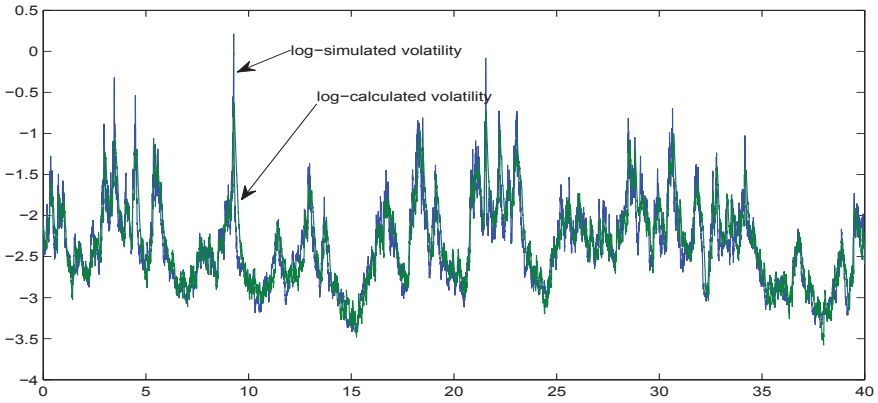

Figure 4.16: Logarithm of the simulated and calculated volatility

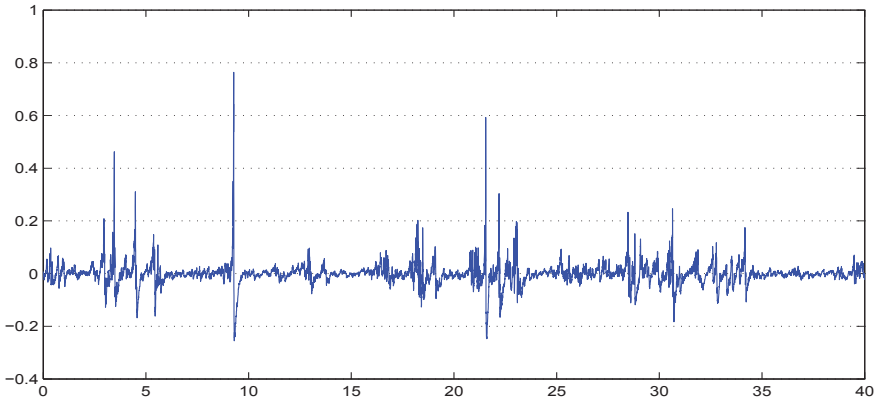

Figure 4.17: Differences of the logarithms of the simulated and calculated volatility

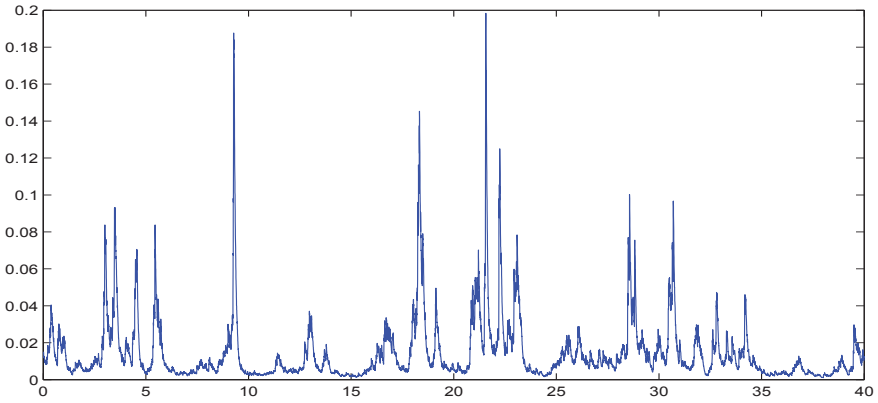

Figure 4.18: Estimated market activity of the simulated index

almost exactly simulated. This is significantly better in accuracy over long time periods than what discrete time approximations, in the sense of Kloeden & Platen (1999), can deliver.

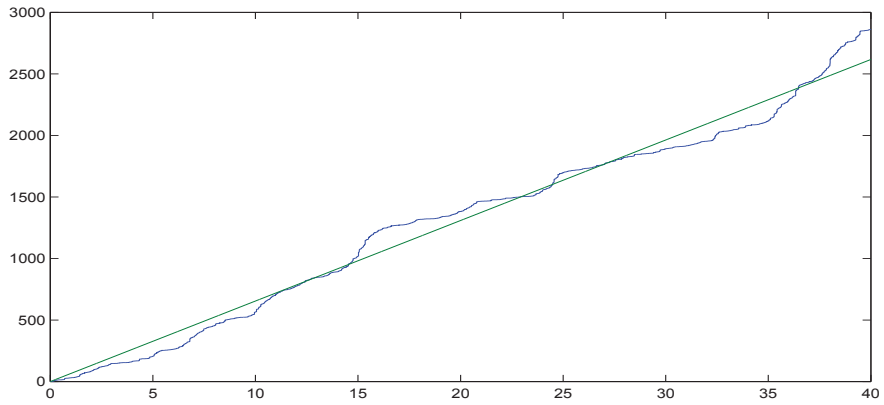

Figure 4.19: Quadratic variation of the square root of the inverse of estimated market activity

For simplicity, we choose in the simulation the initial value $Y_0 = 1$. The resulting trajectory of the normalized index process Y is exhibited in Fig. 4.14, which shows strong similarity with Fig. 4.2. Interesting is that one observes from time to time major sharp drawdowns in the simulated path. The accurate modeling of such drawdowns is crucial for risk measurement.

4. Calculating the Volatility Process

One obtains the trajectory of the simulated volatility σ_t of the index by formula (4.14). The resulting simulated trajectory of the volatility is displayed in Fig. 4.15. For comparison, the same figure shows also the simulated "scaled volatility" given by the formula $\sqrt{\frac{2\epsilon}{\gamma Y_{\tau_t}}}$, which is the volatility resulting purely from the feedback mechanism of the market with respect to "average" τ-time. One notes that the simulated volatility deviates significantly from the "scaled simulated volatility". One notes in the simulated volatility the "exaggerated" values above and below the "scaled simulated volatility". This deviation is modeling the way how market activity is evolving. It models overreactions (underreactions) in response to each downward (upward) move of the index.

For comparison, Fig. 4.16 shows the logarithm of the simulated volatility displayed in Fig. 4.15 together with the logarithm of the calculated volatility according to formula (4.32) by using the estimated market activity, which is first calculating $\tilde{Q}$ from (4.30) and (4.31). These trajectories are visually similar but not identical. The calculated volatility is smoother and lagging behind since it results from the

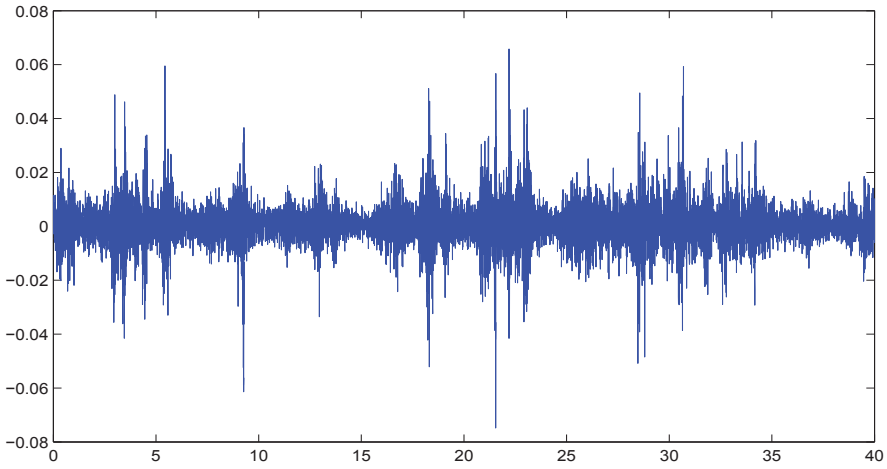

Figure 4.20: Log-returns of the simulated index

smoothing we use to obtain the estimated market activity process. Additionally, in Fig. 4.17 we show the differences between the logarithms of the simulated and calculated volatility.

Fig. 4.18 shows the estimated market activity from the simulated index. We see that the extreme spikes of the simulated M in Fig. 4.12 are substantially smoothed in the estimated market activity shown in Fig. 4.18. Fig. 4.18 resembles well the estimated market activity of the MCI in Fig. 4.3. Moreover, we calculate the quadratic variation of the square root of the inverse of the estimated market activity from the simulated index and plot it in Fig. 4.19. Note the strong similarity of this figure to Fig. 4.4. It is important to take the effects of the smoothing into account when interpreting the dynamics of the index and these of the observed quantities.

The above simulation has shown that the proposed stylized model generates trajectories with visually very similar properties as observed from the MCI. Below we will confirm that all listed stylized empirical facts, which are typical for most diversified stock indices, are captured by the stylized model.

For the proposed stylized model we will now check the seven stylized empirical facts, listed in Section 4.2. We will see whether its properties would allow one to falsify the model. We employ the same standard statistical and econometric techniques as used in Section 4.2. The daily observed index is simulated, as described above, using the estimated parameters for the MCI. Fig. 4.20 displays the daily log-returns of the simulated index, these resemble visually well those of Fig. 4.2.

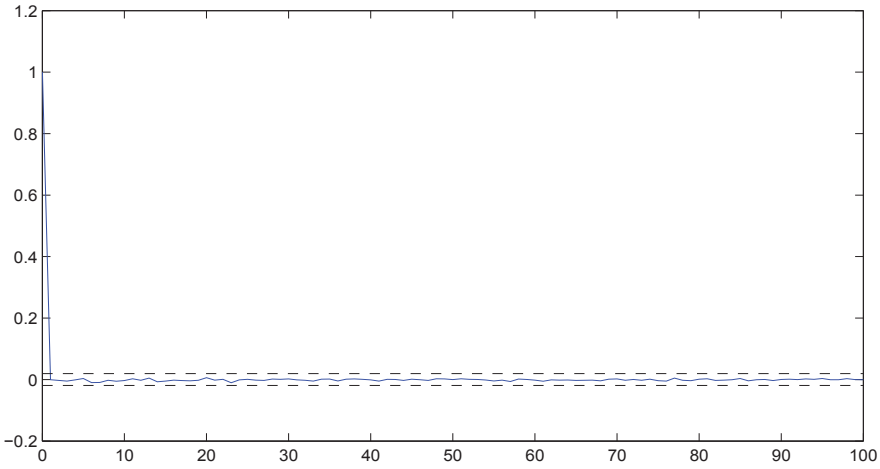

Figure 4.21: Average autocorrelation function for log-returns of the simulated index

(i) Uncorrelated Returns

Daily observed simulations of the equity index under the proposed model provide typical graphs of the Y process similar to those shown in Fig. 4.14. As one should expect, due to the construction of the model there is practically no correlation detectable between the log-returns over time. Fig. 4.21 displays the average over 26 estimated autocorrelation functions for log-returns of the simulated index, which rapidly falls to the level zero with more than 95% significance, similarly as observed in Fig. 4.3.

(ii) Correlated Absolute Returns

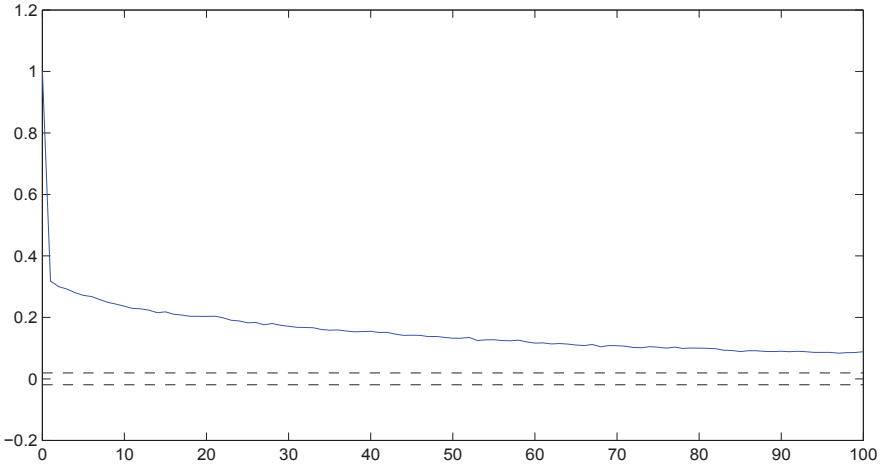

Figure 4.22: Average autocorrelation function for the absolute log-returns of the simulated index

In a similar manner also the average autocorrelation for the absolute log-returns of the simulated index has been estimated. In Fig. 4.22 the average of 26 estimated autocorrelation functions of absolute log-returns is displayed. Similarly, as in Fig. 4.4, one observes that this average of sample autocorrelations does not die away fast. Moreover, as in Fig. 4.4, the decay of the autocorrelation does not seem to be exponential. Even for large lags of more than one month there is still some significant autocorrelation present.

(iii) Student-t Distributed Returns

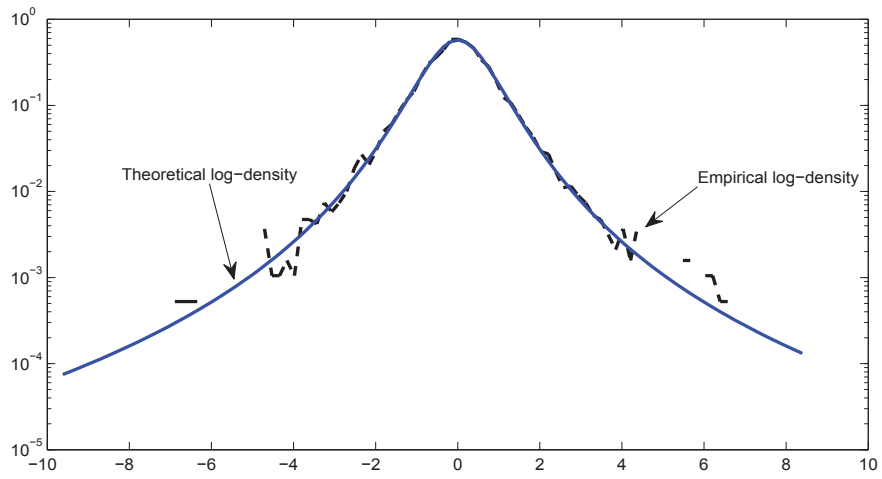

Figure 4.23: Logarithm of empirical density of normalized log-returns of the simulated index and Student-t density with 3.2 degrees of freedom

By the same maximum likelihood estimation method as employed earlier we analyzed 40 years of daily log-returns of the simulated index. As expected from the design of the model these appear to be distributed according to a Student-t distribution with approximately four degrees of freedom. Fig. 4.23 displays the logarithm of the empirical density of simulated normalized index returns displayed in Fig. 4.20. This figure shows also the logarithm of the density of the Student-t distribution, which in this case was estimated with 3.2 degrees of freedom. The fit to the simulated data seems visually similarly good as in Fig. 4.5. Note that even though the simulation was performed with $\delta = 4$ the estimated degrees of freedom of the Student-t density fitted to the simulated data can easily be almost one degree different as in Fig. 4.23.

Additionally, Table 4.3 provides, similarly to Table 4.2, the test statistics for

Table 4.3: Log-Maximum likelihood test statistic for different outcomes of the simulated normalized log-returns

Simulation	Student-t	NIG	Hyperbolic	VG	ν
1	0.008934	37.474149	102.719638	131.240780	4.012850
2	11.485226	11.175028	96.457136	132.916256	3.450916
3	0.000000	100.928524	244.190151	294.719960	3.734148
4	9.002421	35.759464	347.060676	331.014904	2.579009
5	8.767003	11.551178	121.190482	144.084964	3.170449
6	0.401429	60.570898	205.788160	252.591737	3.432435
7	12.239056	4.354888	46.411554	78.273485	3.957696
8	1.693411	23.910523	94.408789	130.623174	3.849691
9	1.232454	47.830407	202.073144	237.168411	3.236322
10	0.000000	43.037206	128.807757	162.582353	3.774957
11	0.433645	47.782681	172.736397	208.847632	3.431803
12	0.000000	56.019354	146.077121	185.624888	3.899403
13	7.137154	48.219756	579.922931	477.383441	2.293363
14	5.873948	16.515390	107.770531	135.508299	3.388307
15	0.000000	54.718046	184.112794	217.304105	3.402049
16	6.982560	3.991610	29.192198	47.105125	4.268740
17	2.966916	22.914863	108.513143	138.044416	3.553629
18	0.000000	52.066364	129.790856	160.373085	3.959605
19	0.006909	39.568695	111.398645	143.914350	3.982892
20	0.000001	56.845664	169.915512	211.260626	3.651091
21	1.674578	17.681088	61.710576	90.679738	4.265834
22	14.010840	3.279722	47.433693	73.789313	3.770825
23	11.198940	12.074044	114.888817	143.800553	3.257146
24	0.455557	27.676102	86.841704	114.452947	4.006528
25	6.658059	13.403411	90.314637	119.523602	3.502497
26	19.812830	1.360441	37.443467	59.435233	3.821602

the Student-t fit and several other symmetric generalized hyperbolic distributions for 26 cases of 40 years simulated log-returns. We note that our estimation provided in many simulations less than four degrees of freedom. This means that there is variation in the estimation of the degrees of freedom and the stylized model based on the theoretical understanding of aggregate wealth dynamics cannot be easily falsified on this ground.

(iv) Volatility Clustering

Obviously, under the proposed model volatility is stochastic. A moving average estimation of volatility from the simulated index log-returns results in the volatility shown in Fig. 4.24. One clearly notes that this observed volatility is indeed stochastic, and it shows clusters of higher values for certain random time periods of time. Note that the graph in Fig. 4.24 is not exactly the volatility that we can calculate under the model, because it is estimated as a standard moving average from the index log-returns. The logarithm of the via moving average estimated volatility is displayed in Fig. 4.26.

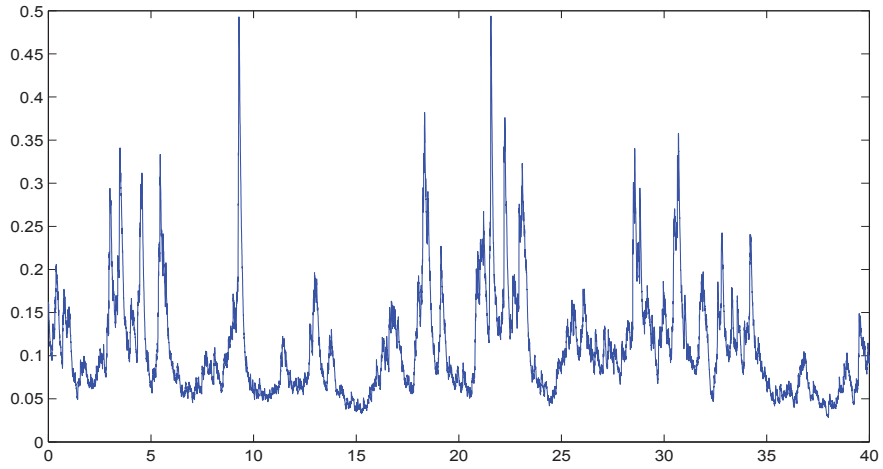

Figure 4.24: Estimated volatility of the simulated index

The proposed model agrees with the stylized empirical fact (iv), which expresses the growing consensus among practitioners and academic researchers that volatility should be modeled as a stochastic process that generates occasional outbursts of volatility clusters. Note that, in reality, one cannot fully observe the hidden exact volatility. Fig. 4.17 illustrates the potential difference between the simulated volatility and our calculated volatility. In Fig. 4.27 we show the differences between the logarithm of the simulated and the estimated volatility from squared index log-returns. When comparing Fig. 4.17 and Fig. 4.27 one notes that the calculated volatility, we propose in Section 4.4 is, in general, closer to the simulated volatility.

(v) Long Term Exponential Growth

In Fig. 4.25 the logarithm of the simulated index has been displayed with its fit to a straight line. Here the linear regression of the logarithm of the simulated index provided the estimates with $A = 64.07$, $\frac{2a\epsilon}{\gamma} = 0.05$.

The logarithm of the simulated normalized index, $\ln(Y)$, is displayed as upper graph in Fig. 4.26. As requested, Y has a stationary density by construction.

(vi) Leverage Effect

Fig. 4.26 is plotting the logarithm of the simulated normalized index together with the logarithm of its via standard moving average from squared log-returns estimated volatility. By construction, under the proposed model the simulated

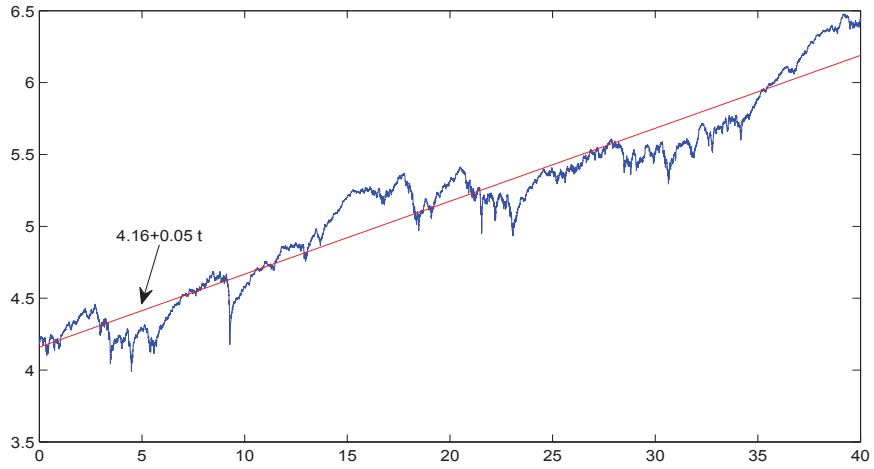

Figure 4.25: Logarithm of simulated index with linear fit

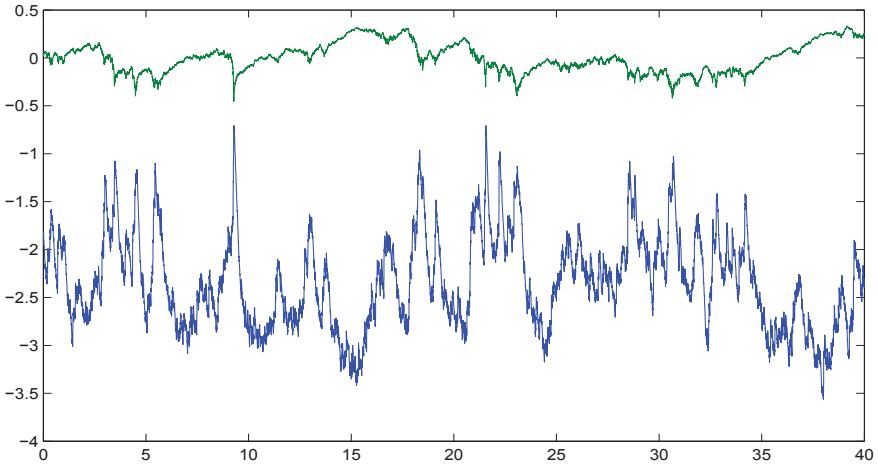

Figure 4.26: Logarithms of simulated normalized index and its estimated volatility

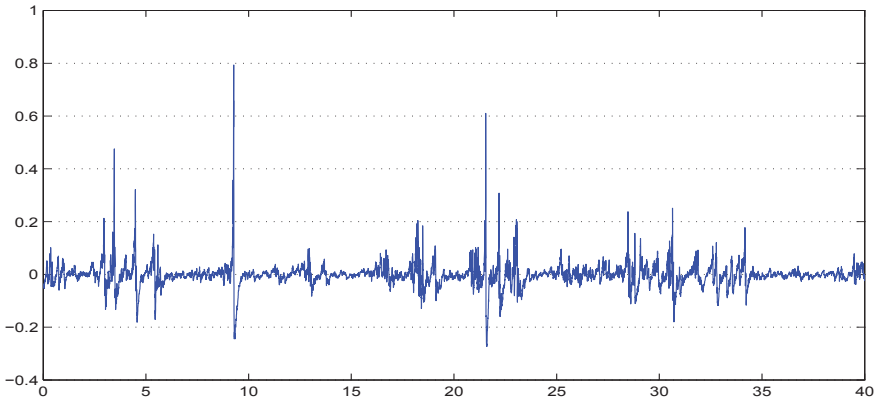

Figure 4.27: Differences of the logarithms of the simulated and estimated volatility

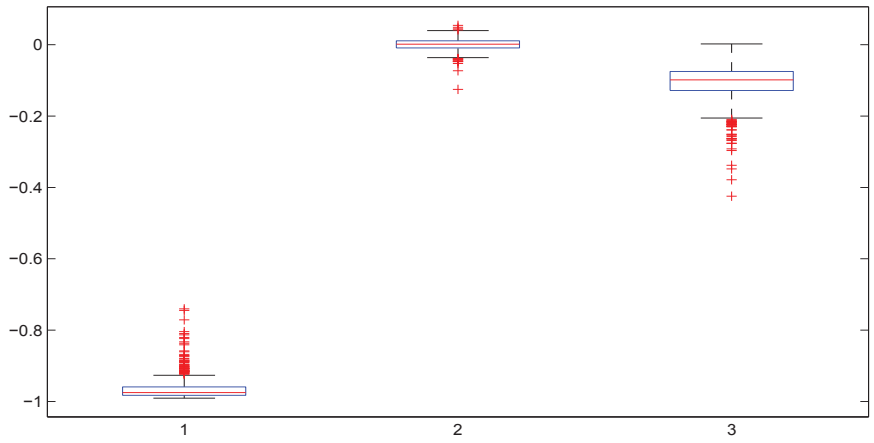

Figure 4.28: Boxplots of correlation coefficient obtained by methods 1-3

volatility, as given in formula (4.14), of the simulated index is perfectly negatively "dependent" on the normalized index. We pointed already out that our calculated volatility lags slightly behind the simulated volatility, as shown in Fig. 4.16.

Let us now perform the following study: We simulate 1000 trajectories of the index and its volatilities and calculate the correlation coefficient between the increments of the logarithm of the simulated index and the increments of the logarithms of the three types of volatilities. These volatilities are obtained by the following three methods:

1. The simulated volatility is obtained from the proposed model by first simulating the market activity process M and then the normalized index Y. The volatility is then obtained via formula (4.14).
2. The estimated volatility is obtained from exponential smoothing of squared index log-returns via a moving average.
3. The calculated volatility is obtained from formula (4.32) by first calculating $\tilde{Q}$ from (4.30) and (4.31).

Fig. 4.28 shows the boxplots corresponding to the correlation coefficient obtained for the volatilities calculated by the three methods. It is clear that the correlation coefficient depends clearly on the method used for the calculation of the volatility even that graphs appear to be visually very similar. We

observe on average $-0.9656(-0.9679, -0.9632)$, $0.000605(-0.000645, 0.0019)$ and $-0.1073(-0.1112, -0.1035)$ of correlation (with the 99% confidence interval shown in the brackets) between the increments of the logarithm of the simulated index and increments of the logarithm of the volatility obtained by the methods 1-3, respectively. This important effect has been pointed out in the work by Ait-Sahalia, Fan & Li (2012). We argue in this chapter that there is no need to introduce an extra independent Brownian motion as driver of stochastic volatility of a diversified index even if one does not seem to observe perfect negative correlation between log-returns and estimated volatility increments.

(vii) Extreme Volatility at Major Downward Moves

In Fig. 4.26 one observes that the logarithm of the volatility increases more than the logarithm of the simulated index moves down at major downward moves of the index. In periods of "normal" index fluctuations the magnitudes of their movements have been more in line. This means, in times of crisis the increase of the volatility is stronger under the proposed model than could be typically generated by, say, a local volatility function model. The proposed more complex interplay between market activity and index fluctuations reflects well the extreme volatilities at major downward moves observed in the market.

Conclusion

In summary, one can say that the proposed model captures well all seven stylized empirical facts, listed in Section 4.2, and cannot be easily falsified on these grounds. This list of empirical properties and the conjecture of affine aggregate wealth dynamics were sufficient to identify the proposed stylized model as a two-component diffusion model with six parameters. The thesis has shown that it is possible to identify a parsimonious model for diversified equity indices in currency denominations. Crucial has been the idea that by considering a benchmark, in form of a diversified index, one can disentangle the factors driving the index dynamics. The conjectured affine nature of the aggregate wealth dynamics could be empirically not falsified. Moreover, the market activity has been identified as an important quantity, which

models the human behavior that exaggerates the volatility movements by making extreme volatilities more extreme than the affine nature of wealth dynamics would suggest. It turns out that it is a single source of uncertainty, which is sufficient to model the nondiversifiable uncertainty of the market such that it drives the entire index dynamics. It has to be emphasized that a two-component model with constant parameters and one driving Brownian motion has been able to reflect well the dynamics of diversified stock indices in currency denomination over long periods of time. Important is the fact that this model lives outside the classical no-arbitrage paradigm and clearly signals the need for a more general financial market modeling framework in the direction of the benchmark approach.

There is significant potential to generalize the proposed model so that it can potentially match reality even better. Most obvious is the possible inclusion of jumps in the modeling of market activity, triggered by major drawdowns of the index. Furthermore, one could use a local volatility function in the process Y to give more flexibility to the implied volatility surface of index options. One could make some of the parameters time dependent or even regime switching stochastic. However, it is questionable whether the available data are sufficient to falsify the proposed stylized model with constant parameters when compared with such potential generalizations. Finally, one could apply the proposed methodology to the modeling of commodities, stocks and other securities by exploiting the fact that not only an exchange rate but also commodity and stock prices in currency denomination can be expressed as the ratio of two denominations of a diversified index.

Bibliography

Abramowitz, M. & I. A. Stegun (Eds.) (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

Ait-Sahalia, Y., J. Fan, & Y. Li (2012). The leverage effect puzzle: Disentangling sources of bias at high frequency. *Journal of Financial Economics*. To appear.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control* **19**(6), 716–723.

Alcock, J. T. & K. Burrage (2006). A note on the balanced method. *BIT Numerical Mathematics* **46**(4), 689–710.

Alfonsi, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11**(4), 355–384.

Alparslan, A., A. Tessitore, & N. Usmen (2012). Stable Paretian vs. Student's t stock market hypothesis. *Journal of Statistical Theory and Practice*. To appear.

Andersen, L. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Finance* **11**(3), 1–42.

Ané, T. & H. Geman (2000). Order flow, transaction clock and normality of asset returns. *J. Finance* **60**(5), 2259–2284.

Angel Canela, M. & E. Pedreira Collazo (2012). Modelling dependence in Latin American markets using copula functions. *Journal of Emerging Market Finance*. To appear.

Bajeux-Besnainou, I. & R. Portait (1997). The numeraire portfolio: A new perspective on financial theory. *The European Journal of Finance* **3**(4), 291–309.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London, Series*

A **353**(1674), 401–419.

Barndorff-Nielsen, O. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.* **5**, 151–157.

Barndorff-Nielsen, O. (1995). Normal-Inverse Gaussian processes and the modelling of stock returns. Technical report, University of Aarhus. 300.

Barndorff-Nielsen, O. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.* **24**, 1–13.

Barndorff-Nielsen, O. & N. Shephard (2001). Modelling by Lévy processes for financial econometrics. In O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick (Eds.), *Lévy Processes - Theory and Applications*, pp. 283–318. Birkhäuser, Boston.

Basel (1996). Amendment to the capital accord to incorporate market risks. *Basel Committee on Banking Supervision , Basel, Switzerland*.

Basle (1996). *Amendment to the Capital Accord to Incorporate Market Risks*. Basle Committee on Banking and Supervision, Basle, Switzerland.

Becherer, D. (2001). The numeraire portfolio for unbounded semimartingales. *Finance Stoch.* **5**, 327–341.

Berkaoui, A., M. Bossy, & A. Diop (2005). Euler schemes for SDEs with non-Lipschitz diffusion coefficient: Strong convergence. INRIA working paper no.5637.

Black, F. (1976). Studies in stock price volatility changes. In *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association*, pp. 177–181.

Black, F. & M. Scholes (1973). The pricing of options and corporate liabilities. *J. Political Economy* **81**, 637–654.

Blair, B. J., S. Poon, & S. Taylor (2001). Forcasting s&p 100 volatility: The incremental information content of implied volatilities and high frequency index returns. *Journal of Econometrics* **105**(1), 5–26.

Blattberg, R. C. & N. Gonedes (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *J. Business* **47**, 244–280.

Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. University of California Press, Berkeley, CA.

Bossy, M. & A. Diop (2004). An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha$, $\alpha \in [1/2, 1)$. INRIA working paper no. 5396.

Breiman, L. (1961). Optimal gambling systems for favorable games. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Volume I, pp. 65–78.

Breymann, W., A. Dias, & P. Embrechts (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3**, 1–14.

Broadie, M. & O. Kaya (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.* **54**, 217–231.

Browne, S. (1999). The risks and rewards of minimizing shortfall probability. *J. Portfolio Manag.* **25**(4), 76–85.

Bru, M.-F. (1991). Wishart processes. *J. Theoret. Probab.* **4**(4), 725–751.

Bruti-Liberati, N. & E. Platen (2008). Strong predictor-corrector Euler methods for stochastic differential equations. *Stochastics and Dynamics* **8**(3), 561–581.

Bühlmann, H. & E. Platen (2003). A discrete time benchmark approach for insurance and finance. *Astin Bulletin* **33**(2), 153–172.

Campbell, J. & L. Viceira (2002). *Strategic Asset Allocation, Portfolio Choice for Long Term Investors*. Oxford, Oxford University Press.

Carr, P., H. Geman, D. Madan, & M. Yor (2004). From local volatility to local Lévy models. *Quant. Finance.* **4**, 581–588.

Carr, P. & J. Sun (2007). A new approach for option pricing under stochastic volatility. *Rev. Deriv. Research* **10**, 87–150.

Clark, J. M. C. & R. J. Cameron (1980). The maximum rate of convergence of discrete approximations for stochastic differential equations. In B. Grigelionis (Ed.), *Stochastic Differential Systems*, Volume 25 of *Lecture Notes in Control and Inform. Sci.*, pp. 162–171. Springer.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–159.

Cont, R. (2010). *Encyclopedia of Quantitative Finance*. Wiley.

Cover, T. (1991). Universal portfolios. *Math. Finance* **1**, 1–29.

Cox, J. C. (1975). Notes on option pricing I: constant elasticity of variance diffusions. Stanford University, Working paper.

Cox, J. C., J. E. Ingersoll, & S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.

Cox, J. C. & S. A. Ross (1976). The valuation of options for alternative stochastic processes. *J. Financial Economics* **3**, 145–166.

Deelstra, G. & F. Delbaen (1998). Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis* **14**, 77–84.

Delbaen, F. & W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300**, 463–520.

Delbaen, F. & H. Shirakawa (1997). Squared Bessel processes and their applications to the square root interest rate model. Preprint. Department of Industrial Engineering and Management, Tokio Institute of Technology.

DeMiguel, V., L. Garlappi, & R. Uppal (2009). Optimal versus naive diversification: How inefficient is the $1/n$ portfolio strategy? *Rev. Financial Studies* **22**(5), 1915–1953.

Derman, E. & I. Kani (1994a). Riding on a smile. *Risk* **7**, 32–39.

Derman, E. & I. Kani (1994b). The volatility smile and its implied tree. *Goldman Sachs Quantitative Strategies Research Notes*.

Dias, A. (2004). Copula inference for finance and insurance. *Doctoral Thesis*.

Dias, A. & P. Embrechts (2004). Dynamic copula models for multivariate high-frequency data in finance. *Working paper. ETH Zurich. Switzerland*.

Dias, A. & P. Embrechts (2008). Modelling exchange rate dependence at different time horizons. *Working paper. ETH Zurich. Switzerland*.

Diop, A. (2003). *Sur la discrétisation et le comportement á petit bruit d'EDS multidimensionnalles dont les coefficients sont á dérivées singuliéres*. Ph. D. thesis, INRIA.

Du, K., E. Platen, & R. Rendek (2012). Modeling of oil prices. *Working paper. University of Technology, Sydney, Australia*.

Duffie, D. & R. Kan (1994). Multi-factor term structure models. *Philos. Trans. Roy. Soc. London Ser. A* **347**, 577–580.

Dupire, B. (1992). Arbitrage pricing with stochastic volatility. In *Proceedings of AFFI Conference, Paris*.

Dupire, B. (1993). Model art. *Risk* **6**, 118–124.

Eberlein, E. & U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli* **1**, 281–299.

Elliott, R. J., L. Aggoun, & J. B. Moore (1995). *Hidden Markov Models: Estimation and Control*, Volume 29 of *Appl. Math.* Springer.

Embrechts, P., A. McNeil, & D. Straumann (2001). Correlation and dependency in risk management: properties and pitfalls. In U. Press (Ed.), *In Risk Management: Value at Risk and Beyond*. Cambridge: M. Dempster and H. Moffatt.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**(4), 987–1007.

Fang, H., K. Fang, & S. Kotz (2002). The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis* **82(1)**, 1–16.

Fergusson, K. & E. Platen (2006). On the distributional characterization of log-returns of a world stock index. *Appl. Math. Finance* **13**(1), 19–38.

Fernholz, E. R. (2002). *Stochastic Portfolio Theory*, Volume 48 of *Appl. Math.* Springer.

Filipović, D. (2001). A general characterization of one factor affine term structure models. *Finance Stoch.* **5**, 389–412.

Filipović, D. & E. Platen (2009). Consistent market extensions under the benchmark approach. *Math. Finance* **19**(1), 41–52.

Fischer, P. & E. Platen (1999). Applications of the balanced method to stochastic differential equations in filtering. *Monte Carlo Methods Appl.* **5**(1), 19–38.

Fleming, J., B. Ostdiek, & R. E. Whaley (1995). Predicting stock market volatility: a new measure. *Journal of Futures Markets* **15**, 265–302.

Frahm, G., M. Junker, & A. Szimayer (2003). Elliptical copulas: applicability and limitations. *Statistics and Probability Letters* **63**, 275–286.

Frey, R. (1997). Derivative asset analysis in models with level-dependent and stochastic volatility. Mathematics of Finance, Part II. *CWI Quarterly* **10**(1), 1–34.

Fujisaki, M., G. Kallianpur, & H. Kunita (1972). Stochastic differential equations for the non-linear filtering problem. *Osaka J. Math.* **9**, 19–40.

Gatheral, J. (2006). *The Volatility Surface. A Practitioner's Guide*. John Wiley & Sons, Inc.

Geman, H., D. Madan, & M. Yor (2001). Asset prices are Brownian motion: Only in business time. In *Quantitative Analysis in Financial Markets*, pp. 103–146. World Sci. Publishing.

Genest, C. & L. Rivest (2002). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* **88(423)**, 1034–1043.

Ghysels, E., A. Harvey, & E. Renault (1996). Stochastic volatility. In *Statistical Methods in Finance*, Volume 14 of *Handbook of Statist.*, pp. 119–191. North-Holland.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*, Volume 53 of *Appl. Math.* Springer.

Gouriéroux, C. & R. Sufana (2004). Derivative pricing with multivariate stochastic volatility: Application to credit risk. Working paper CREST.

Hagan, P. S., D. Kumar, A. S. Lesniewski, & D. Woodward (2002). Managing smile risk. *Willmot*, 84–108.

Hakansson, N. H. (1971). Capital growth and the mean-variance approach to portfolio selection. *J. Financial and Quantitative Analysis* **6**(1), 517–557.

Harrison, J. M. & D. M. Kreps (1979). Martingale and arbitrage in multiperiod securities markets. *J. Economic Theory* **20**, 381–408.

Hernandez, D. B. & R. Spigler (1993). Convergence and stability of implicit Runge-Kutta methods for systems with multiplicative noise. *BIT* **33**, 654–669.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies* **6**(2), 327–343.

Higham, D. J. (2000). Mean-square and asymptotic stability of numerical methods for stochastic ordinary differential equations. *SIAM J. Numer. Anal.* **38**, 753–769.

Higham, D. J., X. Mao, & C. Yuan (2007). Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations. *SIAM J. Numer. Anal.* **45**, 592–609.

Higham, N. J. (2005). The scaling and squaring method for the matrix exponential revisited. *SIAM J. Matrix Anal. Appl.* **26**(4).

Hu, L. (2006). Dependence patterns across financial markets: a mixed copula approach. *Applied Financial Economics* **16**, 717–729.

Hult, H. & F. Lindskog (2001). Multivariate extremes, aggregation and dependence in elliptical distributions. *Working paper, Risklab, Switzerland*.

Hurst, S. R. & E. Platen (1997). The marginal distributions of returns and volatility. In Y. Dodge (Ed.), L_1*-Statistical Procedures and Related Topics*, Volume 31 of *IMS Lecture Notes - Monograph Series*, pp. 301–314. Institute of Mathematical Statistics Hayward, California.

ICB (2008). Industry classification benchmark. http://rcbenchmark.com/index.html.

Ignatieva, K., E. Platen, & R. Rendek (2011). Using dynamic copulae for modeling dependency in currency denominations of a diversified world stock index. *Journal of Statistical Theory and Practice* **5**(3), 425–452.

Jäckel, P. (2002). *Monte Carlo Methods in Finance*. Wiley.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis* **46(2)**, 262 – 282.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall.

Kahl, C. (2004). *Positive numerical integration of stochastic differential equations.* Ph. D. thesis, Bergische University Wuppertal.

Kahl, C. & H. Schurz (2005). Balanced Milstein methods for ordinary SDEs. Technical report, Department of Mathematics, Southern Illinois University.

Kahneman, D. & A. Tversky (1979). Prospect theory: an analysis of decision under risk. *Econometrica* **47**(2), 263–292.

Kallianpur, G. (1980). *Stochastic Filtering Theory*. Springer.

Karatzas, I. & C. Kardaras (2007). The numeraire portfolio in semimartingale financial models. *Finance Stoch.* **11**(4), 447–493.

Kardaras, C. & E. Platen (2008). On financial markets where only buy-and-hold trading is possible. Technical report, University of Technology, Sydney. QFRC Research Paper 213.

Kelly, J. R. (1956). A new interpretation of information rate. *Bell Syst. Techn. J.* **35**, 917–926.

Kendall, W. S. (1989). The diffusion of Euclidean shape. Technical report, Research Report 161, University of Warwick, Dept. of Statistics.

Klauder, J. R. & W. P. Petersen (1985). Numerical integration of multiplicative-noise stochastic differential equations. *SIAM J. Numer. Anal.* **6**, 1153–1166.

Kloeden, P. E. & E. Platen (1992). Higher order implicit strong numerical schemes for stochastic differential equations. *J. Statist. Phys.* **66**(1/2), 283–314.

Kloeden, P. E. & E. Platen (1999). *Numerical Solution of Stochastic Differential Equations*, Volume 23 of *Appl. Math.* Springer. Third printing, (first edition (1992)).

Kloeden, P. E., E. Platen, & H. Schurz (1993). Higher order approximate Markov chain filters. In e. S. Cambanis (Ed.), *Stochastic Processes - A Festschrift in Honour of Gopinath Kallianpur*, pp. 181–190. Springer.

Kloeden, P. E., E. Platen, & H. Schurz (2003). *Numerical Solution of SDEs Through Computer Experiments*. Universitext. Springer. Third corrected printing, (first edition (1994)).

Kou, S. G. (2002). A jump diffusion model for option pricing. *Management Science* **48**, 1086–1101.

Küchler, U., K. Neumann, M. Sørensen, & A. Streller (1999). Stock returns and hyperbolic distributions. *Math. Comput. Modelling* **29**, 1–15.

Latané, H. (1959). Criteria for choice among risky ventures. *J. Political Economy* **38**, 145–155.

Le, T. & E. Platen (2006). Approximating the growth optimal portfolio with a diversified world stock index. *Journal of Risk Finance* **7**(5), 559–574.

Lewis, A. L. (2000). *Option Valuation Under Stochastic Volatility*. Finance Press, Newport Beach.

Lindskog, F., A. McNeil, & U. Schmock (2001). Kendall's tau for elliptical distributions. *Working paper, Risklab, Switzerland*.

Loewenstein, M. & G. A. Willard (2000). Local martingales, arbitrage, and viability: Free snacks and cheap thrills. *Econometric Theory* **16**(1), 135–161.

Long, J. B. (1990). The numeraire portfolio. *J. Financial Economics* **26**, 29–69.

Lord, R., R. Koekkoek, & D. van Dijk (2006). A comparison of biased simulation schemes for stochastic volatility models. Tinbergen Institute Discussion Paper (TI 2006-046/4).

Luenberger, D. G. (1998). *Investment Science*. Oxford University Press, New York.

MacLean, L. C., E. O. Thorp, & W. Ziemba (2011). *The Kelly Capital Growth Investment Criterion*. World Scientific.

Madan, D. & E. Seneta (1990). The variance gamma (V.G.) model for share market returns. *J. Business* **63**, 511–524.

Mandelbrot, B. (1963). The variation of certain speculative prices. *J. Business* **36**, 394–419. Reprinted in Cootner (1964), Chapter 15, 307-337.

Markowitz, H. (1952). Portfolio selection. *J. Finance* **VII**(1), 77–91.

Markowitz, H. (1976). Investment for the long run: New evidence for an old rule. *J. Finance* **XXXI**(5), 1273–1286.

Markowitz, H. & N. Usmen (1996a). The likelihood of various stock market return distributions, Part 1: Principles of inference. *J. Risk & Uncertainty* **13**(3), 207–219.

Markowitz, H. & N. Usmen (1996b). The likelihood of various stock market return distributions, Part 2: Empirical results. *J. Risk & Uncertainty* **13**(3), 221–247.

McLeish, D. L. & C. G. Small (1988). *The Theory and Applications of Statistical Inference Functions. Lecture Notes in Statistics*, Volume 44. Springer-Verlag.

McNeil, A., R. Frey, & P. Embrechts (2005). *Quantitative Risk Management*. Princeton University Press.

Mehra, R. & E. C. Prescott (1985). The equity premium: A puzzle. *J. Monetary Economy* **15**, 145–161.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica* **41**, 867–888.

Milstein, G. N. (1988). A theorem of the order of convergence of mean square approximations of systems of stochastic differential equations. *Theory Probab. Appl.* **32**, 738–741.

Milstein, G. N. (1995). *Numerical Integration of Stochastic Differential Equations*. Mathematics and Its Applications. Kluwer.

Milstein, G. N., E. Platen, & H. Schurz (1998). Balanced implicit methods for stiff stochastic systems. *SIAM J. Numer. Anal.* **35**(3), 1010–1019.

Nelsen, R. (1998). *An Introduction to Copulas*. Springer-Verlag.

Nelson, D. B. (1990). ARCH models as diffusion approximations. *J. Econometrics* **45**, 7–38.

Newton, N. J. (1986). An asymptotic efficient difference formula for solving stochastic differential equations. *Stochastics* **19**, 175–206.

Newton, N. J. (1991). Asymptotically efficient Runge-Kutta methods for a class of Ito and Stratonovich equations. *SIAM J. Appl. Math.* **51**, 542–567.

Platen, E. (1982). An approximation method for a class of Itô processes with jump component. *Liet. Mat. Rink.* **22**(2), 124–136.

Platen, E. (1997). A non-linear stochastic volatility model. Technical report, Australian National University, Canberra, Financial Mathematics Research Reports. FMRR 005-97.

Platen, E. (2001). A minimal financial market model. In *Trends in Mathematics*, pp. 293–301. Birkhäuser.

Platen, E. (2002). Arbitrage in continuous complete markets. *Adv. in Appl. Probab.* **34**(3), 540–558.

Platen, E. (2004). A benchmark framework for risk management. In *Stochastic Processes and Applications to Mathematical Finance*, pp. 305–335. Proceedings of the Ritsumeikan Intern. Symposium: World Scientific.

Platen, E. (2005a). Diversified portfolios with jumps in a benchmark framework. *Asia-Pacific Financial Markets* **11**(1), 1–22.

Platen, E. (2005b). Diversified portfolios with jumps in a benchmark framework. *Asia-Pacific Financial Markets* **11(1)**, 1–22.

Platen, E. (2006). A benchmark approach to asset management. *J. Asset Management* **6**(6), 390–405.

Platen, E. (2011). A benchmark approach to investing and pricing. in:. *MacLean, L.C. and Thorp, E. O. and Ziemba, W. (2011), The Kelly Capital Growth Investment Criterion. World Scientific.*, 409–425.

Platen, E. & N. Bruti-Liberati (2010). *Numerical Solution of SDEs with Jumps in Finance*. Springer.

Platen, E. & D. Heath (2010). *A Benchmark Approach to Quantitative Finance*. Springer Finance. Springer.

Platen, E. & R. Rendek (2008). Empirical evidence on Student-t log-returns of diversified world stock indices. *Journal of Statistical Theory and Practice* **2**(2), 233–251.

Platen, E. & R. Rendek (2009). Exact scenario simulation for selected multi-dimensional stochastic processes. *Communications on Stochastic Analy-*

sis **3**(3), 443–465.

Platen, E. & R. Rendek (2010). Quasi-exact approximation of hidden Markov chain filters. *Communications on Stochastic Analysis* **4**(1), 129–142.

Platen, E. & R. Rendek (2012a). Affine nature of aggregate wealth dynamics. *Working paper. University of Technology, Sydney, Australia*.

Platen, E. & R. Rendek (2012b). Approximating the numéraire portfolio by naive diversification. *Journal of Asset Management* **13**(1), 34–50.

Platen, E. & R. Rendek (2012c). *Simulation of diversified portfolios in a continuous financial market*. Stochastic Analysis and Applications to Finance. World Scientific. pp. 385–410.

Platen, E. & L. Shi (2008). On the numerical stability of simulation methods for sdes. Technical report, University of Technology, Sydney. QFRC Research Paper 234.

Platen, E. & G. Stahl (2003a). A structure for general and specific market risk. *Computational Statistics* **18**(3), 355 – 373.

Platen, E. & G. Stahl (2003b). A structure for general and specific market risk. *Computational Statistics* **18(3)**, 355–373.

Popper, K. R. (1934). *Logik der Forschung*. Springer, Vienna. Amplified English edition, Popper (1959).

Praetz, P. D. (1972). The distribution of share price changes. *J. Business* **45**, 49–55.

Rachev, S. & S. Han (2000). Portfolio management with stable distributions. *Mathematical methods of operations research* **51**, 341–352.

Rachev, S. & S. Mittnik (2000). *Stable Paretian Models in Finance*. John Wiley & Sons.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.). Wiley, New York.

Revuz, D. & M. Yor (1999). *Continuous Martingales and Brownian Motion* (3rd ed.). Springer.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *J. Economic Theory* **13**, 341–360.

Rubinstein, M. (1976). The strong case for the generalized logarithmic utility model as the premier model of financial markets. *J. Finance* **31**, 551–571.

Saito, Y. & T. Mitsui (1993). T-stability of numerical schemes for stochastic differential equations. *World Sci. Ser. Appl. Anal.* **2**, 333–344.

Samuelson, P. A. (1957). Intertemporal price equilibrium: A prologue to the theory of speculation. *Weltwirtschaftliches Archiv* **79**, 181–221.

Shiryaev, A. N. (1984). *Probability*. Springer.

Simon, H. A. & C. P. Bonini (1958). The size distribution of business firms. *Amer. Econ. Rev.* **48**(4), 607–617.

Smith, R. D. (2007). An almost exact simulation method for the Heston model. *J. Comput. Finance* **11**(1), 115–125.

Stutzer, M. J. (2000). A portfolio performance index. *Financial Analysts J.* **56**(3), 52–61.

Talay, D. (1982). Convergence for each trajectory of an approximation scheme of SDE. *Computes Rendus Acad. Sc. Paris, Séries I Math.* **295**(3), 249–252. (in French).

Thorp, E. O. (1972). Portfolio choice and the Kelly criterion. In *Proceedings of the 1971 Business and Economics Section of the American Statistical Association*, Volume 21, pp. 5–224.

Wang, S. (1997). Aggregation of correlated risk portfolios. *Preprint, Casuality Actuarial Society (CAS)*.

Wonham, W. M. (1965). Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM J. Control Optim.* **2**, 347–369.

Wu, F., A. Valdez, & M. Sherris (2007). Simulating exchangeable multivariate archimedean copulas and its applications. *Communications in Statistics - Simulation and Computation* **36**(5), 1019–1034.

Zakai, M. (1969). On the optimal filtering of diffusion processes. *Z. Wahrsch. Verw. Gebiete* **11**, 230–243.

Ziemba, W. T. & J. M. Mulvey (1998). *Worldwide Asset and Liability Modeling.* Cambridge University Press.