Inference and Intraday Analysis of Diversified World Stock Indices

by

Leah KELLY

Submitted to the University of Technology, Sydney

for the degree of Doctor of Philosophy.

Submitted February, 2004

Certificate

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the 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 the thesis. In addition I certify that all information sources and literature used are indicated in the thesis.

(Signed) Production Note:
Signature removed prior to publication.

Acknowledgements

I would like to thank my supervisor, Professor Eckhard Platen for his assistance, guidance and support. I would also like to thank Dr David Heath, Dr Wolfgang Breymann, Mrs Katrin Platen, Dr Mark Craddock and Shane Miller.

I would also like to acknowledge financial support from the Department of Mathematical Sciences and the Quantitative Finance Research Centre.

I would also like to extend my thanks to Michael Kelly, Jan Kelly and Kate Kennedy for their constructive criticism and ongoing support. Finally, I would like to thank Justin den Hertog for his support, helpful comments and good humour.

Chapter 3 is an extended version of a paper written with Professor Michael Sørensen and Professor Platen. Chapter 5 is based on work undertaken with Dr Wolfgang Breymann and Professor Eckhard Platen. The high–frequency data used in Chapters 4, 5 and 6 was made available to me by Dr Breymann from Olsen's Data. Daily data was made available from Thomson Financial. Long term data was made available from Global Financial Data.

Contents

List of Tables

List of Figures

Abstract

The benchmark framework provides an alternative paradigm for financial market modelling. Firstly, approaches to parameter estimation of discretely observed diffusion processes are examined, with particular emphasis on estimating function techniques. A new estimating function technique, called the transform function method, is introduced and applied to a class of stochastic differential equations. The advantage of the transform function method is that explicit information is not required about conditional moments and the existence of stationary transition densities. Despite the flexibility of the transform function technique, it suffers the same drawbacks as existing estimation methods with respect to drift estimation of financial data. The observation period required to estimate the drift function is much greater than what is available in financial markets. Notwithstanding the lack of available data for drift estimation, the diffusion function can be reliably estimated by the transform function method from short periods of frequently observed data. This thesis highlights that the benchmark approach of Platen (2002), where only estimates for diffusion coefficients are required, resolves the issue of drift estimation from financial market data.

Secondly, we extend the benchmark approach to incorporate modelling aspects particular to intraday data. This requires the construction of high–frequency diversified portfolios to approximate the growth optimal portfolio, the central building block of the benchmark model. Three different high–frequency indices are considered and it is demonstrated that an index with proportions based on total world market capitalisation provides the best proxy for the growth optimal portfolio. The benchmark model is then extended to intraday data through the introduction of a market activity process. This process is shown to be readily observable and as such, properties of market activity can be characterised. The analysis includes the consideration of the high–frequency indices in US Dollars and a further twenty currency denominations. As such, we reveal the pairwise co-

variation structures between the currency denominations considered. It is shown that the intraday benchmark model, which has market activity as the main parameter of interest, provides a largely accurate intraday description of financial markets.

Chapter 1

Introduction

Modelling of financial markets using continuous time stochastic processes is central to modern finance, see Merton (1992). The identification and calibration of a financial market model is a key step to risk measurement, the pricing of derivative securities and portfolio optimisation. The aim of this thesis is two–fold. Firstly, to identify an appropriate *statistical method* to estimate and calibrate a financial market model and secondly, to extend the *benchmark model*, see Platen (2002), to incorporate *intraday data*.

Inference for Financial Markets

The estimation and calibration of market models is a difficult task. Many different market models have evolved and alternative estimation techniques have been suggested, see, for example, Gallant & Tauchen (1997), Overbeck & Rydén (1997), Chan, Karolyi, Longstaff & Sanders (1992) and Christensen, Poulsen & Sørensen (2001). A problem with parameter estimation for stochastic processes is that, typically, to estimate the drift coefficient, the data used in the estimation must be observed over an extremely long period of time. For example, consider the classical Black–Scholes–Merton model, see Black & Scholes (1973), with constant appreciation rate μ and volatility parameter σ. We wish to estimate the rate μ. To achieve a reliable estimate $\hat{\mu}$, such that $|\mu - \hat{\mu}| < 0.01$, within a 95% confidence interval, the length of the observation time $T > (196\,\sigma)^2$ is required. Thus, if $\sigma = 15\%$ then the required length of time is $T > 865$ years. Furthermore, the greater the volatility parameter σ, the longer the time period necessary to estimate $\hat{\mu}$ reliably. That is, if $\sigma = 30\%$ then the length of time increases to $T > 3,458$ years, see also Davis (2003) or Rogers (2001). The volatility σ is

simpler to estimate since it is observable as the square root of the slope of the quadratic variation of the logarithm of the sample path. The estimate of σ can be achieved with considerable precision. For example, if we consider the above model with $\sigma = 20\%$ the time required to estimate the volatility reliably from daily observations is approximately $T = 3$ years. However, if we use hourly observations the time required to estimate σ is $T = 0.5$ years. This is because the estimate based on quadratic variation becomes exact as the time between observations approaches zero. This demonstrates that a financial model which only requires estimation of the diffusion function is beneficial for practical applications.

Many existing techniques used for the statistical analysis of financial markets rely on *stationarity*. This suggests that even if a sufficiently long financial data series exists, it may be inappropriate to base results on this data set since financial markets have changed markedly over the years. Furthermore, a true picture of long term financial market evolution during the last century is difficult to grasp due to events such as World War One and World War Two, which raises questions about the reliability of the data series over these time periods. Additionally, many estimation techniques require the drift and diffusion parameters to be constant. This is an unreasonable assumption over long observation periods for a time series of financial data.

Chapter 2 reviews various approaches related to parameter inference for *discretely observed diffusion processes*. Maximum likelihood estimation requires the diffusion coefficient to be known and approximations to the likelihood function are required if the transition density is not explicitly available. An alternative are *estimating function* techniques, which provide extremely flexible methods of parameter estimation and are well suited to discretely observed diffusion processes. Whilst these techniques can be used to estimate both the drift and diffusion coefficient, stationarity is usually required and a long observation interval is still necessary for reliable drift estimation. Additionally, whereas the diffusion coefficient benefits from frequently observed data, such high–frequency data can be

detrimental to the estimation of the drift functions when using certain estimating techniques, see Bibby, Jacobsen & Sørensen (2003).

To increase the flexibility of estimating functions, the *transform function* method is developed in Chapter 3. As an example of the method, we consider the estimation of affine stochastic differential equations (SDEs) with unknown drift and diffusion parameters. This example results in explicit estimators when a power transform function is used. For the simulated processes, the results confirm the applicability of the method. Although the standard problem with inference about the drift function from realistic observation periods is again apparent.

The Benchmark Approach and Stochastic Volatility

A distinct advantage of the benchmark approach of Platen (2002) is that the SDE for the discounted *growth optimal portfolio* (GOP), the building block of the modelling framework, is completely determined by its diffusion coefficient. This implies that an accurate method for estimating the diffusion coefficient is sufficient to obtain estimates for the drift coefficient. Thus, by providing a link between the diffusion and drift coefficients, a potential solution to the problem of reliable drift estimation for financial data is provided by the benchmark approach.

Essentially, the benchmark model consists of several risky assets and one riskless asset, known as primary security accounts. To form the GOP, a self–financing portfolio of primary security accounts is created that achieves the maximum logarithmic growth rate. Identification of the general dynamics of the GOP shows that when used as benchmark or numeraire, benchmarked portfolios are local martingales under the real world probability measure. This gives rise to the concept of fair pricing for derivative securities under the real world probability measure, where benchmarked fair price processes are martingales.

The practical use of the benchmark model relies on the identification and calibra-

tion of the GOP dynamics. The optimal proportions used to form this portfolio depend on the matrix of all volatilities and correlations between the primary security accounts, in addition to the corresponding market prices for risk. The market price for risk appears in the drift coefficient, which is independent of the diffusion coefficient of the primary security accounts and hence cannot be reliably estimated using standard techniques. In addition, estimation of a large volatility matrix is a cumbersome task and in practice, the exact GOP can be extremely difficult to form, see, for example, Long (1990). As an alternative, we concentrate on the analysis of *approximate GOPs*. An approximate GOP is any broadly diversified portfolio. This result is provided by a limit theorem in Platen (2003). In this thesis, we construct *world stock indices* to provide intraday approximations to the GOP.

In principle, proxies for the GOP, observed at daily intervals at least, are already available in the form of broadly diversified accumulation indices. We show in Figure 1.1 the evolution of the logarithm of a discounted accumulated world stock index created by Global Financial Data over the last 77 years. The data is observed monthly from 1925 until 2003. It is clear that the discounted world stock index has experienced significant growth over the years. A trendline with slope of 0.043, which approximates the growth of the logarithm of the index, is shown. This implies that on average, the growth of the index appears to be exponential with an approximate net growth rate of $\eta = 4.3\%$. Note that any reasonable estimation technique will give an approximate growth rate in this vicinity.

Without restricting the generality of the benchmark approach, Platen (2002) shows that the dynamics of the discounted GOP $V(t)$ satisfy the SDE

$$dV(t) = \alpha(t)dt + \sqrt{\alpha(t)V(t)}\, dW(t)$$

for $t \in [0, T]$. The initial value $V(0)$ is assumed to be strictly positive. Here W is a standard Wiener process. An analysis of available market data shows that the *discounted GOP drift process* $\alpha(t) = \xi \exp\{\eta t\}$ can be simply taken

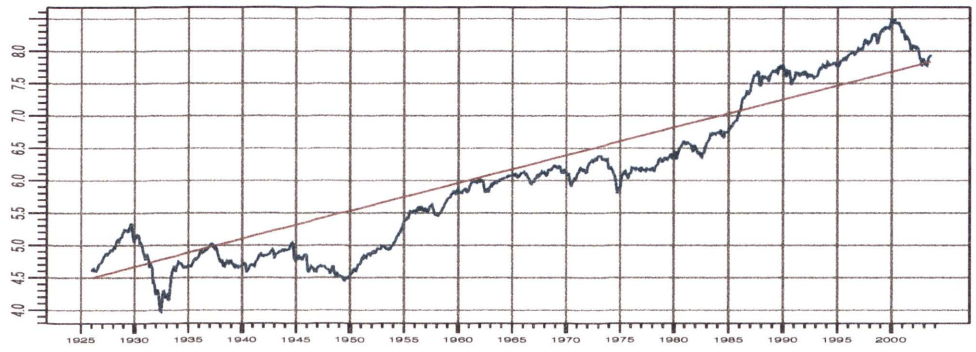

Figure 1.1: Long term dynamics of the logarithm of a discounted accumulated world stock index.

as a parameter process, which relies on the net growth rate η. Throughout the thesis, the net growth rate is assumed to be constant. The resulting dynamics of the discounted GOP are that of a time transformed squared Bessel process of dimension four, see Revuz & Yor (1999). To obtain a stationary process, a squared Bessel process can be transformed into a square root process. This results in the *normalised GOP* $Y(t) = \frac{V(t)}{\alpha(t)}$, which satisfies the SDE

$$dY(t) = \eta \left(\frac{1}{\eta} - Y(t) \right) dt + \sqrt{Y(t)} dW(t)$$

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

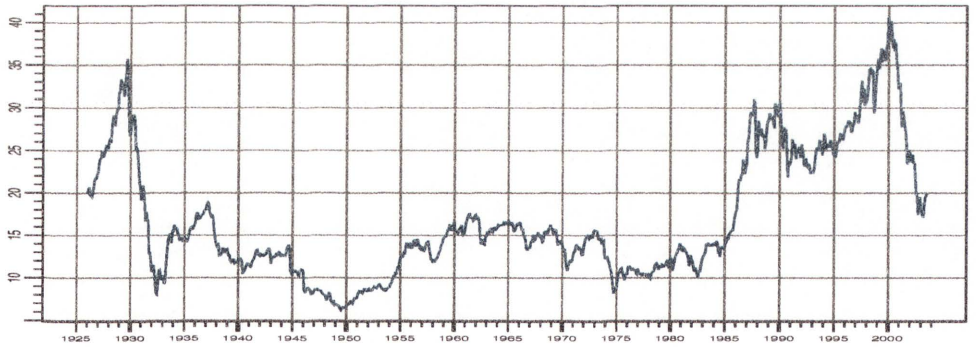

Figure 1.2: Long term dynamics of a discounted normalised world stock index.

Under the above assumptions, we obtain the dynamics of the normalised GOP as shown in Figure 1.2. This allows the calculation of the GOP volatility $|\theta(t)|$,

which is given by

$$|\theta(t)| = \sqrt{\frac{\alpha(t)}{V(t)}} = \sqrt{\frac{1}{Y(t)}}.$$

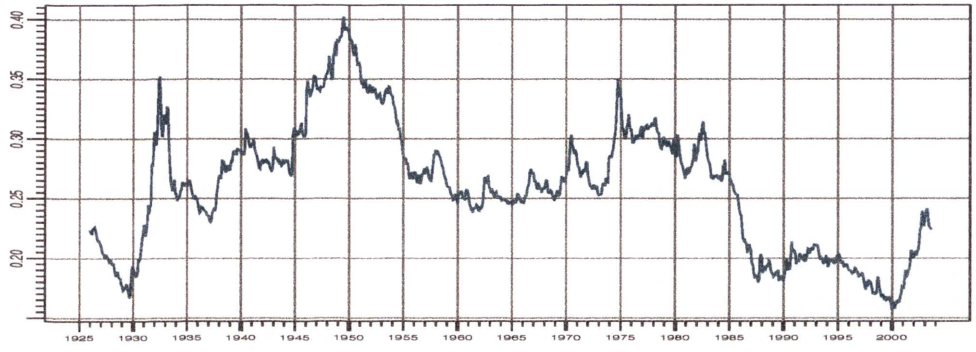

Figure 1.3: Long term dynamics of the volatility of the world stock index.

The resulting volatility process in Figure 1.3 is not constant nor can it be described by a deterministic function of time. This highlights the need for stochastic volatility to be included in any reasonable model. In the benchmark model by Platen (2001, 2002, 2003), the dynamics of volatility are implied through the structure of the GOP, resulting in an endogenous stochastic volatility model. This is in contrast to models commonly referred to in the literature, where an external stochastic volatility process is introduced, see, for instance, Hull & White (1987) or Heston (1993). This specification of the benchmark model makes the identification of the stochastic volatility of the GOP straightforward.

Intraday Analysis

Advances in data technology have made the recording and analysis of high–frequency data possible. As such, the intraday empirical analysis of financial market data has been considered by numerous authors, for example, Dacorogna, Gençay, Müller, Olsen & Pictet (2001) and references therein. However, the modelling and analysis of intraday data is more difficult than, for instance, daily data. This is due mainly to the presence of intraday *seasonal patterns* in high–frequency data. These seasonalities must be properly accounted for in order to

prevent important empirical features of the data from being overshadowed by intraday patterns.

To account for the residual stochastic and seasonal features caused by intraday data, the standard benchmark model is extended in Chapter 4 to incorporate *market activity* $m(t)$. Market activity is obtained as the slope of the quadratic variation of the square root of the normalised GOP and captures much of the short term seasonal and stochastic features of the squared GOP volatility. The introduction of market activity gives rise to a *market activity time* scale. The only assumption is that the expected market activity time approaches physical time over longer observation periods. This is a natural assumption since seasonal patterns do not greatly affect the empirical analysis of data observed over longer time intervals.

Diversified World Stock Indices

To examine the properties of the intraday benchmark model, three high–frequency world stock indices, denominated in US Dollars, are constructed. It is shown that of the high–frequency indices considered, the world stock index based on total world market capitalisation displays the largest growth over the period and also provides statistical properties that best reflect the benchmark model. Therefore, it is interpreted as the closest approximation to the exact GOP of those considered.

Chapter 5 assesses the properties of the three intraday world stock indices. We concentrate entirely on the diffusion coefficient estimation. For these indices, we analyse the properties of the resulting market activity time. Following this, a model for the market activity $m(t)$ is deduced and examined. Market activity is shown to have a multiplicative diffusion coefficient in addition to a strongly mean–reverting drift function with different seasonal patterns in the drift and diffusion coefficients. The seasonal pattern in its diffusion coefficient, captured by the *activity volatility* $\beta(t)$, implies that an *activity volatility time* scale τ can

be introduced that deseasonalises the fluctuations of market activity.

Complete flexibility in the modelling of the drift coefficient of market activity is permitted, which is specified via a non–parametric drift function. It is shown that in activity volatility time τ, the market activity derived from the three indices considered has a stationary density. Since the multiplicative diffusion coefficient of market activity is already identified, this implies a particular form for the market activity drift function. The empirical stationary density of the logarithm of market activity is well described by a gamma density. This leads naturally to a linear mean–reverting drift function for the logarithm of market activity. The logarithm of market activity in activity volatility time m_τ satisfies the SDE

$$d\ln(m_\tau) = \left(\frac{p-1}{2} - \frac{\gamma}{2}m_\tau\right)d\tau + d\tilde{W}_\tau,$$

where the parameters p and γ characterise the stationary gamma density. Here $\tilde{W}$ is a standard Wiener process.

Currency Denominations of Diversified World Stock Indices

A key feature of the benchmark approach is that the choice of denomination does not change the structure of the dynamics for the GOP. In Chapter 6 the analysis of the market capitalisation based world stock index is conducted for a further twenty currency denominations. The market activities derived are shown to behave in a similar manner as that for the world stock index when denominated in US Dollars. It is also shown that the majority of denominations appear to have logarithmic market activity that has approximately a gamma density as stationary density. Again, this implies a particular form of the drift function.

The pairwise relationships of the driving noises of normalised world stock indices and market activities for the different currency denominations are also investigated. We are able to calculate *covariation coefficients* between different denominations. Using *cluster analysis* we are able to clearly identify interesting groupings of currencies that display similar behaviour. The groupings that result

from the fluctuations of the normalised world stock indices seem to be predominantly motivated by economic factors, whereas the groupings that result from the market activities appear to reflect regional effects. Note, that these groupings change over time as new economic alliances are formed and trading technologies and habits change.

The thesis proposes the use of the intraday benchmark framework for the modelling of financial markets. This is different to standard approaches in the financial literature. An alternative parametrisation is provided by the intraday benchmark model through the introduction of the market activity process, which appears to be more effective than the volatility parametrisation proposed in the widely used Black–Scholes–Merton model and its extensions. The suggested concept of market activity in this thesis, models the short term stochastic and seasonal effects particular to high–frequency data. It allows the consistent modelling of the GOPs and provides a largely accurate description of the fluctuations of a diversified world stock index in different currencies.

Chapter 2

Parameter Estimation of Discretely Observed Diffusion Processes

In this chapter we review the problem of statistical inference for stochastic processes and the various approaches that have been proposed. We examine techniques that have been developed for discretely observed diffusion processes, concentrating on estimating function techniques with a brief overview of inference for continuously observed diffusion processes. Common definitions in the area of parameter estimation are given in Appendix A.

2.1 Background and Notation

Consider, for a parameter $\vartheta \in \Theta \subseteq R^p$ where $p \in \{1, 2, \ldots\}$, the one-dimensional diffusion process, $X = \{X(t), t \geq 0\}$ defined as the solution to the following class of stochastic differential equations (SDEs)

$$dX(t) = a(X(t); \vartheta)dt + b(X(t); \vartheta)dW(t) \tag{2.1.1}$$

where $X(0) = x(0)$ and $W = \{W(t), t \geq 0\}$ is a standard Wiener process on the filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P_\vartheta)$, fulfilling the usual conditions, see Karatzas & Shreve (1991). We assume that the drift function $a(X(t); \vartheta)$ and the diffusion function $b(X(t); \vartheta)$ are known with the exception of the parameter vector ϑ. Moreover, the coefficient functions are assumed to satisfy the usual Lipshitz and linear growth conditions, ensuring that a unique solution of (2.1.1) exists for all $\vartheta \in \Theta$, see Karatzas & Shreve (1991). The goal is to estimate the parameter ϑ from a *discretely observed diffusion process.* This situation is more realistic, since in practice, it is impossible to observe data continuously. It is

assumed that the continuous time process X has been observed at discrete time points $t_0 < t_1 < t_2, \ldots$, with values $X(t_0), X(t_1), X(t_2) \ldots$. We observe a sequence of random variables $(X(t_n))$ for $n \in \{1, 2, \ldots\}$, which are values of the process X, rather than the entire path of the process. Of course, this sequence will not contain all the information about the path of X. However, it may be sufficient to obtain reasonable parameter estimates. Throughout this chapter, $\Delta_n = t_n - t_{n-1}$ denotes the time between observations, where $n \in \{1, 2, \ldots, n_t\}$. Here

$$n_t = \max\{n \in \{0, 1, \ldots\} : t_n \leq t\}. \tag{2.1.2}$$

When the time between observations is fixed, we will simply write Δ for Δ_n. For simplicity, throughout the following discussion, we take X to be one-dimensional but many results can be generalised to the d-dimensional case.

The problem of inference for a continuously observed diffusion process is considered in, for instance, Liptser & Shiryaev (2001), Prakasa Rao (1999) and Kutoyants (1984). When $b(X(t); \vartheta) = b(X(t))$, that is, the diffusion coefficient is independent of the parameter to be estimated, *maximum likelihood estimation* can be used for inference about the drift parameter. Maximum likelihood estimation involves forming a likelihood function and finding the maximum of this function. To illustrate the basic idea we let $a(X(t); \vartheta) = \vartheta\, a(X(t))$ and $b(X(t)) = 1$ in (2.1.1). The *likelihood function* is given as

$$L_t(\vartheta) = \exp\left\{\vartheta \int_0^t a(X(s))dX(s) - \frac{1}{2}\int_0^t \vartheta^2 a^2(X(s))ds\right\} \tag{2.1.3}$$

see, Liptser & Shiryaev (2001). Note that the log-likelihood function is simply the logarithm of (2.1.3). The maximum likelihood estimator can be found by solving

$$U(\vartheta) = \frac{L_t'(\vartheta)}{L_t(\vartheta)} = 0. \tag{2.1.4}$$

Here $U(\vartheta)$ is known as the *score function*. Note that $L_t'(\vartheta)$ denotes the first derivative of $L(\cdot)$ with respect to ϑ. The estimators that result are consistent, asymptotically normal and efficient. Results regarding more general inference

problems using maximum likelihood estimation are available, see, for instance, Kutoyants (1984). If the diffusion coefficient depends on the parameter ϑ, it may be determined using the quadratic variation of the process X at time t, denoted by $\langle X \rangle_t$, see (2.1.6) below. Maximum likelihood estimation may then be applied, see, for instance, Florens-Zmirou (1993). Note that if the diffusion coefficient is not constant, the estimator that results is biased, see Liptser & Shiryaev (2001).

Throughout this thesis, $\langle L \rangle_t$ denotes the *quadratic variation* of a continuous stochastic process $L = \{L(t), t \in [0,T]\}$, which can be defined as the limit in probability

$$\langle L \rangle_t \stackrel{P}{=} \lim_{\Delta \to 0} \langle L \rangle_{\Delta,t} \tag{2.1.5}$$

where $\langle L \rangle_{\Delta,t}$ is the *approximate quadratic variation*

$$\langle L \rangle_{\Delta,t} = \sum_{n=1}^{n_t} (L(t_n) - L(t_{n-1}))^2 \tag{2.1.6}$$

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

2.2 Approximate Maximum Likelihood Estimation for Discretely Observed Diffusion Processes

Ideally, the estimation of parameters from a discretely observed diffusion processes would be estimated using maximum likelihood estimation. When the transition densities are known explicitly, the maximum likelihood estimator can be characterised via the log–likelihood function, which can be used to estimate the parameter ϑ. In the case of equidistant time steps, Dacunha-Castelle & Florens-Zmirou (1986), prove consistency and asymptotic normality of the parameter estimate $\hat{\vartheta}_{n_t}$ as the number of observations n_t in (2.1.2) becomes large, regardless of the time between observations. Generally, the estimation requires long time periods to estimate the drift coefficient.

However, when the transition densities of X are not known in an explicit form, it is necessary to approximate the transition densities in some manner in order to apply maximum likelihood estimation. This provides an approximation for the likelihood function for ϑ.

Many alternatives for the approximation of the likelihood function have been suggested. An intuitively appealing approach is to discretise the continuous likelihood function using Itô and Riemann integrals. As in the case of continuous observations, this assumes that the diffusion function is independent of the parameter ϑ. However, estimators for discrete observations resulting from this discretisation are inconsistent if the time between observations is fixed, see Florens-Zmirou (1989) and Pedersen (1995b). This problem can be solved by suitable modifications. For example, H. Sørensen (2001) suggests an approximation to the continuous time score function that depends on the infinitesimal generator and the invariant density for X. This approximation is an example of a simple estimating function, see Kessler (2000). Bibby & Sørensen (1995) suggest approximating the score function by using a martingale estimating function, since the score function is itself a martingale. Martingale estimating functions are discussed in more detail in Subsection 2.4.1.

In addition to explicit approximations to the likelihood function, numerical approximations can also be performed although these techniques can be computationally intensive.

2.2.1 Approximations through Simulation

Pedersen (1995b) and Brandt & Santa-Clara (2002) independently derive a *Simulated Maximum Likelihood* technique. The idea of this method is to derive a sequence of approximations to the log–likelihood function. This is done by approximating the true, yet unknown, transition densities of X by first discretising the process X and producing a sequence of approximate transition densities. The

Euler scheme is used to simulate the diffusion process in this interval. Any scheme with closed form transition densities can be used to do this, such as the Milstein scheme, see Elerian (1998). By assumption, the one–step ahead transition densities of the Euler discretisation are Gaussian. Using this, the integral of the transition density of X is interpreted as an expectation, which can be evaluated using draws from this distribution.

The approximate likelihood function can then be shown to converge in probability to the true likelihood function as the number of observations increases for all $\vartheta \in \Theta$, see Pedersen (1995a, 1995b) and Brandt & Santa-Clara (2002). This implies that the parameters, which maximise the approximate likelihood function, converge to the parameters that maximise the true likelihood function.

An alternative simulated approximation method, based on results in Florens-Zmirou (1989), is suggested by Nicolau (2002).

2.2.2 Approximation using Numerical Solutions of the Kolmogorov Forward Equation

Lo (1988) and Poulsen (1999) suggest solving the *Kolmogorov forward equation*, also known as the *Fokker-Planck equation*, for the transition density of a diffusion process. Provided that the drift and diffusion functions are smooth functions, the transition density satisfies the Kolmogorov forward equation. This approximation of the transition density can then be used to form an *approximate maximum likelihood* (AML) function.

Poulsen (1999) uses the Crank-Nicolson finite difference technique to solve the Kolmogorov forward equation for the transition density and then forms the AML function. It is shown that under certain assumptions, the estimator that results from AML estimation is consistent and asymptotically normal. Moreover, under additional assumptions, the AML estimator is asymptotically equivalent to the maximum likelihood estimator.

2.2.3 Approximation using Non–parametric Techniques

A further method of approximating the transition density is given in Aït-Sahalia (2002). A closed form sequence of approximations to the log–likelihood function is constructed. The approximation is based on a sequence of Hermite approximations to the true transition density. This results in a proxy for the true log–likelihood function, which involves the same parameters as the original model. Aït-Sahalia (2002) shows that the sequence of approximating log–likelihood functions converges to the true log–likelihood function as the number of approximations in the sequence increases. Additionally, maximising the approximate likelihood function results in an estimator that converges to the true maximum likelihood estimator. Aït-Sahalia (2003) extends this work to consider closed form expansions for the transition density and likelihood functions of arbitrary multivariate diffusion processes.

Jensen & Poulsen (1999) give a review and comparison of various numerical techniques that have been implemented to approximate the transition densities of a diffusion process. The results indicate that, when speed and accuracy are taken into account, the Hermitian expansion is superior to other methods considered.

2.2.4 Bayesian Analysis

Eraker (1998) and Elerian, Chib & Shephard (2001) discuss the Bayesian analysis of discretely observed diffusion processes. The idea is to treat the values of the diffusion between any two discrete measurements as missing data and then to apply tuned Markov Chain Monte Carlo (MCMC) methods to "learn" about the missing data and the parameters.

Using MCMC, the posterior density can be sampled by simulating a Markov chain that has an invariant and limiting density equal to the posterior density. After an initial transient stage, the path of the Markov chain provides a sequence of draws from the marginal distribution of the posterior and can be used to conduct

inferences about ϑ. This allows likelihood inference for the stochastic process despite the fact that the likelihood function is not explicitly evaluated.

2.3 Contrast Functions

Discretising the continuous log–likelihood function can also result in a *contrast function* that can be maximised to estimate ϑ, see Yoshida (1992). An alternative construction was introduced by Florens-Zmirou (1989) using the Euler scheme to discretise the diffusion process X. The log–likelihood function can be formed for this discretisation, which results in a contrast for the estimation of ϑ. The joint estimation of the drift and diffusion coefficient functions was also considered by Yoshida (1992). This was performed under two restrictive assumptions involving the form of the diffusion coefficient and the rate at which the time between observations approaches zero. Essentially, the diffusion coefficient is required to be multiplicative so that a diffusion estimator can be based on quadratic variation.

Kessler (1997) relaxes these assumptions and suggests a single contrast to estimate both the drift and diffusion function. To construct the contrast function, approximations of the first two conditional moments are derived. This requires the observation that the contrast function is the log–likelihood of X, provided the transition density is Gaussian. It is shown that the desired estimator is asymptotically efficient. No assumption regarding multiplicative noise is required.

2.4 Estimating Functions

Estimating functions are functions that consist of the parameter and the observations. A flexible method of estimating parameters of discretely observed diffusion processes is provided. An estimator is obtained by setting the estimating function equal to zero and solving for the parameter. Estimating functions provide a natural method for inference of diffusion processes since the maximum likelihood estimator, discussed in Section 2.1, is the zero of the score function. Thus, the

score function is a particular estimating function.

2.4.1 Martingale Estimating Functions

Bibby & Sørensen (1995) consider a *martingale estimating function* of the form

$$G_{n_t}(\vartheta) = \sum_{n=1}^{n_t} g(\Delta, X(t_{n-1}), X(t_n); \vartheta),$$

where n_t is given in (2.1.2). This is the form of the score function discussed in Section 2.1. It is argued in Bibby & Sørensen (1995) that a natural approximation of the score function is of the general form

$$g(\Delta, x, y; \vartheta) = \sum_{j=1}^{N} a_j(\Delta, x; \vartheta) h_j(\Delta, x, y; \vartheta)$$

where $h_j(\Delta, x, y; \vartheta)$ for $j = 1, 2, \ldots, N$ are given real valued functions satisfying the martingale condition

$$\int_l^r h_j(\Delta, x, y; \vartheta) p(\Delta, x, y; \vartheta) dy = 0$$

for all $\Delta > 0$, $x \in (l, r)$ and $\vartheta \in \Theta$. Here $p(\Delta, x, y; \vartheta)$ is the transition density of the process X for a transition from x to y over the time period Δ. The a_j are weights that are given to the functions h_j. These are derived in Bibby & Sørensen (1995) using *optimal estimating function* theory developed by Heyde (1988) and Godambe & Heyde (1987). The weights a_j determine the contribution of the functions h_j to the estimation of the parameters of the diffusion process. The choice of functions h_j is more difficult and depends predominantly on the diffusion process to be estimated.

Bibby & Sørensen (1995) suggest the *linear estimating function*

$$h_1(\Delta, x, y; \vartheta) = y - F(\Delta, x; \vartheta) \tag{2.4.1}$$

where $N = 1$ and $F(\Delta, x; \vartheta) = E_\vartheta(X(t_n)|X_0 = x)$ is the first conditional moment. Linear estimating functions are useful in cases where the diffusion coefficient is

known. When this is not the case the linear estimating function is too simplistic and does not contain adequate information to estimate the diffusion coefficient.

A natural extension to the linear estimating function, so that the diffusion coefficient can be estimated, is the *quadratic estimating function.* Here, h_1 is given as in (2.4.1) and

$$h_2(\Delta, x, y; \vartheta) = (y - F(\Delta, x; \vartheta))^2 - \phi(\Delta, x; \vartheta)$$

where $N = 2$ and

$$\phi(\Delta, x; \vartheta) = E((X(t_n) - F(\Delta, x; \vartheta))^2 | X_0 = x),$$

is the conditional variance of the process X. Provided that the conditional mean and variance are known, the estimators that result from the above estimating functions are consistent. However, the estimators can be biased if approximations have been made to the conditional moments. Kessler & Paredes (2002) investigate the asymptotic properties, discussed in Subsection 2.4.3, of the estimators obtained from martingale estimating functions when the conditional moments are approximated by simulation. It is shown that, provided the number of simulations used is sufficiently large and the time step in the simulation is sufficiently small, approximations of conditional moments do not greatly affect the asymptotic properties of the estimators.

Both the linear and quadratic estimating functions are examples of *polynomial estimating functions.* These are not necessarily the most effective class of martingale estimating functions. Estimating functions, in some instances, can be fitted more closely to the given type of diffusion process by using the eigenfunctions of a generator of the diffusion, see Kessler & Sørensen (1999). The martingale property of the estimating function remains by using the martingale property of the eigenfunction of the generator of diffusions, see Karlin & Taylor (1981). The resulting estimators are consistent and asymptotically normal. This holds without the assumption that the time between observations approaches zero, which

is often unrealistic. Further results on the identification of martingale estimating functions are derived in H. Sørensen (2002).

A generalisation of martingale estimating functions is given by *prediction–based estimating functions*, see Sørensen (2000). These estimating functions use unconditional moments as opposed to conditional moments, which are required by martingale estimating functions. Prediction–based estimating functions are useful for estimating non–Markovian processes, such as stochastic volatility models. Under certain regularity conditions, asymptotic normality and consistency are ensured.

2.4.2 Simple Estimating Functions

Kessler (2000) considers a *simple estimating function* of the form

$$F_{n_t}(\vartheta) = \sum_{n=1}^{n_t} f(X(t_{n-1}); \vartheta).$$

These estimating functions do not have the martingale property. The estimating function relies only on $X(t_{n-1})$, which implies that the nature of the dependence structure of the diffusion process can be difficult to characterise. Moreover, the estimating function relies heavily on the Markovian structure of the diffusion process. It is assumed that

$$\int f(x;\vartheta)\bar{p}(x;\vartheta)dx = 0,$$

where $\bar{p}(x;\vartheta)$ denotes the stationary density of an ergodic Markov process. Consistency and asymptotic normality are proven for one-dimensional diffusion processes. Kessler (2000) notes that optimality in the sense of Godambe & Heyde (1987) and Heyde (1997), characterises simple optimal estimating functions. The functions are chosen in a somewhat adhoc manner. However, Kessler (2000) suggests that when a polynomial is used for the function f, it should be of low order. H. Sørensen (2001), as mentioned previously, applies a simple estimating function to approximate the continuous time score function.

2.4.3 Asymptotics of Estimating Functions

When the estimating function is a martingale, consistency and asymptotic normality can be verified using the ergodic theorem and the central limit theorem for martingales.

We consider a martingale estimating function of the form

$$G_{n_t}(\vartheta) = \frac{1}{n_t}\sum_{n=1}^{n_t} g(\triangle, X(t_{n-1}), X(t_n); \vartheta), \tag{2.4.2}$$

for $t \geq 0$, where G_{n_t} and $g = (g_1, g_2, ..., g_p)^\top$ are p-dimensional and n_t is given in (2.1.2). We denote by ϑ_0 the true parameter value where ϑ_0 is an interior point of Θ. The true probability measure is denoted by P_{ϑ_0}. For a treatment of more general settings, see Sørensen (1999) and Bibby, Jacobsen & Sørensen (2003). Provided that X is ergodic with stationary density $\bar{p}(x;\vartheta)$ for $\vartheta \in \Theta$ and certain regularity conditions are satisfied, see Sørensen (1999), the following theorem can be proved.

Theorem 2.4.1 (Sørensen (1999)) *Under certain conditions, for every n_t, there exists an estimator $\hat{\vartheta}_{n_t}$ that solves the estimating equation $G_{n_t}(\hat{\vartheta}_{n_t}) = 0$ with a probability tending to one as $n_t \to \infty$. Moreover, we have the limit in P_{ϑ_0}-probability*

$$\lim_{n_t \to \infty} \hat{\vartheta}_{n_t} \overset{P_{\vartheta_0}}{=} \vartheta_0$$

and under P_{ϑ_0} the limit in distribution

$$\lim_{n_t \to \infty} \sqrt{n_t}(\hat{\vartheta}_{n_t} - \vartheta_0) \overset{d}{=} R,$$

where

$$R \sim N(0, A(\vartheta_0)^{-1} V_0(\vartheta_0)(A(\vartheta_0)^{-1})^\top) \tag{2.4.3}$$

is a zero mean Gaussian distributed random variable with covariance matrix $A(\vartheta_0)^{-1}V_0(\vartheta_0)(A(\vartheta_0)^{-1})^\top$ where

$$V_0(\vartheta_0) = q^{\triangle}_{\vartheta_0}(g(\triangle, \vartheta_0)g(\triangle, \vartheta_0)^\top) \tag{2.4.4}$$

and the $p \times p$ matrix

$$A(\vartheta_0) = \left\{ q^{\Delta}_{\vartheta_0} \left(\frac{\partial}{\partial \vartheta_j} g_i(\Delta, \vartheta_0) \right) \right\}^{p}_{i,j=1} \tag{2.4.5}$$

is assumed to be invertible.

This theorem ensures the existence of a $\sqrt{n_t}$–consistent estimator provided that G_{n_t} is a martingale estimating function. The proof is given in Sørensen (1999), where

$$q^{\Delta}_{\vartheta_0}(x, y) = \bar{p}(x; \vartheta_0) p(\Delta, x, y; \vartheta_0)$$

is the joint probability density of each pair of neighbouring observations. It can also be assumed that A in (2.4.5) is positive definite, however, this stronger condition is usually only satisfied by optimal estimating functions and estimating functions that are found by maximising a contrast function, see Sørensen (1999). Given that G_{n_t} is a martingale and X is ergodic, the combination of an ergodic theorem and central limit theorem for martingales is required for the proof of the above theorem, see Sørensen (1999). For completeness, we state this theorem below where the state space is the open interval (l, r) where $-\infty \leq l < r \leq \infty$. For a function $f : (\ell, r)^2 \to \mathbb{R}$ we use the notation

$$q^{\Delta}_{\vartheta_0}(f) = \int_{\ell}^{r} \int_{\ell}^{r} f(x, y)\, p(\Delta, x, y; \vartheta_0)\, \bar{p}(x; \vartheta_0)\, dy\, dx \tag{2.4.6}$$

assuming the integral exists.

Theorem 2.4.2 *Suppose the process X has a unique invariant probability measure under P_{ϑ_0} with density $\bar{p}(x; \vartheta_0)$ such that for all values of $x \in (l, r)$, the transition density $p(\Delta, x, y; \vartheta_0)$ is assumed to be zero for all those $y \in (l, r)$ for which $\bar{p}(y; \vartheta_0)$ equals zero. Furthermore, suppose that the function $f : (l, r)^2 \to \mathbb{R}$ is such that $q^{\Delta}_{\vartheta_0}(|f|) < \infty$, then*

$$\lim_{n_t \to \infty} \frac{1}{n_t} \sum_{n=1}^{n_t} f(X_{t_{n-1}}, X_{t_n}) \overset{a.s.}{=} q^{\Delta}_{\vartheta_0}(f)$$

under P_{ϑ_0}. Furthermore, suppose that $q^{\Delta}_{\vartheta_0}(f^2) < \infty$ and that

$$\int_{l}^{r} f(x, y) p(\Delta, x, y; \vartheta_0) dy = 0$$

for all $x \in (l, r)$. *Then we have the following convergence in distribution*

$$\lim_{n_t \to \infty} \frac{1}{\sqrt{n_t}} \sum_{n=1}^{n_t} f(X_{t_{n-1}}, X_{t_n}) \stackrel{d}{=} \tilde{F}$$

under P_{ϑ_0}, *where*

$$\tilde{F} \sim N(0, q^{\triangle}_{\vartheta_0}(f^2))$$

denotes a Gaussian random variable with mean zero and variance $q^{\triangle}_{\vartheta_0}(f^2)$.

In the case when the estimating function is not a martingale, additional results are required. This problem is dealt with in the following chapter. For additional details see Bibby, Jacobsen & Sørensen (2003) and references therein.

The methods discussed in this chapter are not implemented due to the nature of the benchmark model discussed in Chapter 4. However, they provide valuable tools for parameter inference in circumstances where there is no link from the diffusion to drift coefficient.

Chapter 3

Estimation of Discretely Observed Diffusions using Transform Functions

This chapter is an extended version of Kelly, Platen & Sørensen (2004). Here we seek to obtain a simple, yet general, estimation method, which provides flexibility in the estimation of discretely observed diffusion processes via the use of transform functions. The approach presented here lies somewhere between the method of contrast functions, see Section 2.3, and that of martingale estimating functions, see Subsection 2.4.1. Nevertheless, the method is based on the foundations of estimating function theory. One advantage of the method presented here is that information is not required regarding the conditional and unconditional moments. In principle, stationarity is also not required.

3.1 Transform Function for a Diffusion Process

Consider a class of one–dimensional diffusion processes defined by the following SDE

$$dX(t) = a(t, X(t); \vartheta)dt + b(t, X(t); \vartheta)dW(t) \tag{3.1.1}$$

for $t \geq 0$. The initial value $X(0) = x(0)$ is assumed to be $\mathcal{A}_0$–measurable. Here W denotes a standard Wiener process given on the filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P_\vartheta)$, where the filtration $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\geq 0}$ satisfies appropriate conditions, see Karatzas & Shreve (1991) or Jacod & Shiryaev (2003). We assume that the SDE (3.1.1) has a unique solution, see Kloeden & Platen (1999), for all parameter values ϑ in a given open subset $\Theta \subseteq \mathbb{R}^p$, $p \in \{1, 2, \ldots\}$. The drift and diffusion coefficient functions $a(\cdot, \cdot; \vartheta) : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ and $b(\cdot, \cdot; \vartheta) : [0, \infty) \times \mathbb{R} \to \mathbb{R}$,

respectively, are assumed to be known with the exception of the parameter vector $\vartheta = (\vartheta^1, \ldots, \vartheta^p)^\top \in \Theta$. Throughout the discussion, $A^\top$ denotes transposition of A.

It is our aim to estimate the unknown parameter vector ϑ from observations of the diffusion process $X = \{X(t), t \geq 0\}$. For simplicity, an equidistant time discretisation with observation times t_n, where $0 = t_0 < t_1 < \ldots < t_n < t_{n+1} < \ldots$, is assumed to be such that the time step size $\Delta = t_n - t_{n-1} \in (0, 1)$. For $t \geq 0$, the integer n_t is introduced as the largest integer n for which t_n does not exceed t. That is,

$$n_t = \max\{n \in \{0, 1, \ldots\} : t_n \leq t\} = \left[\frac{t}{\Delta}\right], \tag{3.1.2}$$

where $[x]$ denotes the integer part of the real number x, see (2.1.2).

To provide sufficient flexibility for this approach we consider, at the observation times $t_0, t_1, t_2, \ldots$, the original data

$$X(t_0), X(t_1), X(t_2), \ldots \tag{3.1.3}$$

and the transformed data

$$U(t_0, X(t_0); \lambda_i), U(t_1, X(t_1); \lambda_i), U(t_2, X(t_2); \lambda_i), \ldots \tag{3.1.4}$$

for $i \in \{1, 2, \ldots, p\}$. Here $U(\cdot, \cdot; \cdot) : [0, \infty) \times \mathbb{R} \times \Lambda \to \mathbb{R}$ is a smooth real valued function with respect to $t \in [0, \infty)$ and $x \in \mathbb{R}$, where $\Lambda \subseteq \mathbb{R}$.

The function $U(\cdot, \cdot; \lambda_i)$ for $i \in \{1, 2, \ldots, p\}$, is referred to as the *ith transform function*, and is used to transform the data in a manner that allows us to obtain estimates of the unknown parameters. In principle, for each $i \in \{1, 2, \ldots, p\}$, a different function could be used to estimate the parameters. For fixed $\lambda_i \in \Lambda$ we obtain, by the Itô formula, the following SDE for the transformed data

$$dU(t, X(t); \lambda_i) = L_\vartheta^0 U(t, X(t); \lambda_i)dt + L_\vartheta^1 U(t, X(t); \lambda_i)dW(t) \tag{3.1.5}$$

for $t \in [0, \infty)$. Here we have used the operators

$$L_\vartheta^0 u(t, x) = \left(\frac{\partial}{\partial t}u(t, x) + a(t, x; \vartheta)\frac{\partial}{\partial x}u(t, x) + \frac{1}{2}b^2(t, x; \vartheta)\frac{\partial^2}{\partial x^2}u(t, x)\right) \tag{3.1.6}$$

and

$$L^1_\vartheta u(t,x) = b(t,x;\vartheta)\frac{\partial}{\partial x}u(t,x). \tag{3.1.7}$$

For $n \in \{1,2,\ldots,n_t\}$, $i \in \{1,2,\ldots,p\}$ and $\lambda_i \in \Lambda$ we introduce the normalised difference

$$D_{\lambda_i,n,\Delta} = \frac{1}{t_n - t_{n-1}}(U(t_n, X(t_n);\lambda_i) - U(t_{n-1}, X(t_{n-1});\lambda_i)) \tag{3.1.8}$$

and the normalised squared increment

$$Q_{\lambda_i,n,\Delta} = \frac{1}{t_n - t_{n-1}}(U(t_n, X(t_n);\lambda_i) - U(t_{n-1}, X(t_{n-1});\lambda_i))^2. \tag{3.1.9}$$

By a truncated Wagner-Platen expansion, see Kloeden & Platen (1999), the increment of U in (3.1.8) and (3.1.9) can be expressed in terms of multiple stochastic integrals. Thus, we obtain

$$\begin{aligned}
&U(t_n, X(t_n);\lambda_i) - U(t_{n-1}, X(t_{n-1});\lambda_i) \\
&\quad = L^1_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)(W(t_n) - W(t_{n-1})) + L^0_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)(t_n - t_{n-1}) \\
&\quad + L^1_\vartheta L^1_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)\frac{1}{2}((W(t_n) - W(t_{n-1}))^2 - (t_n - t_{n-1})) \\
&\quad + L^0_\vartheta L^0_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)\frac{(t_n - t_{n-1})^2}{2} \\
&\quad + L^1_\vartheta L^0_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{s} dz dW(s) \\
&\quad + L^0_\vartheta L^1_\vartheta U(t_{n-1}, X(t_{n-1});\lambda_i)\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{s} dW(z) ds \\
&\quad + R_{\lambda_i,n,\vartheta}(t_n, t_{n-1}, X(t_n)),
\end{aligned} \tag{3.1.10}$$

where $R_{\lambda_i,n,\vartheta}(t_n, t_{n-1}, X(t_n))$ is the corresponding remainder term, as follows from Kloeden & Platen (1999). The expansion (3.1.10) can also be obtained by application of the Itô formula to U and then repeated to $L^0_\vartheta U$ and $L^1_\vartheta U$. The first term in (3.1.10) has mean zero and is the leading term of the expansion. Note that the third, fifth and sixth terms also have mean zero but are of a higher order than the first term. In (3.1.10) the order of the first term is $\sqrt{\Delta}$, that of the second

and third term is Δ and that of the fourth term is Δ^2. The order of the fifth and sixth term is $\Delta^{\frac{3}{2}}$. The remainder term has mean and variance of order Δ^3.

Using (3.1.8) and (3.1.9), we can construct estimating functions that exploit the structure of the first and second term of the above increment (3.1.10). To do this, we define

$$F_n(\vartheta) = (F_n^{(1)}(\vartheta)^\top, F_n^{(2)}(\vartheta)^\top)^\top$$

with $F_n^{(j)}(\vartheta)^\top = (F_{1,n}^{(j)}(\vartheta), \ldots, F_{q,n}^{(j)}(\vartheta))$, $j \in \{1, 2\}$ where

$$F_{i,n}^{(1)}(\vartheta) = D_{\lambda_i,n,\Delta} - L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_i), \tag{3.1.11}$$

$i \in \{1, 2, \ldots, q\}$ and

$$F_{i,n}^{(2)}(\vartheta) = Q_{\lambda_i,n,\Delta} - (L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i))^2, \tag{3.1.12}$$

$i \in \{1, 2, \ldots, q\}$ for $\vartheta \in \Theta$ and suitably chosen values of $\lambda_i \in \Lambda$. It is not necessary that the number of λ_i's be the same for $F_n^{(1)}(\vartheta)$ and $F_n^{(2)}(\vartheta)$ and they need not have the same value for the two functions. This assumption simplifies the exposition.

A class of estimating functions is then given by

$$K(\vartheta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} M(\vartheta)\, F_n(\vartheta), \tag{3.1.13}$$

where the $p \times 2q$ matrix valued function $M(\vartheta) = M(\vartheta, t_{n-1}, X(t_{n-1}), \Delta)$ is free to be chosen appropriately. Throughout the chapter the dependence of a weighting matrix and its elements on t_{n-1}, $X(t_{n-1})$ and Δ will be suppressed. The estimating function $K(\vartheta_0, t, \Delta)$, where ϑ_0 is taken to be the true parameter value, has expectation of order Δ. Thus, when the observation interval Δ is sufficiently small, the expectation of $K(\vartheta_0, t, \Delta)$ is approximately zero. Essentially, the approach adopted here is to approximate the conditional moments of the transformed diffusion process using the expansion in (3.1.10). This is similar to the approach in Kessler (1997), where closed form approximations for the first

two conditional moments are derived and used to construct a contrast estimator for the parameters in the drift and diffusion functions.

The estimating function given in (3.1.13) is slightly biased. To determine the optimal weighting matrix $M^*(\vartheta)$, we consider the unbiased estimating function $K^\circ(\vartheta, t, \Delta)$ obtained by compensating $K(\vartheta, t, \Delta)$ such that

$$K^\circ(\vartheta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} M(\vartheta) \left(F_n(\vartheta) - \bar{F}_n(\vartheta)\right). \tag{3.1.14}$$

Here $\bar{F}_n(\vartheta) = E_\vartheta(F_n(\vartheta)|X(t_{n-1}))$ is the compensator for $F_n(\vartheta)$ and is of order Δ. The optimal choice for the weighting matrix $M(\vartheta)$ in the unbiased estimating equation (3.1.14), in the sense of Godambe & Heyde (1987), can be derived using the method outlined in Heyde (1997). For details regarding the case of diffusion processes, see Sørensen (1997). Here the optimal weighting matrix is given by

$$M^*(\vartheta) = B^*(\vartheta)\, V^*(\vartheta)^{-1}, \tag{3.1.15}$$

where $V^*(\vartheta)$ is the $2q \times 2q$ conditional covariance matrix

$$V^*(\vartheta) = V^*(\vartheta, t_{n-1}, X(t_{n-1})) = E_\vartheta\left((F_n(\vartheta) - \bar{F}_n(\vartheta))(F_n(\vartheta) - \bar{F}_n(\vartheta))^\top \middle| X(t_{n-1})\right)$$

and $B^*(\vartheta) = (B^{*(1)}(\vartheta), B^{*(2)}(\vartheta))$ with $B^{*(k)}(\vartheta)$, $k \in \{1, 2\}$, denoting the $p \times q$ matrix where the (i,j)th entry is

$$B^{*(k)}_{i,j}(\vartheta) = B^{*(k)}(\vartheta, t_{n-1}, X(t_{n-1}))_{i,j} = E_\vartheta\left(\frac{\partial}{\partial \vartheta^i}\left[F^{(k)}_{j,n}(\vartheta) - \bar{F}^{(k)}_{j,n}(\vartheta)\right] \middle| X(t_{n-1})\right).$$

The value for each λ_i should be chosen in such a way that the conditional covariance matrix $V^*(\vartheta)$ is invertible.

Keeping only the leading terms, we obtain

$$\tilde{K}(\vartheta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} \tilde{M}(\vartheta)\, F_n(\vartheta), \tag{3.1.16}$$

where $\tilde{M}(\vartheta) = B(\vartheta)V(\vartheta)^{-1}$. Here $B(\vartheta) = (B^{(1)}(\vartheta),\ B^{(2)}(\vartheta))$, where the (i,j)th entry of the $p \times q$ matrices $B^{(1)}(\vartheta)$ and $B^{(2)}(\vartheta)$ are

$$B^{(1)}(\vartheta)_{i,j} = \frac{\partial}{\partial \vartheta^i} L^0_\vartheta\, U(t_{n-1}, X(t_{n-1}); \lambda_j) \tag{3.1.17}$$

and

$$B^{(2)}(\vartheta)_{i,j} = \frac{\partial}{\partial \vartheta^i} \left(L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \right)^2, \tag{3.1.18}$$

respectively. Moreover,

$$V(\vartheta) = \left\{ \begin{array}{cc} V^{11}(\vartheta) & V^{12}(\vartheta) \\ V^{21}(\vartheta) & V^{22}(\vartheta) \end{array} \right\},$$

where the (i,j)th entry of the $q \times q$ matrices $V^{11}(\vartheta)$, $V^{22}(\vartheta)$ and $V^{12}(\vartheta)$ are

$$V^{11}(\vartheta)_{i,j} = \frac{1}{\Delta} L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) \, L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j), \tag{3.1.19}$$

$$V^{22}(\vartheta)_{i,j} = 2 \left[L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \right]^2, \tag{3.1.20}$$

and

$$\begin{aligned} & V^{12}(\vartheta)_{i,j} \qquad (3.1.21) \\ & = 2 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \, L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_j) \\ & + 2 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \, L_\vartheta^1 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \\ & + L_\vartheta^1 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) \left[L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \right]^2, \end{aligned}$$

while $V^{21}(\vartheta) = (V^{12}(\vartheta))^\top$.

The weighting matrix $\tilde{M}(\vartheta)$ in (3.1.16) is fixed sample optimal, see Godambe & Heyde (1987) and Heyde (1997). The weighting matrix results in an estimating function (3.1.16), that is, to the order of approximation used, closest within the class of estimating functions of the form (3.1.13) to the corresponding, usually unknown, score function. For example, if the diffusion process is ergodic, it can be proved that a fixed sample optimal martingale estimating function is also asymptotically optimal or Heyde–optimal, see Heyde (1997). The estimating functions proposed in this chapter are approximations of martingale estimating functions to the order Δ.

For given transform functions U, with parameters $\lambda_i \in \Lambda$, $i \in \{1, ..., q\}$, a p-dimensional estimating equation is obtained,

$$\tilde{K}(\vartheta, t, \Delta) = 0$$

for $t \geq t_1$, see (3.1.16). Assuming that the resulting system of p equations has a unique solution, we obtain for the particular SDE (3.1.1), an estimator $\hat{\vartheta} = (\hat{\vartheta}^1, ..., \hat{\vartheta}^p)^\top$ for the parameter vector ϑ. Note that the vector of estimators, $\hat{\vartheta}$, depends on t, Δ, $\lambda_1, ..., \lambda_p$ and the observed data. Appropriate values of λ_i for $i \in \{1, 2, ..., p\}$ can be found by exploiting asymptotic properties of the estimating functions as described in the next section. The choice of the λ_i for $i \in \{1, 2, ..., p\}$ determines the resulting system of equations.

A simpler, although less efficient, estimation procedure can be used when the parameter ϑ is written as $\vartheta = (\varrho, \zeta)^\top$. Here it is assumed that the p_1-dimensional parameter ϱ appears only in the drift coefficient, while the diffusion coefficient depends only on the p_2-dimensional parameter ζ. In this case we first estimate ζ by solving $\tilde{H}(\hat{\zeta}, t, \Delta) = 0$, where

$$\tilde{H}(\zeta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} B^{(2)}(\zeta)\, V^{22}(\zeta)^{-1} F_n^{(2)}(\zeta) \tag{3.1.22}$$

with the $p_2 \times q$ matrix $B^{(2)}(\zeta)$ given by

$$B^{(2)}(\zeta)_{i,j} = \frac{\partial}{\partial \zeta^i} \left(L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_j) \right)^2, \tag{3.1.23}$$

and $V^{22}(\zeta) = V^{22}(\vartheta)$ given by (3.1.20). Note that $\tilde{H}(\zeta, t, \Delta)$ does not depend on ϱ. Next ϱ can be estimated by solving $\tilde{G}(\hat{\varrho}, \hat{\zeta}, t, \Delta) = 0$, where $\hat{\zeta}$ is the estimator of ζ previously obtained, and

$$\tilde{G}(\varrho, \zeta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} B^{(1)}(\varrho, \zeta) V^{11}(\varrho, \zeta)^{-1} F_n^{(1)}(\varrho, \zeta) \tag{3.1.24}$$

with the $p_1 \times q$ matrix $B^{(1)}(\varrho, \zeta)$ given by

$$B^{(1)}(\varrho, \zeta)_{i,j} = \frac{\partial}{\partial \varrho^i} L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_j). \tag{3.1.25}$$

and $V^{11}(\varrho, \zeta) = V^{11}(\vartheta)$ given by (3.1.19). The estimating functions $\tilde{G}(\varrho, \zeta, t, \Delta)$ and $\tilde{H}(\zeta, t, \Delta)$ are, to the order of approximation used, optimal within the classes of estimating functions

$$G(\varrho, \zeta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} M^{(1)}(\varrho, \zeta)\, F_n^{(1)}(\varrho, \zeta) \tag{3.1.26}$$

and

$$H(\zeta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} M^{(2)}(\zeta) F_n^{(2)}(\zeta) \tag{3.1.27}$$

for the estimation of ϱ and ζ, respectively. The optimal martingale estimating function for the form (3.1.14) is the optimal combination of the optimal martingale estimating functions, which (3.1.22) and (3.1.24) approximate, see Heyde (1997) and Bibby (1994).

3.2 Asymptotics

The proposed transform function method is designed to encompass both stationary and nonstationary diffusion processes. Despite this, it is advantageous to analyse the asymptotic behaviour and bias of the parameter estimators for some given class of real valued diffusion processes. We assume in this section that X is ergodic and described by the SDE (3.1.1) with time homogeneous coefficient functions

$$a(t, x; \vartheta) = a(x; \vartheta) \tag{3.2.1}$$

and

$$b(t, x; \vartheta) = b(x; \vartheta) \tag{3.2.2}$$

for $t \geq 0$, $x \in \mathbb{R}$ and $\vartheta \in \Theta$. We use as state space the interval (ℓ, r) where $-\infty \leq \ell < r \leq \infty$. For given parameter vector $\vartheta \in \Theta$, the density of the scale measure $s : (\ell, r) \to [0, \infty)$ is given by the expression

$$s(x; \vartheta) = \exp\left(-2 \int_{y_0}^{x} \frac{a(y; \vartheta)}{b^2(y; \vartheta)} dy \right) \tag{3.2.3}$$

for $x \in (\ell, r)$ with some reference value $y_0 \in (\ell, r)$.

If the following two conditions

$$\int_{y_0}^{r} s(x; \vartheta)\, dx = \int_{\ell}^{y_0} s(x; \vartheta)\, dx = \infty \tag{3.2.4}$$

and

$$\int_{\ell}^{r} \frac{1}{s(x; \vartheta) b^2(x; \vartheta)} dx < \infty \tag{3.2.5}$$

are satisfied, then it is well known that X is ergodic with stationary density

$$\bar{p}(x;\vartheta) = \frac{C(\vartheta)}{b^2(x;\vartheta)} \exp\left(2\int_{y_0}^{x} \frac{a(u;\vartheta)}{b^2(u;\vartheta)} du\right) \tag{3.2.6}$$

for $x \in (\ell, r)$ and $\vartheta \in \Theta$. The constant $C(\vartheta)$ results from the normalisation condition

$$\int_{\ell}^{r} \bar{p}(x;\vartheta) dx = 1. \tag{3.2.7}$$

To prove the existence, consistency and asymptotic normality of the estimators we introduce the following conditions and notation where, essentially, we follow Sørensen (1999). We denote by ϑ_0 the true parameter value, where ϑ_0 is an interior point of Θ. The true probability measure is denoted by P_{ϑ_0}. Additionally, let $p(\Delta, x, y; \vartheta_0)$ be the true transition density of the observed diffusion process X for a transition from x to y over a time period of length $\Delta > 0$. Throughout the remainder of this section we take Δ to be fixed. We consider estimating functions of the form

$$G_t(\vartheta) = \frac{1}{n_t} \sum_{n=1}^{n_t} g(\Delta, X(t_{n-1}), X(t_n); \vartheta), \tag{3.2.8}$$

for $t \geq 0$ and where G_t and $g = (g_1, g_2, \ldots, g_p)^\top$ are p-dimensional. Furthermore, we assume that X is stationary and impose the condition that X is geometrically α–mixing. For a definition of this concept, see, for instance, Doukhan (1994). For a given one–dimensional, ergodic diffusion process X, there are a number of relatively simple criteria ensuring α-mixing with exponentially decreasing mixing coefficients. We cite the following straightforward and rather weak set of conditions used in Genon-Catalot, Jeantheau & Laredo (2000) on the coefficients a and b that are sufficient to ensure geometric α-mixing of X.

Condition 3.2.1

(i) The function a is continuously differentiable and b is twice continuously differentiable with respect to $x \in (\ell, r)$, $b(x;\vartheta_0) > 0$ for all $x \in (\ell, r)$, and there exists a constant $K > 0$ such that $|a(x;\vartheta_0)| \leq K(1 + |x|)$ and $b^2(x;\vartheta_0) \leq K(1 + |x|^2)$ for all $x \in (\ell, r)$.

(ii) $b(x;\vartheta_0)\bar{p}(x,\vartheta_0) \to 0$ *as* $x \downarrow \ell$ *and* $x \uparrow r$.

(iii) $\frac{1}{\gamma(x;\vartheta_0)}$ *has a finite limit as* $x \downarrow \ell$ *and* $x \uparrow r$, *where*

$$\gamma(x;\vartheta_0) = \partial_x b(x;\vartheta_0) - \frac{2a(x;\vartheta_0)}{b(x;\vartheta_0)}.$$

Each pair of neighboring observations $(X(t_{n-1}), X(t_n))$ has the joint probability density

$$q^{\Delta}_{\vartheta_0}(x,y) = \bar{p}(x;\vartheta_0)\, p(\Delta,x,y;\vartheta_0) \tag{3.2.9}$$

on $(\ell,r)^2$. For a function $f : (\ell,r)^2 \to \mathbb{R}^k$, $k \in \{1,2,\ldots\}$, where we assume that the following integral exists, we introduce the vector valued functional

$$q^{\Delta}_{\vartheta_0}(f) = \int_\ell^r \int_\ell^r f(x,y)\, p(\Delta,x,y;\vartheta_0)\, \bar{p}(x;\vartheta_0)\, dy\, dx. \tag{3.2.10}$$

For our purposes, we cannot assume that the estimating function in (3.2.8) is unbiased. Instead, we make the following assumption.

Condition 3.2.2 *There exists a unique parameter value* $\bar{\vartheta}$ *that is an interior point of* Θ *such that*

$$q^{\Delta}_{\vartheta_0}\left(g(\Delta,\bar{\vartheta})\right) = 0. \tag{3.2.11}$$

Note that the equation in Condition 3.2.2 is a vector equation with the zero-vector on the right hand side. We can now impose conditions on the estimating function (3.2.8), that are similar to those used by Barndorff-Nielsen & Sørensen (1994) and Sørensen (1999).

Condition 3.2.3

(i) The function $g_i(\Delta,x,y;\cdot) : \Theta \to \mathbb{R}$ *is twice continuously differentiable with respect to* $\vartheta \in \Theta$ *for all* $x,y \in (\ell,r)$, *and* $i \in \{1,2,\ldots,p\}$.

(ii) The function $g_i(\Delta,\cdot,\cdot;\vartheta) : (\ell,r) \times (\ell,r) \to \mathbb{R}$ *is such that there exists a* $\delta > 0$ *with* $q^{\Delta}_{\vartheta_0}(g_i(\Delta,\vartheta)^{2+\delta}) < \infty$ *for all* $\vartheta \in \Theta$ *and* $i \in \{1,2,\ldots,p\}$.

(iii) For the partial derivatives

$$\frac{\partial}{\partial \vartheta^j} g_i(\Delta, x, y; \vartheta) \quad \text{and} \quad \frac{\partial^2}{\partial \vartheta^j \partial \vartheta^k} g_i(\Delta, x, y; \vartheta), \tag{3.2.12}$$

where $i, j, k \in \{1, 2, \ldots, p\}$, *there exists for every* $\vartheta^* \in \Theta$ *a neighborhood* $\mathcal{N}(\vartheta^*) \subset \Theta$ *of* ϑ^* *and a non-negative random variable* $L(\vartheta^*)$ *with* $E_{\vartheta_0}(L(\vartheta^*)) < \infty$ *such that*

$$\left| \frac{\partial}{\partial \vartheta^j} g_i(\Delta, x, y; \vartheta) \right| \leq L(\vartheta^*) \quad \text{and} \quad \left| \frac{\partial^2}{\partial \vartheta^j \partial \vartheta^k} g_i(\Delta, x, y; \vartheta) \right| \leq L(\vartheta^*) \tag{3.2.13}$$

for all $\vartheta \in \mathcal{N}(\vartheta^*)$, $(x, y) \in (\ell, r)^2$, *and* $i, j, k \in \{1, 2, \ldots, p\}$.

(iv) The $p \times p$ *matrix*

$$A(\vartheta_0, \bar{\vartheta}) = \left\{ q_{\vartheta_0}^{\Delta} \left(\frac{\partial}{\partial \vartheta^j} g_i(\Delta, \cdot, \cdot; \bar{\vartheta}) \right) \right\}_{i,j=1}^{p} \tag{3.2.14}$$

is invertible.

Theorem 3.2.4 *Suppose Conditions 3.2.2 and 3.2.3 are satisfied. Then for every* $t > \Delta$, *there exists an estimator* $\hat{\vartheta}_{n_t}$ *that solves the system of estimating equations*

$$G_t(\hat{\vartheta}_{n_t}) = 0$$

with a probability tending to one as $n_t \to \infty$. *Moreover, we have the limit in* P_{ϑ_0}*-probability*

$$\lim_{n_t \to \infty} \hat{\vartheta}_{n_t} \overset{P_{\vartheta_0}}{=} \bar{\vartheta} \tag{3.2.15}$$

and under P_{ϑ_0} *the limit in distribution*

$$\lim_{n_t \to \infty} \sqrt{n_t}(\hat{\vartheta}_{n_t} - \bar{\vartheta}) \overset{d}{=} R, \tag{3.2.16}$$

where

$$R \sim N\left(0, A(\vartheta_0, \bar{\vartheta})^{-1} v(\vartheta_0, \bar{\vartheta}) (A(\vartheta_0, \bar{\vartheta})^{-1})^{\top}\right) \tag{3.2.17}$$

is a p-dimensional, zero mean Gaussian distributed random variable with covariance matrix $A(\vartheta_0, \bar{\vartheta})^{-1} v(\vartheta_0, \bar{\vartheta}) (A(\vartheta_0, \bar{\vartheta})^{-1})^{\top}$, *where*

$$\begin{aligned} v(\vartheta_0, \bar{\vartheta}) &= q_{\vartheta_0}^{\Delta} \left(g(\Delta, \bar{\vartheta}) g(\Delta, \bar{\vartheta})^{\top} \right) \\ &+ \sum_{k=1}^{\infty} \Big\{ E_{\vartheta_0} \left(g(\Delta, X_{t_0}, X_{t_1}; \bar{\vartheta})\, g(\Delta, X_{t_k}, X_{t_{k+1}}; \bar{\vartheta})^{\top} \right) \\ &\qquad + E_{\vartheta_0} \left(g(\Delta, X_{t_k}, X_{t_{k+1}}; \bar{\vartheta})\, g(\Delta, X_{t_0}, X_{t_1}; \bar{\vartheta})^{\top} \right) \Big\}. \end{aligned} \tag{3.2.18}$$

Under the conditions imposed, the covariances in the infinite sum in (3.2.18) tend to zero exponentially fast as $k \to \infty$, so the sum converges quickly and can usually be well approximated by a finite sum with relatively few terms.

The theorem can be proved in complete analogy with the proof of Theorem 3.6 in Sørensen (1999). The only difference is that the martingale limit theory used in that paper must be replaced by results for α-mixing processes because the estimating functions considered here are not martingales. Let us briefly outline the necessary limit results. By the ergodic theorem, one has

$$\lim_{n_t \to \infty} \frac{1}{n_t} \sum_{n=1}^{n_t} g(\Delta, X(t_{n-1}), X(t_n); \vartheta) \stackrel{a.s.}{=} q^{\Delta}_{\vartheta_0}(g(\Delta, \vartheta)). \tag{3.2.19}$$

Note, from Condition 3.2.2 the limit is zero for $\vartheta = \bar{\vartheta}$.

Under the conditions imposed, we have the following central limit theorem:

$$\lim_{n_t \to \infty} \frac{1}{\sqrt{n_t}} \sum_{n=1}^{n_t} g(\Delta, X(t_{n-1}), X(t_n); \bar{\vartheta}) \stackrel{d}{=} \tilde{F} \tag{3.2.20}$$

under P_{ϑ_0}, where

$$\tilde{F} \sim N(0, v(\vartheta_0, \bar{\vartheta})) \tag{3.2.21}$$

denotes a p-dimensional Gaussian random variable with mean zero and covariance matrix $v(\vartheta_0, \bar{\vartheta})$ given by (3.2.18), provided that the matrix $v(\vartheta_0, \bar{\vartheta})$ is strictly positive definite. This follows from Theorem 1 in Section 1.5 of Doukhan (1994) by application of the Cramér-Wold device. The condition that X is geometrically α-mixing is actually stronger than what is required for the central limit theorem to hold. Minimal, but more technical conditions can be found in Doukhan, Massart & Rio (1994).

It is clearly desirable to use estimating functions for which $\bar{\vartheta}$ is close to the true parameter value ϑ_0. For the estimating function $\tilde{G}$ given by (3.1.24), it follows from (3.1.10), that the leading term in an expansion in powers of Δ of $q^{\Delta}_{\vartheta_0}(g(\Delta, \vartheta))$, from Condition 3.2.2, is $\frac{1}{2}\Delta m(\vartheta_0, \vartheta)$, where

$$m(\vartheta_0, \vartheta) = E_{\vartheta_0}\left(B^{(1)}(\varrho, \zeta)\, V^{11}(\varrho, \zeta)^{-1}\, S_n(\vartheta)\right), \tag{3.2.22}$$

given

$$S_n(\vartheta)_i = L_\vartheta^0 \, L_\vartheta^0 \, U(t_{n-1}, X(t_{n-1}); \lambda_i).$$

We could choose the transform function U and the value of λ_i to make $|m(\vartheta_0, \vartheta_0)|$ as small as possible in an attempt to minimize the bias of the estimating function. However, we can expect to achieve a good approximation to $\bar{\vartheta}$ by solving the equation $m(\vartheta_0, \vartheta) = 0$ with respect to ϑ. By an expansion of m in ϑ around ϑ_0, we find that

$$\vartheta_0 - \bar{\vartheta} \simeq \left(\frac{\partial m(\vartheta_0, \vartheta_0)}{\partial \vartheta^\top} \right)^{-1} m(\vartheta_0, \vartheta_0). \tag{3.2.23}$$

Here $\frac{\partial m(\vartheta_0, \vartheta)}{\partial \vartheta^\top}$ denotes the $p \times p$-matrix, where the (i,j)th entry is $\frac{\partial m_i(\vartheta_0, \vartheta)}{\partial \vartheta^j}$. To reduce the distance between $\bar{\vartheta}$ and ϑ_0, it therefore seems appropriate to choose the transform function U and the values of $\lambda_1, \ldots, \lambda_p$ in a way that makes the right-hand side of (3.2.23) as small as possible. The estimating functions $\tilde{H}$ and $\tilde{K}$ can be treated in a similar manner. However, the resulting expressions for $m(\vartheta_0, \vartheta)$ are more complicated.

A different asymptotic scenario could have been considered, namely, that Δ goes to zero sufficiently fast as n_t tends to infinity, see, for instance, Prakasa Rao (1988) and Kessler (1997). In such a scenario the estimators proposed here would be consistent. However, the kind of asymptotics studied in this section illustrate more clearly the advantage of the transform function approach.

3.3 Affine Diffusions

We now introduce a specific class of *affine diffusions* that will aid us in highlighting the features of the proposed methodology. Consider the affine SDE for the *shifted square root process*

$$dX_t = (\vartheta^1 + \vartheta^2 X(t))dt + \sqrt{\vartheta^3 + \vartheta^4 X(t)} dW(t) \tag{3.3.1}$$

for $t \geq 0$, where both the drift function $a(t, x; \vartheta) = \vartheta^1 + \vartheta^2 x$ and squared diffusion coefficient function $b^2(t, x; \vartheta) = \vartheta^3 + \vartheta^4 x$ are affine. In the following, the

parameter vector $\vartheta = (\vartheta^1, \vartheta^2, \vartheta^3, \vartheta^4)^\top \in \mathbb{R}^4$ shall be chosen such that the process $X = \{X(t), t \geq 0\}$ is ergodic. This happens when either

$$\vartheta^4 = 0, \qquad \vartheta^2 < 0 \qquad \text{and} \qquad \vartheta^3 > 0 \tag{3.3.2}$$

or

$$\vartheta^4 > 0, \qquad \vartheta^2 < 0 \qquad \text{and} \qquad \frac{2}{\vartheta^4}\left(\vartheta^1 - \frac{\vartheta^2\vartheta^3}{\vartheta^4}\right) \geq 1. \tag{3.3.3}$$

In the first case, the Ornstein-Uhlenbeck process, the process X lives on the whole real line and the stationary distribution is Gaussian with mean $-\frac{\vartheta^1}{\vartheta^2}$ and variance $-\frac{\vartheta^3}{(2\vartheta^2)}$. In the latter case, the process X lives on the interval $E = (y_0, \infty)$ with $y_0 = -\frac{\vartheta^3}{\vartheta^4}$. The stationary density for such an ergodic affine diffusion is of the form

$$\bar{p}(x) = \frac{\left(\frac{-2\vartheta^2}{\vartheta^4}\right)^{\frac{2}{\vartheta^4}\vartheta^1 - \frac{\vartheta^2\vartheta^3}{\vartheta^4}} \left(x + \frac{\vartheta^3}{\vartheta^4}\right)^{\frac{2}{\vartheta^4}\vartheta^1 - \frac{\vartheta^2\vartheta^3}{\vartheta^4} - 1} \exp\left(\frac{2\vartheta^2}{\vartheta^4}\left(x + \frac{\vartheta^3}{\vartheta^4}\right)\right)}{\Gamma\left(\frac{2}{\vartheta^4}\left(\vartheta^1 - \frac{\vartheta^2\vartheta^3}{\vartheta^4}\right)\right)} \tag{3.3.4}$$

for $x \in E = (y_0, \infty)$, where $\Gamma(\cdot)$ denotes the Gamma function. In this case the stationary mean is

$$\int_{y_0}^{\infty} x\bar{p}(x)dx = -\frac{\vartheta^1}{\vartheta^2} \tag{3.3.5}$$

and the stationary second moment has the form

$$\int_{y_0}^{\infty} x^2\bar{p}(x)dx = -\frac{(2\vartheta^1 + \vartheta^4)\vartheta^1 - \vartheta^3\vartheta^2}{2(\vartheta^2)^2}. \tag{3.3.6}$$

Note that the stationary density in (3.3.4) is a shifted Gamma distribution.

3.4 Power Transform Functions

To illustrate the transform function method for affine diffusions we need to specify a class of transform functions. Let us consider the *power transform function*, which is one of the most tractable transforms. We set

$$U(t, x; \lambda) = x^\lambda \tag{3.4.1}$$

for $t \in [0,\infty)$, $x \in (\ell, r)$ with $\lambda > 0$. Setting $\varrho = (\vartheta^1, \vartheta^2)$ and $\zeta = (\vartheta^3, \vartheta^4)$ whilst letting $M^{(1)}(\varrho, \zeta)$ and $M^{(2)}(\zeta)$ be the identity matrix, we obtain from (3.1.26) and (3.1.27), the estimating functions

$$H_i(\zeta, t, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} F_{i,n}^{(2)}(\zeta) \tag{3.4.2}$$

with

$$F_{i,n}^{(2)}(\zeta) = \left(Q_{\lambda_i,n,\Delta} - (L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i))^2\right) \tag{3.4.3}$$

for $i \in \{3,4\}$ and

$$G_i(\varrho, \zeta, \Delta) = \frac{1}{n_t} \sum_{n=1}^{n_t} F_{i,n}^{(1)}(\varrho, \zeta), \tag{3.4.4}$$

where

$$F_{i,n}^{(1)}(\varrho, \zeta) = (D_{\lambda_i,n,\Delta} - L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_i)) \tag{3.4.5}$$

for $i \in \{1,2\}$. For the affine diffusions we have by (3.1.6)

$$\begin{aligned} L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_i) &= (\vartheta^1 + \vartheta^2 X(t_{n-1}))\lambda_i X(t_{n-1})^{\lambda_i - 1} \\ &\quad + \frac{1}{2}(\vartheta^3 + \vartheta^4 X(t_{n-1}))\lambda_i(\lambda_i - 1) X(t_{n-1})^{\lambda_i - 2} \end{aligned} \tag{3.4.6}$$

and by (3.1.7)

$$L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) = \lambda_i X(t_{n-1})^{\lambda_i - 1} \sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}. \tag{3.4.7}$$

We obtain from (3.1.9), (3.1.27), (3.4.2) and (3.4.7) the estimating function

$$H_i(\zeta, t, \Delta) = \mathrm{A}_\Delta^{0,1,0}(\lambda_i) - \vartheta^3(\lambda_i)^2 \mathrm{A}_\Delta^{2(\lambda_i - 1)} - \vartheta^4(\lambda_i)^2 \mathrm{A}_\Delta^{2\lambda_i - 1} \tag{3.4.8}$$

for $i \in \{3,4\}$ for two different values λ_3, $\lambda_4 > 0$. Here we have used the notation

$$\mathrm{A}_\Delta^{r,k,j}(\lambda_i) = \frac{1}{n_t} \sum_{n=1}^{n_t} (X(t_{n-1}))^r (Q_{\lambda_i,n,\Delta})^k (D_{\lambda_i,n,\Delta})^j \tag{3.4.9}$$

and $\mathrm{A}_\Delta^r = \mathrm{A}_\Delta^{r,0,0}(\lambda_i)$, which refers to an equidistant time discretisation of step size Δ.

Similarly, we obtain from (3.1.8), (3.1.26), (3.4.4) and (3.4.6) the estimating function

$$G_i(\varrho, \zeta, t, \Delta) = \mathrm{C}_\Delta(\lambda_i, \vartheta^3, \vartheta^4) - \vartheta^1 \lambda_i \mathrm{A}_\Delta^{\lambda_i - 1} - \vartheta^2 \lambda_i \mathrm{A}_\Delta^{\lambda_i} \tag{3.4.10}$$

with

$$C_\Delta(\lambda_i, \vartheta^3, \vartheta^4) = A_\Delta^{0,0,1}(\lambda_i) - \frac{\lambda_i(\lambda_i - 1)}{2}(\vartheta^3 A_\Delta^{\lambda_i - 2} + \vartheta^4 A_\Delta^{\lambda_i - 1}) \tag{3.4.11}$$

for $i \in \{1, 2\}$ for two different values λ_1, $\lambda_2 > 0$.

It follows from (3.1.10), properties of multiple stochastic integrals, see Kloeden & Platen (1999), and the existence of all moments of positive order for X that

$$\begin{aligned}
\lim_{n_t \to \infty} E(H_i(\zeta, t, \Delta)) = \lim_{n_t \to \infty} E\Bigg(& \frac{1}{n_t} \sum_{n=1}^{n_t} \frac{1}{\Delta} \Big[(L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_i))^2 \Delta^2 \\
& + (L_\vartheta^1 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i))^2 \frac{\Delta^2}{2} \\
& + L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) L_\vartheta^1 L_\vartheta^0 U(t_{n-1}, X(t_{n-1}); \lambda_i) \frac{\Delta^2}{2} \\
& + L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) L_\vartheta^0 L_\vartheta^1 U(t_{n-1}, X(t_{n-1}); \lambda_i) \frac{\Delta^2}{2} \Big] \Bigg) \\
& + \Delta^2 E(R_1(\zeta, t, \Delta, \lambda_i)) \\
= \Delta \int_{y_0}^{\infty} \Bigg(& (L_\vartheta^0 U(1, y; \lambda_i))^2 + \frac{1}{2} (L_\vartheta^1 L_\vartheta^1 U(1, y; \lambda_i))^2 \\
& + \frac{1}{2} L_\vartheta^1 U(1, y; \lambda_i) L_\vartheta^1 L_\vartheta^0 U(1, y; \lambda_i) \\
& + \frac{1}{2} L_\vartheta^1 U(1, y; \lambda_i) L_\vartheta^0 L_\vartheta^1 U(1, y; \lambda_i) \Bigg) \bar{p}(y) dy \\
& + \Delta^2 E(R_2(\zeta, 1, \Delta, \lambda_i)).
\end{aligned} \tag{3.4.12}$$

Here $\bar{p}(y)$ is given by (3.3.4), and $E(R_j(\zeta, t, \Delta, \lambda_i))$, for $j \in \{1, 2\}$ and $i \in \{3, 4\}$ are finite. Similarly, we obtain

$$\begin{aligned}
& \lim_{n_t \to \infty} E(G_i(\varrho, \zeta, t, \Delta)) = \\
& \Delta \int_{y_0}^{\infty} \left(\frac{1}{2} (L_\vartheta^0 L_\vartheta^0 U(1, y; \lambda_i)) \right) \bar{p}(y)\, dy + \Delta^{\frac{3}{2}} E(R_4(\varrho, \zeta, 1, \Delta, \lambda_i)),
\end{aligned} \tag{3.4.13}$$

where $E(R_4(\varrho, \zeta, t, \Delta, \lambda_i))$, for $i \in \{1, 2\}$ are finite.

Note that if the time between observations Δ tends to zero, then the expectation of the functions H_i and G_i will approach zero. By setting the estimating functions to zero, we obtain the linear system of four estimating equations

$$0 = A_\Delta^{0,1,0}(\lambda_i) - \hat{\vartheta}^3(\lambda_i)^2 A_\Delta^{2(\lambda_i - 1)} - \hat{\vartheta}^4(\lambda_i)^2 A_\Delta^{2\lambda_i - 1} \tag{3.4.14}$$

for $i \in \{3,4\}$ and

$$0 = \mathrm{C}_\Delta(\lambda_i, \hat{\vartheta}^3, \hat{\vartheta}^4) - \hat{\vartheta}^1 \lambda_i \mathrm{A}_\Delta^{\lambda_i - 1} - \hat{\vartheta}^2 \lambda_i \mathrm{A}_\Delta^{\lambda_i} \tag{3.4.15}$$

for $i \in \{1,2\}$.

As discussed previously, the values of λ_i for $i \in \{1,2,3,4\}$, should ideally be chosen such that the estimator bias is minimized. Here we choose them based on a more practical consideration. Intuitively, the values chosen for the λ_i, $i \in \{1,2,3,4\}$ should remain small, since large values of λ_i would result in transform functions that produce unstable estimating functions with terms that may increase rapidly over time. Furthermore, simple explicit solutions of the system of equations (3.4.14) and (3.4.15) can be obtained by choosing small integer values. A convenient choice is $\lambda_1 = \lambda_3 = 1$, $\lambda_2 = \lambda_4 = 2$. Using these values for λ_i for $i \in \{1,2,3,4\}$ we obtain the following four equations for the estimators,

$$\begin{aligned}
\hat{\vartheta}^1 + \mathrm{A}_\Delta^1 \hat{\vartheta}^2 &= \mathrm{C}_\Delta(1, \hat{\vartheta}^3, \hat{\vartheta}^4) \\
\mathrm{A}_\Delta^1 \hat{\vartheta}^1 + \mathrm{A}_\Delta^2 \hat{\vartheta}^2 &= \frac{1}{2}\mathrm{C}_\Delta(2, \hat{\vartheta}^3, \hat{\vartheta}^4) \\
\hat{\vartheta}^3 + \mathrm{A}_\Delta^1 \hat{\vartheta}^4 &= \mathrm{A}_\Delta^{0,1,0}(1) \\
\mathrm{A}_\Delta^2 \hat{\vartheta}^3 + \mathrm{A}_\Delta^3 \hat{\vartheta}^4 &= \frac{1}{4}\mathrm{A}_\Delta^{0,1,0}(2).
\end{aligned} \tag{3.4.16}$$

This system has the explicit solution

$$\begin{aligned}
\hat{\vartheta}^1 &= \mathrm{C}_\Delta(1, \hat{\vartheta}^3, \hat{\vartheta}^4) - \mathrm{A}_\Delta^1 \hat{\vartheta}^2 \\
\hat{\vartheta}^2 &= \frac{\frac{1}{2}\mathrm{C}_\Delta(2, \hat{\vartheta}^3, \hat{\vartheta}^4) - \mathrm{A}_\Delta^1 \mathrm{C}_\Delta(1, \hat{\vartheta}^3, \hat{\vartheta}^4)}{\mathrm{A}_\Delta^2 - (\mathrm{A}_\Delta^1)^2} \\
\hat{\vartheta}^3 &= \mathrm{A}_\Delta^{0,1,0}(1) - \mathrm{A}_\Delta^1 \hat{\vartheta}^4 \\
\hat{\vartheta}^4 &= \frac{\frac{1}{4}\mathrm{A}_\Delta^{0,1,0}(2) - \mathrm{A}_\Delta^2 \mathrm{A}_\Delta^{0,1,0}(1)}{\mathrm{A}_\Delta^3 - \mathrm{A}_\Delta^1 \mathrm{A}_\Delta^2}.
\end{aligned} \tag{3.4.17}$$

Here we have derived explicit expressions for the estimators of the given class of affine diffusions using power transform functions. The illustrated transform function method can be extended to other classes of diffusions including non–ergodic and multi-dimensional diffusions. If no explicit solution of the system

of estimating equations is available, then numerical methods can be applied to identify their solution.

3.5 Example

To illustrate the practical applicability of the proposed transform function method, we consider an example for the above affine diffusion given in (3.3.1). Sample paths of this process are simulated using two different simulation schemes. Firstly, a *Wagner-Platen order* 1.5 *strong scheme*, see Kloeden & Platen (1999), is used. The scheme has the form

$$\begin{aligned} X(t_n) &= X(t_{n-1}) + (\vartheta^1 + \vartheta^2 X(t_{n-1}))\Delta + \sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}\Delta W_n \\ &+ \frac{1}{2}\vartheta^2(\vartheta^1 + \vartheta^2 X(t_{n-1}))\Delta^2 + \frac{1}{4}\vartheta^4(\Delta W_n^2 - \Delta) + \vartheta^2\sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}\Delta Z_n \\ &+ \left(\frac{\vartheta^4(\vartheta^1 + \vartheta^2 X(t_{n-1}))}{2\sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}} - \frac{(\vartheta^4)^2}{8\sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}}\right)(\Delta W_n \Delta - \Delta Z_n), \quad (3.5.1) \end{aligned}$$

where $\Delta W_n = \sqrt{\Delta}\epsilon_{1_n}$ and $\Delta Z_n = \frac{\Delta^{\frac{3}{2}}}{2}(\epsilon_{1_n} + \frac{\epsilon_{2_n}}{\sqrt{3}})$. Here ϵ_{1_n} and ϵ_{2_n} are independent, standard Gaussian random variables. Additionally, the *balanced implicit scheme*, introduced by Milstein, Platen & Schurz (1999), is also used to simulate the affine diffusion. This scheme has the form

$$\begin{aligned} X(t_n) &= X(t_{n-1}) + (\vartheta^1 + \vartheta^2 X(t_{n-1}))\Delta + \sqrt{\vartheta^3 + \vartheta^4 X(t_{n-1})}\Delta W_n \quad (3.5.2) \\ &+ C_n(X(t_{n-1}) - X(t_n)) \end{aligned}$$

with ΔW_n defined as above. Here the function C_n is chosen to be

$$C_n = (\vartheta^3 + \vartheta^4)|\Delta W_n|.$$

Unknown Drift and Diffusion Functions for T = 20

The affine diffusion is simulated with $20,000$ steps over the time period $[0, T]$, with the parameters set to $\vartheta^1 = 0.01$, $\vartheta^2 = -0.01$, $\vartheta^3 = 0.01$, $\vartheta^4 = 0.01$, $T = 20$

and $X(0) = 1$, see Figure 3.1. There is no significant difference in the paths obtained when simulated by the different numerical schemes. The parameters are estimated from the path shown in Figure 3.1 by application of the estimators given in (3.4.17), using every tenth observation only in the estimation procedure. Thus $\Delta = 0.009$, so we expect the estimator to be consistent.

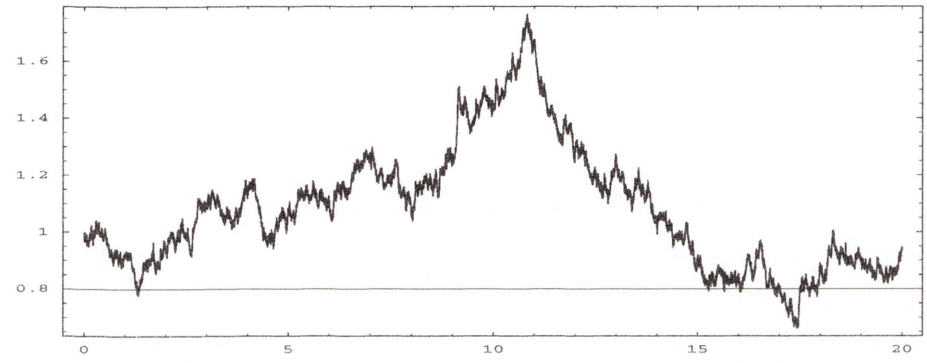

Figure 3.1: Simulated sample path of the affine diffusion with $\vartheta^1 = 0.01$, $\vartheta^2 = -0.01$, $\vartheta^3 = 0.01$ and $\vartheta^4 = 0.01$.

The evolution of the estimators through time is shown in Figures 3.2 and 3.3. There is substantial variability of the drift estimators as they evolve over time. This is to be expected since the observed trajectory in Figure 3.1, shows only one substantial change in direction. The estimates of the diffusion parameters are relatively stable as can be seen from Figure 3.3.

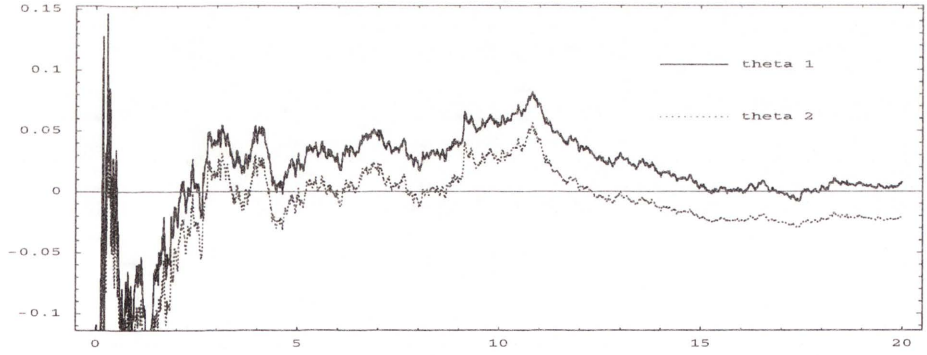

Figure 3.2: Estimates of the drift parameters $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$.

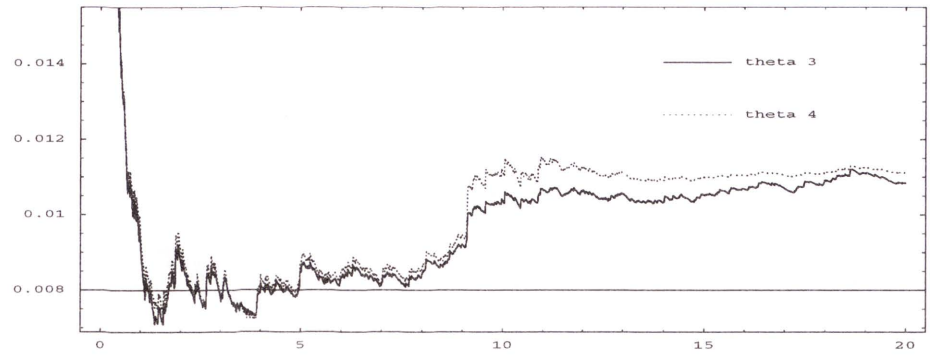

Figure 3.3: Estimates of the diffusion parameters $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$.

To study the variability of the estimators, results from the estimation of 1,000 simulated paths are shown in Table 3.1. As before, we have used every tenth observation and for comparison, we have also used every twentieth and fiftieth observation for the estimation where $\Delta = 0.019$ and $\Delta = 0.049$, respectively. The corresponding number of observations used are given in the first column of Table 3.1. The mean and standard deviation of the estimators are given in the corresponding columns of Table 3.1. It is clear that, on average, the estimates are reasonably accurate. We see that the accuracy of the estimators for the drift parameters increases slightly as the time between observations is decreased. More clearly, there is a notable decrease in the variance of the estimates for the parameters of the diffusion coefficient as the time between observations decreases. Note that the standard deviation for the estimates of the drift parameters is substantially greater than that of the estimates of the diffusion parameters.

The estimation is also performed for the shorter observation period $T = 10$. The diffusion process in (3.3.1) is simulated using 10,000 steps with the parameter settings as discussed above. In order for the time between observations to remain unchanged, the number of simulated points is halved. Results using the estimating equations in (3.4.17) for 1,000 simulated paths are given in Table 3.2. It is clear that there is little change in the accuracy of the estimates for the diffusion coefficients. The standard deviation and mean of the estimators of the diffusion

n_t	$\hat{\vartheta}^1$	$\hat{\vartheta}^2$	$\hat{\vartheta}^3$	$\hat{\vartheta}^4$
2220	0.0101	-0.02087	0.0080	0.0081
	(0.0286)	(0.0394)	(0.0006)	(0.0009)
1050	0.0102	-0.0291	0.0090	0.0090
	(0.0301)	(0.0411)	(0.0010)	(0.0012)
408	0.0102	-0.0292	0.0097	0.0096
	(0.0309)	(0.0420)	(0.0016)	(0.0018)

Table 3.1: Mean and standard deviation of estimators for the drift and diffusion coefficients when $T = 20$.

n_t	$\hat{\vartheta}^1$	$\hat{\vartheta}^2$	$\hat{\vartheta}^3$	$\hat{\vartheta}^4$
1110	0.0104	-0.0192	0.0086	0.0088
	(0.0407)	(0.0482)	(0.0009)	(0.0008)
525	0.0109	-0.0194	0.0094	0.0091
	(0.0421)	(0.0506)	(0.0014)	(0.0018)
204	0.0110	-0.0198	0.0096	0.0096
	(0.0442)	(0.0521)	(0.0019)	(0.0023)

Table 3.2: Mean and standard deviation of estimators for the drift and diffusion coefficients when $T = 10$.

parameters are relatively unchanged. However, there is an increase in the standard deviation for the drift estimates. The need for a long observation interval to reliably estimate the drift coefficients is apparent from this study.

The estimators in Figures 3.2 and 3.3 and Tables 3.1 and 3.2 appear to have some bias. In particular, the diffusion coefficient parameters are underestimated. In theory, this could be rectified by simply using the compensated estimating functions in (3.1.14). However, in practice, when Δ is small, the effect on the estimated parameter values is minimal due to the presence of Δ in the compen-

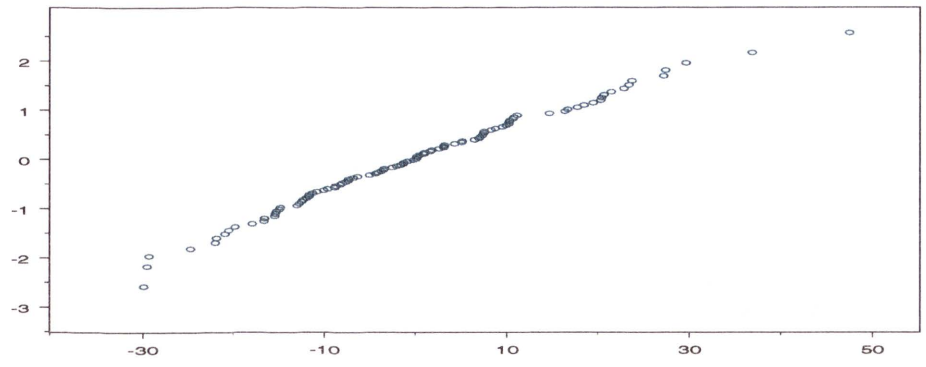

Figure 3.4: Gaussian quantile plot for $\hat{\vartheta}^1$.

sators.

Alternatively, it may be possible to eliminate the leading error term, which is a multiple of the time between observations. This is done by combining estimating functions in a way that the resulting expectation is small. This was tested and the results suggest that such bias–reduced estimators overcompensate for the bias present in the estimators in the given example.

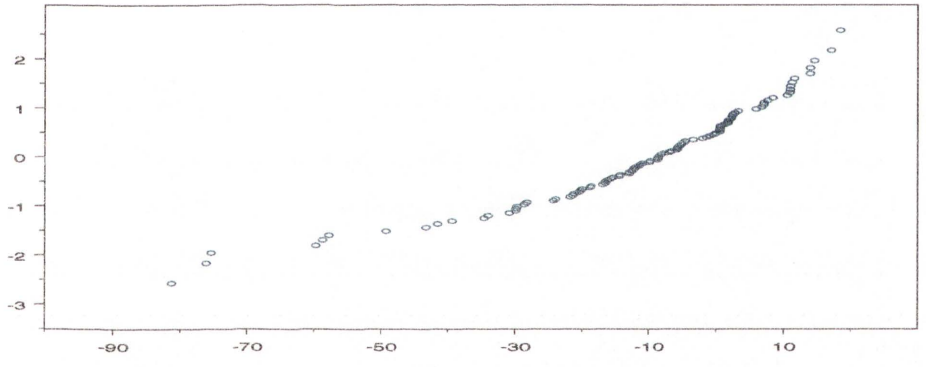

Figure 3.5: Gaussian quantile plot for $\hat{\vartheta}^2$.

Another simple way of correcting for the bias in the method is to artificially simulate the affine diffusion with the biased parameter estimates. By correcting for the observed bias, new estimates are obtained that can be used for an improved simulation. This procedure can be repeated until the biased parameter estimates are reasonably matched by the simulated results. The parameter estimates of the artificially simulated diffusion process may then be interpreted as good proxies

for the true parameters.

The normality of the estimators is illustrated in Gaussian quantile plots shown in Figures 3.4, 3.5, 3.6 and 3.7 for the case $T = 20$. The quantile plots indicate that the distributions of the parameter estimates $\hat{\vartheta}^1$, $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ are close to Gaussian. The distribution of $\hat{\vartheta}^2$ exhibits some larger deviations from a Gaussian distribution. These deviations are also reflected in the estimates given in Tables 3.1 and 3.2.

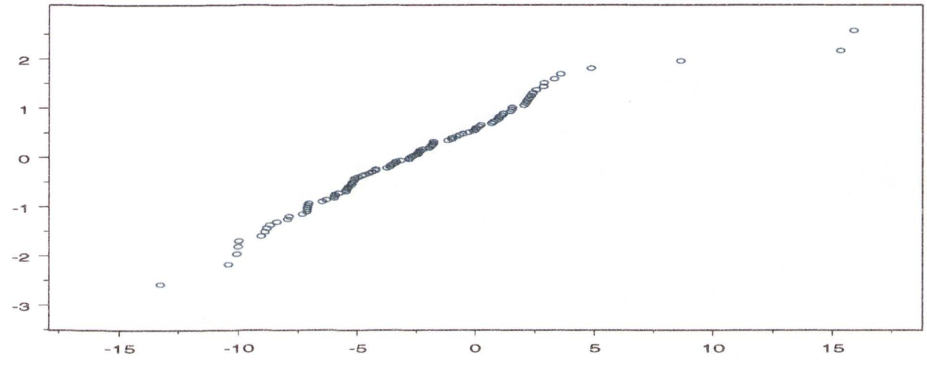

Figure 3.6: Gaussian quantile plot for $\hat{\vartheta}^3$.

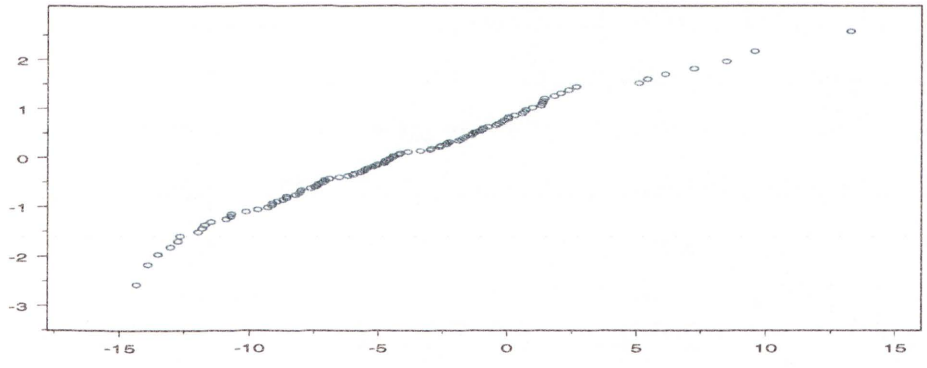

Figure 3.7: Gaussian quantile plot for $\hat{\vartheta}^4$.

Unknown Drift and Diffusion Functions for T = 1,000

Of interest is the length of the observation period T required such that the drift estimators also converge to the true parameter values using the system in (3.4.17).

For the affine diffusion process in (3.3.1) with the same parameters as above, Figure 3.8 shows a simulated path with 100,000 steps where $T = 1,000$. Note that in contrast to the simulated path shown in Figure 3.1, this path shows a number of changes in direction.

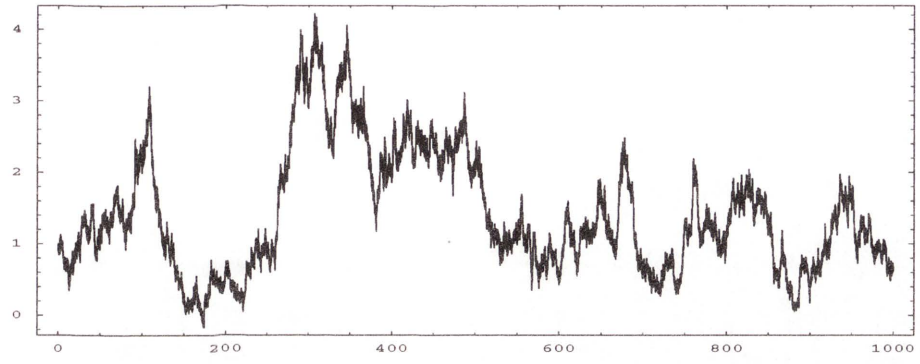

Figure 3.8: Simulated sample path of the affine diffusion when $T = 1,000$ with $\vartheta^1 = 0.01$, $\vartheta^2 = -0.01$, $\vartheta^3 = 0.01$ and $\vartheta^4 = 0.01$.

The evolution of the estimators that result for this path are shown in Figures 3.9 and 3.10. It is clear that the diffusion coefficients are estimated reasonably soon with little variability about the estimates. Note that the distance between tick marks on the vertical axis in Figure 3.10 is 0.00025. However, the drift estimates display significant variability and take much longer than the $T = 20$ years, as shown in the previous example, before convergence to the true parameter values is apparent. In fact, it appears that at least $T = 500$ years is required to be reasonably confident that the drift estimates are reliable. This may reflect the fact that the estimators are the solution of a system of estimating equations. In this sense, the value of an individual estimate relies on those of the other estimated parameter values. If one estimator does not stabilise, neither will the remaining estimators. This occurs despite the relatively fast stabilisation of the diffusion estimators.

Figure 3.11 and 3.12 show the Gaussian quantile plots for the estimators shown in Figure 3.9. It is clear that with the increased time, the estimators for the drift

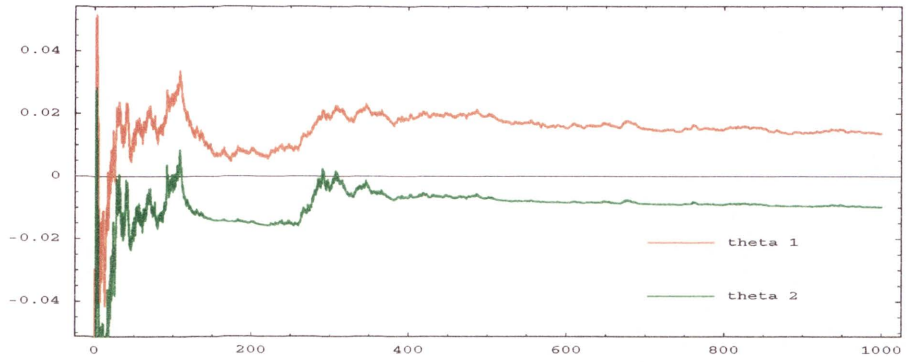

Figure 3.9: Estimates of the drift parameters $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$ when $T = 1,000$.

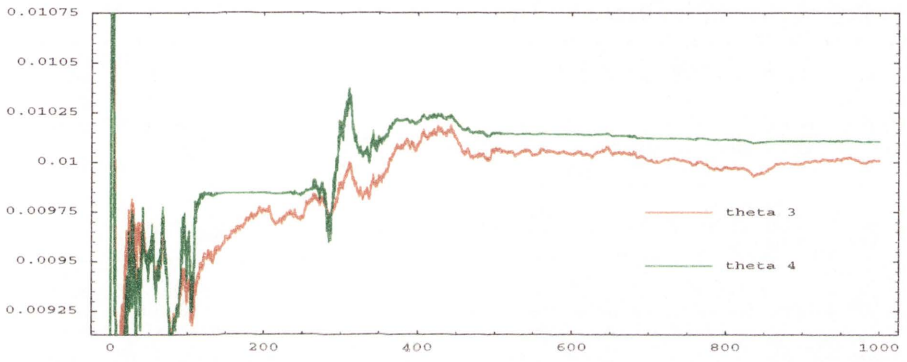

Figure 3.10: Estimates of the diffusion parameters $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ when $T = 1,000$.

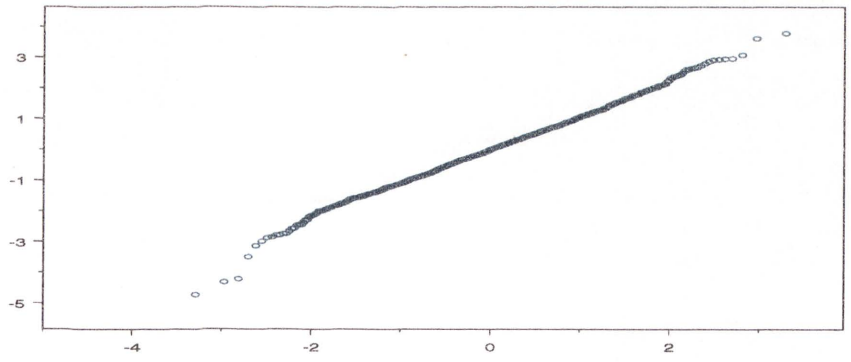

Figure 3.11: Gaussian quantile plot for $\hat{\vartheta}^1$ with $T = 1,000$.

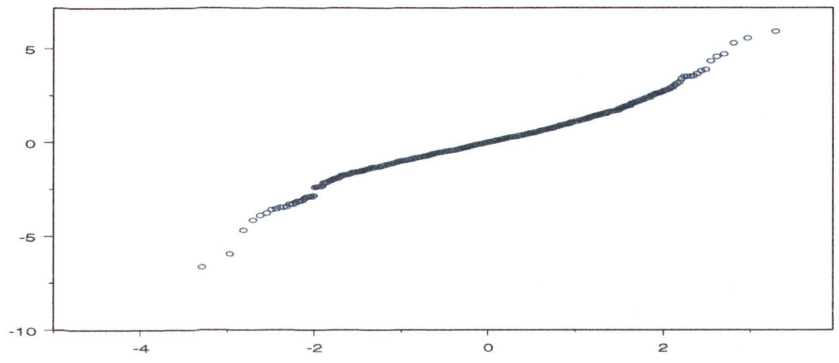

Figure 3.12: Gaussian quantile plot for $\hat{\vartheta}^2$ with $T = 1,000$.

coefficients $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$ are reasonably close to Gaussian.

Known Drift Function and Unknown Diffusion Function for T = 1

It appears that the time required to estimate the diffusion function is far shorter than $T = 20$ years. To illustrate how the estimators in (3.4.17) of the diffusion function perform, we assume that the estimators of the drift function have converged to the correct values and estimate only the diffusion function. Note that this is not a study where optimal estimators are designed. Our aim is to clarify the length of the observation period required T, such that the estimators stabilise.

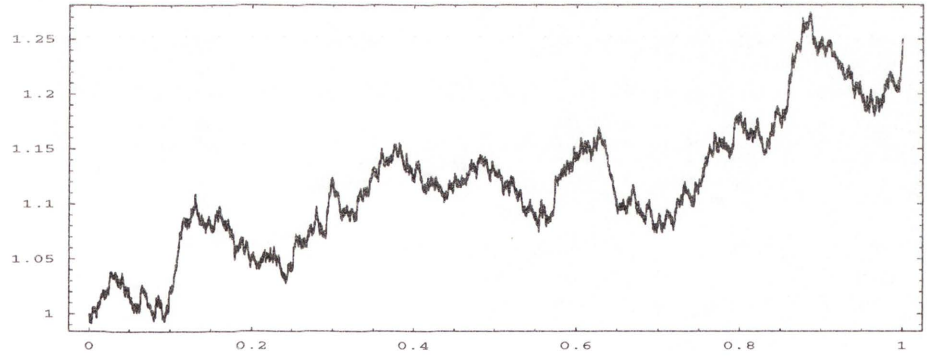

Figure 3.13: Simulated sample path of the affine diffusion when $T = 1$ with $\vartheta^1 = 0.01$, $\vartheta^2 = -0.01$, $\vartheta^3 = 0.01$ and $\vartheta^4 = 0.01$.

We simulate a path of the solution of the affine SDE (3.3.1) with 100,000 simulated points over the comparatively short observation period $T = 1$. This is equivalent to performing the estimation with financial data observed at five minute intervals. The drift and diffusion coefficients have the same values as in the previous examples. An example of the simulated path is shown in Figure 3.13.

The resulting estimators for $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ are shown in Figure 3.14. Here we see that the estimators stabilise extremely quickly and a reasonably reliable estimate of the diffusion coefficients $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ is available after only six months. The reason for this is clear once it is recognised that in (3.4.17)

$$\mathrm{A}_{\Delta}^{0,1,0}(2) = \frac{1}{n_t \Delta} \langle X^2 \rangle_{\Delta,t} \tag{3.5.3}$$

and

$$\mathrm{A}_{\Delta}^{0,1,0}(1) = \frac{1}{n_t \Delta} \langle X \rangle_{\Delta,t}, \tag{3.5.4}$$

see (2.1.6).

Thus, the expressions for $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ can be written as

$$\begin{aligned} \hat{\vartheta}^3 &= \frac{1}{n_t \Delta} \langle X \rangle_{\Delta,t} - \mathrm{A}_{\Delta}^1 \hat{\vartheta}^4 \\ \hat{\vartheta}^4 &= \frac{\frac{1}{4}\frac{1}{n_t \Delta} \langle X^2 \rangle_{\Delta,t} - \mathrm{A}_{\Delta}^2 \frac{1}{n_t \Delta} \langle X \rangle_{\Delta,t}}{\mathrm{A}_{\Delta}^3 - A_{\Delta}^1 A_{\Delta}^2}. \end{aligned} \tag{3.5.5}$$

Hence, it is apparent that the estimators rely on the quadratic variation of X and the transforms of X. This provides an explanation of the short time period required for the estimators to stabilise. We remark that in the case of the smaller time between observations Δ, the evolution of the diffusion estimates is practically indistinguishable.

Unknown Drift Function and Known Diffusion Function for T = 1,000

Given that the diffusion coefficients can be estimated relatively quickly, we examine the length of time required such that the drift function can be reliably

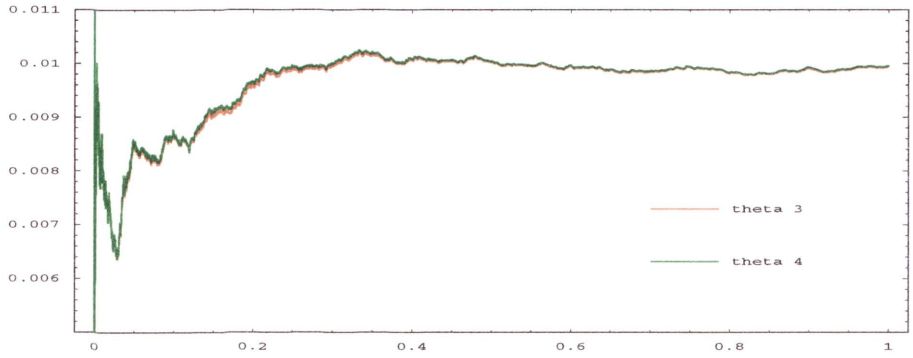

Figure 3.14: Estimates of the diffusion parameters $\hat{\vartheta}^3$ and $\hat{\vartheta}^4$ when $T = 1$.

estimated assuming the parameter estimates for the diffusion function have converged to the correct values. Again note that this is not a study of the design of optimal estimators. We simulate a path of the solution to the SDE in (3.3.1) with 100,000 simulated points over the observation period $T = 1{,}000$. An example of this path with the same parameter values as above is shown in Figure 3.15.

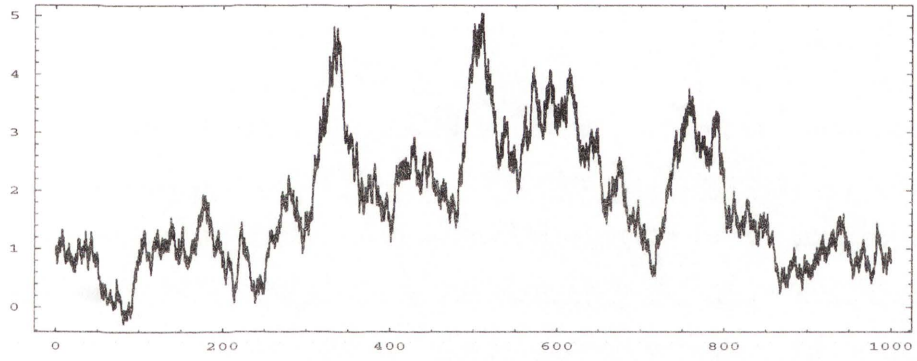

Figure 3.15: Simulated sample path of the affine diffusion when $T = 1{,}000$ with $\vartheta^1 = 0.01$, $\vartheta^2 = -0.01$, $\vartheta^3 = 0.01$ and $\vartheta^4 = 0.01$.

The evolution of the drift estimates $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$ are shown in Figure 3.16. It is clear that in this case at least $T = 600$ years is required to be reasonably confident that the drift estimators $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$ have stabilised. Thus, even when we assume that the diffusion function has completely stabilised, the drift estimates still require a long observation period before convergence to the true parameter values is appar-

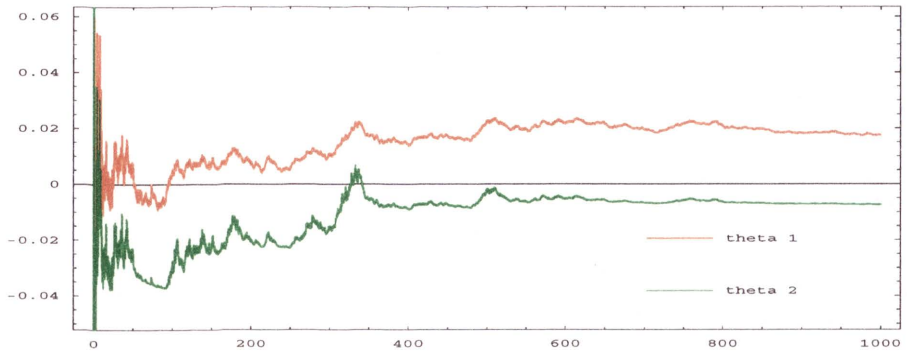

Figure 3.16: Estimates of the drift parameters $\hat{\vartheta}^1$ and $\hat{\vartheta}^2$ when $T = 1,000$.

ent. A long observation period would still be necessary even if optimal maximum likelihood estimators for the drift parameters, where the diffusion parameters are given, were used.

With the transform function method, we propose a simple and rather general estimation technique that is capable of estimating parameters in the drift and the diffusion coefficients of discretely observed diffusions. This method has the advantage of being simple to implement and, in principle, does not require stationarity. Explicit expressions for moments, conditional moments or transition densities of the original process and the transformed observations are not required. This method has been demonstrated for simulated affine diffusion processes, where estimated parameter values for the diffusion function are found quickly and reliably. However, it is shown for the transform function method, that the time required to reliably estimate the drift coefficients is significantly greater than even the longest available data series available for financial markets. Thus, it appears that a financial market model with the drift function determined by the diffusion function is advantageous for practical inference.

Chapter 4

A Continuous Benchmark Model

This chapter reviews the continuous benchmark model of Platen (2002) and extends the modelling framework to incorporate intraday data. The essence of the benchmark model is to form a self–financing portfolio of risky assets and maximise the logarithmic growth rate to find the growth optimal portfolio (GOP). Here for ease of notation, it is assumed that all assets are denominated in the domestic currency and the dependence on the currency is suppressed unless otherwise stated. The exact GOP is difficult to form in practice and as a result, theory regarding the approximation of the GOP is required and presented in Section 4.3. In Section 4.4, three alternative constructions of high–frequency indices are presented. A goal is to find the best approximation to the GOP using the high–frequency indices considered. The remainder of the chapter extends the benchmark framework to intraday modelling. This is done by introducing a market activity process which captures changes in the index due to short term fluctuations. Section 4.7 briefly reviews empirical properties of the GOP.

4.1 A Continuous Market

4.1.1 Model Setup

We consider a continuous financial market with $d+1$ *primary assets*, d of which are risky assets, with the remaining asset being riskless. The d primary risky assets, denominated in the domestic currency, can be shares, foreign savings accounts, indices, commodities or other security contracts. The uncertainty in this market is generated by d independent, standard Wiener processes $W^1, \ldots, W^d$ defined on a filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$ that satisfies the usual conditions. The

filtration $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$ is the augmentation under P of the natural filtration $\mathcal{A}^W$, generated by the vector $W = \{W(t) = (W^1(t), \ldots, W^d(t))^\top, t \in [0,T]\}$ of independent Wiener processes, see Karatzas & Shreve (1991) and Protter (1990). Here $\mathcal{A}_t$ is the information available up until time t. Note that P is the real world probability measure.

4.1.2 Primary Security Accounts

A *primary security account* contains only units of a given primary asset. It accumulates all payments and losses that are generated from holding this asset. For example, the BHP primary security account contains the risky BHP spot share price in addition to the accumulated dividends. The All Ordinaries Index is formed as a portfolio of primary assets and the corresponding portfolio of primary security accounts including accumulated dividends is shown in Figure 4.1. We assume that the units of primary assets are infinitely divisible such that continuous trading is possible.

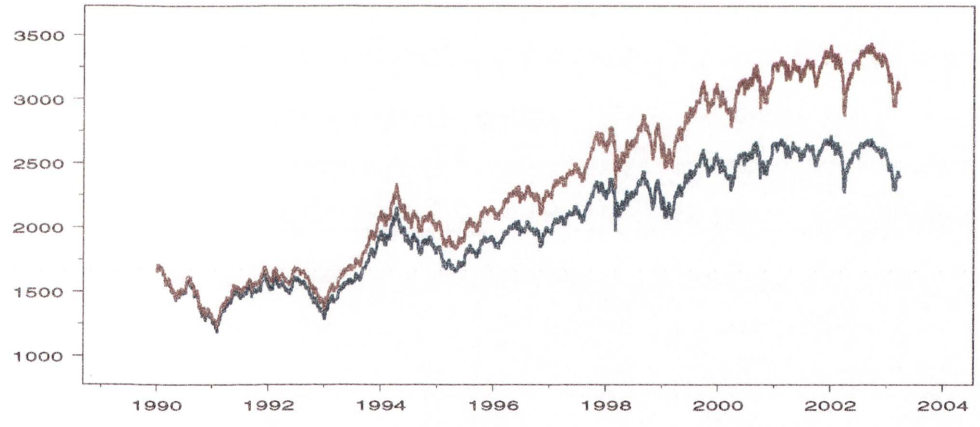

Figure 4.1: An example of an index of primary assets (blue) and corresponding portfolio of primary security accounts (red) for the All Ordinaries Index.

The value of the jth *primary security account process* at time t, is denoted by $S^{(j)}(t)$, when expressed in units of the domestic currency. It is assumed in the

following that $S^{(j)}(t)$ is the unique solution of the SDE

$$dS^{(j)}(t) = S^{(j)}(t)\left\{a^j(t)dt + \sum_{k=1}^{d} b^{j,k}(t)dW^k(t)\right\} \tag{4.1.1}$$

for $t \in [0,T]$ and $j \in \{0,1,\ldots,d\}$ with $S^{(j)}(0) > 0$. Here the jth *appreciation rate* $a^j(t)$ is the expected return at time t that an investor receives for holding the jth primary security account. The (j,k)th *volatility* $b^{j,k}(t)$ measures, at time t, the proportional fluctuations of the value of the jth primary security account with respect to the kth Wiener process. Some technical assumptions regarding the appreciation rate and volatility are required. Namely, that the jth appreciation rate and (j,k)th volatility are considered to be predictable stochastic processes such that

$$\int_0^T \left[|a^j(s) + (b^{j,k}(s))^2|\right] ds < \infty$$

almost surely for $j\,, k \in \{1,2,\ldots,d\}$, see Protter (1990). The *savings account* or riskless asset is denoted as $S^{(0)}(t)$ and has

$$b^{0,k}(t) = 0$$

for $t \in [0,T]$ and $k \in \{1,\ldots,d\}$. It is clear from (4.1.1) that

$$dS^{(0)}(t) = S^{(0)}(t)a^0(t)dt \tag{4.1.2}$$

where $S^{(0)}(0) = 1$. Here $r(t) = a^0(t)$ is the adapted short rate at time t.

Furthermore, we assume that no security can be expressed as a linear combination of other securities, see Platen (2002). That is, no primary security account is *redundant* in the sense that its value cannot be replicated by forming a portfolio of other primary security accounts. This can be written by introducing the *volatility matrix* $b(t) = [b^{j,k}(t)]_{j,k=1}^{d}$ and assuming that it is *invertible*. That is,

$$b^{-1}(t) < \infty.$$

In such an *invertible market* all uncertainties are securitised. This implies that the uncertainties for individual markets are sufficiently different. In practice, this

does not restrict the modelling of the dynamics of primary security accounts since each additional primary security account permits the use of an additional Wiener process for modelling.

By introducing the appreciation rate vector $a(t) = (a^1(t), \ldots, a^d(t))^\top$ and the unit vector $\mathbf{1} = (1, 1, \ldots, 1)^\top$, the *market price for risk* vector is obtained as

$$\theta(t) = (\theta^1(t), \ldots, \theta^d(t))^\top = b^{-1}(t)[a(t) - r(t)\mathbf{1}] \tag{4.1.3}$$

for $t \in [0, T]$. The kth market price for risk $\theta^k(t)$ is proportional to the excess return that an investor expects at time t for taking the risk that is modelled by the kth Wiener process. The market price for risk (4.1.3) allows the SDE (4.1.1) to be written in the form

$$dS^{(j)}(t) = S^{(j)}(t) \left\{ r(t)dt + \sum_{k=1}^{d} b^{j,k}(t)[\theta^k(t)dt + dW^k(t)] \right\} \tag{4.1.4}$$

for $t \in [0, T]$ and $j \in \{0, 1, \ldots, d\}$.

4.1.3 Portfolios and Strategies

To form the GOP, it is necessary to first form a self–financing portfolio. This portfolio is formed by determining a self–financing *strategy*. A predictable stochastic process $\delta = \{\delta(t) = (\delta^{(0)}(t), \delta^{(1)}(t), \ldots, \delta^{(d)})^\top, t \in [0, T]\}$ is a strategy if δ is S-integrable, see Protter (1990). The jth component $\delta^{(j)}(t) \in (-\infty, \infty)$ of the strategy δ denotes the number of units of the jth primary security account that are held in the corresponding portfolio at time $t \in [0, T]$. For a given strategy δ, let $S^{(\delta)}(t)$ be the value of the corresponding portfolio at time t. This means that

$$S^{(\delta)}(t) = \sum_{j=0}^{d} \delta^{(j)}(t) S^{(j)}(t) \tag{4.1.5}$$

for $t \in [0, T]$. A portfolio and the corresponding strategy δ are *self–financing* if

$$dS^{(\delta)}(t) = \sum_{j=0}^{d} \delta^{(j)}(t) dS^{(j)}(t) \tag{4.1.6}$$

for all $t \in [0,T]$. Thus, all changes in the value of a self–financing portfolio are due to corresponding gains from trade in the primary security accounts. Unless otherwise stated, we consider only self–financing strategies and corresponding self–financing portfolios.

For a given strategy δ, the jth proportion $\pi_\delta^{(j)}(t)$ of the value of the jth primary security account held at time t in a strictly positive portfolio $S^{(\delta)}(t)$, is given by

$$\pi_\delta^{(j)}(t) = \delta^{(j)}(t)\frac{S^{(j)}(t)}{S^{(\delta)}(t)} \tag{4.1.7}$$

for all $t \in [0,T]$ and $j \in \{0,1,\ldots,d\}$. Naturally, the sum of the proportions invested in each primary security is one. That is,

$$\sum_{j=0}^{d} \pi_\delta^{(j)}(t) = 1 \tag{4.1.8}$$

for all $t \in [0,T]$. A strictly positive portfolio value $S^{(\delta)}(t)$ satisfies, according to (4.1.1) and (4.1.6), the SDE

$$dS^{(\delta)}(t) = S^{(\delta)}(t)\left\{ r(t)dt + \sum_{k=1}^{d} \beta_\delta^k(t)(\theta^k(t)dt + dW^k(t)) \right\} \tag{4.1.9}$$

with kth *portfolio volatility*

$$\beta_\delta^k(t) = \sum_{j=0}^{d} \pi_j^{(j)}(t) b^{j,k}(t) \tag{4.1.10}$$

for $t \in [0,T]$. The portfolio volatility $\beta_\delta^k(t)$, given in (4.1.10), is a linear combination of the security account volatilities and their corresponding proportions. By application of the Itô formula, it follows from (4.1.9), that the logarithm of a strictly positive portfolio $S^{(\delta)}(t)$ satisfies the SDE

$$d\log\left(S^{(\delta)}(t)\right) = (r(t) + g_\delta(t))dt + \sum_{k=1}^{d} \beta_\delta^k(t)dW^k(t) \tag{4.1.11}$$

with *portfolio net growth rate*

$$g_\delta(t) = \sum_{k=1}^{d} \beta_\delta^k(t)\left(\theta^k(t) - \frac{1}{2}\beta_\delta^k(t)\right) \tag{4.1.12}$$

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

4.2 Growth Optimal Portfolio

Central to the benchmark model is the *growth optimal portfolio* (GOP), see Kelly (1956), Long (1990), Karatzas & Shreve (1998) or Platen (2002). We denote its value at time $t \in [0,T]$ by $S^{(\delta_*)}(t)$. The GOP is the portfolio that maximises the portfolio net growth rate $g_\delta(t)$. Thus, the optimal proportions $\pi^{(\delta_*)} = \{\pi^{(\delta_*)}(t) = (\pi^{(1)}_{\delta_*}(t), \pi^{(2)}_{\delta_*}(t), \ldots, \pi^{(d)}_{\delta_*}(t))^\top, t \in [0,T]\}$ are obtained from the solution of the first order equations for the corresponding quadratic maximisation problem for the portfolio net growth rate $g_\delta(t)$ in (4.1.12). The solution to this problem yields the optimal proportions

$$\pi^{(\delta_*)}(t) = (b^{-1}(t))^\top \theta(t) \tag{4.2.1}$$

for $t \in [0,T]$. Using (4.1.9) together with (4.2.1), the GOP can then be shown to satisfy the SDE

$$dS^{(\delta_*)}(t) = S^{(\delta_*)}(t) \left[r(t)dt + \sum_{k=1}^{d} \theta^k(t)(\theta^k(t)dt + dW^k(t)) \right] \tag{4.2.2}$$

for $t \in [0,T]$, see also Karatzas & Shreve (1998) and Platen (2002). It is assumed that $S^{(\delta_*)}(0) > 0$.

From (4.2.2) and (4.1.3), it follows that the kth GOP volatility $\theta^k(t)$ is the kth corresponding market price for risk. Note that the *risk premium* of the GOP appearing in (4.2.2) equals the square of the *total market price for risk*

$$|\theta(t)| = \sqrt{\sum_{k=1}^{d} (\theta^k(t))^2} \tag{4.2.3}$$

for $t \in [0,T]$. The risk premium of the portfolio in (4.1.9) is given by the correlation of the returns of the portfolio with the returns of the GOP.

In applications to areas in quantitative finance, the GOP can be used as *benchmark* or numeraire, see Platen (2002). Security prices expressed in units of the GOP are known as *benchmarked prices* and can be shown to be local martingales

under the real world probability measure. Note that within the benchmark framework, derivative pricing also takes place under the real world probability measure using the concept of *fair pricing* rather than risk–neutral pricing. An advantage of fair pricing is that it encompasses and extends the commonly used risk–neutral pricing methodology. This is useful since risk–neutral pricing relies on the existence of an equivalent risk–neutral measure and that the Radon–Nikodym derivative is a martingale.

4.3 Approximate Growth Optimal Portfolios

In practice, forming the GOP can be difficult, see, for example, Long (1990). From (4.2.1), it is clear that the optimal proportions depend on the volatility matrix and the market price for risk vector. The market price for risk appears in the drift coefficient of the SDE for the primary security account, see (4.1.4). This implies that this parameter cannot be reliably estimated. In addition, estimation of large volatility matrices can be cumbersome and as such, it is simpler to approximate the GOP by using a suitably diversified portfolio. According to Platen (2003), in order to approximate the GOP using a diversified portfolio, the following definitions are required.

Definition 4.3.1 *A strictly positive portfolio process $S^{(\delta)}$ is called a* diversified portfolio (DP) *if finite constants $K_1 > 0$, $K_2 > 0$ and $K_3 \in \{1, 2, \ldots\}$ exist, independent of d, such that*

$$|\pi_\delta^{(j)}(t)| \leq \frac{K_1}{d^{\frac{1}{2}+K_2}} \tag{4.3.1}$$

almost surely for all $j \in \{0, 1, \ldots, d\}$, $d \in \{K_3, K_3 + 1, \ldots\}$ and $t \in [0, T]$.

This means that the proportion $\pi_\delta^{(j)}(t)$ of the value of a DP, which is invested in the jth primary security account, must decrease slightly faster than $d^{-\frac{1}{2}}$ as $d \to \infty$. This occurs, for example, if equal proportions are used.

When a given portfolio $S^{(\delta)}(t)$ is expressed in units of the GOP, then the *benchmarked portfolio* $\check{S}^{(\delta)} = \{\check{S}^{(\delta)}(t), t \in [0,T]\}$ is the ratio

$$\check{S}^{(\delta)}(t) = \frac{S^{(\delta)}(t)}{S^{(\delta_*)}(t)}. \tag{4.3.2}$$

By application of the Itô formula in conjunction with (4.3.2), (4.1.9), (4.1.10) and (4.2.2), it follows that the benchmarked portfolio satisfies the SDE

$$d\check{S}^{(\delta)}(t) = -\check{S}^{(\delta)}(t) \sum_{k=1}^{d} \sum_{j=0}^{d} \pi_{\delta}^{(j)}(t) \sigma^{j,k}(t) dW^k(t). \tag{4.3.3}$$

The (j,k)th *specific volatility*

$$\sigma^{j,k}(t) = b^{j,k}(t) - \theta^k(t) \tag{4.3.4}$$

for $t \in [0,T]$, $j \in \{0,1,\ldots,d\}$ and $k \in \{1,2,\ldots,d\}$ measures the jth diversifiable risk with respect to the kth Wiener process. The kth *total specific volatility* can then be introduced as

$$\check{\sigma}^k(t) = \sum_{j=0}^{d} |\sigma^{j,k}(t)| \tag{4.3.5}$$

for $t \in [0,T]$ and $k \in \{1,2,\ldots,d\}$.

Definition 4.3.2 *A benchmark model is called* regular *if there exists finite constants, K_3 and K_4, independent of d, such that*

$$E\left(\left(\check{\sigma}^k(t)\right)^2\right) \leq K_4 \tag{4.3.6}$$

for all $t \in [0,T]$, $k \in \{1,2,\ldots,d\}$ and $d \in \{K_3, K_3+1, \ldots\}$.

In a regular benchmark model, each independent source of uncertainty influences a restricted number of benchmarked primary security accounts. This is a property that can be assumed for the world stock market consisting of all stocks traded on the existing exchanges.

In order to form an approximate GOP, a measure of the difference between the exact GOP and a DP is required. Hence, Platen (2003) introduces the notion of

the *tracking rate*. By (4.1.9) and (4.2.2), the difference between the logarithm of the GOP $S^{(\delta_*)}$ and the logarithm of a given portfolio $S^{(\delta)}$ satisfies the SDE

$$d(\log(S^{(\delta_*)}(t)) - \log(S^{(\delta)}(t))) = \frac{1}{2}R_\delta^d(t)dt - \sum_{k=1}^{d}\sum_{j=0}^{d}\pi_\delta^{(j)}(t)\sigma^{j,k}(t)dW^k(t) \quad (4.3.7)$$

with tracking rate

$$R_\delta^d(t) = \sum_{k=1}^{d}\left(\sum_{j=0}^{d}\pi_\delta^{(j)}(t)\sigma^{j,k}(t)\right)^2 \quad (4.3.8)$$

for $t \in [0,T]$. Note that the tracking rate equals the squared diffusion coefficient of the SDE in (4.3.7). It can be interpreted as a measure of the "distance" between a given portfolio $S^{(\delta)}(t)$ and the GOP $S^{(\delta_*)}(t)$ at time $t \in [0,T]$.

Definition 4.3.3 *For an increasing number d of risky primary security accounts we call a strictly positive portfolio $S^{(\delta)}$ an* approximate GOP *if the corresponding sequence of tracking rates $R_\delta^d(t)$ $d \in \{1,2,\ldots\}$ vanishes in probability. That is, for each $\epsilon > 0$ we have*

$$\lim_{d\to\infty} P(R_\delta^d(t) > \epsilon) = 0 \quad (4.3.9)$$

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

This is a natural definition for an approximate GOP. As the number of assets increase, the approximate GOP approaches the exact GOP. Thus, the difference between the values of these indices must approach zero. Under the above assumptions the following limit theorem is proved in Platen (2003).

Proposition 4.3.4 (Platen (2003)) *For a regular benchmark model, a diversified portfolio is an approximate GOP.*

This result implies that a diversified world stock portfolio or index, which is a DP, approximates the exact GOP.

4.4 Construction of Intraday World Stock Indices

The results in Section 4.3 imply that the GOP can be approximated by a sufficiently diversified *world stock index* (WSI). To our knowledge there is no readily available high–frequency WSI. Therefore, in Breymann, Kelly & Platen (2004), three different intraday WSIs are constructed, for which key characteristics of their dynamics will be compared.

Each WSI is constructed as a self–financing portfolio consisting of local stock market indices taken from all major financial markets throughout the world. See Appendix B for a list of those indices included. The WSIs, denominated in US Dollars (USD), are constructed from $d = 34$ local stock market accumulation indices. Many of the local indices used are obtained as spot price indices. We denote by $V^{(i,j)}(t)$, the value of the local stock spot price index at time t of stock market j, when denominated in currency i, for $j \in \{1, 2, \ldots, d\}$ and $i \in \{USD, CHF, \ldots\}$. As a first step, we transform the original local spot price index into an accumulation index $S^{(i,j)}_{Index}(t)$ by forming the expression

$$S^{(i,j)}_{Index}(t) = V^{(i,j)}(t) \exp\left\{\int_0^t y^{(j)}(u)du\right\} \tag{4.4.1}$$

for $t \in [0, T]$, $j \in \{1, 2, \ldots, d\}$ and $i \in \{USD, CHF, \ldots\}$. Here $y^{(j)}(t)$ is the continuously compounded jth dividend yield for the jth local stock index at time t. Secondly, in order to form a WSI in US Dollars, all local accumulation indices must be denominated in US Dollars. The corresponding foreign exchange rate is denoted by $X^{(i,k)}(t)$ at time t. That is, $X^{(i,k)}(t)$ is the price at time t of one unit of the kth currency in terms of units of the ith currency. The jth local accumulation index, denominated in US Dollars, is given by

$$S^{(USD,j)}_{Index}(t) = S^{(i,k)}_{Index}(t) X^{(USD,k)}(t) \tag{4.4.2}$$

for $t \in [0, T]$. A WSI at time t, denominated in US Dollars, denoted by

$S^{(USD,\delta_{(WSI)})}(t)$, is obtained by forming the portfolio

$$S^{(USD,\delta_{(WSI)})}(t) = \sum_{j=1}^{d} \delta^{(j)}_{(\mathrm{WSI})}(t)\, S^{(USD,j)}_{Index}(t). \tag{4.4.3}$$

Here $\delta^{(j)}_{(\mathrm{WSI})}(t)$ denotes the number of units of the jth local accumulation index held in the WSI at time $t \in [0,T]$. The values for $\delta^{(j)}_{(\mathrm{WSI})}(t)$ relate to the corresponding jth proportion $\pi^{(\delta)}_{(\mathrm{WSI})}(t)$, see (4.1.7), at time t by the formula

$$\pi^{(j)}_{(\mathrm{WSI})}(t) = \delta^{(j)}_{(\mathrm{WSI})}(t) \frac{S^{(USD,j)}_{Index}(t)}{S^{(USD,\delta_{(WSI)})}(t)}, \tag{4.4.4}$$

for $t \in [0,T]$ and $j \in \{1,2,\ldots,d\}$. Throughout our analysis we separate the US Dollar short rate evolution from the index movements examined, by studying the *discounted* WSI $\bar{S}^{(USD,\delta_{(WSI)})}(t)$ at time t, which is computed as,

$$\bar{S}^{(USD,\delta_{(WSI)})}(t) = \frac{S^{(USD,\delta_{(WSI)})}(t)}{S^{(USD,0)}(t)} \tag{4.4.5}$$

for $t \in [0,T]$ where $S^{(USD,0)}(t)$ is US Dollar savings account given by

$$S^{(USD,0)}(t) = \exp\left\{\int_0^t r^{USD}(s)ds\right\}. \tag{4.4.6}$$

Here $r^{USD}(t)$ denotes the short rate at time t for the US Dollar market.

We consider three WSIs: the approximately *Equal Weighted Index* (EWI), the *Market Capitalisation Weighted Index* (MCI) and the *Gross Domestic Product Weighted Index* (GDPI). The values of the strategies $\delta^{(j)}_{(\mathrm{WSI})}(t)$ for each WSI are adjusted such that the corresponding proportion $\pi^{(j)}_{(\mathrm{WSI})}(t)$ for the MCI and GDPI correspond to their respective weights. The proportions for the MCI are shown in Figure 4.2. The proportions of the GDPI are omitted since they are similar to those of the MCI. Despite its name, the EWI is not equally weighted for all markets. Developed markets were given an equal weight of 0.0357 and emerging markets a weight of 0.0179. The markets were classified as either developed or emerging to be consistent with the classification used by Morgan Stanley Capital International. The stock markets considered in the MCI account for more than 95% of the total world market capitalisation, while the GDPI weights account for

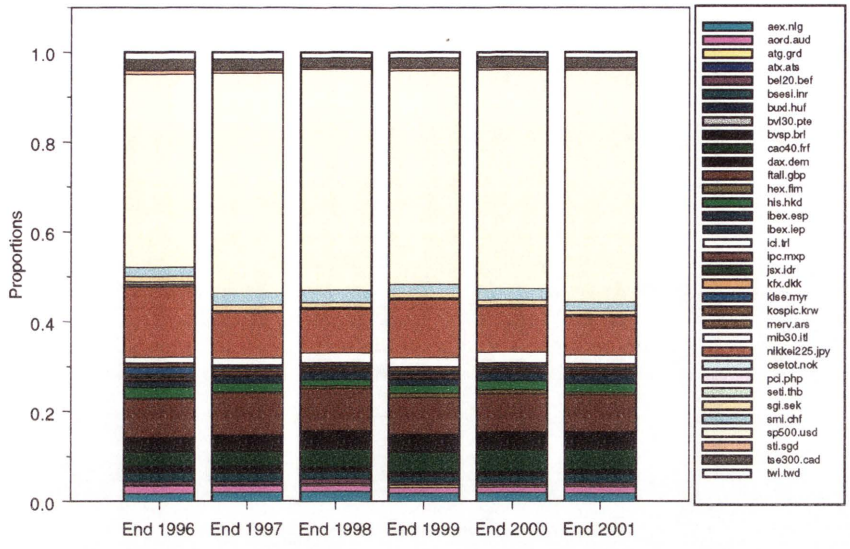

Figure 4.2: Proportions of each local stock index calculated by using market capitalisation.

more than 85% of the total world gross domestic product. In each of the WSIs, the number of units of each of the accumulation indices is kept piecewise constant. The WSIs are rebalanced after one year has elapsed or when a local stock index is to be added to the WSI, whichever occurs first. The rebalancing dates are 05/04/1996, 17/12/1996, 31/01/1997, 30/01/1998, 20/10/1998, 20/10/1999 and 20/10/2000.

To ensure that the WSIs are comparable, each index must have the same initial value. Each WSI is rescaled to have the same initial value as the Morgan Stanley Capital World Index (MSCI) at the starting date t_0 of our sample to enable the convenient comparison of the WSIs to a readily available market index. That is,

$$\tilde{S}^{(USD,\delta_{(WSI)})}(t) = \frac{S^{(USD,\delta_{(WSI)})}(t)}{S^{(USD,\delta_{(WSI)})}(t_0)} S^{(USD,\delta_{(MSCI)})}(t_0) \tag{4.4.7}$$

with

$$t_0 = 05/04/1996\ 00{:}00{:}00\ \text{GMT}. \tag{4.4.8}$$

Note that all analysis of the WSIs is performed using Greenwich Mean Time (GMT).

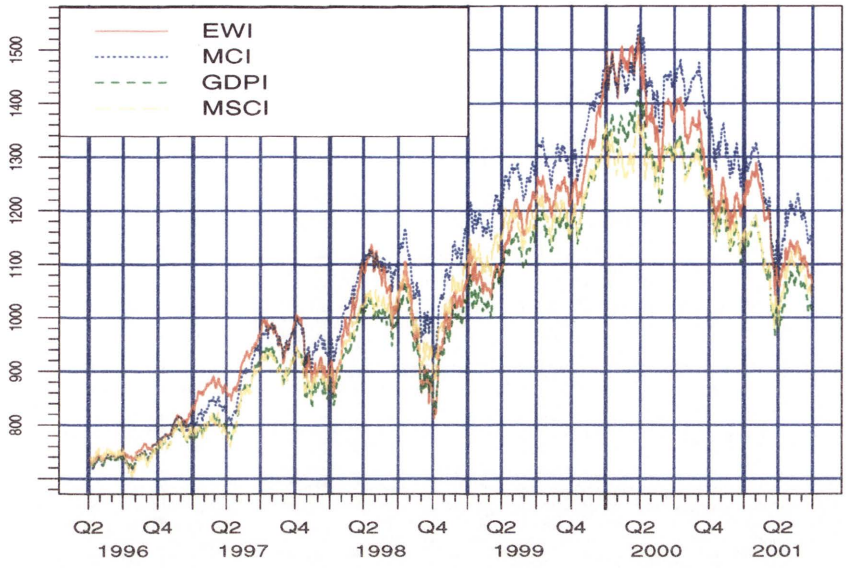

Figure 4.3: World stock indices with equal weights (EWI), market capitalisation adjusted weights (MCI), GDP adjusted weights (GDPI) and MSCI, where initial values are matched to the MSCI.

As any single local accumulation index represents a portfolio consisting of a cross section of the stocks in that local market, a WSI can be regarded as a portfolio containing the stocks of all local stock markets considered. To include as many stocks as possible in the WSIs constructed, all local accumulation indices for which we could obtain high–frequency data were included. As a result of this, the three WSIs are based on between two and three thousand stocks.

In addition to local stock indices, we rely on high–frequency foreign exchange spot data to transform the value of the different local stock indices into US Dollars. The high–frequency index and foreign exchange data consists of tick–by–tick data originating mainly from Reuters, which was collected and filtered by Olsen Data. The period explored is from 4 April 1996 until 29 June 2001. Intraday data for a number of local stock indices start at later dates and are included into the WSIs as soon as they are continuously available. Both the original high–frequency local stock indices and foreign exchange time series are irregularly spaced. Previous tick interpolation was used to transform the observed data to regularly spaced time series with an observation interval of five minutes. In order to form the

discounted accumulated WSIs, daily local dividend yields and US Dollar interest rates obtained from Thomson Financial were used. Omitting the high–frequency information in dividends and short rates is justified because only the exponentials of the integrals of these quantities, not the values themselves, enter the formulae for the relevant quantities in our construction of the WSIs, see (4.4.1) and (4.4.6).

Figure 4.3 displays the three different WSIs in addition to the daily observed MSCI. Of note is that, despite significant differences in the weights, all WSIs appear to be very similar. In particular, as suggested in Platen (2003), the similar fluctuations of all four indices are apparent. The three WSIs are analysed in detail in Chapter 5.

4.5 Discounted GOP

To discount the GOP value $S^{(\delta_*)}(t)$ at time t, see (4.2.2), we introduce the savings account

$$S^{(0)}(t) = \exp\left\{\int_0^t r(s)\, ds\right\}, \tag{4.5.1}$$

see (4.1.2), where $r(t)$ is the short rate at time t of the domestic market. The *discounted* GOP

$$\bar{S}^{(\delta_*)}(t) = \frac{S^{(\delta_*)}(t)}{S^{(0)}(t)} \tag{4.5.2}$$

satisfies by application of the Itô formula, (4.2.2) and (4.5.1) the SDE

$$d\bar{S}^{(\delta_*)}(t) = \bar{S}^{(\delta_*)}(t)\, |\theta(t)|\, (|\theta(t)|\, dt + d\hat{W}(t)) \tag{4.5.3}$$

for $t \in [0, T]$. Here the standard Wiener process $\hat{W} = \{\hat{W}(t),\, t \in [0, T]\}$ has the stochastic differential

$$d\hat{W}(t) = \frac{1}{|\theta(t)|} \sum_{k=1}^{d} \theta^k(t)\, dW^k(t), \tag{4.5.4}$$

where $|\theta(t)|$ is the total market price of risk given in (4.2.3). By Levy's Theorem, see Karatzas & Shreve (1991), $\hat{W} = \{\hat{W}(t), t \in [0, T]\}$ is a standard Wiener process on $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$.

The *discounted GOP drift* $\alpha(t)$ at time $t \in [0,T]$, has the form

$$\alpha(t) = \bar{S}^{(\delta_*)}(t)\,|\theta(t)|^2. \tag{4.5.5}$$

When taken as a parameter process, $\alpha = \{\alpha(t),\, t \in [0,T]\}$, leads to a *GOP volatility* of the form

$$|\theta(t)| = \sqrt{\frac{\alpha(t)}{\bar{S}^{(\delta_*)}(t)}}. \tag{4.5.6}$$

Thus by (4.5.3) and (4.5.5), we obtain the following SDE for the discounted GOP

$$d\bar{S}^{(\delta_*)}(t) = \alpha(t)dt + \sqrt{\alpha(t)\bar{S}^{(\delta_*)}(t)}d\hat{W}(t) \tag{4.5.7}$$

for $t \in [0,T]$. Note that (4.5.7) is a time transformed squared Bessel process of dimension four, see Revuz & Yor (1999).

By application of the Itô formula to (4.5.3), the logarithm of the discounted GOP satisfies the SDE

$$d\log\left(\bar{S}^{(\delta_*)}(t)\right) = \frac{1}{2}|\theta(t)|^2 dt + |\theta(t)|d\hat{W}(t) \tag{4.5.8}$$

for $t \in [0,T]$. This process has quadratic variation, see Section 2.1,

$$\left\langle \log\left(\bar{S}^{(\delta_*)}\right)\right\rangle_t = \int_0^t |\theta(s)|^2\, ds. \tag{4.5.9}$$

Thus, the GOP volatility can be observed as

$$|\hat{\theta}(t)| = \sqrt{\frac{d}{dt}\left\langle \log\left(\bar{S}^{(\delta_*)}\right)\right\rangle_t}. \tag{4.5.10}$$

Figure 4.4 shows the GOP volatility calculated for the MCI over several weeks in April/May 1996. Here we see that volatility is neither constant nor a deterministic function of time. When the GOP volatility process is calculated on a five minute basis it is difficult to see the leverage effects that are exhibited in the medium term. This is because a relatively small change in the index produces a large change in the GOP volatility, which is distorted by the intraday activity. This mixture of different dynamics makes it difficult to directly analyse the behaviour of observed volatility under the Black–Scholes–Merton type models.

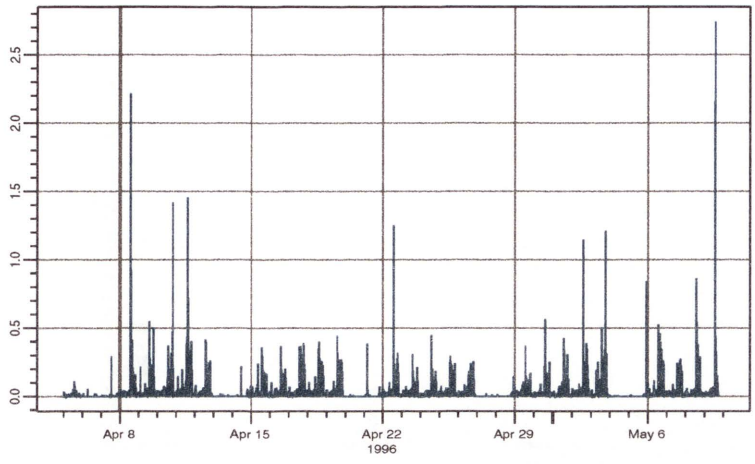

Figure 4.4: Volatility process of the MCI for several weeks in April/May 1996.

4.6 Normalised GOP

The WSIs displayed in Figure 4.3, on average, appear to grow exponentially. In principle, this can be accounted for by modelling the discounted GOP drift using an exponential function.

More precisely, throughout this thesis, we take the model for the discounted GOP drift process α as

$$\alpha(t) = \xi \exp\left\{\eta \int_0^t m(s)\, ds\right\} m(t) \tag{4.6.1}$$

for $t \in [0, T]$. Here $\eta > 0$ is the *net growth rate* of the market and the $\underline{A}$-adapted, nonnegative *market activity* process is denoted by $m = \{m(t),\, t \in [0, T]\}$.

The parameterisation of the discounted GOP drift in (4.6.1) can be justified by consideration of the quadratic variation of the square root of the discounted GOP. That is, by (4.5.7) and the Itô formula, $\sqrt{\bar{S}^{\delta_*}(t)}$ satisfies the SDE

$$d\sqrt{\bar{S}^{\delta_*}(t)} = \frac{3}{8}\frac{\alpha(t)}{\sqrt{\bar{S}^{\delta_*}(t)}}dt + \frac{1}{2}\sqrt{\alpha(t)}d\hat{W}(t) \tag{4.6.2}$$

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

$$\left\langle \sqrt{\bar{S}^{\delta_*}} \right\rangle_t = \frac{1}{4}\int_0^t \alpha(t)\, dt. \tag{4.6.3}$$

Note that this quantity defines so–called *GOP time*, see Platen (2004). Figure 4.5 shows the approximate quadratic variation of $\sqrt{\bar{S}^{\delta_*}(t)}$, which appears to be fairly smooth. Note also that an exponential curve is somewhat evident. This is also true for longer time periods explored in Platen (2004). The discounted GOP drift should incorporate this feature. Furthermore, the inclusion of market activity reflects periods of low discounted GOP drift due to market closures and periods of high discounted GOP drift due to active markets. This can be seen if the quadratic variation is viewed over a shorter observation period than the five years shown in Figure 4.5.

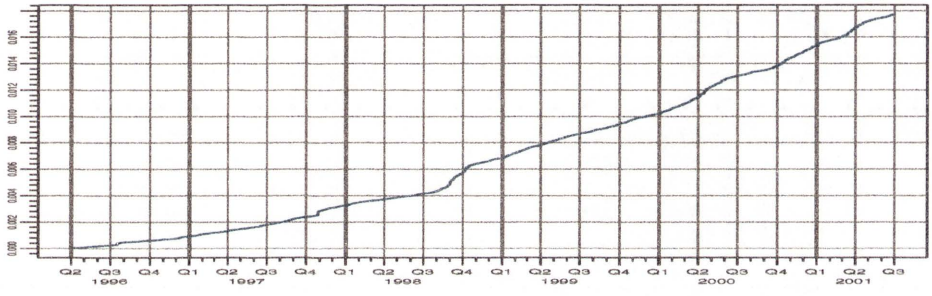

Figure 4.5: Approximate quadratic variation of the square root of the discounted MCI.

It is straightforward to show via the Itô formula, in conjunction with (4.5.7) and (4.6.1), that the *normalised GOP* $Y = \{Y(t),\, t \in [0,T]\}$ with

$$Y(t) = \frac{\bar{S}^{(\delta_*)}(t) m(t)}{\alpha(t)} = \frac{\bar{S}^{(\delta_*)}(t)}{\xi \exp\left\{\eta \int_0^t m(s)\, ds\right\}}, \tag{4.6.4}$$

satisfies the SDE

$$dY(t) = \eta \left(\frac{1}{\eta} - Y(t)\right) m(t)\, dt + \sqrt{Y(t)\, m(t)}\, d\hat{W}(t) \tag{4.6.5}$$

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

$$Y(0) = \frac{\bar{S}^{(\delta_*)}(0) m(0)}{\xi}. \tag{4.6.6}$$

From (4.5.6) and (4.6.4), the GOP volatility can be written as

$$|\theta(t)| = \sqrt{\frac{m(t)}{Y(t)}} \tag{4.6.7}$$

for $t \in [0, T]$. Thus the GOP volatility can also be calculated simply once the normalised GOP and market activity process are known. Relation (4.6.7) highlights the short term effects on the GOP volatility caused by seasonal patterns and intraday trading activity. In the medium to long term, where market activity becomes less important, the GOP volatility will no longer reflect these short term features. This implies that to model the market activity, an exogenous process for this quantity appears to be adequate. Market activity will be modelled in Chapter 5, in accordance with the empirical properties that are observed.

4.6.1 Market Activity Time

Based on the market activity $m(t)$ introduced in the previous section, *market activity time* can be defined as $\psi = \{\psi(t), t \in [0, T]\}$ with

$$\psi(t) = \int_0^t m(s)\, ds \tag{4.6.8}$$

for $t \in [0, T]$. Here, $\psi(0)$ is the starting point of the market activity time scale and 0 years corresponds to the starting date of our sample 05/04/1996 00:00:00 GMT. Additionally, $T = 5.25$ years, which is equivalent to 30/05/2001 00:00:00 GMT, is the terminal date of the data available. Note that (4.6.8) requires proper normalisation of the market activity. It is reasonable to normalise $m(t)$ such that, on average, the market activity time scale elapses approximately as fast as physical time. Furthermore, we assume that

$$\lim_{T \to \infty} \frac{1}{T} E\left(\psi(T)\right) = 1. \tag{4.6.9}$$

This assumption is natural, given that market activity and seasonal patterns do not have a significant affect on the analysis of daily data, see Platen (2004). That is, it is assumed that as time increases, market activity time approaches physical time.

With the notion of market activity time, see (4.6.8), the SDE (4.6.5) for the normalised GOP in market activity time $Y_{\psi(t)} = Y(t)$ can be written as

$$dY_\psi = \eta \left(\frac{1}{\eta} - Y_\psi \right) d\psi + \sqrt{Y_\psi}\, d\hat{W}_\psi \tag{4.6.10}$$

for $\psi(t) \in [0, \psi(T)]$ and

$$d\hat{W}_{\psi(t)} = \sqrt{m(t)}\, d\hat{W}(t) \tag{4.6.11}$$

for $t \in [0, T]$. The normalised GOP in market activity time in (4.6.10) is a *square root process of dimension four*, see Revuz & Yor (1999). The solution of (4.6.10) has a long term mean of $\frac{1}{\eta}$ and a speed of adjustment parameter η. Most importantly, the only parameter is the net growth rate η of the market. Note that the SDE in (4.6.10) is an affine SDE of the type discussed in Section 3.3. This suggests that with this parametrisation the long term net growth rate η can be estimated using techniques discussed in Chapters 2 and 3. Although, as previously discussed, the length of the observation period is inadequate for drift parameter estimation. The parameter η must be identified through inference about the diffusion coefficient. This is possible due to the structure of the SDE for the discounted GOP, see (4.5.7). In Platen (2004), an estimate of $\eta \approx 0.048$ is obtained from 30 years of daily data using the fact that

$$\alpha(t) = 4 \frac{d}{dt} \langle \sqrt{\bar{S}^{(\delta_*)}} \rangle_t. \tag{4.6.12}$$

This estimate is close to the trend of $\eta \approx 0.043$ obtained in Figure 1.1.

Notwithstanding difficulties related to the estimation of η from the normalised GOP in (4.6.10), the derived dynamics are useful due to the nature of the diffusion coefficient. By (4.6.10) and application of the Itô formula, the square root of the normalised GOP in market activity time satisfies the SDE

$$d\left(\sqrt{Y_\psi}\right) = \left(\frac{3}{8\sqrt{Y_\psi}} - \frac{\eta}{2} \sqrt{Y_\psi} \right) d\psi + \frac{1}{2} d\hat{W}_\psi \tag{4.6.13}$$

for $\psi \in [0, \psi(T)]$. It is crucial to note that the diffusion coefficient in (4.6.13) is constant. Therefore, we obtain in market activity time the quadratic variation of

$\sqrt{Y}$ in the form

$$\left\langle \sqrt{Y} \right\rangle_{\psi} = \frac{\psi}{4} \tag{4.6.14}$$

for $\psi \in [0, \psi(T)]$. Despite its simplicity, relation (4.6.14) is fundamental. It allows direct observation of market activity time and hence market activity. Note that from the market activity time given in (4.6.8) and the quadratic variation of $\sqrt{Y}$ in (4.6.14) the market activity can be calculated as

$$m(t) = \frac{d\psi(t)}{dt} = 4\frac{d\langle \sqrt{Y} \rangle_t}{dt} \tag{4.6.15}$$

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

Relationship (4.6.15) implies that market activity is directly observable provided that the normalised GOP is a square root process of dimension four. To observe market activity, the slope of the quadratic variation of the square root of the normalised GOP is required.

4.7 Empirical Properties of the GOP

An extensive amount of literature exists on the stylised empirical features of financial instruments for both daily and high–frequency data. This literature encompasses distributional properties of returns, the behaviour of volatility and the well documented implied volatility skew or smile. Of interest, is whether the benchmark model and its intraday extension for the GOP are able to capture these empirical features in a robust manner for all time scales.

Typically, the log–returns of financial assets display leptokurtic distributions. That is, the distribution has heavier tails and more pronounced peaks around the mode than described by a Gaussian distribution. Hurst & Platen (1997) analysed the daily log–returns of 27 indices. It was found that the Student t distribution with degrees of freedom four best describes the marginal distribution of log–returns, see also Hurst & Platen (1999). Additionally, Breymann, Fergusson &

Platen (2004) confirmed the results in Hurst & Platen (1997) for daily log–returns using a much larger data set.

The marginal distribution of log–returns in the benchmark model can be explained via the properties of the squared GOP volatility. The squared GOP volatility $|\theta(t)|^2$ according to (4.6.7) is given by the ratio

$$|\theta(t)|^2 = \frac{m(t)}{Y(t)} \tag{4.7.1}$$

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

For daily or longer observation intervals, market activity becomes less important in the dynamics of volatility, which implies that the squared GOP volatility is approximately

$$|\theta(t)|^2 \approx \frac{1}{Y(t)} \tag{4.7.2}$$

for $t \in [0, T]$. The stationary density of the square root process Y is a gamma density, which implies that the GOP volatility (4.7.2) yields an inverse gamma distributed conditional variance for log–returns. Hence, log–returns are Student t distributed.

Furthermore, five minute observations of log–returns were also considered in Breymann, Fergusson & Platen (2004) and it was found that the variance–gamma distribution best describes the marginal distribution of high–frequency log–returns. In this thesis, it is shown that this phenomenon can be explained by the nature of the market activity. It is demonstrated that market activity has a gamma density as stationary density. As market activity $m(t)$ changes rapidly compared to the normalised GOP $Y(t)$, the five minute log–returns are obtained as a normal mixture distribution with gamma distributed conditional variance. This implies a variance–gamma log–return distribution.

A change in distributional properties as the observation interval further increases can be easily documented. We consider the hourly, daily and weekly log–returns of the MCI. Note that the high–frequency returns are not deseasonalised. Gaus-

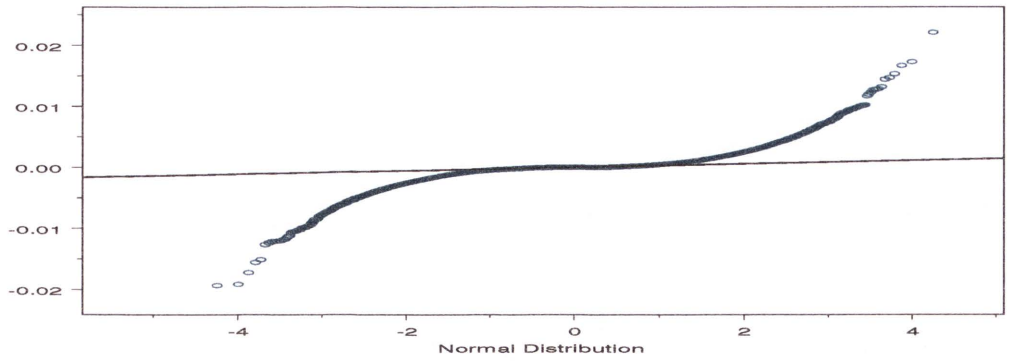

Figure 4.6: Gaussian quantile plot for hourly log–returns of the MCI.

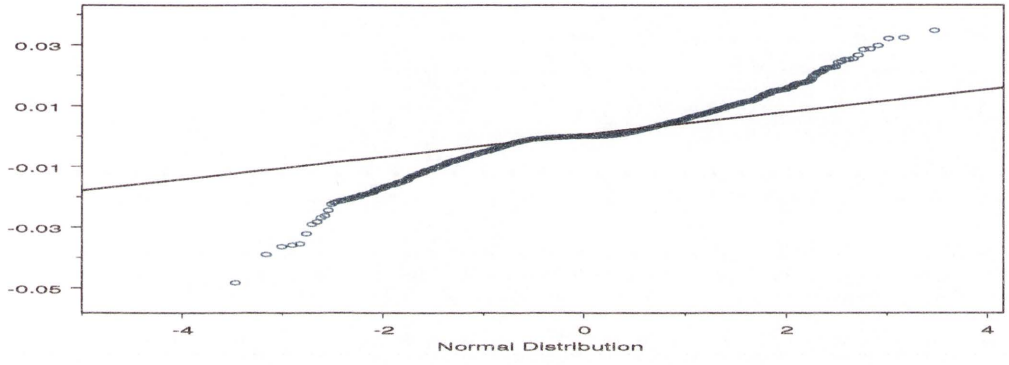

Figure 4.7: Gaussian quantile plot for daily log–returns of the MCI.

sian quantile plots of these returns are shown in Figure 4.6 to 4.8. Clearly, the larger the observation interval, the closer to Gaussian the log–returns appear. When hourly log–returns are considered, the returns display pronounced deviations from normality. This is apparent for various other instruments. For example, Dacorogna, Gençay, Müller, Olsen & Pictet (2001) analyse various currencies and the results suggest that the variance and third moment are finite in the large sample limit. Furthermore, they observe that the fourth moment may not be finite, as is consitent with the Student t distribution of four degrees of freedom.

Additionally, Breymann, Dias & Embrechts (2003) study the bivariate distributional properties of high–frequency foreign exchange returns and find that the Student t copula with four degrees of freedom, see, for instance, Joe (1997), best

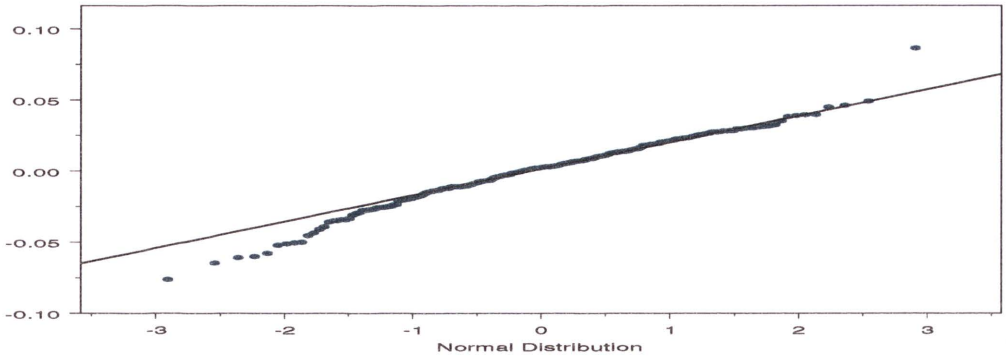

Figure 4.8: Gaussian quantile plot for weekly log–returns of the MCI.

describes the joint dependence of returns for the US Dollar–Deutschmark and US Dollar–Japanese Yen foreign exchange rates.

We emphasise that the intraday distributional properties of the log–returns of the GOP depend on the stationary distribution of the market activity process, provided that this is the dominant factor influencing volatility movements. If market activity has a gamma distributed stationary density, then the log–returns of the GOP are variance–gamma distributed. However, if the market activity has less impact on the squared GOP volatility, such that GOP volatility is an inverse gamma distributed random variable, see Platen, West & Breymann (2004), then the log–returns of the GOP appear to be Student t distributed.

The benchmark model, when examined with daily data, see Platen (2004), incorporates stochastic volatility endogenously without an explicit external volatility process. Furthermore, since the GOP volatilities are proportional to the inverse of the normalised GOP process, see (4.6.7), this generates a negative correlation between the GOP and its volatility. This phenomenon is known as the leverage effect, see Black (1976), which is explained naturally by the benchmark model.

The Black–Scholes–Merton model, see Black & Scholes (1973), suggests that volatilities implied from market traded options are constant with respect to both the moneyness and time to maturity of the option. Deviations from this are

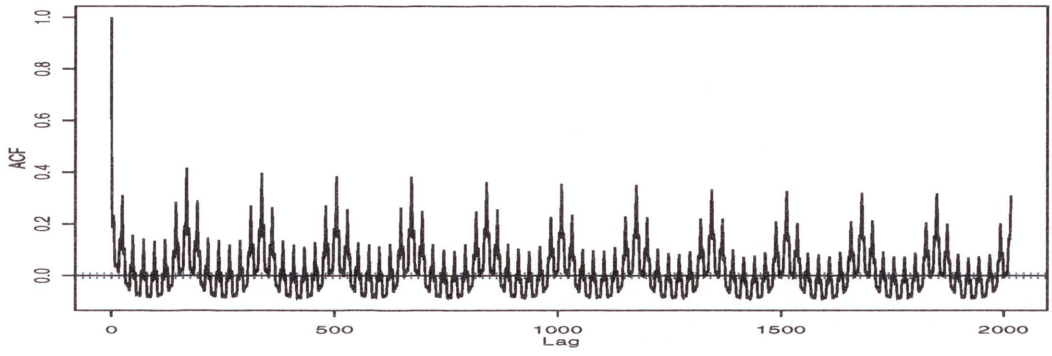

Figure 4.9: Sample autocorrelation function for hourly absolute returns of the MCI.

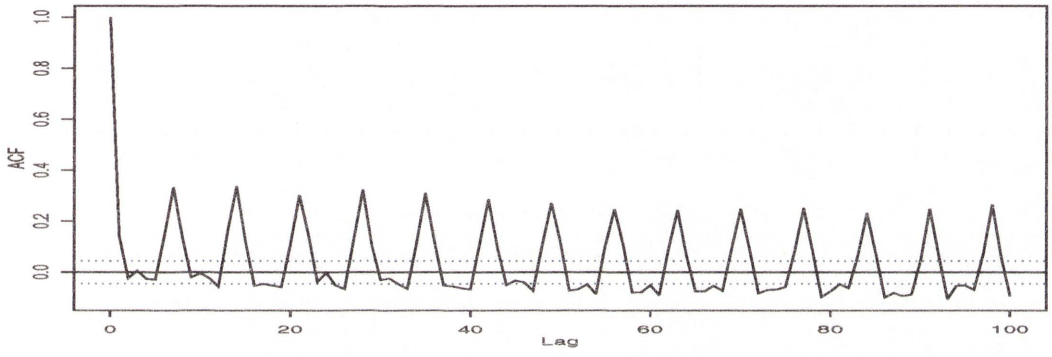

Figure 4.10: Sample autocorrelation function for daily absolute returns of the MCI.

well–documented, see, for example, Derman & Kani (1994), and are commonly referred to as the implied volatility skew or smile and the term structure of implied volatility, respectively. The benchmark model in conjunction with fair pricing, generates empirically observed implied volatility skews and smiles for European options on stock indices, stocks and currencies, see, for instance, Heath & Platen (2002).

Evidence of stochastic volatility has been well documented in the empirical finance literature by features such as volatility clustering and volatility persistence. This is best seen by considering the sample autocorrelation function for the absolute returns. We show this for the MCI for hourly and daily absolute returns in

Figures 4.9 and 4.10, respectively. Volatility persistence is evident in both the daily and hourly absolute returns, which is an indicator of stochastic volatility. This implies that the constant volatility suggested in the classical log–normal model, see Black & Scholes (1973), is incorrect.

Observing the hourly absolute returns, the presence of the periodic peaks in the sample autocorrelation function indicate seasonality. Seasonal patterns in high–frequency data are well–documented, see Dacorogna, Gençay, Müller, Olsen & Pictet (2001) and references therein. The introduction of the market activity process allows the modelling of typical intraday seasonal patterns in a stochastic framework. Studies linking the ideas of market activity, seasonal patterns and deviations from normality have existed for some time. Clark (1973) proposed linking the deviations from normality to the existence of variations in volume during different trading periods. Subordinated processes were introduced to compensate for the fact that physical time may not be the appropriate time scale for financial securities. Thus, Clark (1973) used volume traded as a proxy for market activity. The positive correlation between volume and volatility is well documented, see Karpoff (1987). Note, market activity can be regarded as a model for short term volatility in the standard financial modelling literature. A more recent study was conducted by Xu & Wu (1999). This study considered 141 stocks traded on the New York Stock Exchange. They find that frequency of trades explains return volatility. These results are a justification for the presence of market activity in the diffusion dynamics of the normalised GOP, see (4.6.5), and the drift and diffusion dynamics of the discounted GOP, see (4.5.7). This is as a consequence of the form of the model of the discounted GOP in the intraday benchmark framework.

Advances on Clark (1973) include the work by Ané & Geman (2000) where, using high–frequency observations of Cisco and Intel shares, they investigated whether the volume traded or the number of trades executed best defines business time. In the context of Ané & Geman (2000), the "best" business time is the time scale on which, when returns were examined, they were closest to Gaussian. For

these two stocks, the best time scale was found to be that which is determined by the number of trades executed. The claim is that all processes that define asset returns can be represented as time changed Brownian motion and that this time change is the key element in taking information arrival and market activity into account.

In the intraday benchmark model, the introduction of new market activity based time scales allows a natural, consistent method for the deseasonalisation of financial market data. This refers to allowing for seasonal patterns caused by regular information arrival and trading activity. Deseasonalisation in general, is necessary since the presence of seasonal patterns can mask other empirical features. The time scales considered are similar to that in Dacorogna, Gençay, Müller, Olsen & Pictet (2001) for volatility deseasonalisation, where they suggest a new business time scale based on polynomial approximations to the seasonal pattern present in volatility. Other deseasonalisation techniques have also been considered in the literature. For example, a business time scale can be created by the factorisation of volatility into different components. The seasonal components are then assumed to be deterministic, as opposed to the stochastic component, which is assumed to be essentially free of seasonalities, see, for instance, Andersen & Bollerslev (1997). Furthermore, wavelet techniques for deseasonalisation were suggested by Gençay, Selçuk & Whitcher (2001).

In the existing literature, the majority of approaches focus on volatility as the key parameter to be modelled. This is based on the logarithmic transformation of the underlying security. Under the benchmark approach, one transforms the underlying security, in this case an index, by using the square root, or more generally, a certain power function. This reveals the market activity, which in the benchmark approach, captures the higher order effects of intraday trading on volatility. The advantage is that much of the leverage effect is captured automatically by the square root transformation and as will be seen in the following discussion, market activity is relatively independent of the noise that drives the underlying

index itself.

The remaining chapters investigate the properties of the intraday WSIs introduced in Section 4.4, with particular emphasis on the modelling and empirical characteristics of market activity in both US Dollars and other currency denominations.

Chapter 5

Intraday Analysis of World Stock Indices

In order to assess the performance of the benchmark model, we consider the three world stock indices constructed in the previous chapter, denominated in US Dollars (USD). A key feature of the benchmark model is that the drift function is completely determined by the diffusion coefficient function. This is an advantage of the benchmark model since, as seen in Chapter 3, the drift coefficient is difficult to estimate with a satisfactory level of reliability. For this reason, we concentrate on the estimation and calibration of the diffusion coefficient. Section 5.1 considers properties of the normalised world stock index (WSI) in physical time. Section 5.2 concentrates on empirical features and the modelling of market activity. Finally, we consider the normalised WSI in market activity time. In the following chapter, we suppress the dependence on the currency for ease of exposition.

5.1 Empirical Properties of World Stock Indices

We consider the three high–frequency WSIs, denominated in US Dollars, introduced in Chapter 4. The indices shown in Figure 4.3 appear to grow exponentially. To compensate for this, we introduce the *normalisation function*

$$\bar{\alpha}(t) = \xi \exp(\eta t),$$

for $t \in [0, T]$, where $\xi > 0$ is the *normalisation factor* and $\eta > 0$ is the net growth rate of the market, see Section 4.6. An empirical study conducted by Dimson, Marsh & Staunton (2002), suggests that over the last 100 years, a discounted WSI denominated in US Dollars has experienced an average rate of growth of

approximately $\eta = 4.8\%$. This is in the range of that deduced in a simple manner for the long term WSI shown in Figure 1.1.

We introduce the *normalised* WSI in the form

$$Y^{(\mathrm{WSI})}(t) = \frac{\bar{S}^{(\mathrm{WSI})}(t)}{\bar{\alpha}(t)} \tag{5.1.1}$$

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

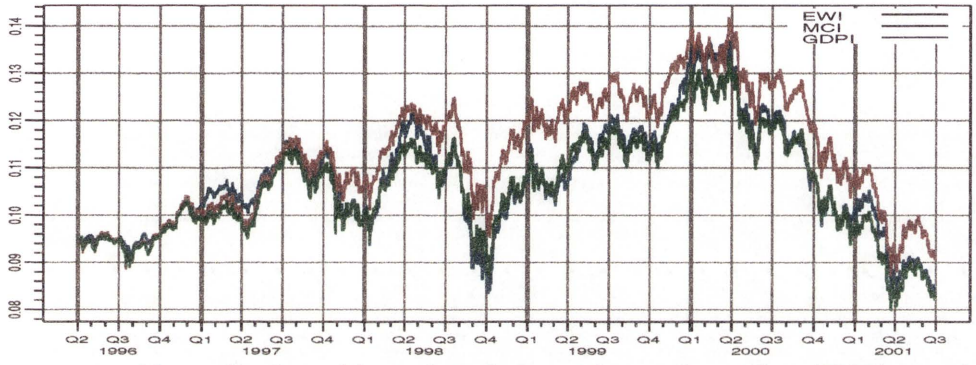

Figure 5.1: Normalised world stock indices with equal weights (EWI), market capitalisation adjusted weights (MCI), GDP adjusted weights (GDPI). All are shown in physical time.

Figure 5.1 shows the normalised WSIs in physical time. The normalised WSIs all display similar dynamics. A theoretical argument for this stylised fact is given in Section 4.3, where the GOP is shown to be approximated by any diversified portfolio.

We show the hourly absolute returns of the normalised MCI and the square root of the normalised MCI in Figure 5.2. Here we see that the square root reduces the fluctuations of returns when the level of the index is high and increases the fluctuations when the level of the index is low. In Section 4.6, it was shown that the quadratic variation of the square root of the normalised GOP is proportional to market activity time. However, we expect that as the observation interval increases, the effects of market activity should decrease and as such, market activity time and physical time should become very similar. It is shown in Platen (2004) that the MSCI observed on a daily basis displays the property $\langle \sqrt{Y} \rangle_t = \frac{t}{4}$.

However, seasonal patterns can be observed for more frequent observation intervals. Figure 5.3 shows four times the quadratic variation of the square root of the normalised MCI, or market activity time, see (4.6.14). One notes market activity time is monotonically increasing and appears to be almost linear over certain periods. It is clear that the process cannot be completely described by a linear function of time. This suggests that the introduction of market activity time is necessary to account for the seasonalities present in high–frequency indices. The presence of seasonal patterns in the intraday indices is confirmed by the sample autocorrelation functions of the hourly absolute returns of the normalised MCI in Figure 5.4. Note that the seasonal patterns evident in Figure 4.9 are also apparent here and the normalisation does not affect the seasonal behaviour. Properties of the normalised WSIs in market activity time are examined in Section 5.3.

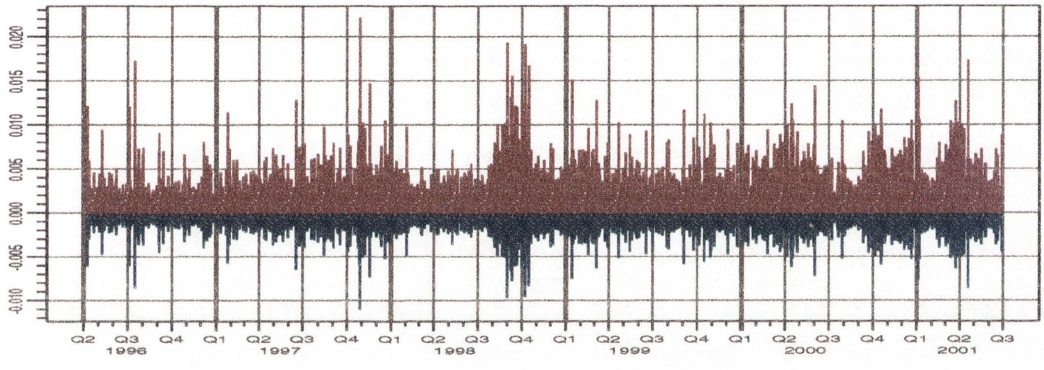

Figure 5.2: Hourly absolute returns of the normalised MCI (red) and the hourly absolute returns of the square root of the normalised MCI (blue).

5.2 Market Activity of the WSIs

5.2.1 Empirical Behaviour of Market Activity

The ease with which market activity time can be observed, allows for the direct calculation of the market activity. We calculate market activity via the numerical derivative corresponding to (4.6.15) using five minute time steps. Note that the

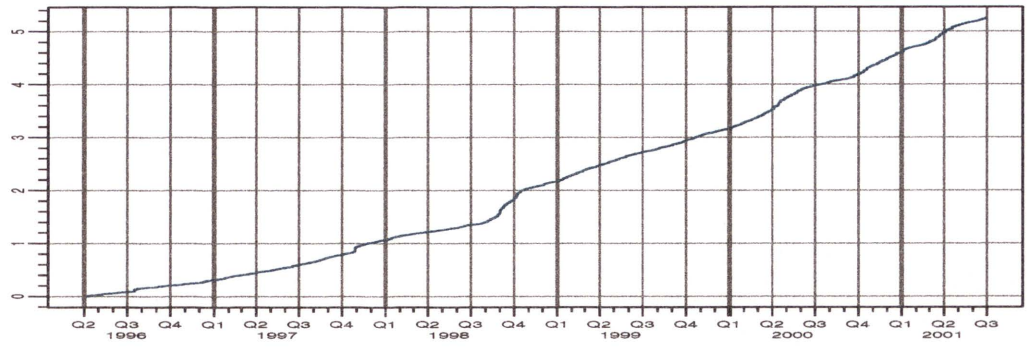

Figure 5.3: The market activity time for the MCI.

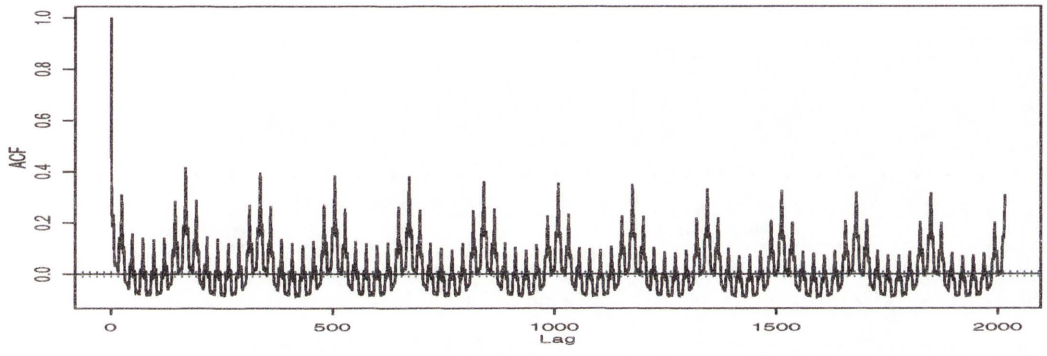

Figure 5.4: Sample autocorrelation function for the hourly absolute returns of the normalised MCI.

market activity is positive since market activity time is a non–decreasing function of physical time. The empirically observed market activity fluctuates over a wide range. Therefore, we show the logarithm of market activity $\ln(m(t))$ in Figure 5.5 for the MCI spanning several weeks in April/May 1996. It appears that the observed market activity process displays some seasonal patterns and reverts quickly back to a reference level. The observed market activity processes for the EWI and GDPI appear to be very similar and therefore are omitted.

To some degree, seasonal patterns are evident in the average level of market activity in addition to the average fluctuations. The presence of seasonal patterns within the fluctuations of market activity is further confirmed by the empirical quadratic variation $\langle \ln(m) \rangle_{\Delta,t}$ of the logarithm of the market activity. This is

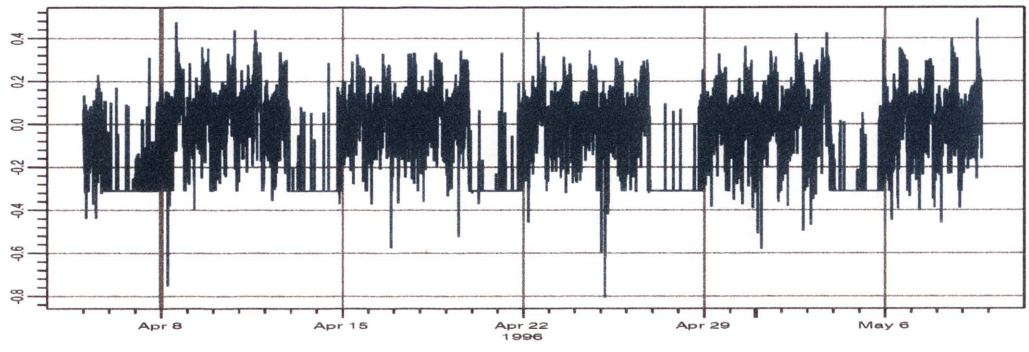

Figure 5.5: Logarithm of market activity, $\ln(m(t))$, for the MCI.

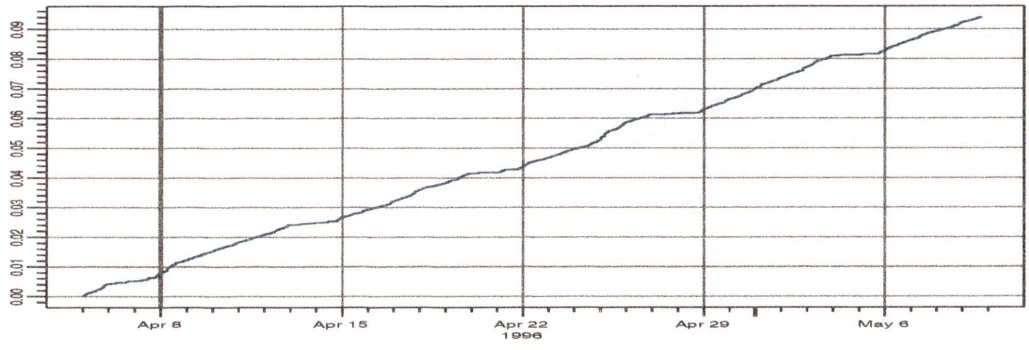

Figure 5.6: Quadratic variation of the logarithm of the market activity process, $\langle \ln(m) \rangle_{\Delta,t}$, for the MCI.

shown in Figure 5.6, where the period displayed is the same as in Figure 5.5. Here we see that the weekends are characterised by a plateau in the quadratic variation. Notably, despite the seasonal pattern, the empirical quadratic variation in Figure 5.6 appears to be almost linear for the periods when the markets are actively trading. This indicates market activity has a multiplicative diffusion coefficient for active trading periods. In addition, we are able to recover the GOP volatility from relation (4.6.7) from the observed market activity and normalised WSI.

In order to model the market activity it is necessary to characterise the relationship between its fluctuations and those of the normalised WSI. For each WSI, the empirical covariation of the square root of the normalised WSI with the logarithm

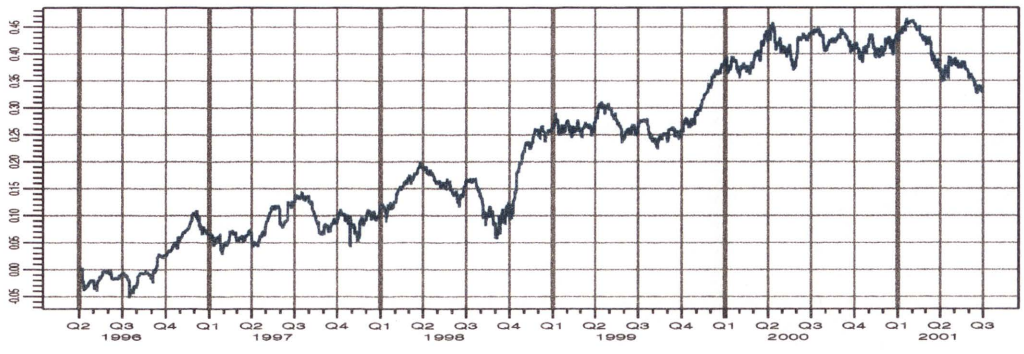

Figure 5.7: Covariation of the logarithm of the market activity with the square root of the normalised WSI, $\langle \ln(m), \sqrt{Y} \rangle_{\Delta,t}$, for the MCI.

of the corresponding market activity process is analysed.

The *covariation* $\langle L, K \rangle = \{\langle L, K \rangle_t, t \in [0,T]\}$ of two continuous processes $L = \{L(t), t \in [0,T]\}$ and $K = \{K(t), t \in [0,T]\}$, can be defined as the limit in probability

$$\langle L, K \rangle_t \stackrel{P}{=} \lim_{\Delta \to 0} \langle L, K \rangle_{\Delta,t} \tag{5.2.1}$$

where $\langle L, K \rangle_{\Delta,t}$ denotes the *approximate covariation*

$$\langle L, K \rangle_{\Delta,t} = \sum_{n=1}^{n_t} (L(t_n) - L(t_{n-1}))(K(t_n) - K(t_{n-1})) \tag{5.2.2}$$

for $t \in [0,T]$, where n_t is given in (2.1.2).

For the MCI, the empirical covariation is shown in Figure 5.7. We observe that the empirical covariation remains reasonably close to zero. There appears to be some evidence of a slight positive trend, although it seems reasonable to neglect this for the normalised WSIs when denominated in US Dollars. In Chapter 6 a possible method for refining the analysis is discussed, such that there is practically no covariation between the noise of the WSI and that of the resulting market activity process.

5.2.2 Model for Market Activity

The *expected market activity* $\hat{m}(t)$ at time t, which is defined as the expectation

$$\hat{m}(t) = E\left(m(t)\right), \tag{5.2.3}$$

is used to extract the seasonal pattern in the drift. For the full observation period, which covers 258 weeks in total, we use the law of large numbers to obtain an estimate for $\hat{m}(t)$ for each five minute interval from the arithmetic average of the corresponding observed market activity. This implies that the expectation of the market activity at time t is estimated by assuming the same pattern for each week and sampling over all weeks in the observation period. Figure 5.8 displays the estimated weekly pattern of the expected market activity $\hat{m}(t)$ when daylight savings time (DST) is taken into account. Red represents the expected summertime market activity in the Northern Hemisphere, whilst the black represents wintertime. Here we note that the daily pattern shown in Figure 5.8 is very similar for each working day irrespective of the DST period. Note, that Figure 5.8 is calculated from market activity and not its logarithm. The overall daily pattern is composed of several U-shaped patterns of different magnitude. The individual patterns are characteristic of localised, exchange-traded markets. The presence of similar U-shaped patterns in intraday return volatility is well documented. Wood, McInish & Ord (1985) and Harris (1986) provide early evidence of this pattern. This shape indicates that market activity is high at the open and close of trading and relatively low around the middle of the day. Pronounced intraday volatility patterns have also been documented in foreign exchange markets, see Müller, Dacorogna, Olsen, Pictet, Schwarz & Morgenegg (1990). As expected, in Figure 5.8, the magnitude of the average market activity is highest when the European and American markets are simultaneously active. It is lowest during the Pacific gap, that is, 21:00 to 00:00 GMT, and on the weekends.

Figure 5.6 suggests that the market activity $m(t)$ is likely to have multiplicative

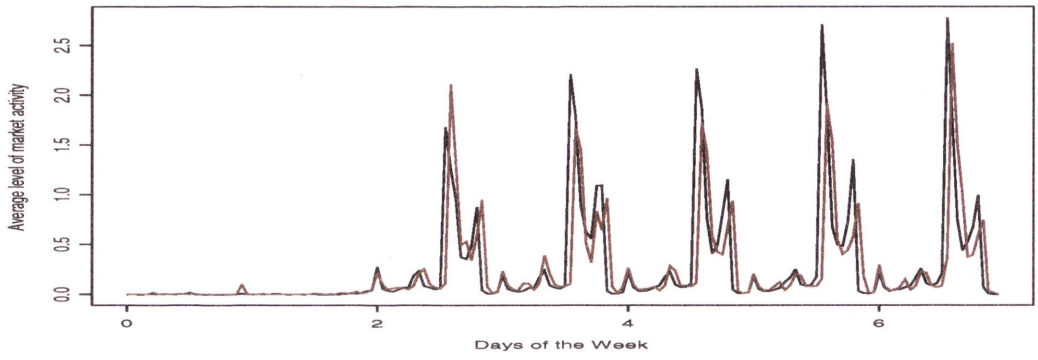

Figure 5.8: Weekly pattern of expected market activity $\hat{m}(t)$. Red represents summertime in the Northern Hemisphere, black represents wintertime in the Northern Hemisphere.

noise with some seasonal activity volatility since the empirical quadratic variation of its logarithm shows a piecewise linear pattern. A possible model for $m(t)$ with multiplicative noise is given by the SDE

$$dm(t) = A(m(t))\beta^2(t)dt + \beta(t)m(t)d\tilde{W}(t) \tag{5.2.4}$$

for $t \in [0, T]$, where $\tilde{W}$ denotes a standard Wiener process on the filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$ that satisfies the usual conditions. Here $A(\cdot)$ is a drift function that is specified in Section 5.2.3. The *activity volatility* $\beta(t) > 0$ is assumed to exhibit some weekly periodic seasonal patterns. Squared activity volatility is estimated by averaging over the weekly observations of $\beta^2(t)$, obtained from the slope of $\langle \ln(m) \rangle_{\Delta,t}$. Figure 5.9 shows the estimated values for $\beta(t)$ for the MCI. These estimates are obtained from five minute observations. Here we have set the activity volatility to zero during the weekends, since spurious observations do not allow a meaningful calculation of $\hat{\beta}(t)$ at these times. Note that during the active trading days the activity volatility is almost constant and slightly larger than one. We see that the activity volatility spikes when the Asia–Pacific markets open.

The seasonal activity volatility $\beta(t)$ allows us to introduce an alternative time

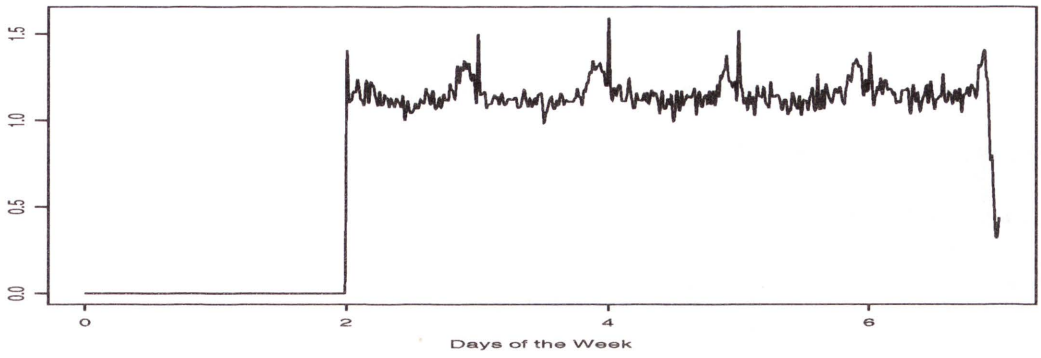

Figure 5.9: Average weekly pattern of activity volatility $\hat{\beta}(t)$.

scale. We define *activity volatility time* $\tau = \{\tau(t),\, t \in [0,T]\}$ to be

$$\tau(t) = \langle \ln(m) \rangle_t = \int_0^t (\beta(u))^2 \, du \tag{5.2.5}$$

for $t \in [0,T]$. Again, as with market activity time (4.6.8), the activity volatility time (5.2.5) requires normalisation.

It is useful to consider market activity in activity volatility time $m_{\tau(t)} = m(t)$. By (5.2.4) and (5.2.5), we obtain the SDE

$$dm_\tau = A(m_\tau)d\tau + m_\tau d\tilde{W}_\tau, \tag{5.2.6}$$

for $\tau \in [0, \tau(T)]$, where

$$d\tilde{W}_{\tau(t)} = \beta(t) d\tilde{W}(t) \tag{5.2.7}$$

for $t \in [0,T]$. Note that the drift function $A(\cdot)$ is the market activity drift in activity volatility time. In this new time scale, we capture the changes in activity volatility that occur both overnight and on weekends, as well as changes due to periods where there are unusually high levels of trading in the world financial markets.

By considering the logarithm of market activity in activity volatility time and using Itô's formula together with (5.2.6) we obtain the SDE

$$d\ln(m_\tau) = \tilde{A}(\ln(m_\tau))d\tau + d\tilde{W}_\tau \tag{5.2.8}$$

where

$$\tilde{A}(\ln(m_\tau)) = \left(\frac{1}{m_\tau} A(m_\tau) - \frac{1}{2}\right) \tag{5.2.9}$$

for $\tau \in [0, \tau(T)]$. Note that the diffusion coefficient in (5.2.8) is equal to one.

The empirical quadratic variation of the logarithm of market activity in activity volatility time is shown in Figure 5.10. It confirms the theoretical slope, which is equal to one, for the quadratic variation $\langle \ln(m) \rangle_\tau = \tau$ that follows from (5.2.8).

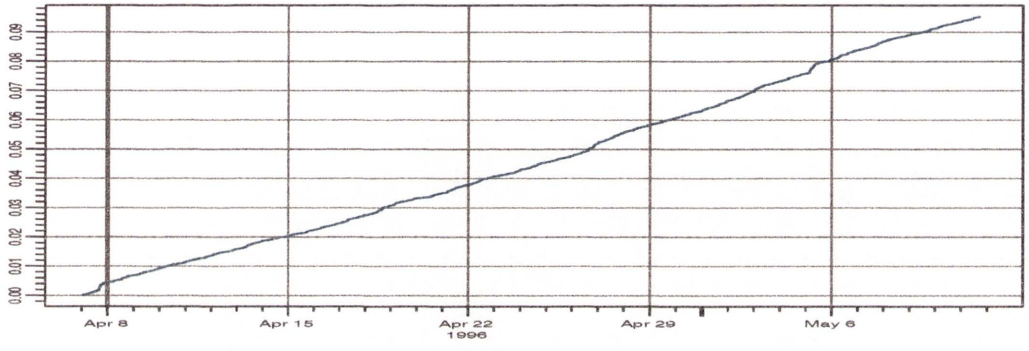

Figure 5.10: Empirical quadratic variation $\langle \ln(m) \rangle_{\Delta,\tau}$ of the logarithm of the market activity process in activity volatility time.

5.2.3 Estimation of Drift Function

The identification of the drift function of market activity remains. For the US Dollar denomination of the GOP the logarithm of market activity $l_\tau = \ln(m_\tau)$ has the empirical stationary density that is shown in Figure 5.11 via its histogram formed from 523,585 observations. For the mode and right tail of the histogram of l_τ we obtain an outstanding fit from the gamma density, which is shown in Figure 5.11. The gamma density for the logarithm of market activity is given by

$$\bar{p}_l(l; \gamma, p) = \frac{(\gamma)^p}{\Gamma(p)} \exp\left\{-\gamma e^l\right\} e^{l(p-1)} \tag{5.2.10}$$

for $\tau \in [0, \tau(T)]$, where $\gamma > 0$ is the *speed of adjustment*, $p > 0$ is the *reference level* and $\Gamma(\cdot)$ is the Gamma function. By the stationary solution of the Fokker–Plank equation, this implies a corresponding form of the drift function for the

logarithm of market activity given as

$$\tilde{A}(l_\tau) = \frac{\gamma}{2}\left(\frac{p-1}{\gamma} - e^{l_\tau}\right) \tag{5.2.11}$$

for $\tau \in [0, \tau(T)]$.

The speed of adjustment parameter γ and the reference level p, in (5.2.11), remain to be estimated. A very distinct deviation from the gamma density is shown in the histogram, where there is a concentration of negative spikes around -0.4. These negative spikes plus the thicker left tail, result from the sudden effects of market opening and closing. These are not included in the given dynamics of the model for market activity described by the SDEs (5.2.4), (5.2.6) and (5.2.8) when the reference level is assumed to be constant. For this reason we will ignore values less than $\underline{l} = -0.2$ for the estimation of the drift parameters.

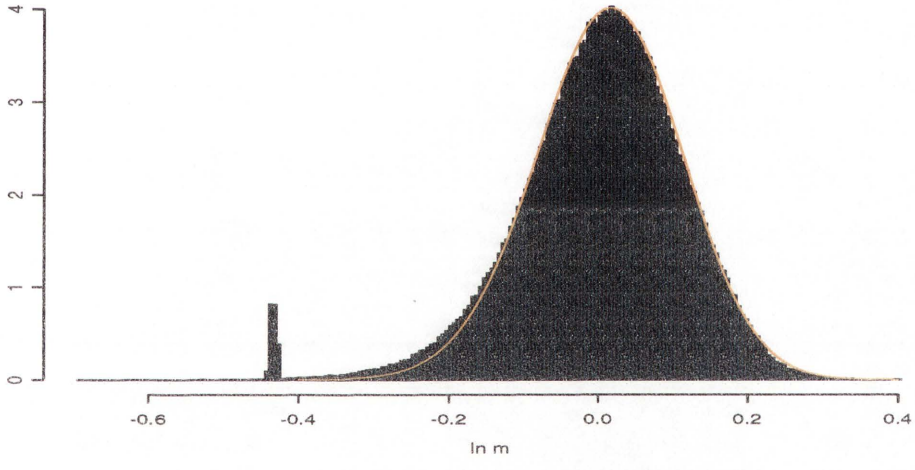

Figure 5.11: Histogram and estimated probability density function of $\ln(m_\tau)$.

In this case, we can perform maximum likelihood estimation. We exclude most of the distortions caused by the effects of market opening and closing from our estimation by forming a restricted log-likelihood function

$$L(\gamma, p) = \sum_{n=1}^{n_T} \mathbf{1}_{\{l_{\tau_n} \geq \underline{l}\}} \ln\left(\bar{p}_l(l_{\tau_n}; \gamma, p)\right). \tag{5.2.12}$$

Here $n_T = 523,585$ is the total number of observations and $\mathbf{1}_{\{l_{\tau_n} \geq \underline{l}\}}$ denotes the indicator function which takes only observations with $l_{\tau_n} \geq \underline{l}$ into account, $n \in \{1, 2, \ldots, n_T\}$. This condition results in $401,656$ observations used in the restricted maximum likelihood estimation. The plot of the estimated probability density function of $l_\tau = \ln(m_\tau)$ based on the resulting maximum likelihood estimates $\hat{\gamma} \approx 103$ and $\hat{p} \approx 106$ is shown in Figure 5.11. We see that the left tail of the histogram is fatter than what is given by the theoretical gamma density from the effects of market opening and closing. However, the right tail and mode of market activity are extremely well described by the gamma density. The estimates for each of the WSIs, together with 99% confidence intervals, are shown in Table 5.1.

Index	$\hat{\gamma}$	$\hat{p}$
MCI	103.2	105.8
	(89.3,116.4)	(92.3,119.2)
GDPI	136.6	138.9
	(120.7,152.5)	(122.5,155.3)
EWI	137.3	139.7
	(121.2,153.4)	(123.5,155.8)

Table 5.1: Estimates for the drift parameters of market activity in activity volatility time with 99% confidence intervals.

5.3 Normalised WSIs in Market Activity Time

In Section 4.6 it was shown that the normalised GOP, when observed in market activity time, is a square root process of dimension four. As explained in the previous chapters, we concentrate entirely on inference from the diffusion coefficient of the normalised WSIs. The trajectory of the normalised MCI in market activity time is shown in Figure 5.12. By considering market activity time, we know that the quadratic variation of the square root of the normalised MCI should be linear

with a slope close to 0.25, see (4.6.14). This relationship is confirmed by performing a simple linear regression of the empirical quadratic variation of the square root of the normalised WSIs in market activity time against the corresponding estimated market activity time. The slope coefficient and R^2 value are given in Table 5.2 for each of the WSIs considered.

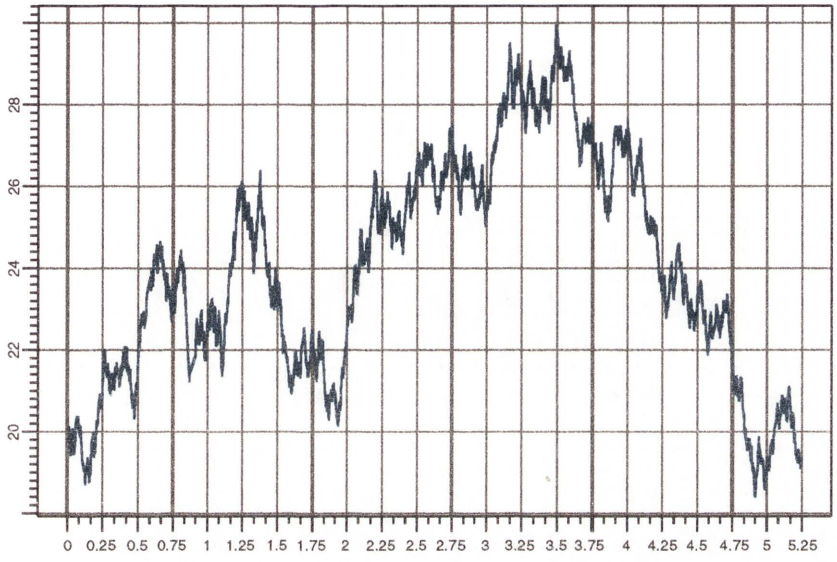

Figure 5.12: Normalised MCI in market activity time.

Index	Slope Coefficient	R^2
MCI	0.245	0.9968
GDPI	0.232	0.9969
EWI	0.198	0.9919

Table 5.2: Slope coefficients and R^2 values for the empirical quadratic variation of the square root of normalised WSIs in market activity time against market activity time.

Observation of the normalised WSIs in market activity time should also remove the seasonal patterns from the absolute returns of the normalised WSIs, see Figure 5.4. We show the sample autocorrelation function of the hourly absolute returns of the normalised MCI in market activity time in Figure 5.13. It is clear,

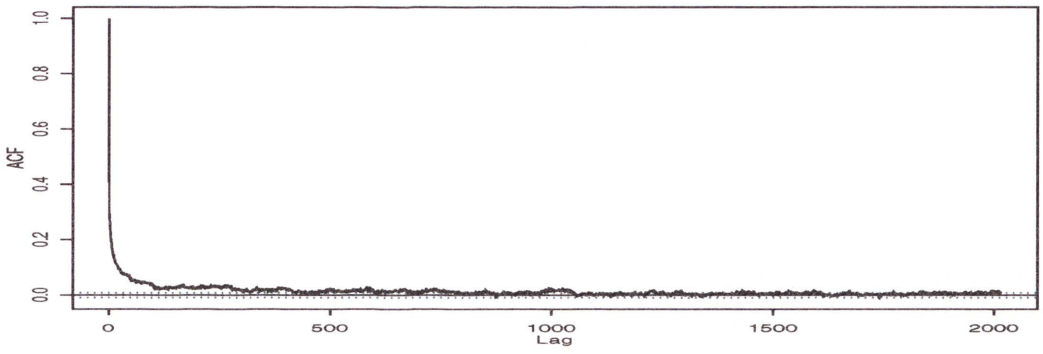

Figure 5.13: Sample autocorrelation function of the hourly absolute returns of the normalised MCI in market activity time.

when comparing the results with those in Figure 5.4, that the seasonal pattern has been largely removed.

This makes the model derived and calibrated above, a largely accurate intraday description of the dynamics of the discounted world stock market index denominated in US Dollars and thus the discounted GOP. The above methodology yields similar results for a range of diversified indices. Hence, all three WSIs can be interpreted as reasonable approximations of the GOP. Table 5.2 shows that the MCI and GDPI fit the model equally well and both outperform the EWI when the quadratic variation processes of the square root of the corresponding normalised index are considered. Additionally, the corresponding market activity processes in activity volatility time are shown to have the hypothesised dynamics when effects of the opening and closing of trading are omitted. It follows from Figure 4.3 that the MCI has the long term maximum value with respect to all other indices considered. This reflects the main property of the GOP. Consequently, we consider the MCI to be the best proxy for both the GOP and the world stock portfolio, on the basis of statistical results and for economic reasons. That is, the MCI can also be considered as the market portfolio, which provides a link to the classical Capital Asset Pricing Model (CAPM), see Merton (1992).

It must be emphasised that the intraday benchmark model centres upon market activity as its parameter process. Market activity models the short term fluctuations of the squared GOP volatility of the security dynamics. The inverse of the normalised GOP, see (4.6.7), captures the medium and long term fluctuations of squared GOP volatility. This is in contrast to the parametrisation used in the Black–Scholes–Merton model, which is built upon volatility as the sole parameter process of interest.

Chapter 6

Intraday Analysis of the MCI in Different Currency Denominations

Within the benchmark framework, the structure of the dynamics for the GOP is unaffected by the chosen units of denomination. To investigate this feature, we examine twenty–one currency denominations of the MCI in this chapter. Properties of the market activity processes that result from different denominations of the MCI are examined in Section 6.1. Section 6.2 considers the co–movements of the normalised GOPs and Section 6.3 considers the pairwise relationships between the market activity processes.

6.1 Market Activity of the MCI

6.1.1 The MCI in Different Denominations

As the MCI is the best proxy for the GOP, the remaining analysis is conducted only on the MCI as the underlying index. Here, we focus on the empirical behaviour of market activity, derived from the MCI denominated in currencies other than US Dollars. The high–frequency MCI in different currencies is formed from the US Dollar MCI by

$$S^{(i,\delta_{(MCI)})}(t) = S^{(USD,\delta_{(MCI)})}(t)\, X^{(i,USD)}(t), \tag{6.1.1}$$

where $X^{(i,USD)}(t)$ is the spot price of one US Dollar at time t when measured in units of the ith currency. In general, the (i,j)th foreign exchange rate at time t can be expressed as the following ratio of GOPs

$$X^{(i,j)}(t) = \frac{S^{(i,\delta_*)}(t)}{S^{(j,\delta_*)}(t)}, \tag{6.1.2}$$

see Platen (2001). Here $S^{(i,\delta_*)}(t)$ and $S^{(j,\delta_*)}(t)$ denote the value at time t of the GOPs denominated in terms of currencies i and j, respectively. Note that the discounted MCI $\bar{S}^{(i,\delta_{(MCI)})}(t)$ at time t is formed by

$$\bar{S}^{(i,\delta_{(MCI)})}(t) = \frac{S^{(i,\delta_{(MCI)})}(t)}{S^{(i,0)}(t)}, \tag{6.1.3}$$

where $S^{(i,0)}(t)$ is the savings account for the ith currency. This depends on the short rate for the ith market at time t, denoted by $r^i(t)$.

The MCI in twenty–one different currency denominations is shown in Figure 6.1. On average, the majority of the denominations of the MCI appear to grow exponentially. As in Chapter 5, an exponential function is used to compensate for the observed long term growth to form the normalised MCI.

6.1.2 Empirical Analysis of Market Activity

The benchmark model outlined in Chapter 4 models the dynamics of the GOP in an idealised framework. Specifically, the assumption that the normalised GOP has the dimension four, is likely to be satisfied only in the case of the most developed markets. However, the benchmark approach is flexible. The dimension of the square root process is specific to individual markets and is permitted to take other values greater than two, see Platen (2001). Circumstances that may result in dimensions other than four include the possible inefficiencies within emerging markets, or when the leverage effect is not reflected in the reactions of the market to new information. In such cases the GOP is formed by a power of a square root process. Furthermore, market activity is assumed to be uncorrelated with the noise driving the normalised GOP. To examine whether the dimension of the MCI when denominated in each currency is four, we consider the empirical covariation of the square root of the normalised MCI with the logarithm of market activity, but calculated as if the dimension is four. If the dimension of the normalised MCI is four for a given currency denomination, then the empirical covariation $\langle \ln(m^i), \sqrt{Y^i} \rangle_{\Delta,t}$ should remain close to zero. From Figure 6.2, it is clear that with

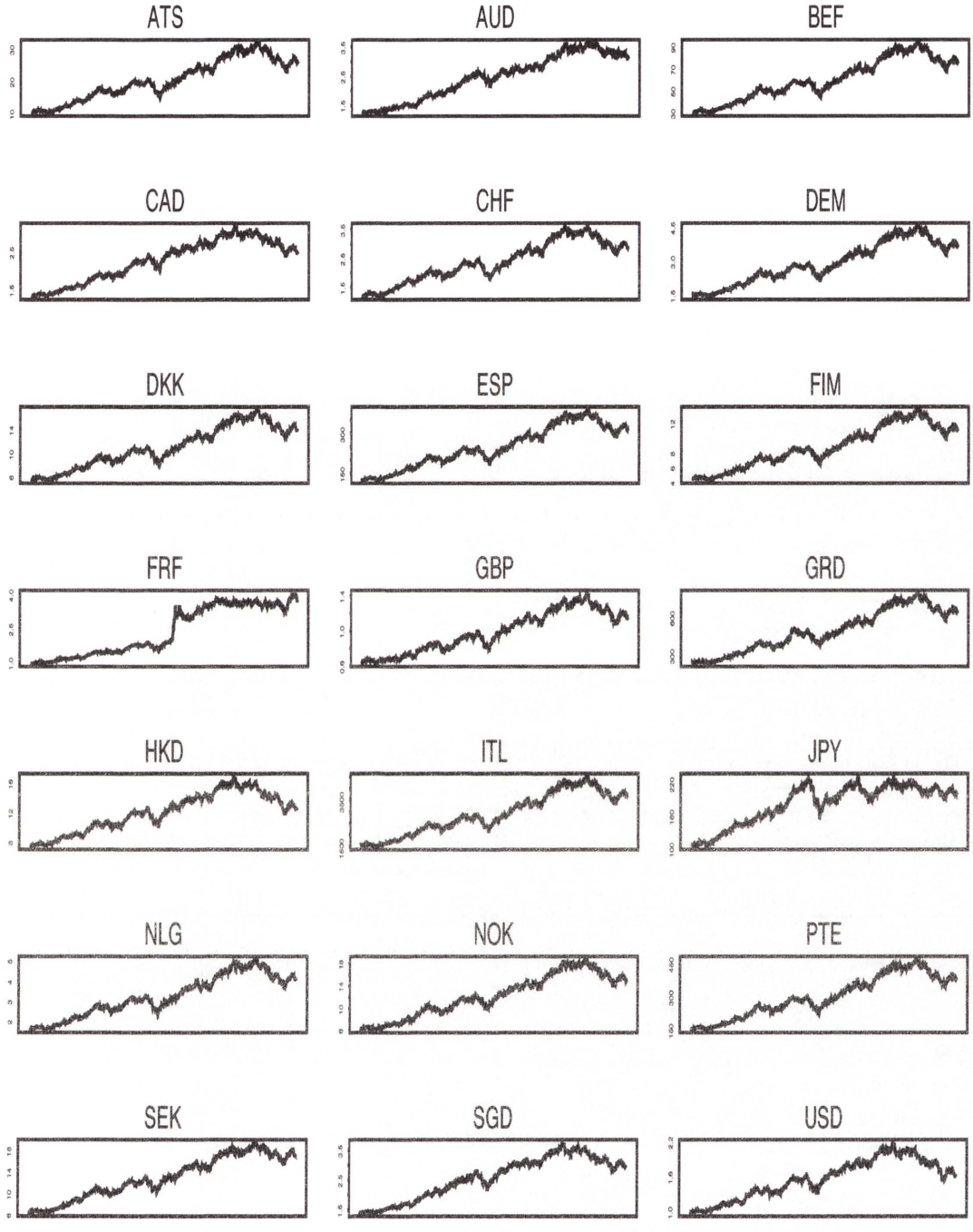

Figure 6.1: The MCI denominated in units of twenty–one different currencies for the five year period.

the exception of the US Dollar denomination, which shows a slight positive trend in the empirical covariation but remains close to zero, the empirical covariations for the remaining twenty indices are not close to zero. More precisely, periods with certain trends are apparent.

In such cases, we transform the denomination of the MCI by a power function such that the resulting covariation between the square root of the transformed MCI and the logarithm of the resulting market activity remains relatively close to zero. The power function transform reflects the minimal market model proposed in Platen (2001) where powers of square root processes are considered to model the normalised GOP.

General Form of the Normalised GOP

Suppose we form the normalised GOP Z^i_ψ in market activity time for each denomination i using a square root process Y^i_ψ via the power transform

$$Z^i_\psi = \left(Y^i_\psi\right)^{\frac{\nu^i}{2}-1}, \tag{6.1.4}$$

for $i \in \{USD, CHF, \ldots\}$ and $\psi^i \in [0, \psi^i(T)]$, see Platen (2001) and Heath, Hurst & Platen (2001). Here ν^i denotes the dimension for the ith denomination. Note also that for each currency i, a corresponding market activity time exists denoted by ψ^i. However, for ease of exposition the index i is omitted from the time scale.

It can be shown using the Itô formula together with (6.1.4) and (4.6.10), that the *general normalised GOP* $Z^i = \{Z^i_\psi, \psi^i \in [0, \psi^i(T)]\}$ satisfies the SDE

$$dZ^i_\psi = \left(\frac{\nu^i}{2} - 1\right)\left(Z^i_\psi\right)^{\frac{\nu^i-4}{\nu^i-2}}\left(\frac{\nu^i}{4} - \eta^i\left(Z^i_\psi\right)^{\frac{2}{\nu^i-2}}\right)d\psi + \left(\frac{\nu^i}{2} - 1\right)\left(Z^i_\psi\right)^{\frac{\nu^i-3}{\nu^i-2}} d\hat{W}_\psi \tag{6.1.5}$$

for $i \in \{USD, CHF, \ldots\}$. Note that the dependence of market activity time ψ on the denomination i is suppressed for ease of notation. It is clear that when $\nu^i = 4$, the SDE in (6.1.5) simplifies to that given in (4.6.10). Recall, that to obtain the market activity process we calculate the slope of the quadratic variation of the

process $\sqrt{Y^i_\psi}$. Note, by (6.1.5) and the Itô formula, the square root of the general normalised GOP $\sqrt{Z^i_\psi}$ satisfies the SDE

$$d\sqrt{Z^i_\psi} = \frac{1}{2}\left(\frac{\nu^i}{2}-1\right)\left(Z^i_\psi\right)^{\frac{\nu^i-6}{2(\nu^i-2)}}\left(\frac{1}{4}\left(\frac{\nu^i}{2}+1\right)-\eta^i\left(Z^i_\psi\right)^{\frac{2}{\nu^i-2}}\right)d\psi \qquad (6.1.6)$$
$$+\frac{1}{2}\left(\frac{\nu^i}{2}-1\right)\left(Z^i_\psi\right)^{\frac{\nu^i-4}{2(\nu^i-2)}}d\hat{W}_\psi$$

for $\psi^i \in [0, \psi^i(T)]$. Thus, the quadratic variation of $\sqrt{Z^i_\psi}$ is

$$\left\langle \sqrt{Z^i} \right\rangle_{\psi(t)} = \frac{1}{4}\int_0^{\psi(t)}\left(\frac{\nu^i}{2}-1\right)^2\left(Z^i_\psi\right)^{\frac{\nu^i-4}{\nu^i-2}}d\psi. \qquad (6.1.7)$$

This quadratic variation is only independent of Z^i if the dimension ν^i is exactly four. If we were to attempt to calculate market activity based on (6.1.7) by applying the relationship in (4.6.15), the resulting process would be related to the process Z^i. Thus, we expect to see some covariation if the dimension of the observed normalised GOP is not exactly four, as displayed in Figure 6.2.

We must find the correct power, which is equivalent to choosing the correct dimension ν^i, to ensure that market activity is not correlated to the noise that drives the normalised GOP. That is, from (6.1.4), we must find the dimension ν^i

$$Y^i_\psi = \left(Z^i_\psi\right)^{\frac{2}{\nu^i-2}}$$

for $i \in \{USD, CHF, \ldots\}$, such that the covariation between the square root of the observed process Z^i and the logarithm of its market activity is minimised. Note that the dimension ν^i is not necessarily constant. However, we assume throughout this chapter that ν^i is a piecewise constant function of time, which simplifies the analysis. From the empirical covariation processes shown in Figure 6.2, it appears that the quantity ν^i could be conveniently described as a continuous time Markov chain.

We deduce the dimension ν^i through a basic analysis of the empirical covariation processes in Figure 6.2. Figure 6.3 shows the estimated piecewise constant dimensions as a function of time for each denomination. The values for ν^i range

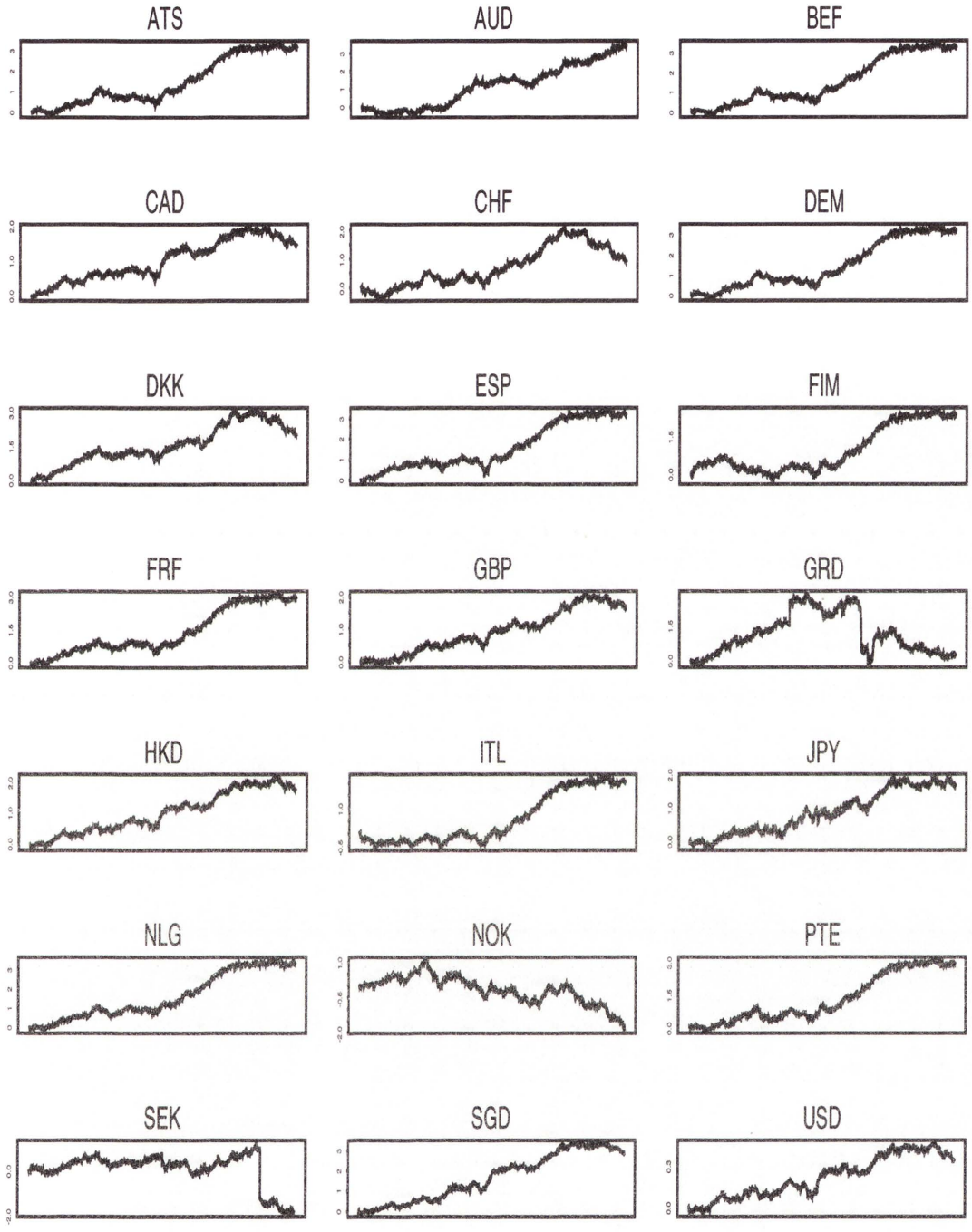

Figure 6.2: The empirical covariation of the square root of the normalised GOP and the logarithm of market activity assuming that the dimension of the normalised GOP is exactly four.

from 2.5 to 6.0 and are on average close to four. Further research is required to estimate the changes in ν^i more objectively.

The resulting logarithm of the market activity process for the different denominations of the MCI for two weeks in April 1996 are shown in Figure 6.4. This segment of the logarithm of market activity is typical of the behaviour for the five year period. In Chapter 5, it was demonstrated that the market activity of three different WSIs, denominated in US Dollars, behaves in a similar fashion despite the different methods used in their construction. Namely, all three WSIs were shown to result in a market activity that appears to be a strongly mean-reverting process displaying seasonal patterns. We see that all denominations of market activity considered behave in the same manner. In particular, it appears that the observed market activity processes exhibit seasonal patterns and revert rapidly back to a reference level. Viewing these segments of market activity illustrates the seasonality apparent in the average level of market activity and to some extent, the seasonality apparent in the fluctuations of market activity.

The presence of seasonal patterns is confirmed by the empirical quadratic variation of the logarithm of market activity shown in Figure 6.5, which is plotted for the same segments as per Figure 6.4. We see that the weekends are characterised by a plateau in the empirical quadratic variation. Despite this, for periods when the market is actively trading, the empirical quadratic variation again appears to be almost linear. This indicates that the assumption of multiplicative noise is reasonable for the twenty–one currency denominations considered.

We can again extract the seasonal patterns in market activity. The expected market activity at time t is defined as the expectation

$$\hat{m}^i(t) = E\left(m^i(t)\right) \tag{6.1.8}$$

as discussed in Chapter 5, see (5.2.3). The weekly expected market activity pattern is shown for summertime in the Northern Hemisphere for each currency denomination in Figure 6.6. Notably, for the majority of denominations we see a

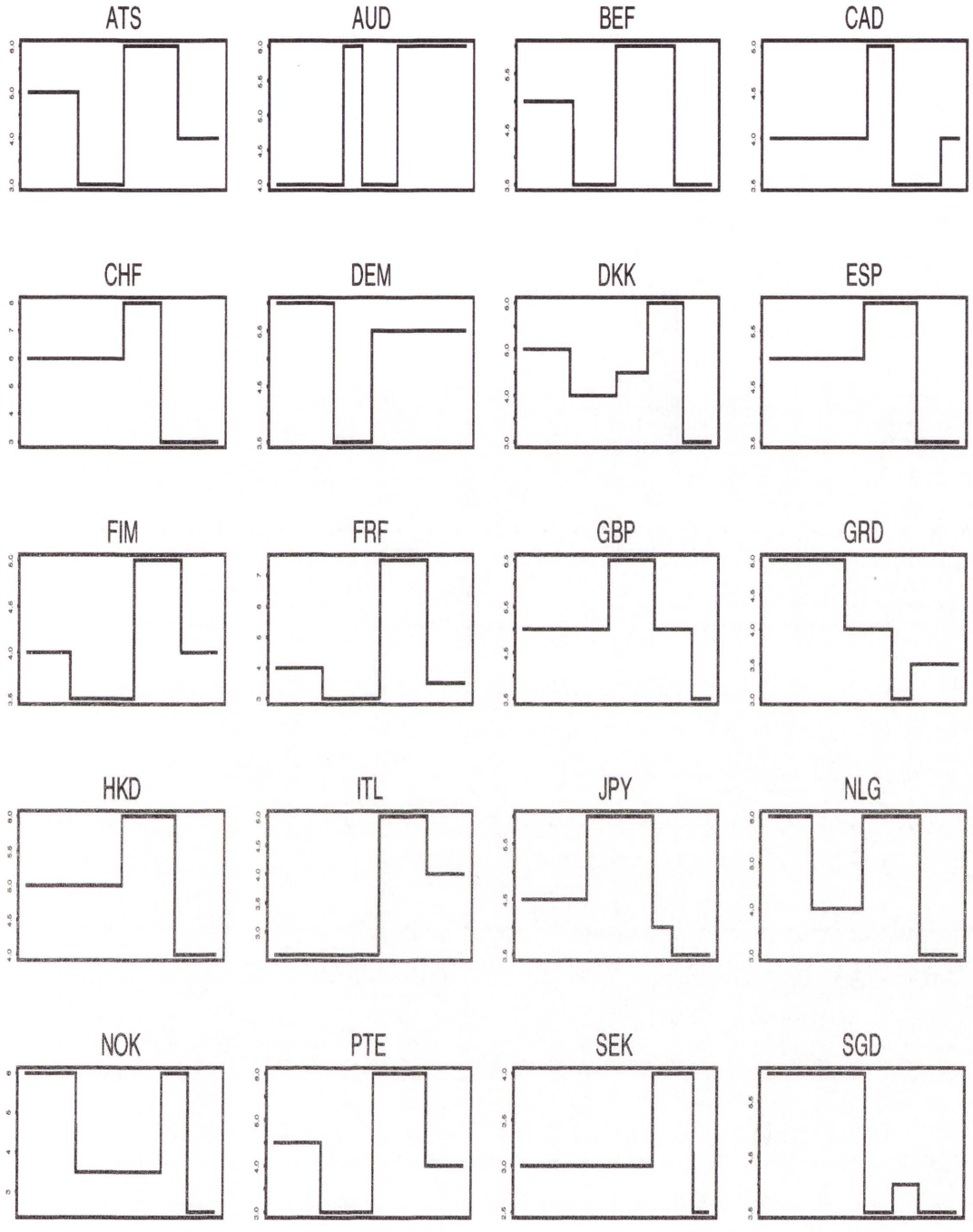

Figure 6.3: Piecewise constant dimension ν^i of the normalised MCI as a function of time.

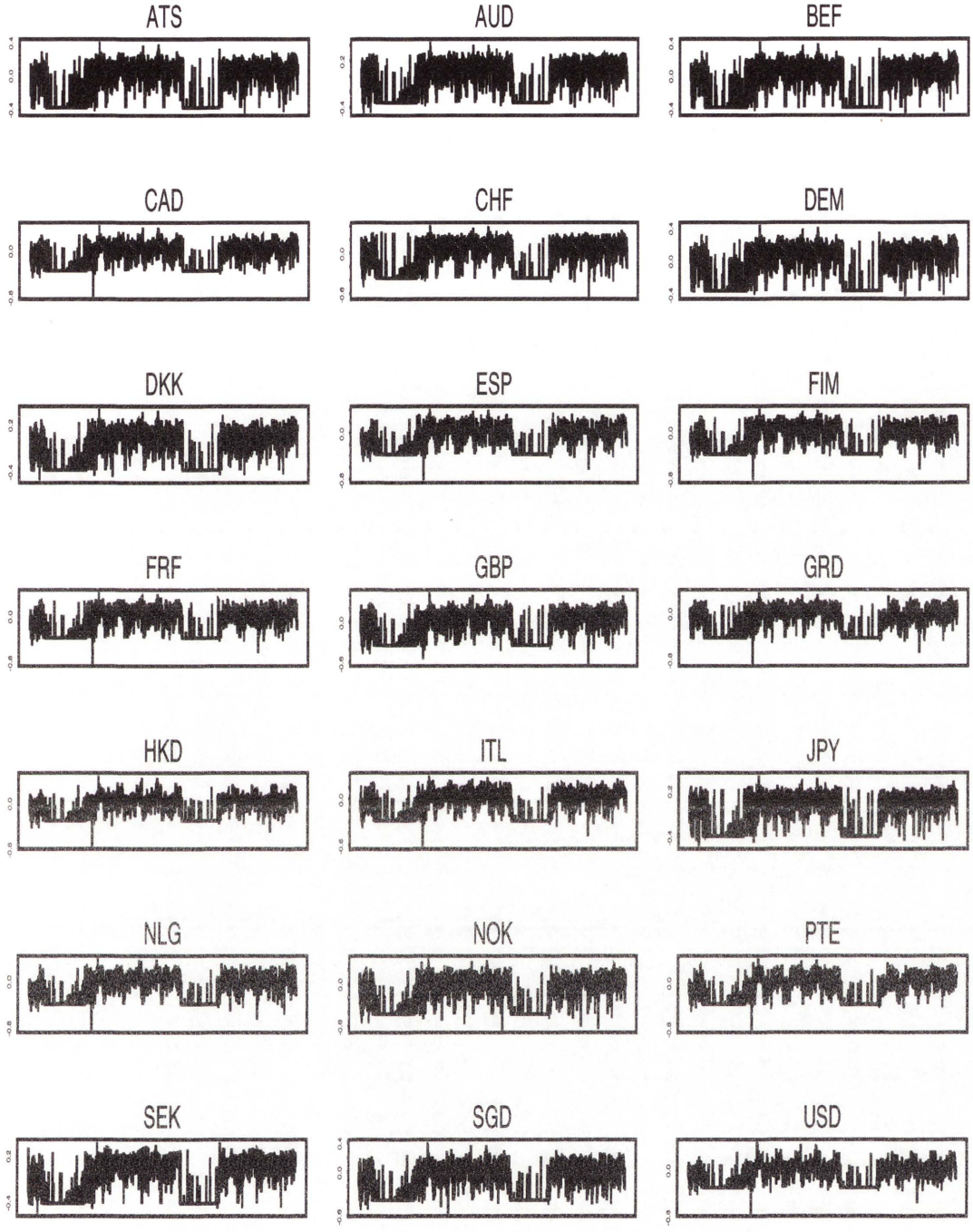

Figure 6.4: The logarithm of market activity denominated in twenty–one different currencies for two weeks during April 1996.

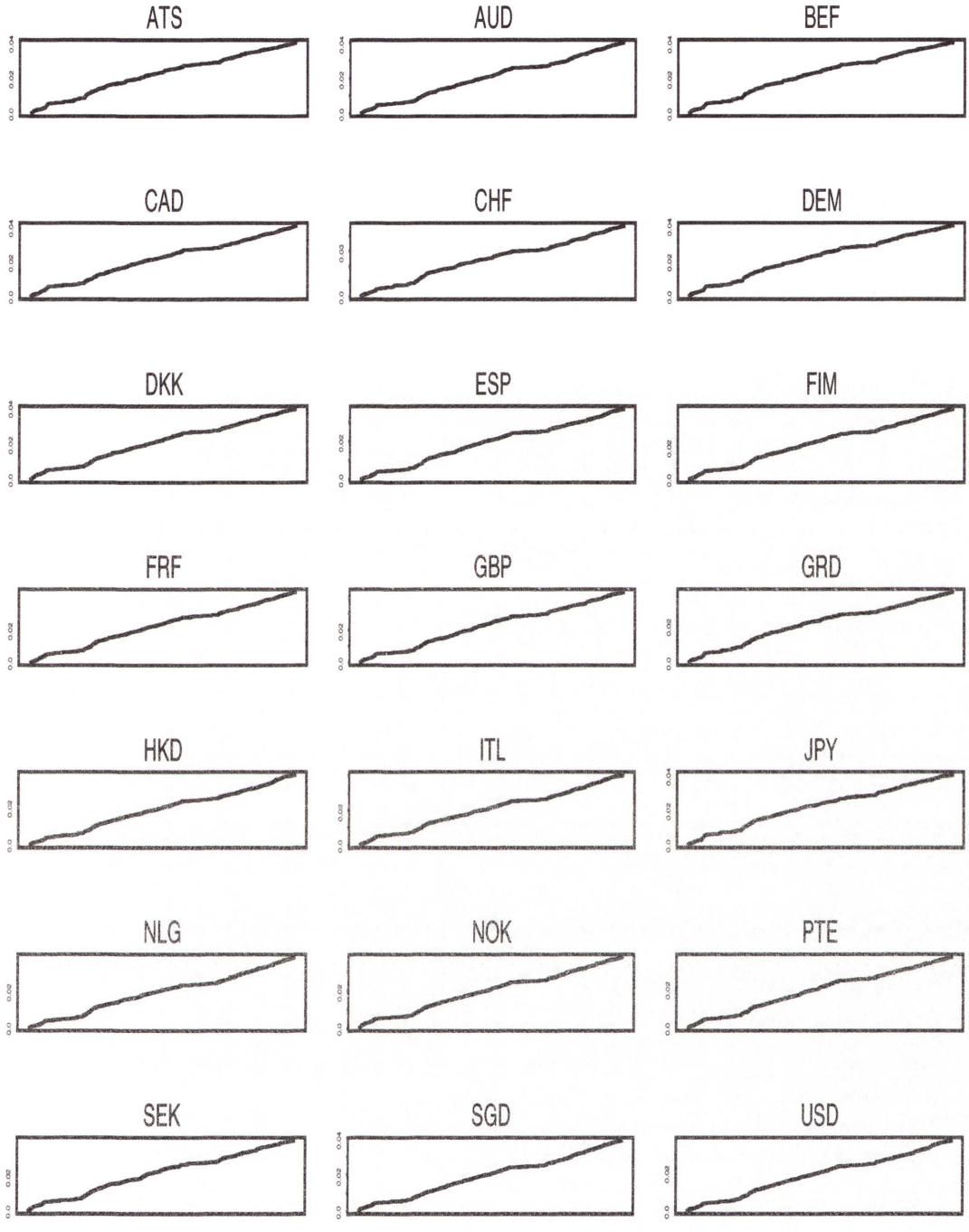

Figure 6.5: Empirical quadratic variation of the logarithm of market activity denominated in twenty–one different currencies for two weeks during April 1996.

largely clean pattern with five spikes that demonstrate the opening and closing of the different regional markets. This pattern is typically referred to as the U-shaped pattern and was originally seen for the return volatility of exchange traded instruments, see Dacorogna, Gençay, Müller, Olsen & Pictet (2001) and references therein. Again, market activity is lowest on weekends and during the Pacific gap. This clean pattern is not quite so evident for some of the less globally traded currencies.

Activity volatility is obtained as described in Section 5.2. It also has a weekly pattern and is estimated by averaging over the weekly observations of the squared activity volatility $(\beta^i(t))^2$. We show in Figure 6.7, the average activity volatility $\hat{\beta}^i(t)$ for the twenty–one denominations of the MCI. Again there is always a spike when the Asia–Pacific markets have opened, which occurs after the Pacific gap. The calculation of activity volatility time, as described in Chapter 5, allows us to observe market activity on this additional time scale. A simple linear regression of the empirical quadratic variation of the logarithm of market activity, in terms of activity volatility time, against activity volatility time, was performed. The resulting R^2 values are shown in Table 6.1 and confirm that the diffusion coefficient is practically constant.

6.1.3 Parametric Drift Estimation

In Figure 6.8, the histograms of all observations from the individual denominations of the logarithm of the market activity $l^i_\tau = \ln(m^i_\tau)$ are shown. One notes a very distinct shape of the stationary density but also a concentration of negative spikes around -0.4 to -0.6, depending on the currency. These result from markets opening and closing, which have not been accounted for in the model for market activity that is employed. Here we assume that irrespective of the denominating currency, that market activity has as stationary density, a gamma density.

The stationary density of the logarithm of the market activity l^i_τ in activity

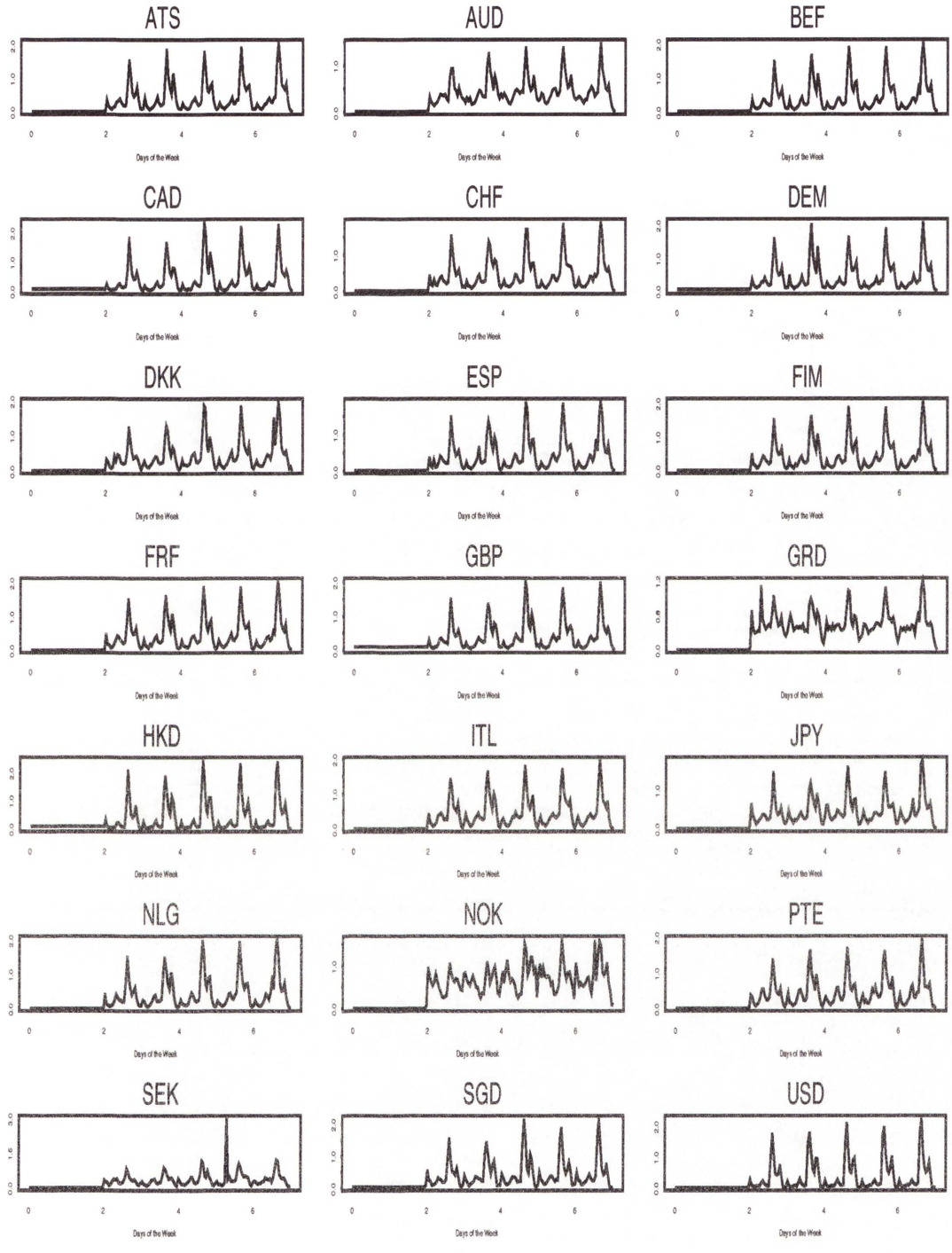

Figure 6.6: The weekly pattern of expected market activity $\hat{m}^i(t)$ for each currency for summertime in the Northern Hemisphere.

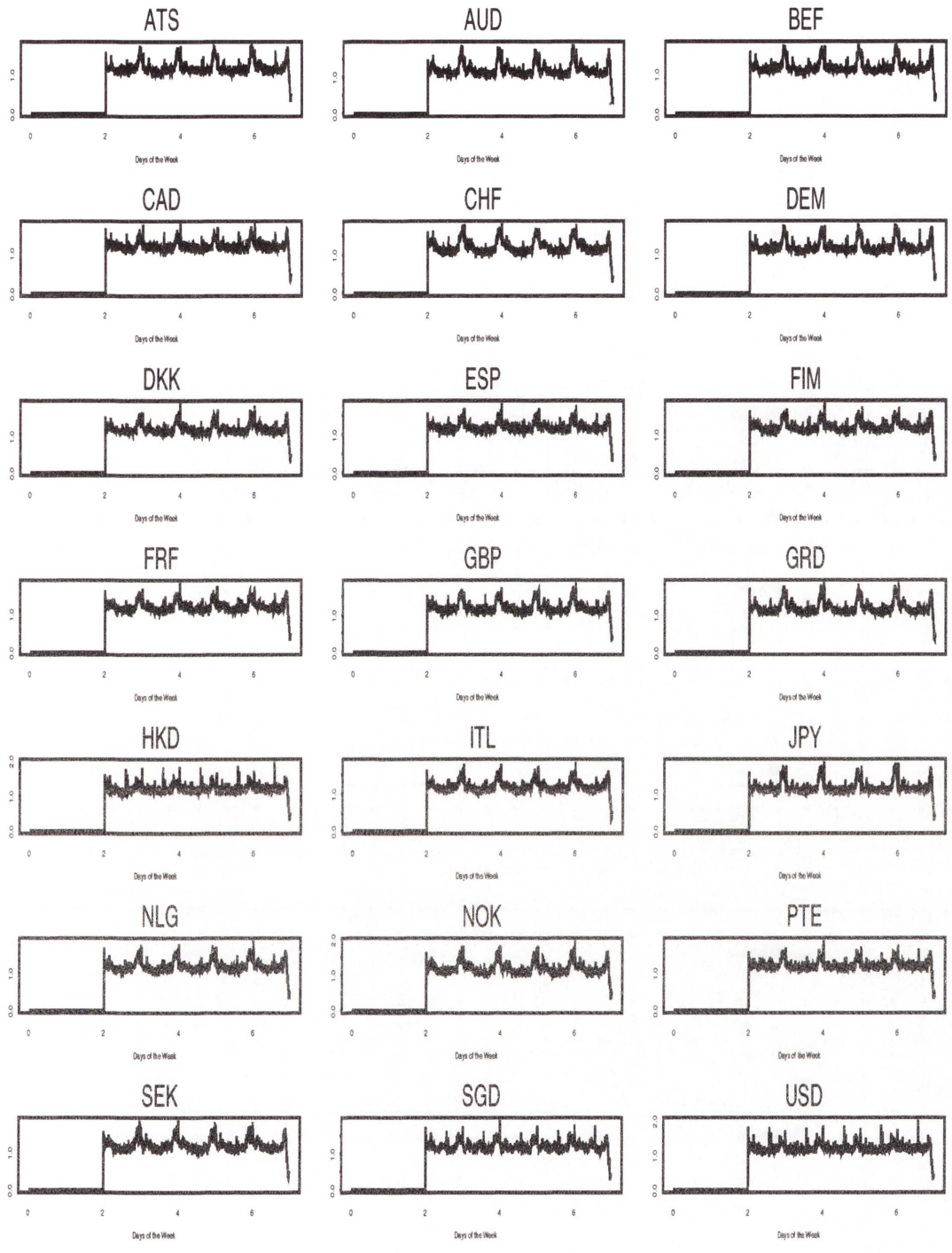

Figure 6.7: Average weekly pattern of activity volatility $\hat{\beta}^i(t)$ for twenty–one currency denominations.

i	R^2	i	R^2	i	R^2
ATS	0.9997	AUD	0.9999	BEF	0.9996
CAD	0.9999	CHF	0.9995	DEM	0.9997
DKK	0.9998	ESP	0.9998	FIM	0.9997
FRF	0.9998	GBP	0.9995	GRD	0.9997
HKD	0.9999	ITL	0.9998	JPY	0.9994
NLG	0.9998	NOK	0.9998	PTE	0.9999
SEK	0.9998	SGD	0.9998	USD	0.9999

Table 6.1: Table of R^2 values from the regression of the empirical quadratic variation of the logarithm of market activity against activity volatility time.

volatility time can be written as

$$\bar{p}_l(l^i;\gamma^i,p^i) = \frac{(\gamma^i)^{p^i}}{\Gamma(p^i)} \exp\left\{-\gamma^i e^{l^i}\right\} e^{l^i(p^i-1)}.$$

Here $\Gamma(\cdot)$ is the Gamma function and $\gamma^i > 0$ and $p^i > 0$ are the only free parameters for each currency i.

As in Chapter 5, we can perform a restricted maximum likelihood estimation. A straightforward maximum likelihood estimation would be adversely affected by the distortions in the negative tail of the histogram. Therefore, we exclude most of the distortions from our estimation by forming the restricted log-likelihood function for each currency denomination,

$$L(\gamma^i,p^i) = \sum_{n=1}^{n_T} \mathbf{1}_{\{l^i_{\tau_n}\geq \underline{l}\}} \ln \bar{p}_l(l^i_{\tau_n};\gamma^i,p^i).$$

Here $n_T = 523,585$ and $\mathbf{1}_{\{l^i_{\tau_n}\geq \underline{l}\}}$ denotes the indicator function which allows only observations with $l^i_{\tau_n} \geq \underline{l}$. The restricted maximum likelihood technique was applied using $\underline{l} = -0.2$. A plot of the estimated gamma density function of $\ln(m^i_\tau)$ based on the resulting maximum likelihood estimate of γ^i and p^i is shown in Figure 6.8 for $i \in \{USD, CHF, \ldots\}$. We see that the effects caused by market

opening and closing make the left tail of the histogram fatter than what would be expected if market activity was a gamma distributed random variable. Although, note that the right tail and mode of the histogram is well described by the gamma density for most denominations. The estimates of γ^i and p^i together with the 99% confidence intervals are shown in Table 6.2.

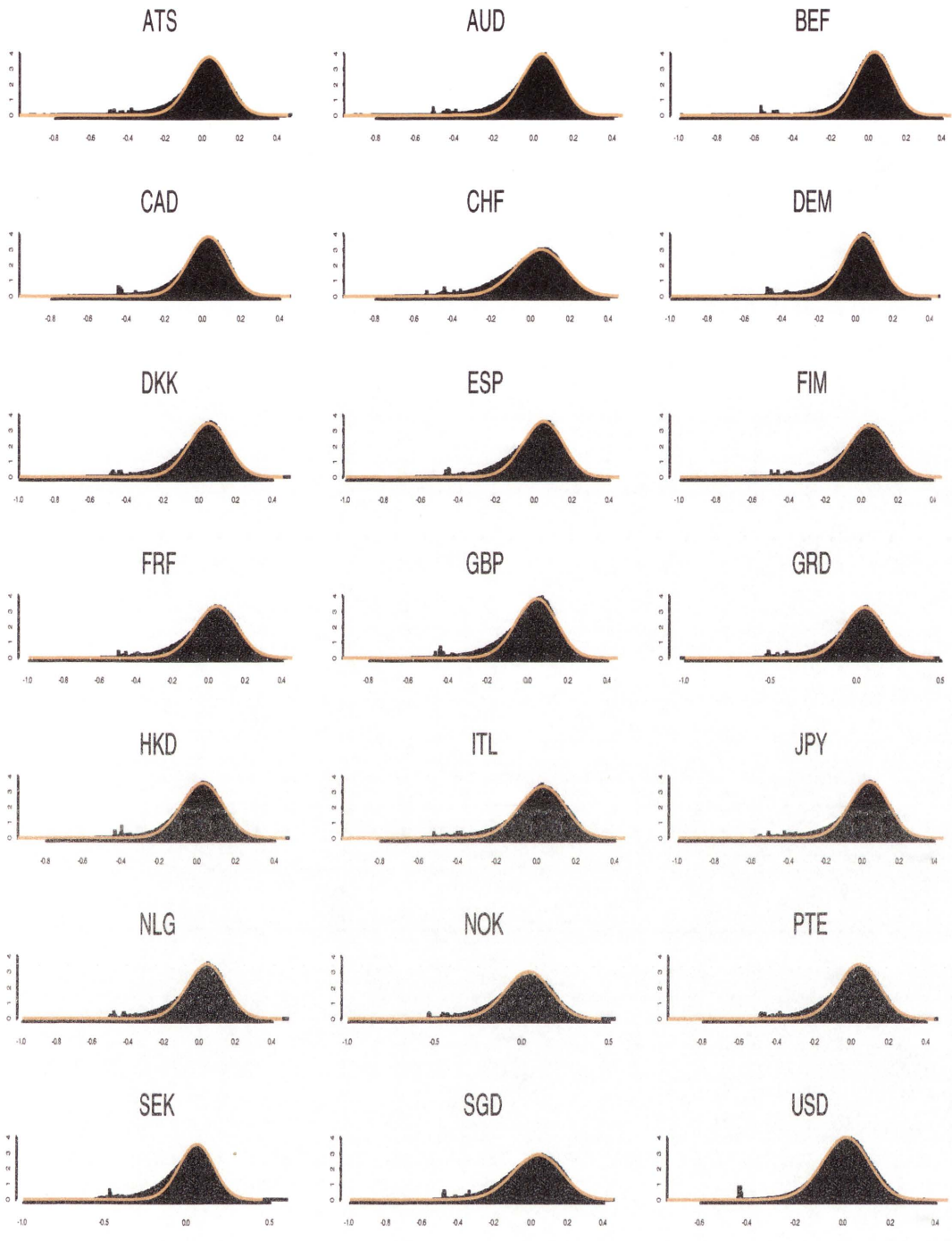

Figure 6.8: Histogram and estimated probability density function $\ln(m_\tau^i)$.

i	$\hat{\gamma}^i$	$\hat{p}^i$	i	$\hat{\gamma}^i$	$\hat{p}^i$	i	$\hat{\gamma}^i$	$\hat{p}^i$
ATS	92	96	AUD	103	108	BEF	108	112
	(80.2,103.8)	(82.4,109.5)		(89.4,118.6)	(97.4,118.6)		(96.8,119.2)	(98.4,125.6)
CAD	95	99	CHF	70	73	DEM	102	107
	(83.2,106.8)	(85.4,112.5)		(56.4,83.5)	(62.4,83.6)		(90.8,113.2)	(93.4,120.8)
DKK	80	85	ESP	85	90	FIM	75	79
	(68.2,91.8)	(71.4,98.6)		(71.4,98.5)	(79.4,100.6)		(63.8,86.2)	(65.4,92.5)
FRF	73	77	GBP	96	101	GRD	70	74
	(61.2,84.7)	(63.4,90.6)		(82.4,109.5)	(90.3,111.6)		(58.8,81.2)	(60.4,87.6)
HKD	83	86	ITL	73	76	JPY	87	91
	(71.2,94.8)	(72.4,99.6)		(59.4,86.6)	(65.3,86.6)		(75.8,98.2)	(77.4,104.6)
NLG	80	84	NOK	59	62	PTE	79	83
	(68.2,91.8)	(70.4,97.5)		(45.4,72.6)	(51.4,72.6)		(51.4,72.6)	(69.4,96.6)
SEK	79	84	SGD	55	58	USD	103	106
	(67.2,90.8)	(70.4,97.6)		(41.4,68.6)	(47.4,69.2)		(89.3,116.4)	(92.3,119.2)

Table 6.2: Estimates of the speed of adjustment and reference level parameters in addition to 99% confidence intervals.

6.1.4 Non–parametric Drift Estimation

It is clear from Figure 6.8, that it may not always be appropriate to assume that market activity has a gamma density as stationary density. There are some discrepancies, for example, the denomination of market activity in terms of the Norwegian Krona (NOK), where it appears that the estimated gamma probability density is more peaked than the histogram of the data. Furthermore, Platen, West & Breymann (2004) find that market activity, when denominated in terms of half–hourly Australian electricity spot prices, has a stationary density that is best described by the inverse gamma density. This indicates that the assumed drift dynamics in (5.2.11) for the logarithm of market activity may be different for commodity denominations of a WSI. To make the model more flexible, we estimate the unknown stationary density of the logarithm of market activity using non–parametric methods.

This allows us to deduce the drift function from the stationary density of $l_\tau^i = \ln(m_\tau^i)$ since the diffusion function is fixed. Using the solution of the stationary Fokker–Plank equation, the stationary density for the logarithm of market activity l^i can be written as

$$\vec{p}_l^{\,i}(l^i) = C \exp\left\{ 2 \int_{l_0^i}^{l^i} \tilde{A}^i(u^i) du^i \right\}, \tag{6.1.9}$$

where $\tilde{A}^i(\cdot)$ is the drift function associated with the logarithm of market activity for each denomination i, see (5.2.8). Here $l_0^i \in \mathbb{R}$ is an arbitrary point and the constant C results from the normalisation condition

$$\int_{-\infty}^{\infty} \vec{p}_l^{\,i}(l^i) dl^i = 1.$$

By taking the derivative of the natural logarithm of the stationary density, the drift function can be directly obtained as

$$\tilde{A}^i(l^i) = \frac{1}{2} \frac{d}{dl^i} \ln\left(\vec{p}_l^{\,i}(l^i)\right).$$

Since there is enough data available, the stationary density can be fitted in a non–parametric manner on the basis of the corresponding histogram. The form of the

drift function can therefore be deduced without making any extra assumptions. We do this for the twenty–one currency denominations considered by estimating the non–parametric density from the histograms of each logarithm of market activity $l_\tau^i = \ln(m_\tau^i)$. In Figure 6.9 the resulting drift function estimates are shown, where l_τ^i refers to the horizontal axis and $\tilde{A}^i(l_\tau^i)$ is shown on the vertical axis for each currency i. We select only the central segment of the drift function since the tail distortions make it difficult to discern the relevant shapes. This is due to the low number of observations in the tails of the stationary densities.

In the case of the gamma density, as identified for the US Dollar denomination of the MCI, a plot of l_τ^i against $\tilde{A}^i(l_\tau^i)$ should be

$$\tilde{A}^i(l^i) = \frac{\gamma^i}{2}\left(\frac{p^i - 1}{\gamma^i} - e^{l^i}\right).$$

Thus around $\tilde{A}^i(l^i) = 0$ an approximate straight line should be evident. This is seen in Figure 6.9 for the majority of denominations. Furthermore, some negative curvature should be apparent away from this area. Overall, the US Dollar market activity drift function appears to show a negative exponential function. However, each market is different with some showing a distinct change in slope, such as the Italian Lira (ITL) and the Japanese Yen (JPY) denominations.

6.2 Co–movements of Normalised WSIs

The pairwise dependence of the normalised WSIs is required so that a complete characterisation of the behaviour of the normalised WSIs is obtained. To do this, the covariation of the square root of the normalised WSI in physical time is considered. That is, from (4.6.5), the covariation of the square root of the ith and jth normalised GOP in physical time can be written as

$$\left\langle \sqrt{Y^i}, \sqrt{Y^j} \right\rangle_t = \frac{1}{4}\int_0^t \sqrt{m^i(s)}\sqrt{m^j(s)}\rho_Y^{i,j}(s)ds$$

where

$$\left\langle \hat{W}^i, \hat{W}^j \right\rangle_t = \int_0^t \rho_Y^{i,j}(s)ds.$$

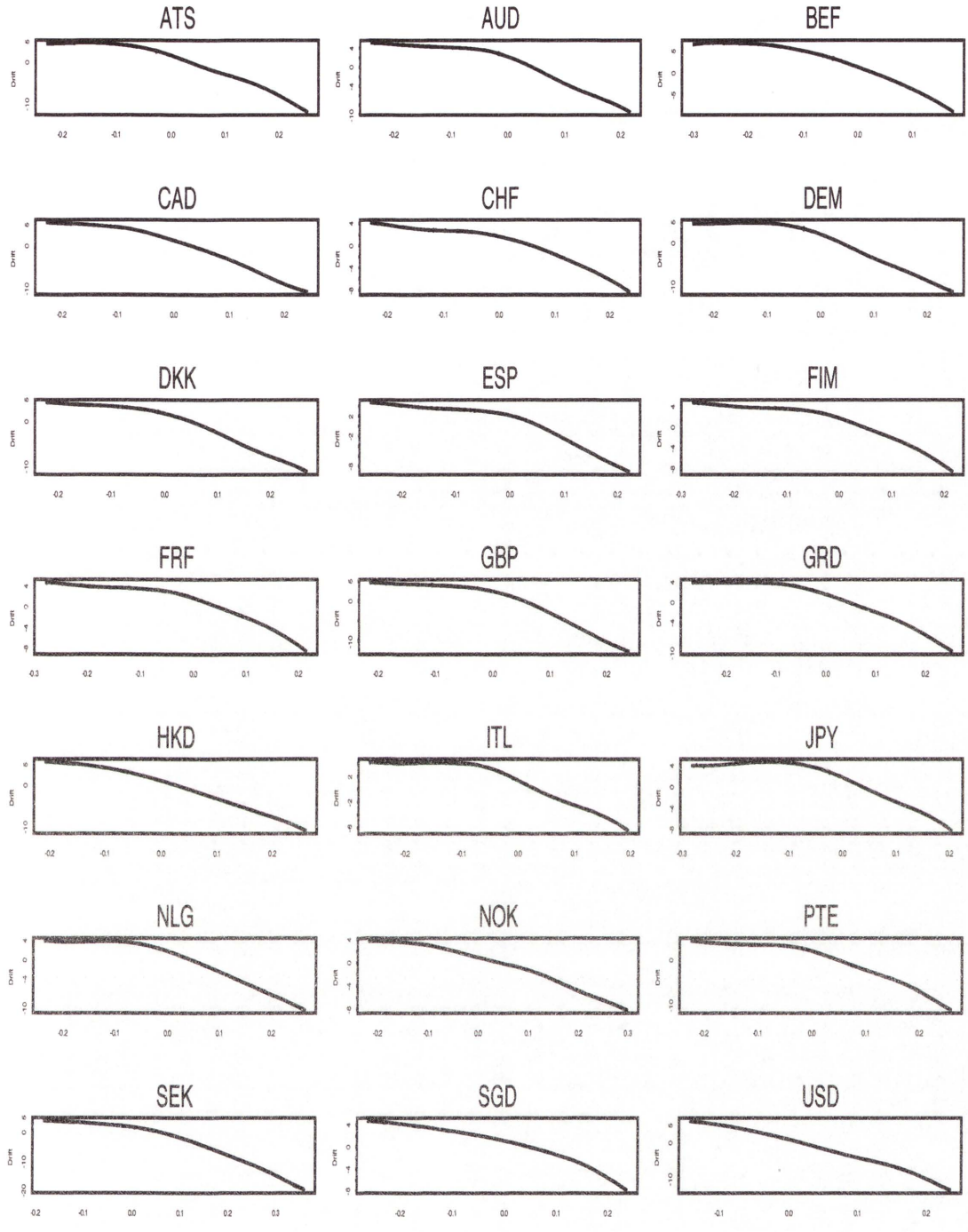

Figure 6.9: Non–parametric drift function estimates for twenty–one denominations.

If the two Wiener processes are uncorrelated then

$$\left\langle \hat{W}^i, \hat{W}^j \right\rangle_t = 0.$$

Provided that the covariation of the square root of the two normalised WSIs is smooth, we can write the *covariation coefficient* between the ith and jth normalised GOP as

$$\rho_Y^{i,j}(t) = \frac{4\frac{d}{dt}\left\langle \sqrt{Y^i}, \sqrt{Y^j} \right\rangle_t}{\sqrt{m^i(s)}\sqrt{m^j(s)}}.$$

From (4.6.15), this can be written as

$$\rho_Y^{i,j}(t) = \frac{\frac{d}{dt}\left\langle \sqrt{Y^i}, \sqrt{Y^j} \right\rangle_t}{\sqrt{\frac{d}{dt}\langle \sqrt{Y^i} \rangle_t}\sqrt{\frac{d}{dt}\langle \sqrt{Y^j} \rangle_t}}. \tag{6.2.1}$$

Due to the large number of pairwise relationships we show only twenty of the empirical covariation processes in Figure 6.10. The empirical covariations shown display patterns typical to the various combinations of the normalised WSIs. It is clear that the empirical covariations of the pairs of the square root of the normalised WSIs shown here, remain close to zero. Notably both the empirical covariation and market activity processes in the denominator of (6.2.1) are not linear. This makes the calculation of $\rho_Y^{i,j}$ more complicated.

To calculate a covariation coefficient for the square root of the normalised WSIs, we assume that the covariation process remains linear for the last year of the observation period. Further research is required to examine the behaviour of the covariation coefficients over time. Figure 6.11 shows the covariation coefficients ordered according to similarity between pairs. In particular, note that the behaviour of the covariation coefficients is very similar for each of the European markets. A graph summarising these relationships, known as a dendogram, is shown in Figure 6.12. The dendogram is typical to cluster analysis, a multivariate statistical technique used to summarise the similarity of a group of variables. Here the cluster analysis is performed using the matrix of covariation coefficients depicted in Figure 6.11. A single linkage agglomerative hierarchical method is

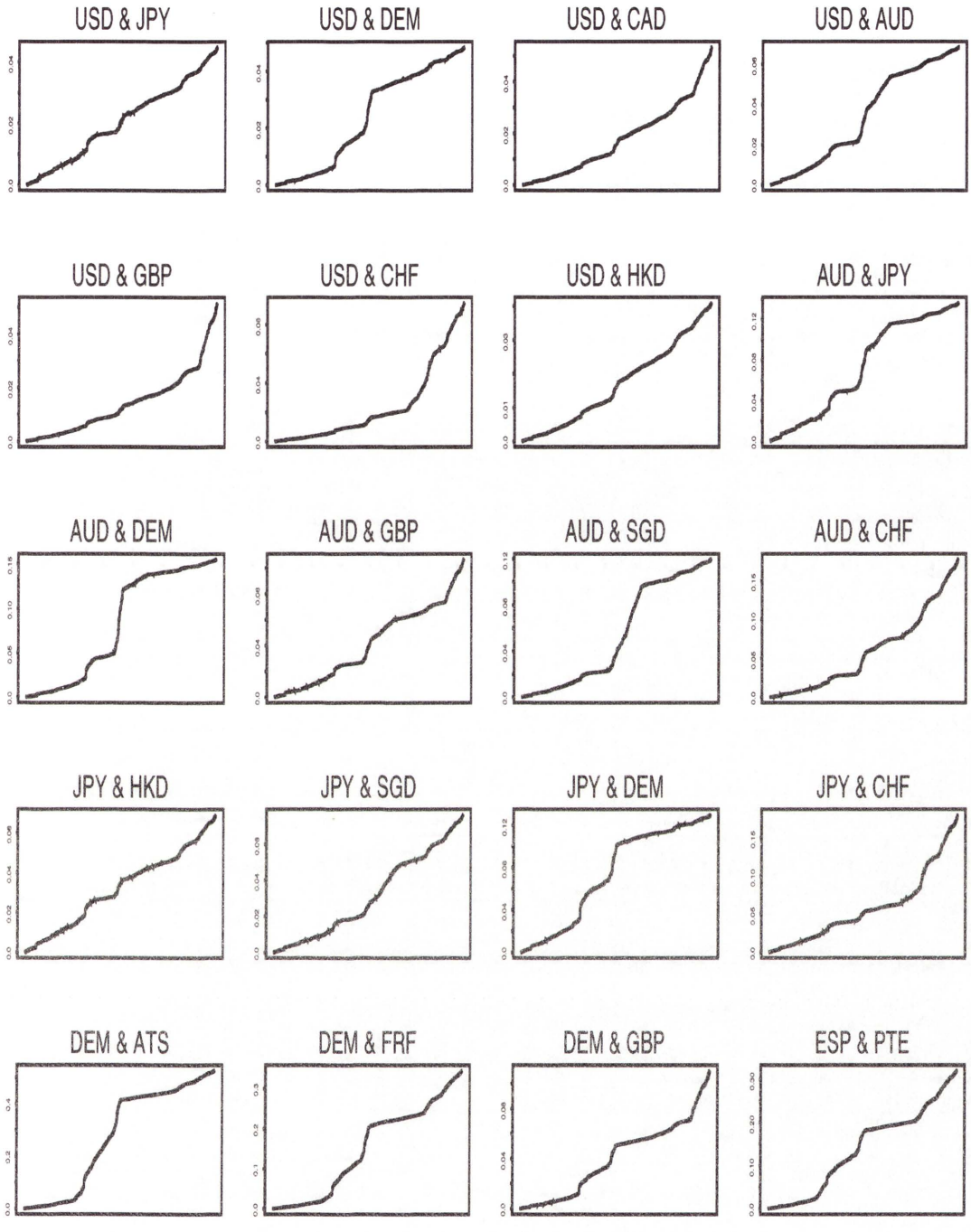

Figure 6.10: Empirical covariation processes for the square root of normalised WSIs. Only twenty pairs of covariation processes are shown.

used to perform the analysis, see Johnson & Wichern (1998). An agglomerative hierarchical method starts with each of the individual denominations. The most similar objects are grouped first and then these initial groups are merged until eventually a single cluster remains. The similarity of each currency denomination is determined via the minimum distance of the covariation coefficients, known as the single linkage method. Different techniques for the cluster analysis were tried, however the groupings shown earlier did not change markedly. The dendogram in Figure 6.12 shows the clusters clearly. A European cluster is apparent, indicating that the normalised WSIs have covariation coefficients very close to one another, as one would expect. This is also seen for the US Dollar (USD) and Hong Kong Dollar (HKD) covariation coefficients. This is not surprising since the Hong Kong Dollar is pegged to the US Dollar. The dendogram indicates an Asia–Pacific cluster, although note that the Australian Dollar (AUD) is not as similar to the other Asian currencies. The Swedish Krona (SEK) is quite different to all currencies, although, there is some evidence of a Scandinavian cluster, comprising the Danish Krona (DKK) and the Norwegian Krona (NOK). The similarity of the normalised WSIs is illustrative of some economic relationships. These will change, for example, if new trade agreements are reached that provide a link between two economies. Further research will investigate how the covariation coefficients behave over time.

6.3 Co–movements of Market Activity

The pairwise behaviour of market activity in different denominations is also of interest. To analyse this dependence, we consider the empirical covariation of pairs of logarithmic market activity processes in physical time. That is, from (5.2.4) the covariation of the logarithm of the ith and jth market activity processes in physical time can be written as

$$\langle \ln(m^i), \ln(m^j) \rangle_t = \int_0^t \beta^i(s)\beta^j(s)\rho_m^{i,j}(s)ds, \tag{6.3.1}$$

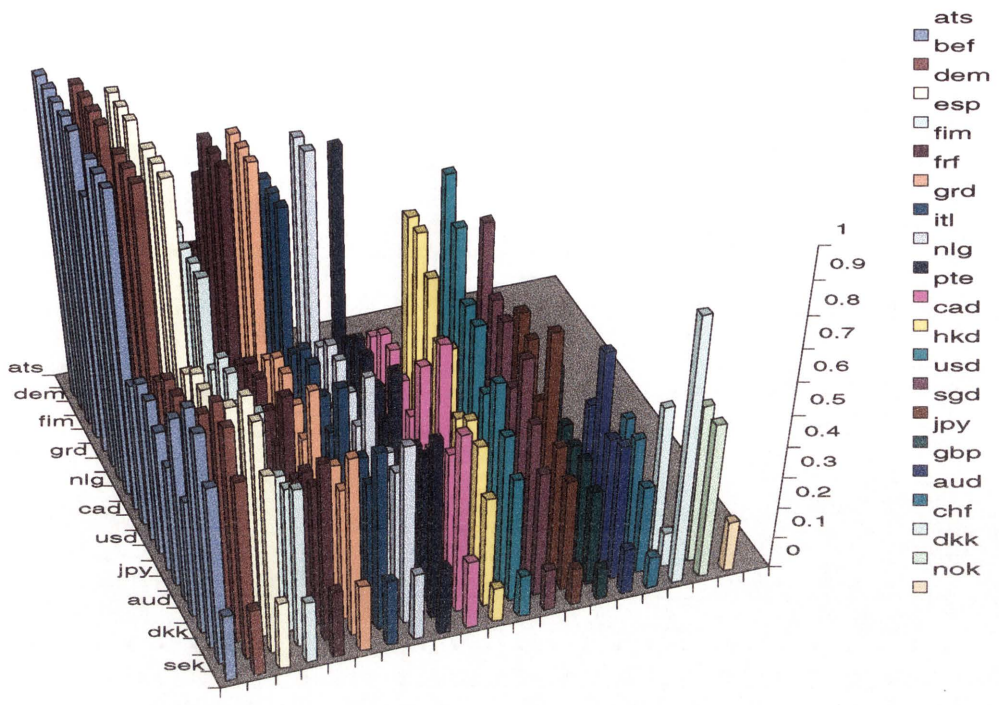

Figure 6.11: Covariation coefficients for the square root of the normalised MCI in different currency denominations.

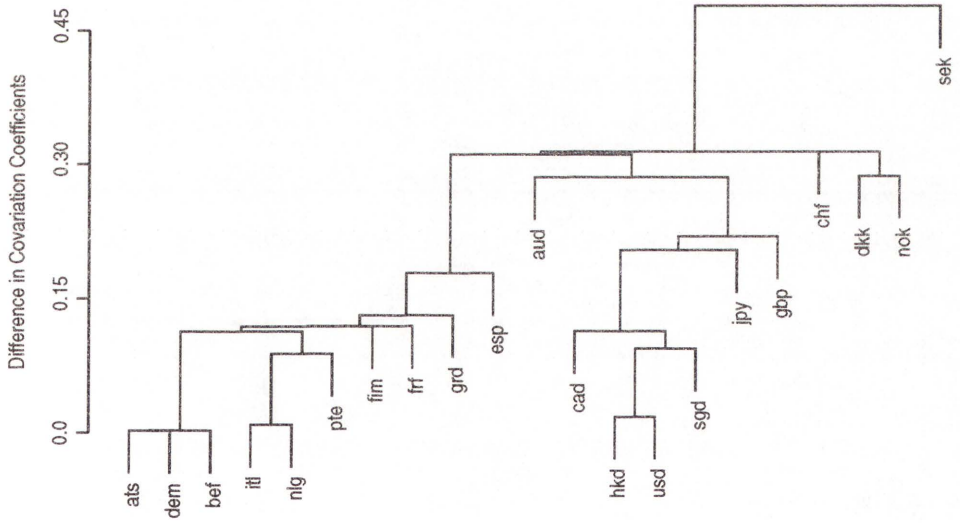

Figure 6.12: Dendogram of covariation coefficients for the square root of the normalised WSIs.

where

$$\langle \tilde{W}^i, \tilde{W}^j \rangle_t = \int_0^t \rho_m^{i,j}(s) ds.$$

If the two Wiener processes are uncorrelated, then

$$\langle \tilde{W}^i, \tilde{W}^j \rangle_t = 0.$$

Again, due to the large number of pairwise relationships, we show only twenty of the empirical covariation processes for the logarithm of market activity in Figure 6.13. It is striking that in most cases, over the entire five year period, the empirical covariation appears almost linear. We see in some cases that there is a change in slope of the otherwise straight line. For example, the German Deutschmark (DEM) and French Franc (FRF) covariation process displays such behaviour. Over a smaller observation period, see Figure 6.14, we see the seasonal pattern in the empirical covariation processes. This is the interaction of the activity volatility processes for the two currencies. Note that the US Dollar (USD) and Hong Kong Dollar (HKD) empirical covariation and the German Deutschmark (DEM) and Austrian Shilling (ATS) empirical covariation behave in a similar manner to the empirical quadratic variation processes in Figure 6.5. This indicates that the activity volatility time scales of the market activity processes are very similar.

Given that we know the activity volatility for each denomination, this implies that the covariation coefficient between the ith and jth market activity processes can be written as

$$\rho_m^{i,j}(t) = \frac{\frac{d}{dt}\langle \ln(m^i), \ln(m^j) \rangle_t}{\sqrt{\frac{d}{dt}\langle \ln(m^i) \rangle_t} \sqrt{\frac{d}{dt}\langle \ln(m^j) \rangle_t}} \tag{6.3.2}$$

for $t \in [0, T]$. For simplicity, we assume that the covariation coefficient is constant and calculate all estimates at the end of the period. The values of $\hat{\rho}_m^{i,j}$ obtained are shown in Figure 6.15. The graphs are grouped according to the similarity between pairs. We see the relationships between the individual pairs of covariation coefficients. Again, note that the behaviour of the European markets are very similar for each pair of covariation coefficients. Interestingly, the logarithmic

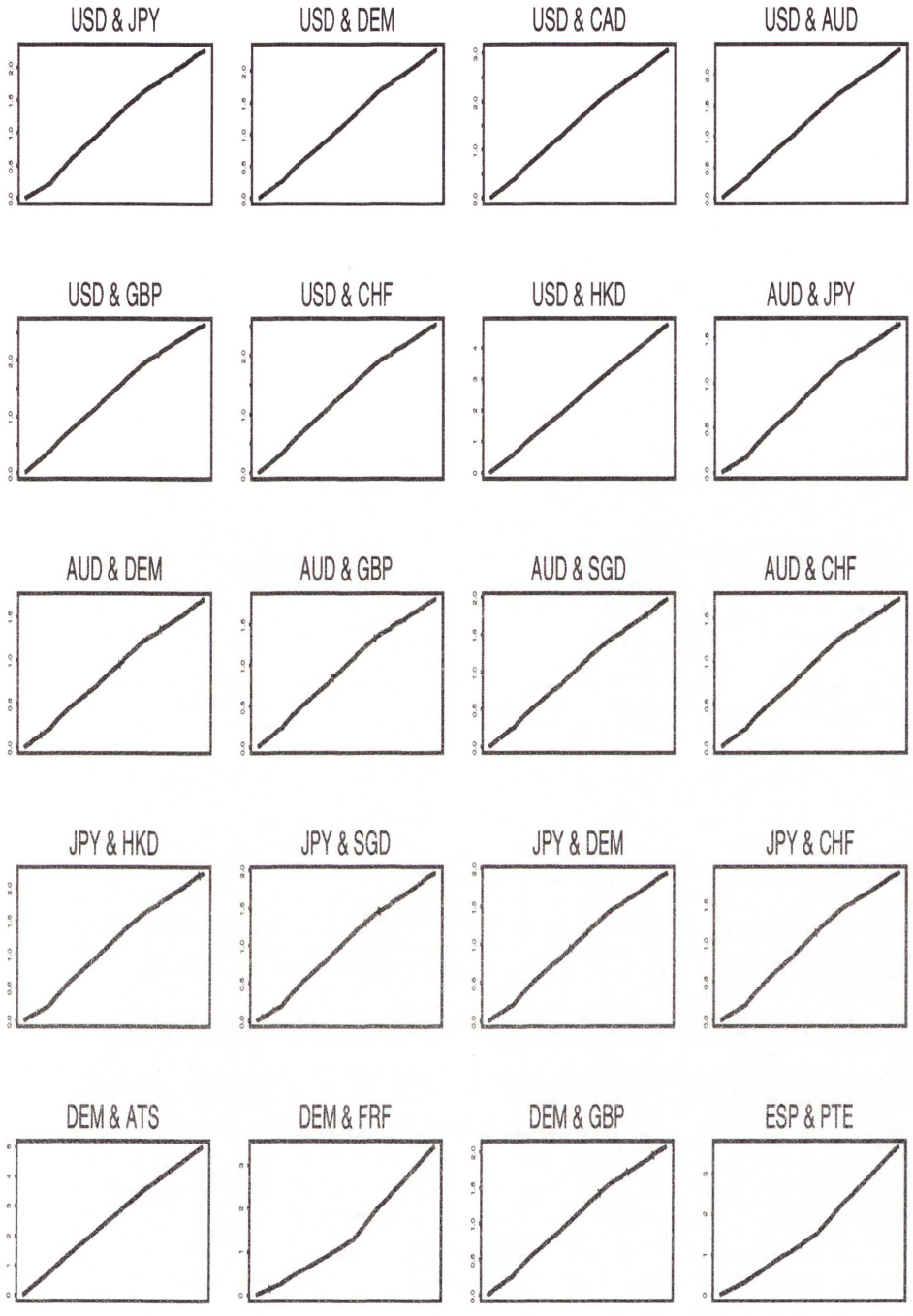

Figure 6.13: Empirical covariation processes for twenty pairs of logarithmic market activity.

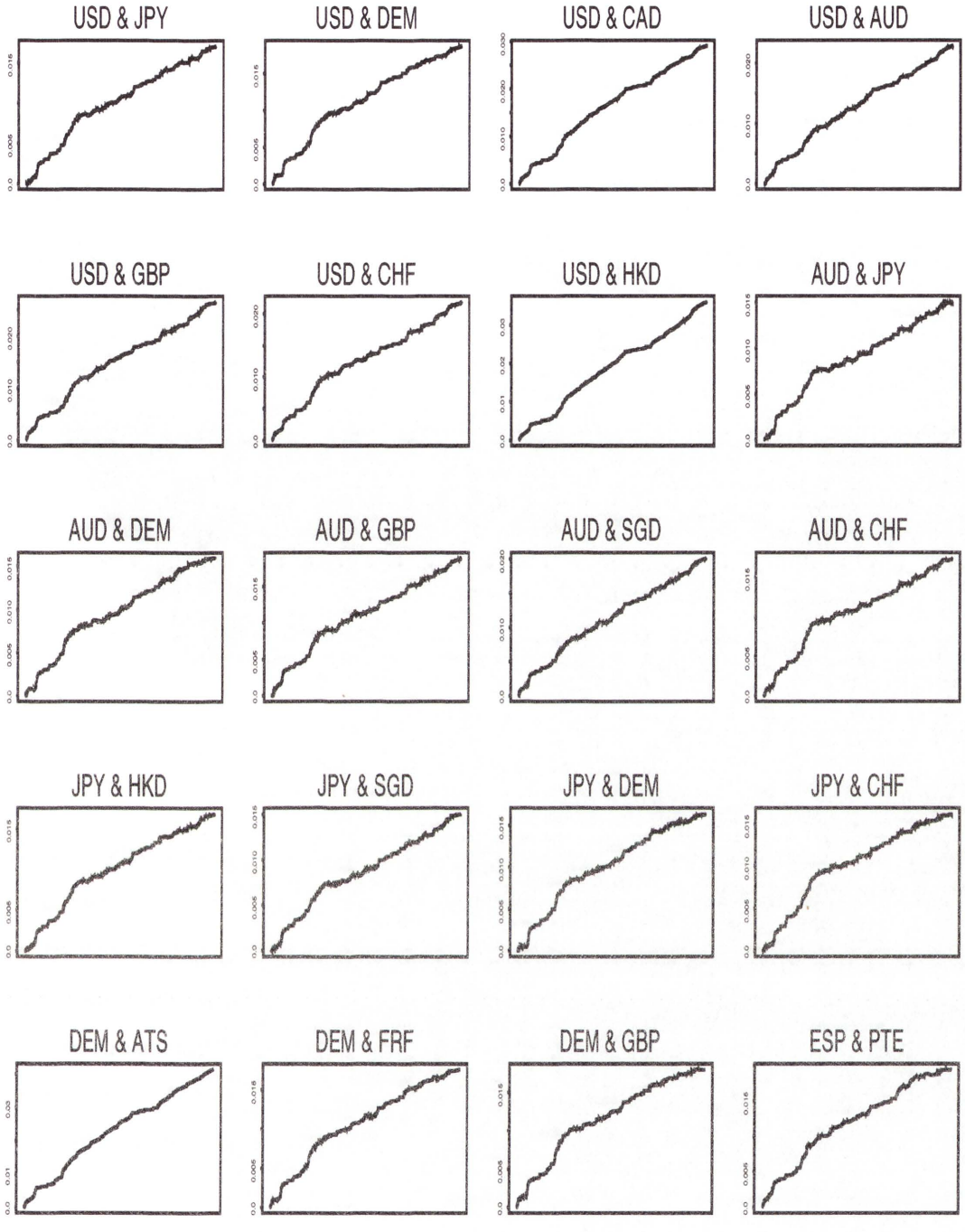

Figure 6.14: Empirical covariation processes covering two weeks for twenty pairs of logarithmic market activity.

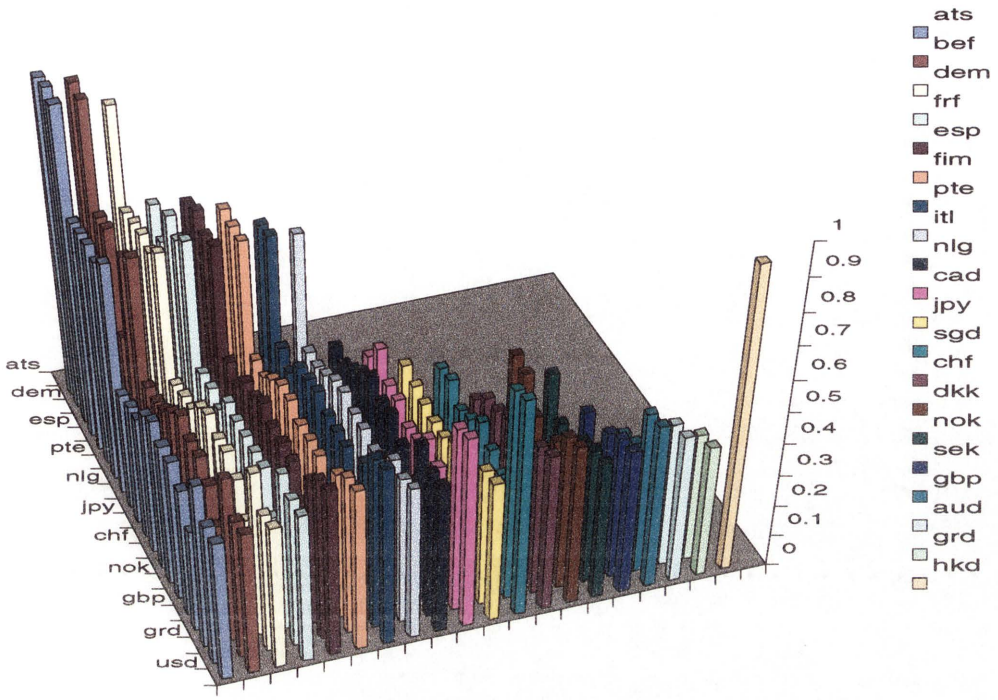

Figure 6.15: Covariation coefficients of the logarithmic market activity in different currency denominations.

market activity denominated in US Dollars has a covariation coefficient of at least 0.4 for every currency denomination considered. The dendogram summarising these relationships is shown in Figure 6.16. As one would expect, the group on the left hand side shows that the European markets are extremely similar with respect to the covariation of market activity. This is seen for the normalised WSIs, see Figure 6.12. Additionally, the Hong Kong Dollar (HKD) and US Dollar (USD) have a high covariation coefficient for market activity. Furthermore, the US Dollar (USD) is highly correlated with the Canadian Dollar (CAD) and Singapore Dollar (SGD). A clear Scandinavian cluster, consisting of the Danish Krona (DKK), the Norwegian Krona (NOK) and the Swedish Krona (SEK) is apparent. Note also that an Asian cluster is no longer evident in the covariation coefficients of market activity. Whilst the denominations contained in the clusters shown in Figures 6.12 and 6.16 are very similar, there are some differences. This reflects the fact that

market activity indicates both trading volume and trading activity, whereas the normalised WSI is an indicator of level and value.

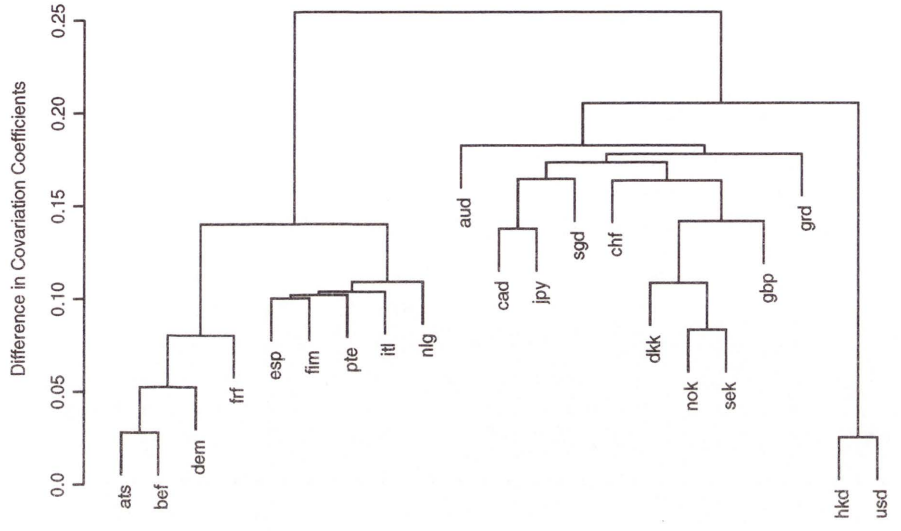

Figure 6.16: Dendogram of covariation coefficients for logarithmic market activity.

We have shown that the MCI in different currency denominations, when analysed after a power transformation, display similar properties in terms of the behaviour of the market activity processes, as seen for the US Dollar. As a result, it appears different denominations of the GOP can be modelled using the same general structure of the SDEs. The co–movements of the square root of the normalised WSIs are examined. This reveals interesting currency groupings indicative of economic relationships. Additionally, this analysis is undertaken for the logarithm of market activities, revealing information about trading behaviour, which in the main appear to be associated with regional effects. This implies that the model described in Chapters 4 and 5 provides a reasonably accurate description of the discounted dynamics of the high–frequency MCI in currencies other than the US Dollar.

Chapter 7

Conclusion

An advantage of the benchmark approach is that estimates of the drift function are not required. The diffusion function, however, is easily estimated from a short period of high–frequency data. This is illustrated using the transform function technique developed. This new technique, by using approximations to martingale estimating functions, does not rely on stationarity nor is the explicit knowledge of moments or transition densities required.

Several approximations of the growth optimal portfolio are formed from high–frequency data. This appears to be the first time that such diversified high–frequency world indices have been constructed and analysed. It is shown that the total world market capitalisation weighted index appears to be the best proxy for the growth optimal portfolio in terms of statistical properties and maximum growth.

This thesis successfully confirms the applicability of the benchmark framework to intraday data. The classical volatility based approaches have difficulties accommodating high–frequency data. We achieve this through the introduction of market activity time, which allows consistent deseasonalisation of the high–frequency indices. Market activity time approaches physical time, implying that the benchmark model is relevant for both micro and macro time scales.

Market activity provides a natural parametrisation of the short term residual fluctuations apparent in the high–frequency indices considered. Through the use of a square root transform, or more generally, a power transform, market activity is directly obtained. This is in contrast to the logarithmic transform commonly used in the standard volatility based analysis of financial markets. The

empirical results documented for the currencies considered, suggest that market activity is a strongly mean–reverting stochastic process with different seasonal patterns in the drift and diffusion coefficients. The diffusion coefficient of the market activity process is shown to be multiplicative. Following this, an activity volatility time scale is developed. From the stationary density of the logarithm of market activity in activity volatility time, a distinct drift function is non–parameterically identified. By a straightforward transformation of the index, we are able to non–parametrically identify; the expected market activity, activity volatility and logarithmic market activity drift. These non–parametric, robust estimates characterise the market activity and thus the dynamics of the index.

The co–movements of the square root of the normalised world stock indices are characterised via their covariations. Such an analysis is also conducted for the logarithm of market activities. Through the use of cluster analysis, it is shown that interesting currency groupings exist. These groupings are indicative of existing economic relationships and regional effects. This provides scope for future research which aims to explain contagion and other behavioural patterns.

The intraday benchmark model provides a largely accurate description of financial markets. This lays the foundation for risk measurement, derivative pricing and various directions of future research.

Appendix A

Definitions

Here definitions from Liptser & Shiryaev (2001) regularly used in the area of the statistics of random processes are given. These terms are referred to throughout Chapters 2 and 3. Here X is a continuous stochastic process on the probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P_\vartheta)$, where $\vartheta \in \Theta$ is the parameter to be estimated.

Definition A.1 *The statistic* $\hat{\vartheta}$ *is called an* unbiased *parameter estimate of* $\vartheta \in \Theta$ *if* $E_\vartheta(\hat{\vartheta}) = \vartheta$.

Here E_ϑ denotes expectation with respect to the probability P_ϑ.

Definition A.2 *The statistic* $\hat{\vartheta}$ *is* sufficient *for* ϑ *if for each* $A \in \underline{\mathcal{A}}$ *a version of the conditional probability* $P(A|\hat{\vartheta})$*, not depending on* ϑ*, can be chosen.*

The *Factorisation Theorem*, see Liptser & Shiryaev (2001), gives necessary and sufficient conditions for $\hat{\vartheta}$ to be sufficient.

Definition A.3 *The sequence of statistics* $\hat{\vartheta}_n$, $n \in \{1, 2, ...\}$ *is called a* consistent *parameter estimate if* $\vartheta_n \xrightarrow{P} \vartheta_0$ *as* $n \to \infty$ *for all* $\vartheta \in \Theta$ *then*

$$P_\vartheta\{|\vartheta_n - \vartheta_0| > \epsilon\} \to 0$$

as $n \to \infty$ *for* $\epsilon > 0$.

Here ϑ_0 denotes the true parameter estimate.

The sequence of statistics ϑ_n, $n \in \{1, 2, ...\}$ is a *strongly consistent* estimator of the parameter $\vartheta \in \Theta$ if $\vartheta_n \xrightarrow{P_\vartheta} \vartheta_0$ with P_ϑ-probability one for all $\vartheta \in \Theta$.

It is possible to investigate the quality of parameter estimates through the use of the *Fisher information matrix*

$$I(\vartheta) = ||I_{i,j}(\vartheta)|| \quad \text{where} \quad I_{i,j}(\vartheta) = E_\vartheta \left[\left\{ \frac{\partial}{\partial \vartheta_i} \ln L \right\} \left\{ \frac{\partial}{\partial \vartheta_j} \ln L \right\} \right].$$

Here L is the likelihood function, see Chapter 2. For unbiased estimates of the parameter $\vartheta \in \Theta \subseteq \mathbb{R}^p$ the *Cramer-Rao matrix inequality*

$$E_\vartheta (\vartheta - \hat{\vartheta})(\vartheta - \hat{\vartheta})^\top \geq I^{-1}(\vartheta)$$

for $\vartheta \in \Theta$ is true under certain conditions of regularity, see Liptser & Shiryaev (2001).

Definition A.4 *The estimator is known as an* efficient *estimator if for all* $\vartheta \in \Theta$ *the Cramer-Rao matrix inequality attains equality.*

Appendix B

List of Countries with Stock Market Indices included in the World Stock Indices

Country	Index	Currency	Country	Index	Currency
Argentina	MERV	ARS	Japan	Nikkei 225	JPY
Australia	AORD	AUD	Korea	KOSPIC	KRW
Austria	ATX	ATS, EUR	Malaysia	KLSE	MYR
Belgium	BEL20	BEF, EUR	Mexico	IPC	MXP
Brazil	BVSP	BRL	The Netherlands	AEX	NLG, EUR
Canada	TSE 300	CAD	Norway	OSETOT	NOK
Denmark	KFX	DKK	Philippines	PCI	PHP
Finland	HEX	FIM, EUR	Portugal	BVL 30	PTE, EUR
France	CAC 40	FRF, EUR	Singapore	STI	SGD
Germany	DAX	DEM, EUR	Spain	IBEX	ESP, EUR
Greece	ATG	GRD, EUR	Sweden	SGI	SEK
Hong Kong	HSI	HKD	Switzerland	SMI	CHF
Hungary	BUXI	HUF	Taiwan	TWI	TWD
India	BSESI	INR	Thailand	SETI	THB
Indonesia	JSX	IDR	Turkey	ICI	TRL
Ireland	ISEQ	IRP, EUR	UK	FTSE ALL	GBP
Italy	MIB 30	ITL, EUR	US	S&P 500	USD

Bibliography

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed form approximation approach. *Econometrica* **70**, 223–262.

Aït-Sahalia, Y. (2003). Closed–form likelihood expansions for multivariate diffusions. Working paper, Department of Economics, Princeton University.

Andersen, T. & T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long–run in high–frequency returns. *Journal of Finance* **52**, 975–1005.

Ané, T. & H. Geman (2000). Order flow, transaction clock and normality of asset returns. *Journal of Finance* **55**, 2259–2284.

Barndorff-Nielsen, O. & M. Sørensen (1994). A review of some aspects of asymptotic likelihood theory for stochastic processes. *International Statistical Review* **62**, 133–165.

Bibby, B. M., M. Jacobsen, & M. Sørensen (2003). Estimating functions for discretely sampled diffusion-type models. In Y. Aït-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Economics.*, pp. 425–440. Amsterdam: North Holland.

Bibby, B. M. & M. Sørensen (1995). Martingale estimating functions for discretely observed diffusions. *Bernoulli* **1**, 17–39.

Bibby, B. M. (1994). Optimal combination of martingale estimating functions for discretely observed diffusions. Research Report No. 298, Department of

Theoretical Statistics, Institute of Mathematics, University of Aarhus.

Black, F. & M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* **81**, 637–654.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics* **3**, 167–179.

Brandt, M. & P. Santa-Clara (2002). Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics* **63**, 161–210.

Breymann, W., A. Dias, & P. Embrechts (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3**(1), 1–14.

Breymann, W., K. Fergusson, & E. Platen (2004). Marginal return distributions of intraday diversified world stock indices. University of Technology Sydney, (working paper).

Breymann, W., L. Kelly, & E. Platen (2004). Intraday empirical analysis and modelling of diversified world stock indices. University of Technology Sydney, (working paper).

Chan, K., G. Karolyi, F. Longstaff, & A. Sanders (1992). An empirical comparison of alternative models of the short–term interest rate. *Journal of Finance* **47**(3), 1209–1227.

Christensen, B., R. Poulsen, & M. Sørensen (2001). Optimal inference for diffusion processes with applications to the short rate of interest. Working Paper No. 102, Centre for Analytical Finance. University of Aarhus.

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–155.

Dacorogna, M., R. Gençay, U. Müller, R. Olsen, & O. Pictet (2001). *An Introduction to High–Frequency Finance*. Academic Press, San Diego, CA.

Dacunha-Castelle, D. & D. Florens-Zmirou (1986). Estimation of the coefficients of a diffusion from discrete observations. *Stochastics* **19**, 263–284.

Davis, M. (2003). Valuation, hedging and investment in incomplete financial markets. In *5th International Congress in Industrial and Applied Mathematics*. Sydney, Australia.

Derman, E. & I. Kani (1994). Riding on a smile. *RISK* **7**, 32–39.

Dimson, E., P. Marsh, & M. Staunton (2002). *Triumph of the Optimists: 101 years of global investment returns*, Chapter 4. Princeton University Press.

Doukhan, P., P. Massart, & E. Rio (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.* **30**, 63–82.

Doukhan, P. (1994). *Mixing, Properties and Examples*. Springer–Verlag, New York.

Elerian, O., S. Chib, & N. Shephard (2001). Likelihood inference for discretely observed non-linear diffusions. *Econometrica* **69**, 959–993.

Elerian, O. (1998). A note on the existence of a closed form conditional transition density for the Milstein scheme. Working Paper, Nuffield College, Oxford.

Eraker, B. (1998). Markov Chain Monte Carlo analysis of diffusion models with applications to finance. Discussion Paper 1998-5, Norwegian School of Economics and Business Administration.

Florens-Zmirou, D. (1989). Approximate discrete-time schemes for statistics of diffusion processes. *Statistics* **20**, 547–557.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability* **30**, 790–804.

Gallant, A. & G. Tauchen (1997). Estimation of continuous time models for stock returns and interest rates. *Macroeconomic Dynamics* **1**, 135–158.

Gençay, R., F. Selçuk, & B. Whitcher (2001). Differentiating intraday seasonalities through wavelet multiscaling. *Physica A* **289**, 543–556.

Genon-Catalot, V., T. Jeantheau, & C. Laredo (2000). Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* **6**, 1051–1079.

Godambe, V. P. & C. Heyde (1987). Quasi-likelihood and optimal estimation. *International Statistical Review* **55**, 231–244.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics* **16**, 99–117.

Heath, D., S. Hurst, & E. Platen (2001). Modelling the stochastic dynamics of volatility for equity indices. *Asia-Pacific Financial Markets* **8**, 179–195.

Heath, D. & E. Platen (2002). Perfect hedging of index derivatives under a minimal market model. *International Journal of Theoretical and Applied Finance* **5**, 747–774.

Heston, S. (1993). A closed–form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**, 327–343.

Heyde, C. (1988). Fixed sample and asymptotic optimality for classes of estimating functions. *Contemporary Mathematics* **80**, 241–247.

Heyde, C. (1997). *Quasi-Likelihood and its Application: A general approach to optimal parameter estimation.* Springer–Verlag, New York.

Hull, J. & A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* **42**, 281–300.

Hurst, S. & 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.

Hurst, S. & E. Platen (1999). On the marginal distributions of trade weighted currency indices. QFRG Research Paper 8, University of Technology, Sydney.

Jacod, J. & A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (Second ed.). Springer, New York.

Jensen, B. & R. Poulsen (1999). A comparison of approximation techniques for transition densities of diffusion processes. Working Paper 38, Centre for Analytical Finance, University of Aarhus, Aarhus School of Business.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman and Hall, London.

Johnson, R. & D. Wichern (1998). *Applied Multivariate Statistical Analysis* (Fourth ed.). Prentice Hall.

Karatzas, I. & S. Shreve (1991). *Brownian Motion and Stochastic Calculus* (Second ed.). Springer, New York.

Karatzas, I. & S. Shreve (1998). *Methods of Mathematical Finance.* Springer, New York. Volume 39 of Appl. Math.

Karlin, S. & H. Taylor (1981). *A Second Course in Stochastic Processes.* Academic Press, New York.

Karpoff, J. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis* **22**, 109–126.

Kelly, J. (1956). A new interpretation of information rate. *Bell Syst. Techn. J.* **35**, 917–926.

Kelly, L., E. Platen, & M. Sørensen (2004). Estimation for discretely observed diffusions using transform functions. *Journal of Applied Probability Special Volume* **41A**, 99–117.

Kessler, M. & S. Paredes (2002). Computational aspects related to martingale estimating functions for a discretely observed diffuion. *Scandinavian Journal of Statistics* **29**, 425–440.

Kessler, M. & M. Sørensen (1999). Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli* **5**, 299–314.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics* **24**, 211–229.

Kessler, M. (2000). Simple and explicit estimating functions for a discretely observed diffusion. *Scandinavian Journal of Statistics* **27**, 65–82.

Kloeden, P. & E. Platen (1999). *Numerical Solution of Stochastic Differential Equations*. Springer–Verlag, Berlin Heidelberg.

Kutoyants, Y. (1984). *Parameter Estimation for Stochastic Processes*. Heldermann–Verlag Berlin.

Liptser, R. & A. Shiryaev (2001). *Statistics of Random Processes* (Second ed.). Springer Verlag, Berlin–Heidelberg. Vol. 5 and 6 of Appl. Math.

Long, J. (1990). The numeraire portfolio. *J. Financial Economics* **26**, 26–69.

Lo, A. (1988). Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory* **4**, 231–247.

Merton, R. (1992). *Continuous–Time Finance*. Blackwell, Oxford.

Milstein, G., E. Platen, & H. Schurz (1999). Balanced implicit methods for stiff stochastic systems. *SIAM Journal of Numerical Analysis* **35**(2), 1010–19.

Müller, U. A., M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, & C. Morgenegg (1990). Statistical study of foreign exchange rates, empirical evidence of a price change scaling law and intraday analysis. *Journal of Banking and Finance* **14**, 1189–1208.

Nicolau, J. (2002). A new technique for simulating the likelihood of stochastic differential equations. *Econometrics Journal* **5**, 91–103.

Overbeck, L. & T. Rydén (1997). Estimation in the Cox–Ingersoll–Ross model. *Econometric Theory* **13**, 430–461.

Pedersen, A. (1995a). Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. *Bernoulli* **1**, 257–279.

Pedersen, A. (1995b). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics* **22**, 55–71.

Platen, E., J. West, & W. Breymann (2004). The benchmark approach to pricing electricity derivatives. University of Technology Sydney, (working paper).

Platen, E. (2001). A minimal financial market model. In *Trends in Mathematics*, pp. 293–301. Birkhäuser Verlag Basel/Switzerland.

Platen, E. (2002). Arbitrage in continuous complete markets. *Advances in Applied Probability* **34**, 540–558.

Platen, E. (2003). Diversified portfolios in a benchmark framework. QFRG Research Paper 87, University of Technology, Sydney.

Platen, E. (2004). Modelling the volatility and expected value of a diversified world index. *International Journal of Theoretical and Applied Finance*. Forthcoming.

Poulsen, R. (1999). Approximate maximum likelihood estimation of discretely observed diffusion processes. Working Paper No. 29. Centre for Analytical Finance, University of Aarhus.

Prakasa Rao, B. L. S. (1988). Statistical inference from sampled data for stochastic processes. *Contemp. Math.* **80**, 249–284.

Prakasa Rao, B. (1999). *Statistical Inference for Diffusion Type Processes*. Arnold, London.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer.

Revuz, D. & M. Yor (1999). *Continuous Martingales and Brownian Motion* (Third ed.). Springer, New York.

Rogers, L. (2001). The relaxed investor and parameter uncertainty. *Finance and Stochastics* **5**, 131–154.

Sørensen, H. (2001). Discretely observed diffusions: Approximation of the continuous–time score function. *Scandinavian Journal of Statistics* **28**, 113–121.

Sørensen, H. (2002). Parametric inference for diffusion processes observed at discrete points in time: a survey. Working paper No. 119, Centre for Analytical Finance. University of Aarhus.

Sørensen, M. (1997). Estimating functions for discretely observed diffusions: A review. In I. V. Basawa, V. P. Godambe, and R. L. Taylor (Eds.), *Selected Proceedings of the Symposium on Estimating Functions, Volume 32 of IMS Lecture Notes - Monograph Series*, pp. 305–325. Institute of Mathematical Statistics, Hayward, CA,.

Sørensen, M. (1999). On asymptotics of estimating functions. *Brazilian Journal of Probability and Statistics* **13**, 111–36.

Sørensen, M. (2000). Prediction based estimating functions. *Journal of Econometrics* **3**, 123–47.

Wood, R., T. McInish, & J. Ord (1985). An investigation of transaction data for NYSE stocks. *Journal of Finance* **40**, 723–739.

Xu, X. & C. Wu (1999). The intraday relation between return volatility, transactions and volume. *International Review of Economics and Finance* **8**, 375–397.

Yoshida, N. (1992). Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis* **41**, 220–242.