UNIVERSITY OF
TECHNOLOGY SYDNEY

UTS

THINK.CHANGE.DO

QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 125　　　　　　**May 2004**

Intraday Empirical Analysis and Modeling of Diversified World Stock Indices

Wolfgang Breymann, Leah Kelly and Eckhard Platen

Intraday Empirical Analysis and Modeling of Diversified World Stock Indices

Wolfgang Breymann
Eidgenössische Technische Hochschule
Zürich
Leah Kelly
University of Technology Sydney,
School of Finance & Economics,
PO Box 123, Broadway, NSW, 2007, Australia
Eckhard Platen
University of Technology Sydney,
School of Finance & Economics and Department of Mathematical Sciences,
PO Box 123, Broadway, NSW, 2007, Australia

May 5, 2004

Abstract. This paper proposes an approach to the intraday analysis of diversified world stock accumulation indices. The growth optimal portfolio (GOP) is used as reference unit or benchmark in a continuous financial market model. Diversified portfolios, covering the world stock market, are constructed and shown to approximate the GOP, providing the basis for a range of financial applications. The normalized GOP is modeled as a time transformed square root process of dimension four. Its dynamics are empirically verified for several world stock indices. Furthermore, the evolution of the transformed time is modeled as the integral over a rapidly evolving mean-reverting market activity process with deterministic volatility. The empirical findings suggest a rather simple and robust model for a world stock index that reflects the historical evolution, by using only a few readily observable parameters.

1 Introduction

The purpose of this paper is the empirical construction of an intraday world stock index (WSI) that provides a good approximation of the *growth optimal portfolio* (GOP), see Kelly (1956) and Long (1990). It is the portfolio that maximizes logarithmic expected utility of terminal wealth. For practical applications, which include portfolio optimization and numeraire based derivative pricing, it is useful to be able to construct approximations of the GOP from observed data. This can be achieved by considering diversified portfolios, see Platen (2004). It is the objective of this paper to study such proxies and model the intraday dynamic properties. More specifically, we will (i) construct several GOP proxies with five minutes data, (ii) compare their empirical quantities with theoretical model predictions, and (iii) select the best one on the basis of theoretical requirements.

Increasingly, diffusion processes are used to model financial markets in continuous time, see Merton (1992). The well-known Black-Scholes (BS) framework, in which volatility is a central concept, continues to provide the basis for quantitative financial techniques in practice. The simplest form of the BS framework implies Gaussian log-return distributions. The deviation of observed log-returns from a Gaussian distribution is an increasingly established stylized empirical fact. This is particularly apparent in the intraday regime, which has been widely documented, see, for example, Dacorogna, Gençay, Müller, Olsen & Pictet (2001).

In the discrete time setting, the most advanced time-series models are able to reflect the main stylized empirical features reasonably well.[1] In general, however, the time-aggregation properties of these models make it difficult to integrate intraday information into a typically daily discrete time framework.[2] Continuous time diffusion type models, on the other hand, can resolve much of this problem. Furthermore, continuous time models benefit from the high observation frequency realized with intraday data, in particular, for the estimation of the diffusion coefficients.

The *benchmark approach* (BA), see Platen (2004), uses the GOP as a central building block and generalizes existing financial market modeling approaches. By using the BA, a GOP model is obtained that is able to accommodate many of the known empirical stylized facts. The practical availability of a suitable proxy for the GOP will be demonstrated. This allows a number of applications of the BA with regards to portfolio optimization, derivative pricing, integrated risk management and the efficient use of intraday data.

Diffusion processes are determined by the form of their drift and diffusion co-efficients. It is known that the drift coefficient of the discounted GOP is fully

---

[1] For a review using ARCH-type models, see for example Bollerslev, Chou & Kroner (1992). For a review on stochastic volatility models, see for example Taylor (1994).

[2] The unsatisfactory behaviour of the coefficients of a GARCH(1,1) model under time aggregation is demonstrated in Dacorogna, Gençay, Müller, Olsen & Pictet (2001).

determined by its diffusion coefficient, see Long (1990). Indeed, the risk premium of the GOP simply equals its squared volatility. A similar property can be shown for any other security under the BA, see Platen (2004). The BA provides an extension to previous approaches insofar as it excludes arbitrage without requiring the existence of an equivalent risk neutral martingale measure. As shown in Platen (2002), the ideal dynamics of the GOP implies Student t distributed log-returns with degrees of freedom four. This is consistent with empirical studies that indicate tail indices around four, see Hurst & Platen (1997), Müller, Dacorogna & Pictet (1998) and Breymann, Dias & Embrechts (2003).

In general, reliable estimation of the drift coefficient or trend using the relatively short period of historical financial data available is extremely difficult for financial market models. However, the diffusion coefficient is relatively simple to determine from short observation periods. In fact, inference about the diffusion coefficient benefits from the use of intraday data, allowing considerable precision. Therefore, a market model that is fully determined by the diffusion coefficient, as provided by the BA, is highly desirable since it allows the information about the drift to be automatically extracted. Analysis of high-frequency data requires special techniques, see Dacorogna, Gençay, Müller, Olsen & Pictet (2001) and Goodhart & O'Hara (1997). This includes the appropriate treatment of daily and weekly seasonal patterns apparent, see Breymann, Dias & Embrechts (2003). Thus, daily and weekly seasonal patterns need to be integrated in the modeling. We will demonstrate that this can be done in a consistent and robust manner. It leads to a reliable observation and estimation of market activity, which under the BA turns out to be a more suitable financial quantity for incorporating stylized empirical features than is volatility in the BS framework.

Initially, the analysis is illustrated using the Morgan Stanley Capital Growth World Index (MSCI). It is shown that the MSCI, despite being observed daily, approximates the GOP reasonably well. Deviations exist although these can be attributed to two shortcomings. Firstly, the lack of intraday information and secondly, the missing data from emerging markets that are not sufficiently accounted for in the MSCI. Therefore, to rectify these problems we construct three new high-frequency WSIs that are observed at five minute intervals. These appear to be more diversified than the MSCI and approximate the GOP well.

Section 2 discusses the empirical features of the MSCI and the construction of the high-frequency WSIs. Section 3 summarizes the BA for an intraday GOP and Section 4 discusses the empirical analysis and models market activity of the WSIs.

2 World Stock Indices

2.1 A Preliminary Analysis of a Daily World Stock Index

Initially we examine the empirical features of the daily observed Morgan Stanley Capital Growth World Index (MSCI), denominated in USD, and denoted by $V^{(MSCI)}(t)$. This is one of the most diversified, readily available accumulation indices. Since we are not interested in the interest rate dynamics we will study the *discounted* MSCI

$$\bar{V}^{(MSCI)}(t) = V^{(MSCI)}(t) \exp\left\{-\int_0^t r(u)du\right\} \tag{2.1}$$

for $t \in [0, T]$. Here $r(t)$ is the *short term interest rate* for the USD market at time t.

The MSCI index includes stocks from 22 countries and is based on approximately 1200 stocks. Figure 1 shows the evolution of the discounted MSCI index over the period from January 1970 until January 2003.
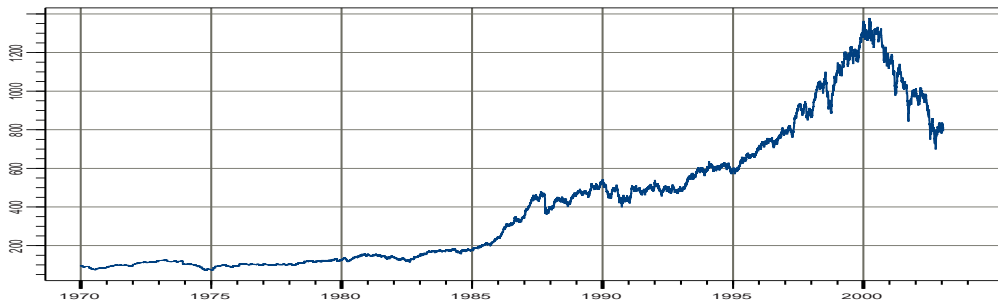

Figure 1: The discounted MSCI World Index 1970-2003.

The discounted index displayed in Figure 1 appears, on average, to grow exponentially. To compensate for this we introduce the *normalisation function*

$$\bar{\alpha}(t) = \xi \exp(\eta t), \tag{2.2}$$

for $t \in [0, T]$, where $\xi > 0$ is a *normalisation constant* and $\eta > 0$ is called the *net growth rate* of the market.

An empirical study by Dimson, Marsh & Staunton (2002) suggests that over the last one hundred years a discounted *World Stock Index* (WSI), denominated in USD, has experienced an average rate of net growth of approximately $\eta = 4.8\%$. This is consistent with the parameter range that other earlier studies detected. Thus, we introduce the *normalized* MSCI

$$Y^{(MSCI)}(t) = \frac{\bar{V}^{(MSCI)}(t)}{\bar{\alpha}(t)} \tag{2.3}$$

for $t \in [0, T]$. In the following illustrations, we have chosen $\eta = 0.048$ and $\xi = 10.5$, as in Platen (2004), which matches the long term growth of a market capitalization weighted WSI, as will be shown in Section 2.2. Figure 2 displays the resulting normalized MSCI over the period of thirty two years. As we will see, the time series shown in Figure 2 is in essence stationary even though fluctuations on all scales are evident and the mean reversion appears to be weak. It is interesting to note that the strength of the fluctuation depends on the level of the normalized MCSI. Periods of high values correspond to periods of large fluctuations and periods of low values correspond to periods of small fluctuations.

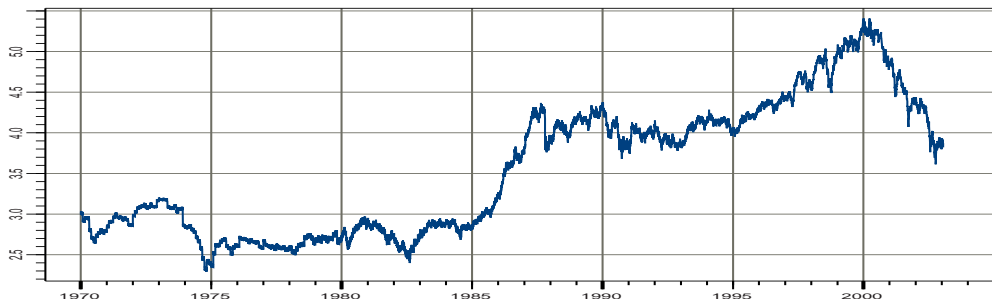Figure 2: Normalized MSCI $Y^{(MSCI)}(t)$ for 1970 - 2003.

Figure 3: Square root $\sqrt{Y^{(MSCI)}(t)}$ of normalized MSCI.

An important observable, when considering the fluctuations of a stochastic process X, is its quadratic variation $\langle X \rangle_t$. It is given by the limit of the sum of the squared increments of this process X when the step size of the underlying time discretization tends to zero. For Brownian motion, the quadratic variation is simply the time itself. Typically, the empirically observed quadratic variation of a security or an index displays behaviour more complicated than a linear function of time. The slope of quadratic variation reflects phenomena like stochastic volatility. Ideally, one would like to transform the time series of the normalized MSCI in a way that results in a process with approximately *level independent* fluctuations. This means that the corresponding quadratic variation should be

approximately proportional to time, which implies that the diffusion term of the transformed time series would then be proportional to a Brownian motion.
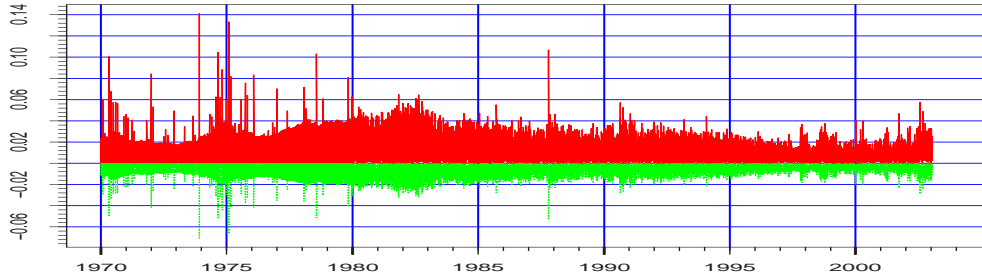

Figure 4: Absolute returns of normalized MSCI $Y^{(MSCI)}(t)$ (above) and $\sqrt{Y^{(MSCI)}(t)}$ (below).
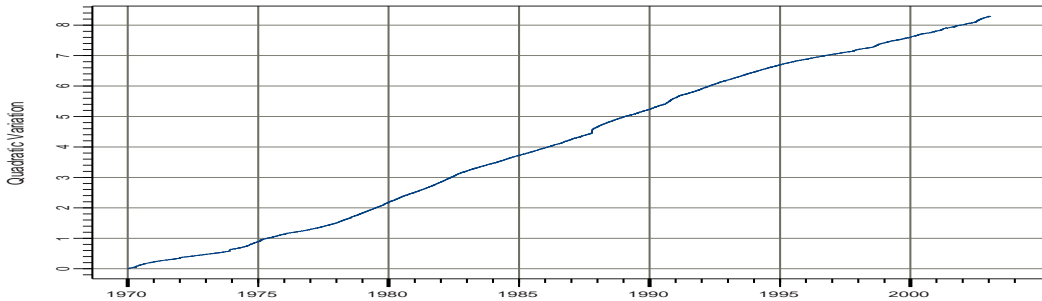

Figure 5: Quadratic variation of $\langle\sqrt{Y^{(MSCI)}}\rangle_t$.

With this goal in mind we consider the square root $\sqrt{Y^{(MSCI)}(t)}$. This particular transformation is motivated by theoretical arguments which will be presented in Section 3. Taking the square root increases fluctuations when the level of the normalized index is low and decreases fluctuations when its level is high. According to Figure 3 and 4, the fluctuations of $\sqrt{Y^{(MSCI)}(t)}$ seem to be reasonably similar for high and low levels. To verify this impression Figure 5 shows the corresponding quadratic variation $\langle\sqrt{Y^{(MSCI)}}\rangle_t$, which indeed appears to be approximately proportional to time. Note that we observe in Figure 5 a slope of approximately $\frac{1}{4}$ per year. This empirical feature will be explained in Section 3.

2.2   Construction of Intraday World Stock Indices

To our knowledge there is no readily available high-frequency WSI that one could use for an intraday analysis that extends over several years. Therefore, we construct three different intraday WSIs, for which we will compare key features of

their dynamics. Each WSI is constructed as a self-financing portfolio consisting of local stock market indices taken from almost all financial markets throughout the world. The WSIs, which will be denominated in USD, are constructed from $d = 34$ local stock market accumulation indices. Many of the local indices used are spot price indices. We denote by $P^{(j,i)}(t)$ the local stock spot price index at time t of stock market $j \in \{1, 2, \ldots, d\}$, when denominated in currency $i \in \{USD, CHF, \ldots\}$. As a first step we transform the original local spot price index into an accumulation index $I^{(j,i)}(t)$ by forming the expression

$$I^{(j,i)}(t) = P^{(j,i)}(t) \exp \left\{ \int_0^t y^{(j)}(u) du \right\} \tag{2.4}$$

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

$$I^{(j,\text{USD})}(t) = I^{(j,i)}(t) \, X^{(i,USD)}(t) \tag{2.5}$$

for $t \in [0, T]$. We then obtain a WSI at time t, denoted by $V^{(\text{WSI})}(t)$, which is denominated in USD, by forming the portfolio

$$V^{(\text{WSI})}(t) = \sum_{j=1}^{d} \delta_{(\text{WSI})}^{(j)}(t) \, I^{(j,\text{USD})}(t). \tag{2.6}$$

Here $\delta_{(\text{WSI})}^{(j)}(t)$ denotes the number of units of the jth local accumulation index held in the WSI at time $t \in [0, T]$. The value used for $\delta_{(\text{WSI})}^{(j)}(t)$ corresponds to the jth *weight* or *proportion* $\pi_{(\text{WSI})}^{(j)}(t)$. Throughout our analysis we separate the USD interest rate evolution from our study by considering the *discounted* WSI $\bar{V}^{(\text{WSI})}(t)$ at time t, which is computed from $V^{(\text{WSI})}(t)$ in the same manner as shown in (2.1).

We construct three WSIs: the approximately *Equal Weighted Index* (EWI), the *Market Capitalization Weighted Index* (MCI) and the *Gross Domestic Product Weighted Index* (GDPI). The values $\delta_{(\text{WSI})}^{(j)}(t)$ for each WSI are adjusted such that the proportions $\pi_{(\text{WSI})}^{(j)}(t)$ for the MCI and GDPI correspond to the respective weights of the jth local accumulation index, $j \in \{1, 2, \ldots, d\}$. Despite its name, the EWI is not equally weighted for all markets. Developed markets were given a weight of 0.0357 and emerging markets only 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 capitalization, while the GDPI weights account for more than 85% of the total

world GDP. 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. Additionally, to enable the convenient comparison of the WSIs to a readily available market index, each WSI is rescaled to have the same initial value as the MSCI at the starting date $t_0 = 05/04/1996$ 00:00:00 Greenwich Mean Time (GMT) of our sample.
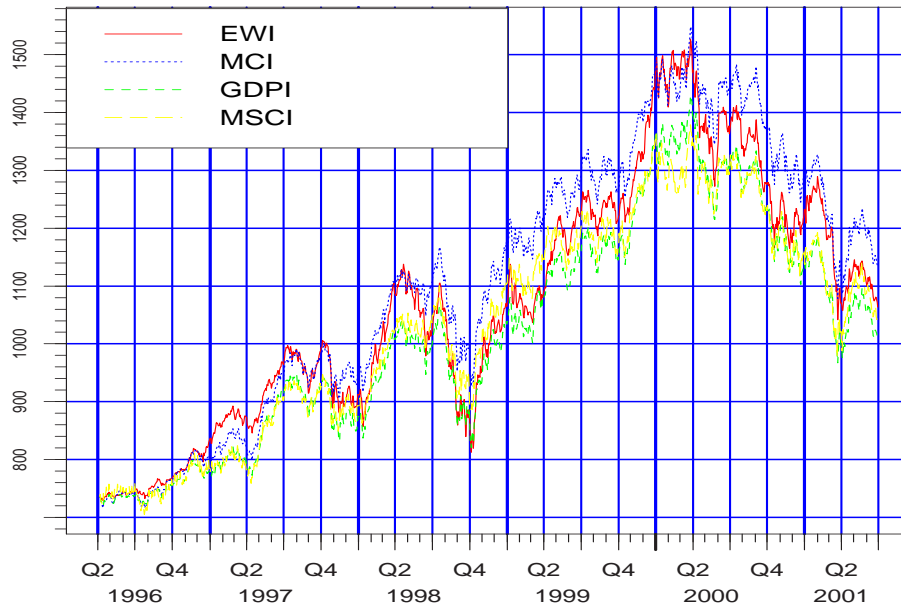

Figure 6: World stock indices with equal weights (EWI), market capitalization adjusted weights (MCI), GDP adjusted weights (GDPI) and MSCI, where initial values are matched to the MSCI.

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 diversified 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 FX spot data to transform the value of the different local stock indices into USD. The high-frequency index and FX data consists of tick-by-tick data, 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 became continuously available. Both the original high-frequency local stock indices and FX time series are irregularly spaced. Previous tick interpolation was used to transform the observed data to

8

regularly spaced time series with an observation time of five minutes. Furthermore, in order to form the discounted accumulated WSIs we used daily dividend rates and USD interest rates. Omitting the high-frequency information in dividends and short rates is justified because only the exponentials of integrals of these quantities, but not the values themselves, enter the formulae for the relevant quantities in our construction of the WSIs, see (2.1) and (2.4).

Figure 6 displays the three different WSIs in addition to the daily observed MSCI. It is striking that, in spite of significant differences in the weights, all WSIs appear to be very similar. In particular, the similar fluctuations of all four indices are apparent. Using the normalisation function $\bar{\alpha}(t)$, given in (2.2), enables us to calculate the *normalized* WSI

$$Y(t) = \frac{\bar{V}^{(\text{WSI})}(t)}{\bar{\alpha}(t)} \tag{2.7}$$

for $t \in [0, T]$, as in (2.3) .

We omit a plot of the normalized WSIs, however the normalized MCI is shown in Figure 15 in an alternative time scale. The normalized WSIs all display similar dynamics. In Section 3, we will give a theoretical argument for this stylized fact. The hourly log-returns of the high-frequency MCI are shown in Figure 7, those of the EWI and GDPI are very similar. On the basis of their similar visual appearance it is reasonable to expect that the statistical characteristics of the intraday WSIs will be similar as well. However, this is not obvious and needs to be explored. To place this type of analysis in a theoretically sound context, the benchmark approach (BA) is introduced in the next section.
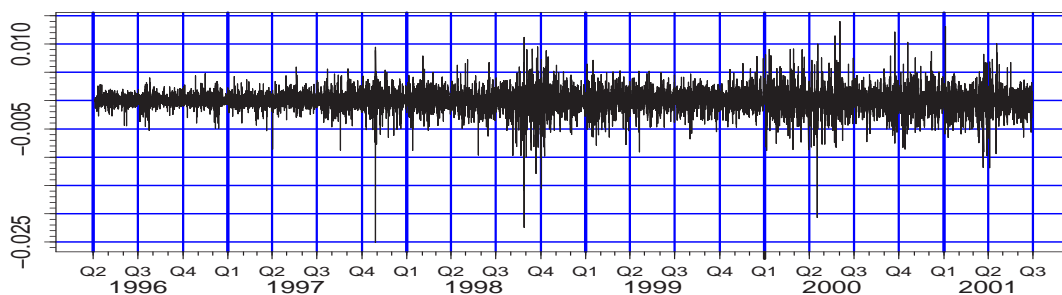

Figure 7: Hourly log-returns of the MCI.

3   Benchmark Approach

In the following we summarize essential facts about the BA in a somewhat informal way. To ensure the readability of this section, mathematical technicalities are omitted. For more detailed information we refer to Platen (2004).

9

3.1 Primary Security Accounts and Portfolios

We consider a continuous financial market with $d + 1$ primary security accounts with values $S^{(0)}(t), \ldots, S^{(d)}(t)$ at time $t \in [0, T]$, where T is finite. A primary security account contains units of a given asset with all accrued income reinvested. Here $S^{(0)}(t)$ models the riskless money market savings account and $S^{(j)}(t)$, the jth accumulation stock index at time t as mentioned in Section 2.2. Note however that a primary security account can consist of any type of asset class. Without loss of generality, the jth primary security account is supposed to satisfy the stochastic differential equation (SDE)

$$dS^{(j)}(t) = S^{(j)}(t) \left\{ r(t)\, dt + \sum_{k=1}^{d} b^{j,k}(t) \left(\theta^k(t) dt + dW^k(t) \right) \right\} \qquad (3.1)$$

for $t \in [0, T]$ with $S^{(j)}(0) > 0$, $j \in \{0, 1, \ldots, d\}$.

The uncertainties are modeled by d independent standard Brownian motions $W^k = \{W^k(t), t \in [0, T]\}$, $k \in \{1, 2, \ldots, d\}$ fulfilling the usual conditions, see Øksendal (1998). Only d sources of uncertainty are necessary in our complete market since one of the assets is riskless. The kth market price for risk $\theta^k(t)$ is proportional to the expected excess return that an investor obtains at time t for taking risk that is modeled by the kth Brownian motion.

The savings account is given by

$$S^{(0)}(t) = \exp \left\{ \int_0^t r(s) ds \right\} \qquad (3.2)$$

where $r(t)$ denotes the interest rate at time t. Zero volatility $b^{0,k}(t)$ applies for $S^{(0)}$ in (3.1) for all $t \in [0, T]$ and $k \in \{1, 2, \ldots, d\}$. The *volatility matrix* $b(t) = [b^{j,k}(t)]_{j,k=1}^{d}$ is assumed to be *invertible* to ensure that the uncertainties of the underlying securities are uniquely securitised.

Central to the BA is a portfolio that is used as benchmark and numeraire. The jth component $\delta^{(j)}(t) \in (-\infty, \infty)$ of the strategy δ denotes the number of units of the jth primary security account, which are held at time $t \in [0, T]$ in the corresponding portfolio, $j \in \{0, 1, \ldots, d\}$. The value $V^{(\delta)}(t)$ of the corresponding *portfolio* at time t can be written as the sum

$$V^{(\delta)}(t) = \sum_{j=0}^{d} \delta^{(j)}(t)\, S^{(j)}(t) \qquad (3.3)$$

for $t \in [0, T]$. The portfolio is called *self-financing* if

$$dV^{(\delta)}(t) = \sum_{j=0}^{d} \delta^{(j)}(t)\, dS^{(j)}(t) \qquad (3.4)$$

for all $t \in [0, T]$. That is, all changes in the value of the portfolio are due to gains from trade in the primary security accounts.

The self-financing portfolio value $V^{(\delta)}(t)$ satisfies according to (3.4) and (3.1) the SDE

$$dV^{(\delta)}(t) = V^{(\delta)}(t) \left\{ r(t) \, dt + \sum_{k=1}^{d} \beta_\delta^k(t) \left(\theta^k(t) \, dt + dW^k(t) \right) \right\} \qquad (3.5)$$

with kth volatility

$$\beta_\delta^k(t) = \sum_{j=0}^{d} \pi_\delta^{(j)}(t) \, b^{j,k}(t) \qquad (3.6)$$

and jth proportion

$$\pi_\delta^{(j)}(t) = \delta^{(j)}(t) \frac{S^{(j)}(t)}{V^{(\delta)}(t)} \qquad (3.7)$$

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

3.2 Growth Optimal Portfolio

From (3.5) it follows by application of the Itô formula for the logarithm of a strictly positive portfolio $V^{(\delta)}(t)$ the SDE

$$d \log \left(V^{(\delta)}(t) \right) = g_\delta(t) \, dt + \sum_{k=1}^{d} \beta_\delta^k(t) \, dW^k(t) \qquad (3.8)$$

with *portfolio growth rate*

$$g_\delta(t) = r(t) + \sum_{k=1}^{d} \beta_\delta^k(t) \left(\theta^k(t) - \frac{1}{2} \beta_\delta^k(t) \right) \qquad (3.9)$$

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

The *growth optimal portfolio* (GOP), see Long (1990), is the portfolio that maximizes the portfolio growth rate (3.9). By solving the first order conditions for the corresponding quadratic maximization problem one obtains the jth GOP proportion

$$\pi_{(\delta_*)}^{(j)}(t) = \sum_{k=1}^{d} [b^{-1}(t)]^{k,j} \theta^k(t) \qquad (3.10)$$

for $t \in [0, T]$ and $j \in \{1, 2, \dots, d\}$. Here $[b^{-1}(t)]^{k,j}$ denotes the (k, j)th element of the inverse $b^{-1}(t)$ of the volatility matrix. The GOP value $V^{(\delta_*)}(t)$ at time t satisfies the SDE

$$dV^{(\delta_*)}(t) = V^{(\delta_*)}(t) \left[r(t) \, dt + \sum_{k=1}^{d} \theta^k(t) \left(\theta^k(t) \, dt + dW^k(t) \right) \right] \qquad (3.11)$$

for $t \in [0, T]$. It can be seen from (3.11) that the GOP volatilities $\theta^k(t)$, $k \in \{1, 2, \ldots, d\}$, are the corresponding market prices for risk. Consequently, the drift of the GOP can be identified via its diffusion coefficients.

To form the GOP using the proportions (3.10) is a difficult task, see for example, Long (1990). However, the GOP can be approximated in a robust manner. It is well-known that well diversified stock portfolios behave in a similar fashion, see Figure 6. In Platen (2004) it has been shown given some weak regularity conditions, that any well diversified world stock portfolio approximates the GOP. This makes it reasonable to assume that the WSIs constructed in Section 2.2 approximate the GOP of the world stock market. The most intuitive property of the GOP is that it systematically outperforms all other portfolios in the long run. From this perspective, the MCI in Figure 6 seems to emerge as the best approximate GOP because it outperforms the other WSIs over the observation period considered.

3.3 Discounted GOP

Let us discount the GOP value $V^{(\delta_*)}(t)$, see (3.11), at time t by the savings account value $S^{(0)}(t)$, see (3.2). By application of the Itô formula together with (3.2) and (3.11), the *discounted* GOP

$$\bar{V}^{(\delta_*)}(t) = \frac{V^{(\delta_*)}(t)}{S^{(0)}(t)} \tag{3.12}$$

satisfies the SDE

$$d\bar{V}^{(\delta_*)}(t) = \bar{V}^{(\delta_*)}(t) \, |\theta(t)| \, (|\theta(t)| \, dt + d\hat{W}(t)) \tag{3.13}$$

where

$$d\hat{W}(t) = \frac{1}{|\theta(t)|} \sum_{k=1}^{d} \theta^k(t) dW^k(t) \tag{3.14}$$

and the GOP volatility $|\theta(t)|$ equals the *total market price for risk*

$$|\theta(t)| = \sqrt{\sum_{k=1}^{d} (\theta^k(t))^2}. \tag{3.15}$$

The *discounted GOP drift* $\alpha(t)$ is defined to be

$$\alpha(t) = \bar{V}^{(\delta_*)}(t) \, |\theta(t)|^2. \tag{3.16}$$

Allowing $\alpha(t)$ to be a, possibly stochastic, parameter process leads to a GOP volatility of the form

$$|\theta(t)| = \sqrt{\frac{\alpha(t)}{\bar{V}^{(\delta_*)}(t)}}. \tag{3.17}$$

We will demonstrate in this paper that one can model the discounted GOP drift in the form

$$\alpha(t) = \bar{\alpha}(t)\, m(t) \tag{3.18}$$

for $t \in [0, T]$, where $\bar{\alpha}(t)$ is given in (2.2). Here $m = \{m(t),\, t \in [0, T]\}$ denotes the nonnegative *market activity* process that fluctuates around one and will be specified below. This process is particularly important when modeling intraday indices as short term fluctuations in GOP volatility are captured by $m(t)$.

3.4 Normalized GOP and Market Activity

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

$$\psi(t) = \int_0^t m(s)\, ds \tag{3.19}$$

for $t \in [0, T]$. Note that (3.19) requires proper normalisation of the market activity. It is reasonable to normalize $m(t)$ in such a way that on average the market activity time scale elapses approximately as fast as physical time. The choice in our case is such that $t = 0$ years corresponds to the starting date of our sample 05/04/1996 00:00:00 GMT and $\psi(T) = 5.25$ years, which is equivalent to 30/05/2001 00:00:00 GMT, the terminal date of the data available. Furthermore, we assume that

$$\lim_{T \to \infty} \frac{1}{T} E\left(\psi(T)\right) = 1. \tag{3.20}$$

The *normalized GOP* $Y = \{Y_\psi,\, \psi \in [0, \psi(T)]\}$, see (2.3) and (2.7), in market activity time is defined to be

$$Y_{\psi(t)} = Y(t) = \frac{\bar{V}^{(\delta_*)}(t)}{\bar{\alpha}(t)}. \tag{3.21}$$

It is straightforward to show via the Itô formula, by using (3.13) and (3.18), that (3.21) satisfies the SDE

$$dY_\psi = \eta\left(\frac{1}{\eta} - \frac{Y_\psi}{m_\psi}\right) d\psi + \sqrt{Y_\psi}\, d\hat{W}_\psi \tag{3.22}$$

for $\psi \in [0, \psi(T)]$. Here we set

$$d\hat{W}_{\psi(t)} = \sqrt{m(t)}\, d\hat{W}(t) \quad \text{and} \quad m_{\psi(t)} = m(t) \tag{3.23}$$

for $t \in [0, T]$. The normalized GOP in market activity time in (3.22), is for $m_\psi = 1$, a *square root process with dimension four*, which is a CIR process, see Cox, Ingersoll & Ross (1985). As such, the solution of (3.22) has a long term mean of $\frac{1}{\eta}$ and a speed of adjustment parameter η. One observes that the only relevant

parameter in (3.22) is the net growth rate η, which is also a key parameter for the economy. If we consider the square root of the normalized GOP, then by (3.22) and application of the Itô formula, it evolves according to the SDE

$$d\left(\sqrt{Y_\psi}\right) = \left(\frac{3}{8\sqrt{Y_\psi}} - \frac{\eta\sqrt{Y_\psi}}{2\,m_\psi}\right) d\psi + \frac{1}{2}\,d\hat{W}_\psi \qquad (3.24)$$

for $\psi \in [0, \psi(T)]$. It is crucial to note that the diffusion coefficient in (3.24) is constant. Therefore we obtain in market activity time the quadratic variation of $\sqrt{Y}$ in the form

$$\left\langle \sqrt{Y} \right\rangle_\psi = \frac{\psi}{4} \qquad (3.25)$$

for $\psi \in [0, \psi(T)]$, see Platen (2004). Relation (3.25) holds under general circumstances since no restrictive assumptions have been imposed on the actual model dynamics. The market activity process can still be freely chosen.

Note that from the market activity time given in (3.19) and the quadratic variation of $\sqrt{Y}$ in (3.25), the market activity can be calculated as

$$m(t) = \frac{d\psi(t)}{dt} = 4\frac{d\langle\sqrt{Y}\rangle_t}{dt} \qquad (3.26)$$

for $t \in [0, T]$. This implies that market activity is *directly observable*. A measure of the slope of the quadratic variation of the square root of the normalized GOP is all that is required. For the daily observed MSCI the quadratic variation of its square root has been shown in Figure 5. The slope of this graph is then the corresponding market activity, which remains to be modeled. However, seasonal patterns are to be expected in intraday market activity since the market experiences active and non-active periods that depend on the time of the week, see, for example, Dacorogna, Gençay, Müller, Olsen & Pictet (2001) and Breymann, Dias & Embrechts (2003).

4   Intraday Market Activity

4.1   Market Activity of the WSIs

As a next step we investigate the three intraday WSIs introduced in Section 2.2. According to (3.25) their market activity time is simple to calculate. In Figure 8 we show the market activity time of the MCI for the period covered by the intraday data. This allows us to calculate the market activity directly. We simply calculate the numerical derivative corresponding to (3.26) using five minute time steps. This derivative fluctuates over a wide range. Therefore, we show the logarithm $\ln(m(t))$ of this derivative in Figure 9 for the MCI over a few weeks
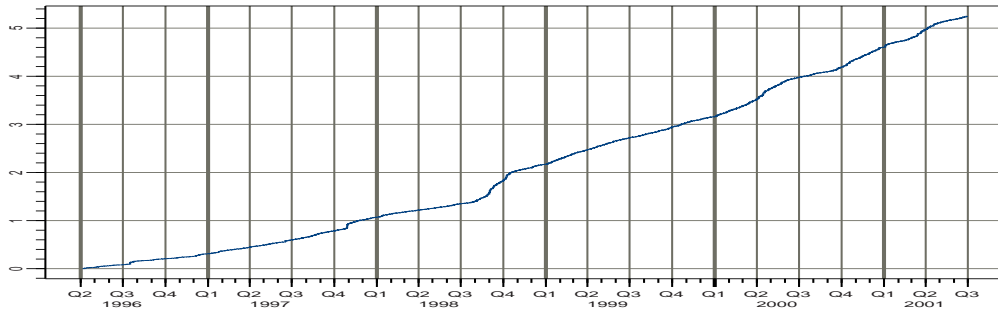
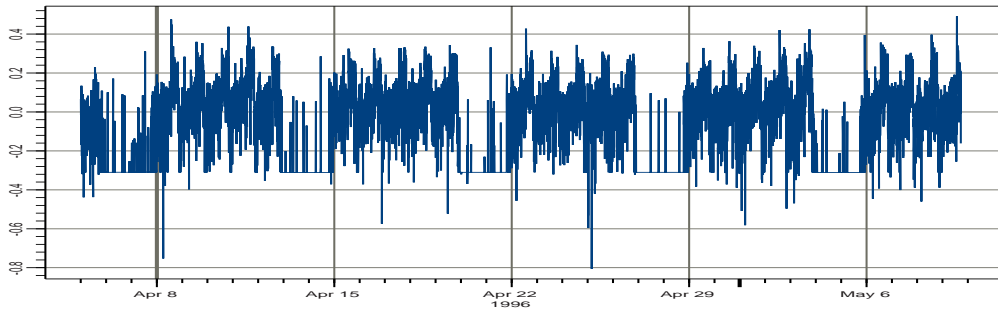Figure 8: The market activity time for the MCI.

Figure 9: Logarithm of market activity, $\ln(m(t))$, for the MCI.

in April/May 1996. It appears that the observed market activity process shows some seasonal patterns and reverts quickly back to a reference level.

The observed market activity processes for the EWI and GDPI look very similar and are therefore omitted. This segment of the logarithm of market activity illustrates the seasonality apparent in the average level of market activity and to some degree also some seasonality present in the average fluctuation. The
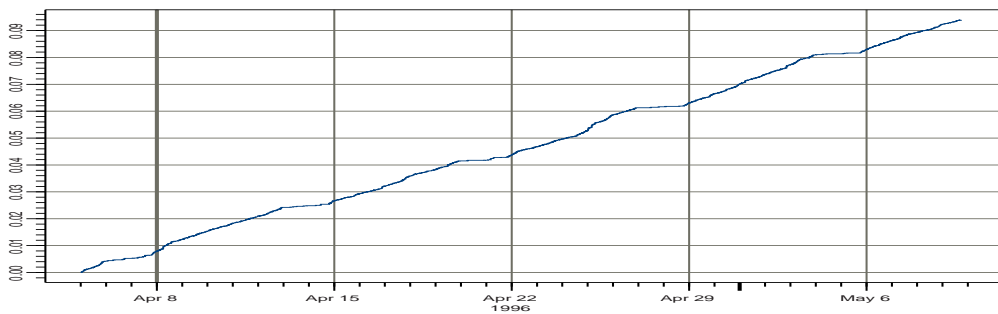

Figure 10: Quadratic variation of the logarithm of the market activity, $\langle \ln(m) \rangle_t$, process for the MCI.

presence of seasonality within the fluctuations of market activity is further confirmed by the quadratic variation $\langle \ln(m) \rangle_t$ of the logarithm of the market activity, which is shown for the above segment in Figure 10. Here we see that the weekends are characterised by a plateau in the quadratic variation. Notably, despite the seasonal pattern, the quadratic variation in Figure 10 appears to be almost linear for the periods when the global market is open and actively trading. Of interest is whether the seasonal patterns in the level and fluctuations of $\ln(m(t))$ can be extracted leaving only an almost stationary residual market activity process. The deseasonalization of $m(t)$ requires a two stage procedure. The first step deseasonalizes the average of the market activity and the second its diffusion coefficient.

In order to model the market activity it is necessary to characterize the relationship between its fluctuations and those of the normalized WSI. For each WSI we analyzed the covariation of the square root of the normalized WSI with the logarithm of the corresponding market. The quadratic covariation is the sum of the product of the increments of the respective processes. We observed in all three cases over the entire period that the observed covariation remains close to zero. There seems to be some evidence of a slight positive trend. However, for simplicity we assume that the noise that drives market activity is independent of that driving the WSI. This assumption will be relaxed in forthcoming work.

4.2   Modeling Market Activity

The expected value of the market activity is deseasonalized by the *average seasonal market activity* $\hat{m}(t)$ at time t, which is defined as the expectation

$$\hat{m}(t) = E\left(m(t)\right). \tag{4.27}$$

By the law of large numbers we obtain an estimate for $\hat{m}(t)$ for each five minute interval of the week during the full observation period, 258 weeks in total, from the arithmetic average of the corresponding observed market activity. This means we simply estimate the expectation of the market activity at time t by assuming the same pattern for each week and sampling over all weeks of observation. Figure 11 displays the estimated Northern Hemisphere summertime weekly pattern of the average market activity $\hat{m}(t)$. The average wintertime market activity is simply shifted by one hour. Note that Figure 11 is calculated from market activity, not the logarithm. The overall daily pattern is composed of several U-shaped patterns of different magnitude. The individual patterns are characteristic of localized, exchange-traded markets. As expected, the magnitude is highest when the European and American market are simultaneously active. It is lowest during the pacific gap, that is 21:00 till 00:00 GMT, and of course on the weekends.

Figure 10 suggests that the market activity $m(t)$ is likely to have multiplicative noise when global markets are open, with some seasonal activity volatility since
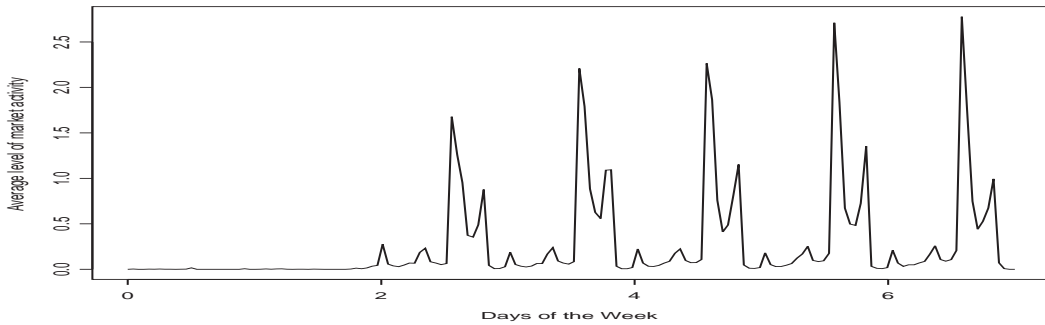
Figure 11: Weekly pattern of average market activity $\hat{m}(t)$ for summertime in the Northern Hemisphere.

the quadratic variation of its logarithm shows a piecewise linear pattern. For the open market regime, a possible model for $m(t)$ with multiplicative noise is given by the SDE

$$dm(t) = \frac{\gamma}{2}\beta^2(t)m(t)\left(\frac{p(t)-1}{\gamma} - m(t)\right)dt + \beta(t)m(t)dW(t) \qquad (4.28)$$

with *speed of adjustment* $\gamma > 0$, *reference level* $p(t) > 0$ and *activity volatility* $\beta(t) > 0$ for $t \in [0,T]$. The reference level $p(t)$ and the activity volatility $\beta(t)$ are assumed to exhibit some weekly periodic seasonal patterns. The activity volatility $\beta(t)$ is estimated by averaging over the weekly observations of $\beta^2(t)$. Figure 12 shows the estimated values for $\beta(t)$ for the MCI during summertime in the Northern Hemisphere. These estimates are also obtained from the 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 $\beta(t)$ outside the open market regime. Note that during the open market trading days the activity volatility is almost constant and slightly larger than one. We see that the activity volatility spikes once the Asia-Pacific markets have opened.
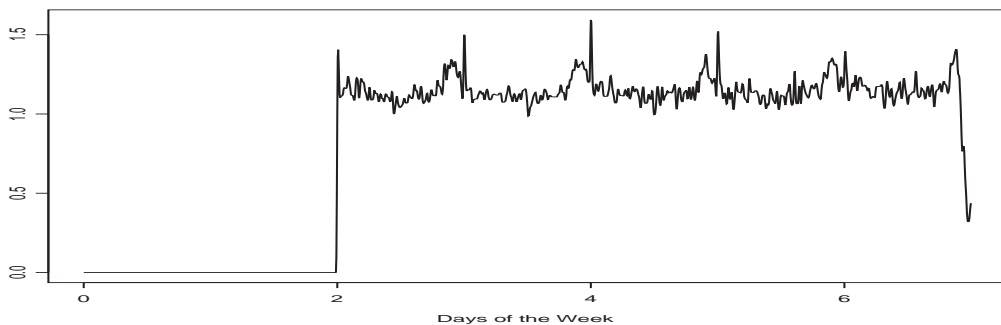

Figure 12: Average weekly pattern of activity volatility $\beta(t)$ for summertime in the Northern Hemisphere.

17

The activity volatility $\beta(t)$ allows us to introduce *activity volatility time* $\tau = \{\tau(t),\, t \in [0, T]\}$ as

$$\tau(t) = \langle \ln(m) \rangle_t = \int_0^t (\beta(u))^2 \, du \tag{4.29}$$

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

It is useful to consider the market activity in activity volatility time $\tau(t)$, denoted by $m_{\tau(t)} = m(t)$, as the presence of seasonal patterns may overshadow empirical features. Then, by (4.29) we obtain the SDE (4.28) in activity volatility time,

$$dm_\tau = \frac{\gamma}{2} m_\tau \left(\frac{p_\tau - 1}{\gamma} - m_\tau \right) d\tau + m_\tau dW_\tau, \tag{4.30}$$

where $\tau \in [0, \tau(t)]$, $p_{\tau(t)} = p(t)$ and

$$dW_{\tau(t)} = \beta(t) dW(t) \tag{4.31}$$

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

By considering the logarithm of market activity in activity volatility time and using Itô's formula together with (4.30) we obtain the SDE

$$d \ln(m_\tau) = \gamma \left(\frac{p_\tau}{\gamma} - m_\tau \right) d\tau + dW_\tau \tag{4.32}$$

for $\tau \in [0, T]$. Note that the diffusion coefficient in (4.32) is constant and equal to one. Furthermore, we see that, for the open market regime, the dynamics of the logarithm of market activity are described by a mean-reverting model with speed of adjustment γ and time dependent reference level $\frac{p_\tau}{\gamma}$. By considering (4.32), we may set $p_\tau \approx \hat{m}_\tau \gamma + \frac{d \ln(\hat{m}_\tau)}{d\tau}$ to account for the seasonal patterns in the reference level.
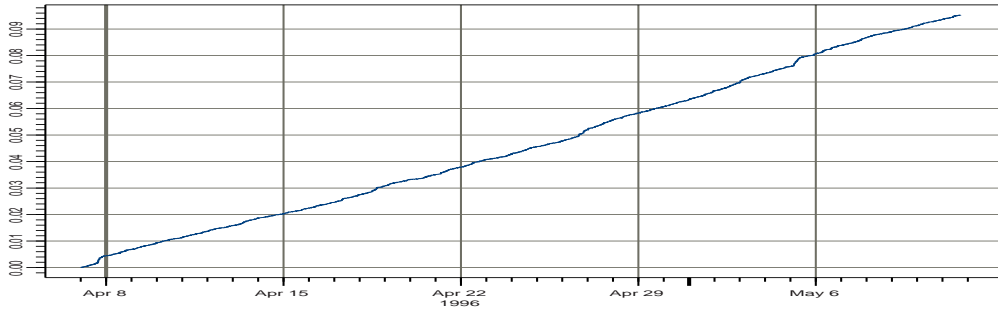

Figure 13: Quadratic variation $\langle \ln(m) \rangle_\tau$ of the logarithm of the normalized market activity process in activity volatility time.

The quadratic variation of the logarithm of market activity is shown in Figure 13. It confirms the theoretical slope of one of the quadratic variation $\langle \ln(m) \rangle_\tau = \tau$ that follows from (4.32).

4.3 Estimation of the Speed of Adjustment

The speed of adjustment parameter γ in (4.30) remains to be estimated. In Figure 14 we plot the histogram of all observations of the logarithm $l_\tau = \ln(m_\tau)$ of the market activity. One notes a very distinct shape of the histogram with a concentration of negative spikes around -0.4. These negative spikes and other negative values result from the effects of market opening and closing, which can be interpreted as the typical levels of market activity in the opening and closing regimes. The given model dynamics for market activity during the open market regime, see (4.28) and (4.30), are not intended to include the opening and closing market regimes. For this reason we will ignore values less than $\underline{l} = -0.2$ in the histogram in Figure 14 for the estimation of the drift parameters for the open market regime that is modeled by (4.32). A more detailed model would need to incorporate the probability mass created in the lower tail by the superposition of the stationary densities for the different regimes.

Under the simplifying assumption that $p_\tau = p$ is constant, the market activity process can be shown to have as stationary density a gamma density. The stationary density of the logarithm l of the corresponding market activity process in activity volatility time can be written as

$$\bar{p}_l(l; \gamma, p) = \frac{(\gamma)^p}{\Gamma(p)} \exp\left\{-\gamma e^l\right\} e^{l(p-1)}. \tag{4.33}$$

Here $\Gamma(\cdot)$ is the Gamma function and γ and p are the only free parameters.

In this case, we can perform a restricted maximum likelihood estimation. As indicated, 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

$$L(\gamma, p) = \sum_{n=1}^{n_T} \mathbf{1}_{\{l_{\tau_n} \geq \underline{l}\}} \ln\left(\bar{p}_l(l_{\tau_n}; \gamma, p)\right). \tag{4.34}$$

Here $n_T = 523585$ 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, $n \in \{1, 2, \ldots, n_T\}$. A plot of the estimated probability density function of $l_\tau = \ln(m_\tau)$ based on the resulting maximum likelihood estimate $\hat{\gamma} \approx 103$ and $\hat{p} \approx 106$ is shown in Figure 14. We see from the effects of market opening and closing that the left tail of the histogram is fatter than what is given by the theoretical gamma density. The estimates for each of the WSIs, together with 99% confidence intervals, are shown in Table 1.

4.4 Analysis of Normalized WSIs in Market Activity Time

In Section 3 it was shown that the normalized GOP, when observed in market activity time, is a square root process of dimension four. The trajectory of the

Index	$\hat{\gamma}$ $[year^{-1}]$	$\hat{p}$ $[year^{-1}]$
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 1: Estimates for the drift parameters of market activity in activity volatility time with 99% confidence intervals.
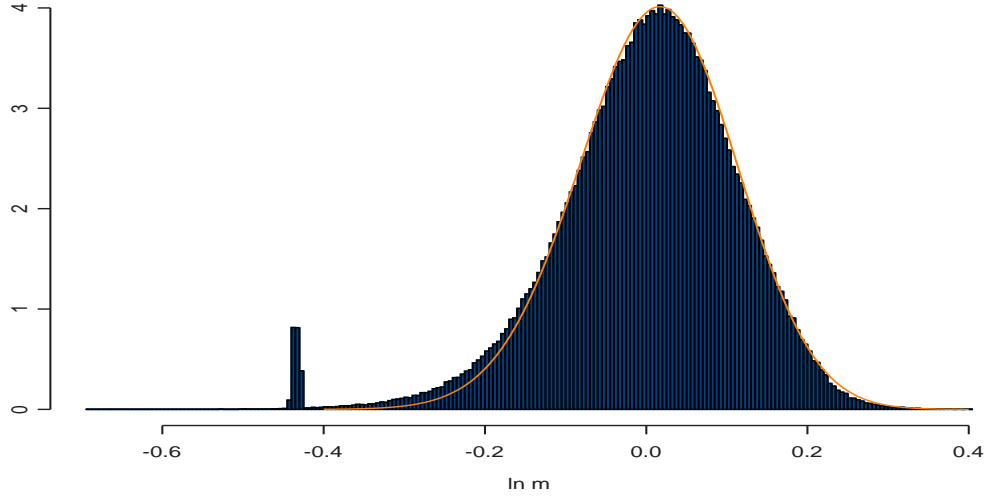

Figure 14: Histogram and estimated probability density function of $\ln(m_\tau)$.

normalized MCI in market activity time is plotted in Figure 15. By using market activity time we know that the quadratic variation of the square root of the normalized MCI should be linear with a slope close to 0.25, see (3.25). This relationship is confirmed by performing a simple linear regression of the quadratic variation of the square root of the normalized WSI against the corresponding estimated market activity time. The slope coefficient and R^2 value are given in Table 2 for all three WSIs.

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

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

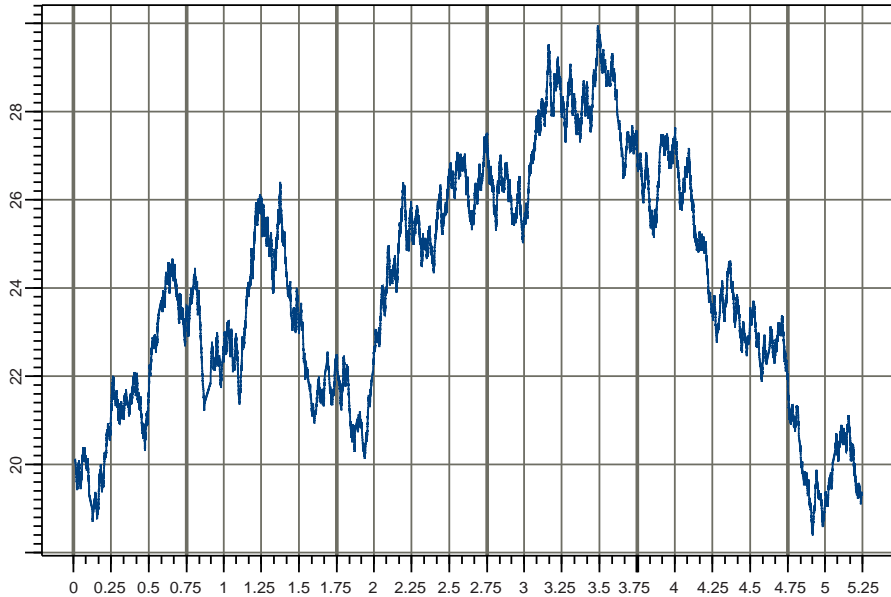Table 2 shows that the MCI and GDPI fit the model well and the MCI outper-

Figure 15: Normalized MCI in market activity time.

forms the EWI and the GDPI, when the quadratic variations of the square root of the corresponding normalized index are considered. Additionally, the corresponding market activity processes in activity volatility time are shown to have the hypothesized dynamics, when the opening and closing regimes are omitted. Moreover, it follows from Figure 6 that the MCI has the maximum value in relation to all other indices considered, at the end of the observation period, which is a main property of the GOP. Consequently, we consider the MCI to be the best proxy for the GOP and the world stock market portfolio on the basis of the above statistical results. The model calibrated above appears to be a largely accurate intraday description of the dynamics of the market capitalization weighted world stock index.

5   Conclusion

We constructed and examined three diversified high-frequency world stock indices. A simple and robust way of calculating and modeling market activity is demonstrated. Market activity is found to contain seasonal patterns in both the drift and the diffusion term. We showed that the market activity can be modeled as a strongly mean reverting process with constant speed of adjustment incorporating the weekly periodicities contained in average market activity and activity volatility. Furthermore, we confirmed that, in market activity time, the normalized indices represent square root processes of dimension four. The world stock portfolio appears to be an excellent proxy for the GOP. Work is in progress on the use of the MCI for applications of the benchmark approach in derivative pricing, portfolio management and integrated risk management.

Acknowledgement

References

Bollerslev, T., R. Y. Chou, & K. F. Kroner (1992). ARCH modeling in finance. *Journal of Econometrics* **52**, 5–59.

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

Cox, J., J. Ingersoll, & S. Ross (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.

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

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

Goodhart, C. & M. O'Hara (1997). High–frequency data in financial markets: Issues and applications. *Journal of Empirical Finance* **4**, 73–114.

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.

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

Long, J. (1990). The numeraire portfolio. *J. Financial Economics* **26**, 26–69.

Merton, R. (1992). *Continuous–Time Finance*. Blackwell, Oxford.

Müller, U., M. Dacorogna, & O. Pictet (1998). Hill, bootstrap and jackknife estimators for heavy tails. In R. Adler, R. Feldman, and M. Taqqu (Eds.), *A practical guide to heavy tails: Statistical techniques for Analysing Heavy Tailed Distributions*, pp. 283–310. Birkhäuser, Boston.

Øksendal, B. (1998). *Stochastic Differential Equations. An Introduction with Applications* (FIFTH ed.). Springer, Berlin.

Platen, E. (2002). Arbitrage in continuous complete markets. *Advances in Applied Probability* **34**, 540–558.

Platen, E. (2004). Modelling the volatility and expected value of a diversified world index. *International Journal of Theoretical and Applied Finance*. Forthcoming.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* **4**(2), 183–204.