Dynamics of a Well-Diversified Equity Index

Eckhard Platen and Renata Rendek
Dynamics of a Well-Diversified Equity Index

January 17, 2019

Eckhard Platen* 1,2,3 and Renata Rendek 4

Abstract: The paper derives a parsimonious model for the long-term dynamics of a well-diversified stock index, the S&P500. The index is modeled as growth optimal portfolio. Its normalized value evolves, in some market time, as a square root process. The derivative of market time is a linear function of the squared derivative of a smoothed proxy of the single driving Brownian motion. The model explains the feedback effects from index moves typically observed for monthly and daily S&P500 values. It is highly tractable, permits almost exact simulation and leads beyond classical assumptions in finance.

Key words and phrases: long-term index model, growth optimal portfolio, square root process, market time, leverage effect puzzle, benchmark approach.

JEL Classification: G10, C10, C15
1991 Mathematics Subject Classification: 62P05, 62P20, 62G05, 62-07, 68U20

1University of Technology Sydney, Finance Discipline Group and School of Mathematical and Physical Sciences, PO Box 123, Broadway, NSW, 2007, Australia, Email: Eckhard.platen@uts.edu.au, Phone: +61295147759.
2University of Cape Town, Department of Actuarial Science.
3College of Business and Economics, Australian National University, Canberra.
4University of Technology Sydney, Quantitative Finance Research Centre, School of Mathematical and Physical Sciences.
1 Introduction

The accurate modeling and theoretical understanding of the long-term dynamics of a well-diversified equity index, e.g. the S&P500, have been challenging tasks. No agreement has so far emerged in the literature what a reasonably accurate, long-term index model should look like. The market index appears to evolve in its own market time, as pointed out in Clark (1973): ‘On days when no new information is available, trading is slow, and the price process evolves slowly. On days when new information violates old expectations, trading is brisk, and the price process evolves much faster’. The challenge is to capture the typical feedback in market activity in a parsimonious model so that it explains estimated volatility.

The traditional focus on modeling volatility as a separate stochastic process seems to have hindered an early solution of this modeling challenge. It encounters the difficulty that the model volatility is usually hidden and only observed indirectly through the estimated volatility. When using estimated volatility for a given observation frequency, it is extremely difficult, if not impossible, to infer any link between the hidden model volatility and the index. In Ait-Sahalia, Fan & Li (2013) these modeling difficulties have been labeled as the leverage effect puzzle. Recently, volatility estimated from high-frequency data revealed that volatility paths are rough with many spikes; see e.g. Bayer, Friz & Gatheral (2016) and Gatheral, Thibault & Rosenbaum (2018). The estimated rough volatility and also observed flash crashes, as the one experienced in 2010 on May 6, request convincing parsimonious modeling.

‘Ad hoc’ models, typically chosen for tractability, dominate the literature. Adding another ‘ad hoc’ model would not provide much progress in our understanding of the dynamics of well-diversified indexes. We aim to change this disappointing situation by deriving a new class of models which we base on three crucial insights. These insights involve the notion of a growth optimal portfolio (GP), which is the portfolio that is maximizing expected logarithmic utility; see Kelly (1956) and Merton (1992). They also exploit the similarity of diversified wealth evolutions with the dynamics of population sizes of birth-and-death-processes, also known as branching processes; see Feller (1971). We capture the three insights by the following assumptions:

- A well-diversified total return index is a proxy of the respective growth optimal portfolio.
- A well-diversified index evolves, in some market time, similarly to the continuous time limit of population sizes of birth-and-death processes.
- The derivative of market time is a function of the derivative of a moving average of a proxy of the single driving Brownian motion.
The first two assumptions determine the ‘normal’ dynamics of the index, which concerns its evolution in some market time. The third assumption models the market activity, the derivative of the market time, which accelerates the ‘normal’ index dynamics when the index moves away from its recent values. In the third assumption we deliberately leave some freedom for the specification of the market activity because it may turn out to be different for different indexes. We propose with these three assumptions a new model class. The new model class is nested in continuous time finance, pioneered by Merton (1973). It is highly tractable and allows almost exact simulation. Most importantly, it leads beyond classical finance assumptions and is, therefore, derived under the benchmark approach; see Platen (2002) and Platen & Heath (2010).

To be directly applicable and comparable with other index models, the current paper focuses on modeling the dynamics of the S&P500, which is arguably the best studied well-diversified equity index. By specifying the market activity as a linear function of the square of the derivative of the moving average of a proxy of the driving Brownian motion, we obtain the model we propose for the S&P500. When fitting the proposed model to the monthly and daily observed S&P500 data, we extract also the paths of not directly observable model components, including that of the through the model postulated single driving Brownian motion. To control the propagation of errors in the extraction of hidden model components from monthly data, we need to employ higher order, implicit Wagner-Platen expansions for the increments of the solution of the model’s stochastic differential equation (SDE); see Platen & Bruti-Liberati (2010) and Kloeden & Platen (1999).

The proposed estimation of the model parameters ‘inverts’, in some sense, scenario simulation for an SDE where the Brownian motion path is the input. The model postulates a single Brownian motion path, which becomes an output of the inference. This allows us to test whether this path cannot be rejected as that of a true Brownian motion, using a variance ratio test; see Chen & Deo (2006). The proposed new model class generates naturally the well-known leverage effect; see e.g. Black (1976). It can also explain from high-frequency data observed roughness of estimated volatility paths, as well as, occasional flash crashes.

The paper is organized as follows: Section 2 describes, illustrates and fits the proposed model to the S&P500. Section 3 briefly reviews the literature on volatility and index modeling from the perspective of the proposed model. Section 4 derives the model from three assumptions. Almost exact simulation of the model is described in Section 5.
2 Long-Term Index Model

2.1 Model

In this subsection we present the proposed model, where we defer the discussion of its links to the literature to Section 3 and its derivation to Section 4. The value $S_t$ at time $t$ of a well-diversified, real-value (inflation adjusted), total return (dividends reinvested) stock index is modeled by the product

$$S_t = A_t Y_t,$$  

for $t \geq 0$. The ‘fundamental value’ $A_t$ is defined as

$$A_t = A_0 \exp(\tau_t + \ell t),$$

where we call $\ell$ the excess growth rate and $A_0 > 0$ the initial ‘fundamental value’. The normalized index $Y_t = \frac{S_t}{A_t}$ follows in some market time $\tau = \{\tau_t, t \geq 0\}$ a square root process with SDE

$$dY_t = (1 - Y_t)d\tau_t + \sqrt{Y_t}dW_{\tau_t},$$

for $t \geq 0$. Here $W = \{W_\tau, \tau \geq 0\}$ is the single driving Brownian motion, which models the nondiversifiable uncertainty of the given stock universe. The market time

$$\tau_t = \int_0^t M_s ds$$

is characterized by its derivative, the market activity

$$M_t = \left(\left(U_t'\right)^2 + \varepsilon\right) \xi.$$  

The latter is a linear function of the square of the derivative

$$U_t' = \frac{dU_t}{dt} = \lambda(2Y_t^{\frac{1}{2}} - U_t)$$

of the, in (2.6) determined, moving average $U_t$ of $2Y_t^{\frac{1}{2}}$ with initial value $U_0 = 2Y_0^{\frac{1}{2}}$.

The three state variables of the proposed model are the ‘fundamental value’ $A_t$, the normalized index $Y_t$ and the moving average $U_t$. $S_0$ and $A_0$ represent initial values and the four remaining parameters consist of the excess growth rate $\ell$, the baseline activity $\varepsilon > 0$, the scaling $\xi > 0$ and the speed of adjustment $\lambda \geq 0$. These few parameters are all economically meaningful. Setting $\lambda$ equal to zero yields the deterministic market time $\tau_t = \varepsilon \xi t$ and the stylized version of the proposed model results. When substituting the function (2.5) by more refined functions, other models from the new model class emerge that may not only fit well the monthly and daily observed S&P500 but also its intraday values.
2.2 Long-Term Index Dynamics

Figure 2.1: Logarithm of standardized index $\ln(\tilde{S}_t)$ and estimated market time $\hat{\tau}_t$.

Figure 2.2: Normalized index $Y_t$.

By employing monthly observed S&P500 data, we illustrate in the following key properties of the proposed model. The S&P500 data after 1963 are obtained from Datastream Thomson Financial and the earlier part from 1871 until 1963 is reconstructed in Shiller (2015). Later we will use also daily observed S&P500 data to demonstrate that the model fitted to monthly data fits also daily data.

The available observation windows for estimating parameters in drift coefficients of SDEs for financial securities are usually too short to provide useful estimates; see e.g. DeMiguel, Garlappi & Uppal (2009). However, as we will see, the 144 years of monthly S&P500 data permit us to estimate a meaningful constant excess growth rate $\ell$ by matching, in a least squares sense, the long-term theoretical
average of the normalized index $Y$ with its observed average. The excess growth rate can be interpreted as the average real interest rate. Since the normalized index is in market time an ergodic process, it has a long-term theoretical average, which is the value 1.0. By searching for the minimum

$$
\min_{A_0, \ell} \left( \frac{1}{N} \sum_{i=1}^{N} Y_{t_i} - 1 \right)^2,
$$

(2.7)

we estimate the excess growth rate $\ell$ and the initial ‘fundamental value’ $A_0$. $N$ is here the number of observations and $0 = t_0 < t_1 < \cdots < t_i < \cdots < t_N$ are the observation times. For the monthly observations of the S&P500 we set $t_i = t_{i-1} + \frac{1}{12}$. During the above minimization we calculate, according to (2.1) and (2.2), the normalized index value $Y_{t_i}$ as the ratio

$$
Y_{t_i} \approx \frac{\hat{S}_{t_i}}{\exp\{\hat{r}_{t_i}\}}
$$

(2.8)

by employing the standardized index value

$$
\hat{S}_{t_i} = \frac{S_{t_i}}{A_0 \exp\{\ell t_i\}}.
$$

(2.9)

The market time $\tau_{t_i}$ can be obtained via the formula

$$
\tau_{t_i} = \ln \left( 4[\sqrt{\hat{S}}]_{t_i} + 1 \right),
$$

(2.10)

which follows by the Itô formula from (2.3) and (2.8). Here $[X]_t$ denotes the quadratic variation of a process $X = \{X_t, t \geq 0\}$, which is the limit of the sum of the squares of its increments when the step size converges to zero. We approximate the quadratic variation on the right hand side of (2.10) by the approximate quadratic variation

$$
[\sqrt{\hat{S}}]_{t_i} \approx \sum_{k=1}^{i} \left( \sqrt{\hat{S}_{t_k}} - \sqrt{\hat{S}_{t_{k-1}}} \right)^2
$$

(2.11)

for $i = 1, 2, \ldots, N$. The estimated market time $\hat{\tau}_{t_i}$ is, thus, obtained via the formula

$$
\hat{\tau}_{t_i} = \ln \left( 4 \sum_{k=1}^{i} \left( \sqrt{\hat{S}_{t_k}} - \sqrt{\hat{S}_{t_{k-1}}} \right)^2 + 1 \right).
$$

(2.12)

From the above least squares fit of the monthly observed S&P500, with initial index value $S_0 = 83.78$, we infer the initial ‘fundamental value’ $A_0 = 157.007$ and the excess growth rate $\ell = 0.011$. Note that the average real interest rate for the US economy was estimated in Dimson, Marsh & Staunton (2002) at about 0.01
for the last century, which matches our findings. Thus, we could have avoided estimating the excess growth rate by simply taking the observed average of the real interest rate or the time-varying observed real interest rate. This would have removed the need for estimating the only drift parameter in the model.

We extract the path of the standardized index $\tilde{S}_t$ and display in Figure 2.1 its logarithm $\ln(\tilde{S}_t)$. Due to (2.2) one can interpret $\ln(A_t \exp(-lt)/A_0) = \ln(\exp(\tau_t)) = \tau_t$ as the logarithm of the 'fundamental value' of the standardized index, it makes sense to display, for comparison, in Figure 2.1 also the estimated market time $\hat{\tau}_t$.

We emphasize the visually confirmed remarkable fact that the model provides an equivalence between market time and the long-term expectation of the logarithm of the standardized index.

The in Figure 2.1 observable mean-reverting dynamics of the index becomes even better visible in Figure 2.2, where we display the normalized index value $Y_t$, which mean-reverts theoretically to the level 1.0. The normalized index process $Y_t$ is by the SDE (2.3) in market time a square root process (often called Cox-Ingersoll-Ross process) of dimension four and, thus, never reaches zero; see Revuz & Yor (1999). It has an explicitly known transition density of non-central chi-square type. Therefore, it belongs to a highly tractable, well-studied class of scalar diffusion processes with a wide range of explicit formulae for many of its functionals; see Baldeaux & Platen (2013).

The mean-reversion rate or speed of adjustment parameter with respect to market time of the normalized index has in the SDE (2.3) the value $\frac{1}{2}$, which indicates a half-life time of $\ln(2) \approx 0.693$ in market time of shocks to the normalized index. This translates into a half-life time of about $\ln(2)/\hat{M} \approx 40$ calendar years, when taking into account the average market activity of about $\hat{M} = 2.5/144 \approx 0.017$; see Figure 2.1. This means that the normalized index evolves very slowly with a half-life time of shocks of about the length of a human working life.

### 2.3 ‘Normal’ Volatility

The model volatility $\sigma_t$ of $S_t$ is by (2.1), (2.2) and (2.3) equivalent to the model volatility of the normalized index $Y_t$. It equals by (2.3) and (2.4) the product

$$\sigma_t = \theta_t \sqrt{M_t}. \quad (2.13)$$

The ‘normal’ volatility $\theta_t=Y_t^{-\frac{1}{2}}$ is the volatility with respect to market time, the $\tau$-time, and shown in Fig.2.3. By comparing Fig.2.3 and Fig.2.2 one can see that the ‘normal’ volatility increases when the normalized index decreases and vice versa, which models naturally the well-known leverage effect; see Black (1976).

The squared ‘normal’ volatility $(\theta_t)^2$ equals by (2.13) the inverse $Y_t^{-1}$ of the normalized index. It satisfies by (2.3) and application of the Itô formula the SDE
Figure 2.3: ‘Normal’ volatility $\theta_t = Y_t^{-\frac{1}{2}}$.

\[ d (\theta_t)^2 = (\theta_t)^2 d\tau_t - ((\theta_t)^2)^{\frac{3}{2}} dW_{\tau_t}. \]

(2.14)

One notes in the diffusion coefficient of the SDE (2.14) the power $3/2$, which identifies the ‘normal’ volatility dynamics as those of the $3/2$-volatility model, originally proposed in Platen (1997) (as early part of the development of the benchmark approach) and independently suggested in Heston (1997) (due to its tractability as inverse of a square root process). Key features of the $3/2$-volatility model are the rather high power of $1.5$ in the diffusion coefficient of the SDE (2.14), which can reduce or increase rapidly the heteroscedasticity of volatility. Note that the $3/2$-volatility model describes the volatility dynamics of the stylized version of the proposed model when setting the market activity constant.

2.4 Market Activity

The model volatility $\sigma_t$, given in (2.13), is proportional to the square root $\sqrt{M_t}$ of the market activity $M_t$. The formula (2.5) for the market activity employs the moving average $U_t$, determined by the differential equation (2.6). For extracting the model market activity under the proposed model, the increments of the solution of this differential equation have to be approximated for the rather large time step size of the monthly observed S&P500. As pointed out in Chapter 12 of Kloeden & Platen (1999) and Chapter 14 of Platen & Bruti-Liberati (2010), substantial truncation errors may arise when using large time step sizes for a Wagner-Platen expansion of increments of the solution of an SDE. When extracting below the model market activity we, essentially, ‘invert’ scenario simulation and, thus, potentially face this type of problems. To overcome these difficulties, the semi-drift-implicit Wagner-Platen expansion

\[ U_{t_i} \approx \left( U_{t_{i-1}} \left( 1 - \frac{\lambda}{2} \Delta \right) + \lambda \Delta \left( Y_{t_i}^{\frac{1}{2}} + Y_{t_{i-1}}^{\frac{1}{2}} \right) \right) \left( 1 + \frac{\lambda}{2} \Delta \right)^{-1} \]  

(2.15)
for $\Delta = t_i - t_{i-1}$, $i \in \{1, 2, \ldots, N\}$, with $U_{t_0} = 2Y^{1/2}_{t_0}$ is employed, suggested for scenario simulation in Section 12.2 of Kloeden & Platen (1999). Since the market activity can move extremely fast, such a higher strong order, implicit stochastic approximation was found to be necessary for successful inference when using monthly observed S&P500 data. In Figure 2.4 we display $2Y^{1/2}_t$ and its moving average $U_t$.

![Figure 2.4: $2Y^{1/2}_t$ (black) and its moving average $U_t$ (red).](image)

Now, we can form the model market activity $M_{t_i} = \xi\lambda^2(2Y^{1/2}_{t_i} - U_{t_i})^2 + \xi \varepsilon$ to obtain the integrated model market activity $\sum_{k=1}^i M_{t_k} \Delta \approx \tau_{t_i}$, where $\Delta = t_k - t_{k-1}$ for $k = 1, \ldots, N$, with $N$ denoting the number of observations; see (2.5). For

![Figure 2.5: Model market activity $M_t$.](image)

estimating the remaining parameters $\lambda$, $\xi$ and $\varepsilon$ we search for the least squares
distance
\[ \min_{\lambda, \xi, \varepsilon} \frac{1}{N} \sum_{i=1}^{N} (\tau_{t_i} - \hat{\tau}_{t_i})^2 \] (2.16)

between the above model market time \( \tau_{t_i} \) and the estimated market time \( \hat{\tau}_{t_i} \); see (2.10). We obtain the parameter estimates \( \lambda = 8, \xi = 0.15 \) and \( \varepsilon = 0.02 \).

The resulting model market activity \( M_t \) is exhibited in Figure 2.5. By formula (2.5) it is the square of the derivative \( U'_t \) that causes clearly visible spikes in the market activity shown in Figure 2.5. These spikes and the obvious roughness of the market activity \( M_t \) are consequences of the fact that \( U'_t \) is the derivative of \( U_t \), which is smoothing the non-differentiable path of \( 2Y_t^{\frac{3}{2}} \).

We note in Figure 2.5 that the market activity becomes sometimes rather high when it spikes, e.g. during the dramatic drawdown of the Great Depression around 1929. Consequently, in Figure 2.2 this drawdown occurs extremely fast, a feature that is difficult to capture by a scalar diffusion. Figure 2.6 shows the integrated model market activity, the model market time \( \tau_t \), together with the estimated market time \( \hat{\tau}_t \). The estimated market time uses the observed approximate quadratic variation in (2.11), whereas the model market time is calculated according to the derived model equations. We note a good fit between the two trajectories over the entire 144 year period. It is clear that there have to be differences between the two trajectories displayed in Figure 2.6 because the random monthly increments of the driving Brownian motion impact the approximate quadratic variation but not the model market activity.

![Figure 2.6: Model market time (black) and estimated market time \( \hat{\tau}_t \) (red).](image)

### 2.5 Model Volatility

In the literature one focuses typically on modeling the estimated volatility \( \hat{\sigma}_t \), which we obtain here by standard volatility estimation, that is, via exponential
smoothing of squared monthly observed index returns (with decay parameter 0.84); see e.g. RiskMetrics (1996). By using equation (2.13) we calculate the model volatility \( \sigma_t \), which we display in Figure 2.7 together with the estimated volatility \( \hat{\sigma}_t \). When calculating \( \hat{\sigma}_t \), the randomness of squared returns and the delay effect, caused by exponential smoothing, create differences between the estimated volatility \( \hat{\sigma}_t \) and the model volatility \( \sigma_t \). The path of the model volatility has spikes and its roughness is caused by the roughness of the model market activity shown in Figure 2.5. Visually one notes that whenever the model volatility spikes, the estimated volatility jumps up and declines slowly afterwards until the next major spike occurs.

![Figure 2.7: Model volatility \( \sigma_t \) (black) and estimated volatility \( \hat{\sigma}_t \) (red)](image)

Since the fluctuations in the market activity are triggered by the same fluctuations that drive the index, only one Brownian motion is driving the index dynamics under the proposed model. For the monthly observed S&P500 the proposed model resolves the previously mentioned ‘leverage effect puzzle’, see Ait-Sahalia, Fan & Li (2013), since it allows us to extract the usually hidden trajectory of the model volatility.

### 2.6 Brownian Motion

The following explicit formula for the Brownian motion value

\[
W_{\tau t} = 2 \left( Y_t^{\frac{1}{2}} - Y_0^{\frac{1}{2}} \right) - \int_0^t \left( \frac{3}{4} Y_s^{-\frac{1}{2}} - Y_s^\frac{1}{2} \right) d\tau_s
\] (2.17)

follows from the SDE (2.3) via an application of the Itô formula. The equations (2.17) and (2.4) allow us to extract the increments of the postulated Brownian motion \( W \) by approximating numerically the integral on the right-hand side of (2.17) by an implicit Wagner-Platen expansion (equivalent to an implicit Euler approximation) as
\[ W_{\tau_{i+1}} - W_{\tau_i} \approx 2 \left( Y_{t_i}^{\frac{1}{2}} - Y_{t_{i-1}}^{\frac{1}{2}} \right) - \left( \frac{3}{4} Y_{t_i}^{\frac{1}{2}} - Y_{t_{i-1}}^{\frac{1}{2}} \right) M_{t_i}(t_i - t_{i-1}). \] (2.18)

We show in Figure 2.8 the resulting estimated path of the postulated Brownian motion \( W_{\tau} \) in market time.

![Figure 2.8: Estimated postulated Brownian motion path \( W_{\tau} \) in market time.](image1)

![Figure 2.9: Quadratic variation of postulated \( W_{\tau} \) in market time.](image2)

Figure 2.8: Estimated postulated Brownian motion path \( W_{\tau} \) in market time.

Figure 2.9: Quadratic variation of postulated \( W_{\tau} \) in market time.

motion \( W_{\tau} \) in market time. Note that there are periods in market time where we have only few observations, e.g. near \( \tau \approx 1 \), because the market activity is extremely high at that time. To check whether the path in Figure 2.8 is possibly that of a Brownian motion with respect to \( \tau \)-time, we plot in Figure 2.9 its quadratic variation \( [W_{\tau}]_{\tau} \) in \( \tau \)-time. It appears to be almost a straight line. Only near \( \tau \approx 1 \) the line curves up for a period, which is most likely due to discretization errors in our expansions of the increments of the components of the derived model.
Given the increments of $W$ we can extract the increments of the related postulated Brownian motion $\tilde{W}$ in $t$-time via the approximate formula

$$\tilde{W}_{t_i} - \tilde{W}_{t_{i-1}} \approx \frac{1}{\sqrt{M_{t_i}}} (W_{\tau_{t_i}} - W_{\tau_{t_{i-1}}}).$$

(2.19)

The path of the postulated Brownian motion $\tilde{W}$ is displayed in $t$-time in Figure 2.10. Additionally, in Figure 2.11 we exhibit the approximate quadratic variation $[\tilde{W}]_t$ in $t$-time, which turns out to be close to a straight line. The slope of this line, estimated by linear regression, equals about 0.89, which is about 11% below the theoretical value of 1.0. As we will see in the next subsection, this bias seems to result from truncation errors of the Wagner-Platen expansions employed for the increments of the components of the system of SDEs for the proposed model. Such biases are common for discrete time approximations of solutions of SDEs; see Platen & Bruti-Liberati (2010).
2.7 Daily Observations

Since the proposed model is characterized by a system of SDEs, we can apply it also to more frequently observed data using the same parameters, which is not easily achieved when employing classical time series models. As we will see below, more frequent observations reduce the above observed bias when applying the proposed model to daily observed S&P500 data for the period from January 1963 until December 2015.

We continue using the above estimated parameters and extract analogously as before the respective path of $\tilde{W}_t - \tilde{W}_{t_q}$ from the daily observed S&P500 data. We display in Figure 2.12 the resulting path in $t$-time, where $t_q$ denotes the starting time at 1 January 1963 of the daily observed S&P500 data. This path exhibits the daily fluctuations of the postulated Brownian motion $\tilde{W}$. Its approximate quadratic variation is displayed in Figure 2.13 with average slope of about 1.07. This slope is closer to its theoretical value of 1.0 than the one we obtained from monthly observed S&P500 data.

The implicit, higher-order Wagner-Platen expansions we employ for the increments of the components of the solution of the system of SDEs of the proposed model seem to handle sufficiently well the fast moving market time around the 1987 market crash. Since the New York Stock Exchange was closed during the crash, intermediate observations are not available for that period, which results in a small jump in the quadratic variation shown in Figure 2.13.

![Figure 2.12: Postulated Brownian motion $\tilde{W}_t - \tilde{W}_{t_q}$ from daily observed S&P500.](image)

Since the Figures 2.11 and 2.13 indicate that a higher observation frequency generates less bias in the estimation of $\tilde{W}$, we use in the next subsection only daily observed S&P500 data.
2.8 Testing the Postulated Brownian Motion

So far we fitted, in some least square sense, the proposed model to the S&P500 data using Wagner-Platen expansions for the increments of the model components. If the model were correct, then it should be difficult to reject the hypothesis that the path of the potential Brownian motion $\tilde{W}$ is that of a true Brownian motion. We already know that the approximate quadratic variation, shown in Figure 2.13, of the extracted path of the postulated Brownian motion $\tilde{W}$ matches well the respective theoretical quadratic variation. It remains to show that we cannot reject the hypothesis that the trajectory of $\tilde{W}_t$ is that of a martingale. By Lévy’s Theorem this would then not allow us to reject the postulated Brownian motion as a true Brownian motion; see e.g. Revuz & Yor (1999).

There are many possibilities for such a test. Below we follow a well-established literature that tests the martingale difference hypothesis; see e.g. Escanciano & Lobato (2009). More precisely, we use the Chen & Deo (2006) joint variance ratio (VR) test, supported by the in Kim (2006) introduced wild bootstrap VR test, which returns bootstrap $p$-values of the Lo & MacKinlay (1988) and Chow & Denning (1993) tests. We refer to Charles & Darné (2009) for a review and details on the test statistics of the variance ratio (VR) test used. Below we describe briefly how the Chen & Deo (2006) test is designed and what kind of null and alternative hypotheses are tested.

The VR test we employ uses the fact that the variance of Brownian motion increments is linear in all sampling intervals. The VR test is then used to test the hypothesis that the sequence of differences $\omega_i - \omega_{i-1}$ of a time series $\omega$ is a collection of i.i.d. observations, which form a martingale difference sequence. We define the variance ratio of the $(i,k)$-th difference $\omega_i - \omega_{i-k}$ as

$$V(i, k) = \frac{\text{Var}(\omega_i - \omega_{i-k})}{k \text{Var}(\omega_i - \omega_{i-1})} = 1 + 2 \sum_{l=1}^{k-1} \left( \frac{k-l}{k} \right) \rho_l,$$

(2.20)
where $\rho_l$ is the $l$th lag autocorrelation coefficient of $\omega_i$. Therefore, the VR test is a test where

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0,$$  

(2.21)

which corresponds to differences $\omega_i - \omega_{i-1}$ that are uncorrelated. The test statistic is based on the estimator of $V$ and equals

$$VR(k) = \frac{\hat{\sigma}^2(k)}{\hat{\sigma}^2(1)},$$  

(2.22)

where $\hat{\sigma}^2(k)$ is the estimator of the $k$ period return variance.

The Chen & Deo (2006) joint VR test is based on its individual power transformed VR statistics. In a joint test, multiple comparisons of VRs using different time horizons are applied. This joint test considers the following null and alternative hypotheses:

$$H_0 : V(i, k) = 1 \text{ for all } i = 1, \ldots, m \text{ vs } H_1 : V(i, k) < 1 \text{ for some } i.$$  

(2.23)

The Chen & Deo (2006) test is a Wald statistic based test which follows a $\chi^2$ distribution with $m$ degrees of freedom.

We apply the above joint VR test to this difference series and obtain a value for the test statistic of 4.47. This value is significantly less than the value of 7.81 for the 5% quantile of the respective $\chi^2$ distribution. Therefore, we cannot reject the $H_0$ hypothesis at the 5% level of significance. We conclude that there is not enough evidence to reject the hypothesis that the extracted path $\tilde{W}$ is that of a true martingale with quadratic variation equal to $t$-time. A continuous martingale with quadratic variation equal to time is by Lévy’s theorem a Brownian motion; see Revuz & Yor (1999). Thus, the proposed model cannot be easily rejected for the daily S&P500 data we examine.

For illustration, Figure 2.14 shows the difference series $\tilde{W}_{t_i} - \tilde{W}_{t_{i-1}}$ and Figure

![Figure 2.14: Increments of $\tilde{W}_t$.](image)
2.15 the autocorrelation function for this difference series, which appears to be typical for a Brownian motion. To round up our study, we finally estimate in a standard way the kurtosis of the increments of $\tilde{W}$. We obtain an estimated value of about 3.2, which is close to the theoretical value of 3.0 for the independent Gaussian increments of a true Brownian motion.

3 Links to the Literature

Historically, the standard continuous time market model for an equity index has been the Black-Scholes model; see Black & Scholes (1973). Its popularity is mostly due to its excellent tractability, however in reality, volatility is stochastic and log-returns have leptokurtic distributions; see e.g. Ghysels, Harvey & Renault (1996). As long-term index model the Black-Scholes model and many of its generalizations are not well-suited because the variance of the logarithm of the index grows linearly, whereas in reality, it appears to remain somehow bounded, as visually indicated by Figure 2.1. Various models have been proposed for the pricing and hedging of index derivatives with maturities of up to about three years. In the following we refer to a few strands of this rich literature to point at links to the proposed model.

Engle (1982) initiated extensive research on autoregressive conditional heteroscedastic time series models, which typically depend on the given observation frequency. Continuous time models, characterized by stochastic differential equations (SDEs), avoid this dependency and remain meaningful under different observation frequencies. They can be conveniently handled via the rules of stochastic calculus. Some time series models have similarities with the market activity model proposed in the current paper. Squares of ‘innovations’ and their moving average appear as key elements in these volatility models. Nelson (1990) shows that some continuous time limit of an ARCH model yields a volatility process.
that is driven by a Brownian motion which turns out to be independent of the one driving the index fluctuations. This independence is the opposite to what the proposed model suggests, where only one single Brownian motion is driving the index value and also its volatility. This means that the proposed model shows perfect correlation between the Brownian motions driving the index and its volatility.

Under the proposed model the index fluctuations act as signals for the changes in market activity, which makes economic sense because investors reallocate stock holdings according to their strategies when the index changes its value. Furthermore, the well-known leverage effect of equity indexes, see Black (1976), is in many papers artificially introduced and not endogenously generated, whereas the ‘normal’ volatility emerges endogenously and generates in a natural manner the leverage effect under the proposed model.

The literature suggests a wide range of continuous time stochastic volatility models, where a systematic overview is given, e.g., in Cont (2010). The proposed model can be interpreted as a stochastic volatility model. Its stylized version, the $3/2$ model, where the market activity is set to a constant, is a local volatility function model because the volatility is here a function of index value and time. Local volatility function models played an important role in the development of quantitative methods for index derivatives; see e.g. Dupire (1993). The stylized version of the proposed model, with constant market activity, can be interpreted as a constant elasticity of variance (CEV) model, which is a class of models that goes back to Cox (1975).

Due to its random market time, the proposed model is also related to the wide class of subordinated models, which can be traced back to Bochner (1955) and Clark (1973). Here the dynamics of the index evolve in some transformed time, typically generalizing the Black & Scholes (1973) model.

The proposed model is different to most models that aim directly at modeling stochastic volatility, as does the widely used Heston (1993) model. The Heston model became popular due to its excellent tractability. However, it revealed over time serious shortcomings; detailed e.g. in Cont (2010). As a consequence, other models became popular among traders, in particular the SABR model; see Hagan et al. (2002). This model evolved among practitioners who aimed for better calibration to market derivative prices, while maintaining reasonable tractability. The SABR model is deemed to reflect better volatility effects encapsulated in index derivatives than the Heston model does. Not by chance has the popular SABR model similarities with the model proposed in the current paper. Both models can be interpreted as CEV models that evolve in some random time. The ‘market activity’ of the SABR model is modeled as a geometric Brownian motion. Consequently, the variance of this ‘market activity’ grows proportionally to time, which is not what one observes when studying index data over long time periods; see again Figure 2.1. In reality, one observes a mean-reverting behavior for market activity. The SABR model remains an ‘ad hoc’ model, designed for pricing
and hedging short-dated derivatives. The derived model, when subordinated by deterministic market time with constant market activity, is equivalent to the 3/2 volatility model or minimal market model of Platen (1997) and Heston (1997). Many of the above mentioned stochastic volatility models, including the SABR model, assume correlated Brownian motions driving the index and its volatility. Ait-Sahalia, Fan & Li (2013) point at the difficulties in estimating the correlation between these two Brownian motions, which they call the leverage effect puzzle. The proposed model resolves this puzzle by demonstrating that only one Brownian motion is needed to explain the paths of the index and its hidden volatility.

In recent years variance swaps and derivatives on the volatility index (VIX) of the S&P500 became heavily traded, which revealed shortcomings in popular index models; see e.g. Mencia & Sentana (2013). As a consequence, versions of the 3/2 volatility model became popular, which are close to the proposed model with constant market activity; see e.g. Carr & Sun (2007). Moreover, Goard & Mazur (2013) show that a 3/2 volatility model provides a better fit to traded VIX and volatility derivatives than most popular models. Furthermore, Mencia & Sentana (2013) point out that a 3/2 volatility model reproduces naturally the observed positive skew of implied volatilities of VIX options, which is considered to be the most relevant stylized empirical feature of VIX derivatives that has been puzzling traders. Since the proposed model is, in market time, a 3/2 volatility model, this feature is consistent with the findings of the current paper.

In Baldeaux, Ignatieva & Platen (2014) a time dependent constant elasticity of variance model has been proposed, which generalizes the 3/2 model. By generalizing the 3/2 and the Heston model, Grasselli (2017) developed an even more general model, the 4/2 volatility model, which has high tractability and fits well derivative data; see Baldeaux, Grasselli & Platen (2014).

In recent years the availability of high-frequency data revealed that with high-frequency estimated index volatilities show trajectories that appear to be much rougher than typical diffusion models would be able to produce; see e.g. Bayer, Friz & Gatheral (2016) and Gatheral, Thibault & Rosenbaum (2018). This recently approached challenge for index modeling becomes resolved by the proposed model. It explains that rough model volatility emerges naturally from rough market activity, which is a response in trading activity to observed fluctuations of the index. More precisely, the index dynamics evolves in market time as that of a diffusion, where in periods of high market activity the volatility spikes when viewed in calendar time. This effect generates the, so called, rough volatility. It can already be noticed in monthly and daily observed index data, as we show in the current paper.

There exists extensive work on volatility models that involve fast and slow moving components; see Fouque, Papanicolaou & Sircar (2000). The proposed model generates such fast and slow moving volatility components endogenously. The slow moving component is the ‘normal’ volatility, which results from the movements of the normalized index. The fast moving component is the market activity, which
is a linear function of the square of the derivative of the moving average of a proxy of the driving Brownian motion. Since the Brownian motion path is not differentiable, the market activity can raise to extremely high values at certain periods.

There is a wide range of models that allow for jumps, including exponential Lévy process models and jump diffusion models; see e.g. Barndorff-Nielsen & Shephard (2001) and Bakshi, Cao & Chen (1997). We demonstrate with the proposed model that many jumps that one observes in estimated volatilities can be explained as being endogenously generated through the roughness of model market activity. Straightforward extensions of the proposed model can easily accommodate jumps caused by particular events. Non-decreasing Lévy processes can be incorporated as substitute for $t$-time, making the time transformed driving Brownian motion a Lévy process.

Most models in the literature assume the existence of an equivalent risk-neutral probability measure under the classical no-arbitrage theory; see e.g. Ross (1976), Harrison & Kreps (1979) and Delbaen & Schachermayer (1994). This assumption becomes too restrictive when modeling long-term risk, as argued in Platen (2002), Platen & Heath (2010) and Baldeaux, Ignatieva & Platen (2018). The benchmark approach avoids this restrictive assumption. It requires instead only the existence of the growth optimal portfolio (GP), which maximizes expected logarithmic utility and goes back to Kelly (1956). The benchmark approach uses the GP as benchmark and numeraire. All benchmarked (in units of the GP denominated) nonnegative securities form supermartingales under the real-world probability measure. Under the stylized version of the proposed model the Radon-Nikodym derivative of the putative risk-neutral measure, the benchmarked savings account, emerges as the inverse of a time transformed squared Bessel process of dimension four, which is a text book example for a local martingale that is not a true martingale and, thus, a strict supermartingale; see Revuz & Yor (1999) and Baldeaux, Ignatieva & Platen (2018). Therefore, it is not a true martingale, which the classical risk-neutral assumptions would require. In the sense of Loewenstein & Willard (2000) this causes a money market bubble (see Baldeaux, Ignatieva & Platen (2018)), which constitutes a form of classical arbitrage. However, under the proposed model there is no economically meaningful arbitrage because the pathwise, in the long-run best performing portfolio, the GP, has finite expected growth rate at any finite time. Thus, no market participant can generate from finite capital infinite wealth at any finite time.

Under the benchmark approach, which generalizes the classical no-arbitrage approach, see Platen & Heath (2010), one can perform consistently pricing, hedging, portfolio optimization, expected utility maximization and other risk-management tasks. New effects appear and can be exploited that are not captured under classical risk-neutral assumptions. In particular, long-term contracts for pensions and life-insurance can be less expensively produced than is possible under the classical no-arbitrage approach; see Platen & Heath (2010).
4 Derivation of the Proposed Model

Within this section we derive the proposed model based on three well-founded assumptions. We emphasize that this model is not another ‘ad hoc’ model that is chosen for tractability. Due to its derivation from ‘first principles’ it becomes parsimonious and a rather accurate reflection of reality.

4.1 Index as Growth Optimal Portfolio

We deliberately model the dynamics of a well-diversified stock index and not that of a stock price or exchange rate because an exchange price depends on two underlyings, reflecting both sides of the exchanged securities. The diversification that is taking place when forming a well-diversified equity index removes the specific or idiosyncratic uncertainty of stocks. It remains the nondiversifiable uncertainty of the respective stock market that drives the index dynamics.

Diversification theorems have been established in Platen (2005), Platen & Heath (2010) and Platen & Rendek (2012). These allow us to conclude under extremely weak assumptions that a well-diversified stock index with a large number of constituents approximates the GP. This model independent property of well-diversified stock indexes can be interpreted as a consequence of the Law of Large Numbers. It provides the link between risk and return in the stochastic differential equation (SDE) for the index.

For a continuous stock market model, where we do not include the risk-free asset in the investment universe, it follows by a combination of Theorem 5.1 and Theorem 3.1 in Filipović & Platen (2009) that the value $S_t$ of the GP satisfies an SDE of the form

$$\frac{dS_t}{S_t} = \ell_t dt + \theta_t (\theta_t d\tau_t + dW_{\tau_t})$$

$t \geq 0, S_0 > 0$. Here $W = \{W_\tau, \tau \geq 0\}$ is a Brownian motion with respect to some strictly increasing market time process $\tau = \{\tau_t, t \geq 0\}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{E}, P)$. $\mathcal{E} = (\mathcal{F}_t)_{t \geq 0}$ is a filtration, modeling the evolution of information, satisfying the usual conditions; see Karatzas & Shreve (1998). $P$ denotes the real-world probability measure. The ‘normal’ volatility $\theta_t$ is the volatility with respect to market time and forms an adapted process. The excess growth rate $\ell_t$ forms an adapted process, representing a Lagrange multiplier. In our case, where we model the real-value total return S&P500, it can be interpreted as virtual real (inflation adjusted) interest rate, which would become the real interest rate when the savings account is included in the investment universe, as explained in Filipović & Platen (2009).

Any drift parameter in an SDE needs an extremely long observation window to be estimated accurately. Therefore, it is reasonable to set the excess growth
rate \( \ell_t \) constant in the proposed model to have a chance to estimate at least the long-term average of the excess growth rate. Furthermore, since real (inflation adjusted) values matter economically in the long-term, we choose the index value \( S_t \) to be the value at time \( t \) of a real-value total return index. This means, the consumer price index is used for denomination. As we show in the current paper, for the real-value S&P500 the average of the virtual real interest rate equals about the historical average of the real interest rate. This means, we are not assuming in our modeling that the virtual real interest rate equals always the real interest rate as set by the Federal Reserve. However, our results indicate that the real interest rate is on average mostly likely set by the Federal Reserve close to the objectively existing virtual interest rate, which will be the topic of forthcoming research. These insights lead us to our first assumption:

**Assumption 4.1** A well-diversified real-value index follows the dynamics of the respective GP, satisfying the SDE (4.1), where we assume a constant excess growth rate.

Note that the SDE (4.1) for the GP shows a crucial link between its drift and its diffusion coefficient, which relates risk and return. This link guides us to form the normalized index

\[
Y_t = \frac{S_t}{A_t} = \frac{\tilde{S}_t}{e^\tau_t},
\]

(4.2)

where we use for our normalization the exponential function

\[
A_t = A_0 \exp\{\tau_t + \ell_t\}
\]

(4.3)

for \( t \geq 0 \), with \( A_0 > 0 \), which gives equation (1.2) of the model. On the right hand side of (4.2) we introduced the standardized index

\[
\tilde{S}_t = \frac{S_t}{A_0} \exp\{-\ell t\},
\]

(4.4)

with SDE

\[
\frac{d\tilde{S}_t}{\tilde{S}_t} = \theta_t (\theta_t d\tau_t + dW_t)
\]

(4.5)

for \( t \geq 0 \). For the normalized index \( Y_t = \tilde{S}_t e^{-\tau_t} \), see (4.2), we obtain via the Itô formula the SDE

\[
dY_t = Y_t (\theta_t^2 - 1) d\tau_t + Y_t \theta_t dW_{\tau_t}
\]

(4.6)

for \( t \geq 0 \).
4.2 Variance of Well-Diversified Wealth Increments

The challenge is now to determine the ‘normal’ volatility $\theta_t$ in (4.6), which is the volatility of a well-diversified index with respect to market time. This leads us to the question: What is the natural evolution of the variance of the increments of a well-diversified index? We make the crucial observation that the value of a well-diversified portfolio evolves similar to the size of a population, where the individuals give independently birth from time to time or die. This interpretation of a diversified wealth evolution appears to be new. It turns out to be very powerful, as the current paper demonstrates.

Birth-and-death processes (also called branching processes) and their continuous time diffusion limits are well studied in the literature; see e.g. Feller (1971). To illustrate the obvious similarity between the evolution of the population size of a birth-and-death process and that of a well-diversified index, let us invest at the beginning of a short investment period the index value (the population size) in wealth units of standard size (the individuals). Each of these wealth units generates independent wealth increments (births and deaths). Due to the independence of these wealth increments, the variance of the increment of the total wealth (the index value) equals the sum of the variances of the individual wealth increments. Hence, the variance of the increment of the total wealth turns out to be proportional to the number of wealth units invested at the beginning of the short investment period. This means, the variance of the index increments is proportional to the index value itself. At the end of each short time period wealth becomes reallocated such that the total wealth is portioned again into standard wealth units that then evolve independently. This reallocation of wealth represents the diversification activity, which is crucial not only for the approximation of the GP but also for the generation of the feedback effect characterizing the variance of the increments of the resulting diffusion process. One can deduce the diffusion coefficient of the SDE for the respective continuous time diffusion limit for the index value (the population size) using, e.g., Theorem 14.1.5 in Kloeden & Platen (1999) as a convergence theorem. Since the variance of the index increments is proportional to the index value, the diffusion coefficient of the SDE for the index is proportional to the square root of the index value. This line of arguments applies also to the above introduced normalized index, leading us to the following assumption:

**Assumption 4.2** The squared diffusion coefficient of a well-diversified normalized index evolves, in some market time, proportional to the normalized index value.

Since the market time has so far not been fixed, under Assumption 4.2 the diffusion coefficient in (4.6) can be set, without loss of generality, to $Y_t \theta_t = \sqrt{Y_t}$. Therefore, the ‘normal’ volatility $\theta_t$, which is the volatility with respect to market
time, emerges as
\[ \theta_t = Y_t^{-\frac{1}{2}} \] (4.7)
for \( t \geq 0 \). By using (4.6) this derives the SDE (2.3) for the normalized index.

4.3 Feedback in Market Activity

The changes of the market time \( \tau_t \) are captured by its derivative, the market activity \( M_t = \frac{d\tau_t}{dt} \); see equation (2.4). From (4.5) we obtain by application of the Itô formula the model market time
\[ \tau_t = \ln \left( \left[2 \tilde{S}_t \right]_t^\frac{1}{2} + 1 \right) \] (4.8)
for \( t \geq 0 \); see (2.10). We show in Figure 2.1 the estimated market time and note that the market activity \( M_t \) is fluctuating considerably, in particular, in periods of major market downturns, as the one during the Great Depression after 1929. Market activity appears to increase significantly during periods when the index moves dramatically. More precisely, we observe that when \( W_{\tau_t} \) deviates from its recent average level, market activity increases. Intuitively, this happens because market participants adjust after index moves their holdings according to their respective investment strategies.

Since \( W_{\tau_t} \) fluctuates like \( 2Y_t^{\frac{1}{2}} \) according to (2.17), we approximate the recent average level of \( W_{\tau_t} \) by the moving average \( U_t \) of \( 2Y_t^{\frac{1}{2}} \), where
\[ U'_t = \frac{dU_t}{dt} = \lambda(2Y_t^{\frac{1}{2}} - U_t) \] (4.9)
with $U_0 = 2Y_0^{\frac{1}{2}}$, $\lambda > 0$. This provides equation (2.6) of the proposed model. The derivative $U'_t$ is an indicator for the magnitude of changes in $W$. The challenge is now to capture the typical feedback effect in market activity, potentially as a function of the derivative $U'_t$. Systematic empirical studies, using polynomials of $U'_t$ as potential functions for the market activity, reveal rather clearly that the market activity is approximately a linear function of the square of the derivative $U'_t$. This crucial discovery leads us to the following assumption:

**Assumption 4.3** The market activity $M_t$ is a linear function of the square $(U'_t)^2$ of the derivative of the moving average of $2Y_t^{\frac{1}{2}}$.

Under Assumption 4.3 we obtain the market activity in the form

$$M_t = \left( (U'_t)^2 + \varepsilon \right) \xi$$

(4.10)

for $t \geq 0$, and some $\varepsilon > 0$ and $\xi > 0$. This formula provides the model equation (2.5) and concludes the derivation of the proposed model.

One could ask whether there exist more accurate functional relationships for the dependence of $M_t$ on $U'_t$? This can be indeed expected, as we will indicate in the next section. We propose in the current paper a standard model for market activity which can be refined and may vary for different markets and observation frequencies. The proposed model aims to open a direction of future research, in particular, for the study of high-frequency index data.

The existence and uniqueness of a strong solution of the SDE (2.3) for the normalized index in market time follows by using the well-known Yamada-condition; see Ikeda & Watanabe (1989). The differential equation (2.6) for the moving average $U$ can be shown to have in market time a bounded drift. Thus, by using a Lipschitz condition it has a unique, strong solution in $\tau$-time; see Ikeda & Watanabe (1989). Consequently, the market activity in equation (2.5) exists, is unique and strictly positive when modeled in $\tau$-time. Since $t$-time and $\tau$-time are linked by the market activity, this ensures that the $t$-time is uniquely determined. Therefore, the system of SDEs describing the proposed model in calendar time has a unique, strong solution in the sense of Ikeda & Watanabe (1989).

## 5 Simulation

It has been pointed out, e.g., in Platen & Rendek (2009), for long-term risk management and for capturing accurately extreme events the use of simple discrete
time approximations, e.g. the Euler scheme, for the simulation of the trajectory of a solution of a given system of SDEs, may be insufficient for large time step sizes. Since the absolute value of the derivative of the moving average of a (non-differentiable) Brownian motion path can become extremely large, major numerical errors can be expected during scenario simulation of the derived model. The truncation errors of Wagner-Platen expansions and the propagation of these errors over the long time periods we are interested in, can lead to paths of discrete time approximations that deviate significantly from those of the theoretical solution, in particular, when the driving Brownian motion moves severely. Fortunately, the proposed model permits exact simulation of the normalized index in τ-time. This avoids discretization errors in long-term scenario simulations of the index value in τ-time and delivers accurate extreme values. For the scenario simulation of the entire system of SDEs describing the proposed model we rely on almost exact simulation. This is a scenario simulation where no diffusion part and only some drift parts of the given system of SDEs are approximated by a Wagner-Platen expansion. Almost exact simulation can usually be performed with extremely high accuracy by reducing the time step size to a desired level; see Platen & Rendek (2009). In our case we have exact simulation in market time and need only to obtain an excellent proxy for the market activity via some almost exact simulation.

5.1 Simulating the Normalized Index in Market Time

We demonstrate below almost exact simulation for the derived model. Note that other models from the indicated new model class can be handled similarly. The square root processes \( Y \), modeling the normalized index, can be simulated exactly in τ-time by sampling from its non-central chi-square transition distribution; see e.g. Platen & Rendek (2009) and Chapter 2 in Platen & Bruti-Liberati (2010). Given an equidistant discretization of τ-time and the most recent value \( Y_{t+i} \) of the normalized index for \( i \in \{0,1,\ldots\} \), we simulate a non-central chi-square distributed random variable of dimension four to obtain

\[
Y_{t+i+1} = \frac{1 - e^{-(\tau_{t+1} - \tau_{t})}}{4} \left( \chi^2_{3,i} + \left( \sqrt{\frac{4e^{-(\tau_{t+1} - \tau_{t})}}{1 - e^{-(\tau_{t+1} - \tau_{t})}} Y_{t+i} + Z_i} \right)^2 \right)
\]

(5.1)

Here \( Z_i \) is an independent standard Gaussian distributed random variable and \( \chi^2_{3,i} \) is an independent chi-square distributed random variable with three degrees of freedom. Using our estimated parameters and 10,000 steps in market time, we simulated a path of the normalized index \( Y \), which is displayed in Figure 5.1 in τ-time. When the same path is displayed in the resulting t-time in Figure 5.2, we notice that during periods of major downward or upward moves of the index the market time evolves so fast in t-time that these moves appear as spikes or flash crashes. In our monthly and daily data sets spikes and flash crashes
were historically rarely observed. The flash crash observed on 6 May in 2010 and similar later flash crashes indicate that increased algorithmic trading on our electronic exchanges is most likely moving the S&P500 dynamics towards a dynamics from the proposed model class. Spikes in volatility and flash crashes become naturally generated by rare but occasional sudden downward excursions of the driving Brownian motion generated by the nondiversifiable uncertainty of the stock market. Since the market activity becomes very high during such downward excursion the crash passes very fast in calendar time and appears as a flash crash.

One may improve the proposed model in many ways, e.g. by making the market activity depending also on the sign of the derivative of the moving average, the calendar time and potentially other seasonal factors. For instance, one could model the feedback to upward moves weaker than the one to downward moves. Such refinements can make very good sense and will be studied in forthcoming work for high-frequency data but are beyond the scope of the current paper.
5.2 Simulating the Market Activity

Note that the differential equation (2.6) is evolving in \( t \)-time and can also be interpreted as a differential equation with respect to \( \tau \)-time when replacing \( dt \) with \( d\tau M_t \). Using the simulated values of the normalized index \( Y \) at the current and next time discretization point, the moving average \( U \) of \( 2Y^{1/2} \) at the next discretization point is obtained via almost exact simulation using the Wagner-Platen expansion described below; see Chapter 12 in Kloeden & Platen (1999).

More precisely, we employ the strong order 1.0 predictor-corrector expansion with corrector

\[
U_{t_i} \approx \frac{U_{t_{i-1}} \left( 1 - \frac{\lambda (\tau_i - \tau_{i-1})}{2M_{t_{i-1}}} \right) + \lambda (\tau_i - \tau_{i-1}) \left( \frac{Y_{t_i}^{1/2}}{N_{t_i}} + \frac{Y_{t_{i-1}}^{1/2}}{M_{t_{i-1}}} \right)}{1 + \frac{\lambda (\tau_i - \tau_{i-1})}{2N_{t_i}}}, \tag{5.2}
\]

and predictor

\[
\bar{U}_{t_{i-1}} = \frac{U_{t_{i-1}} \left( 1 - \frac{\lambda (\tau_i - \tau_{i-1})}{2M_{t_{i-1}}} \right) + \lambda (\tau_i - \tau_{i-1}) \left( \frac{Y_{t_i}^{1/2}}{M_{t_{i-1}}} + \frac{Y_{t_{i-1}}^{1/2}}{M_{t_{i-1}}} \right)}{1 + \frac{\lambda (\tau_i - \tau_{i-1})}{2M_{t_{i-1}}}}. \tag{5.3}
\]

Here

\[
N_{t_i} = \left( \left( \lambda (2Y_{t_i}^{1/2} - \bar{U}_{t_i}) \right)^2 + \varepsilon \right) \xi \tag{5.4}
\]

predicts the market activity, which becomes approximately

\[
M_{t_i} \approx \left( \left( \lambda (2Y_{t_i}^{1/2} - U_{t_i}) \right)^2 + \varepsilon \right) \xi \tag{5.5}
\]

\( i \in \{1, 2, \ldots \} \) with \( U_{t_0} = 2Y_{t_0}^{1/2} \) as initial value for the moving average. For a usual diffusion, one would expect that a simple Euler approximation step would approximate the increments of the moving average well enough. However, for the sometimes extremely large market activity, the use of a strong order 1.0 predictor-corrector expansion becomes necessary. The reason is that the market activity becomes so large in certain periods that the, otherwise, resulting truncation errors would distort the resulting market time too much. It would not be acceptable as a reasonably accurate approximation to the solution of the system of model SDEs.

The simulated path of the market activity \( M \) in \( \tau \)-time is shown in Figure 5.3. In Figure 5.4 we exhibit the market activity with respect to \( t \)-time. One notes
the spikes in the trajectory of the simulated model market activity. The periods with high market activity shrink in $t$-time and, thus, spikes result in $\tau$-time, which creates the visible roughness of the trajectory of the market activity when viewed in $t$-time. By integrating the market activity we obtain in Figure 5.5 the simulated $\tau$-time, which grows visibly faster in periods of high market activity. According to formula (2.13) we obtain the simulated model volatility $\sigma_t$, which we display in Figure 5.6 in $\tau$-time. When displayed in $t$-time, it would look very similar to the market activity, already shown in Figure 5.4, where rare spikes dominate the trajectory.

One can apply the above proposed almost exact simulation method over extremely long time periods, which makes it suitable for long-term risk management tasks, e.g., the pricing of pension and life insurance contracts.
6 Conclusion

Based on well-founded insights, the paper derives a new model class for the long-term dynamics of well-diversified stock indexes. The proposed model fits well the monthly and daily observed evolution of the S&P500. It is described via a three-dimensional system of stochastic differential equations, driven by a single Brownian motion, which can be interpreted as the non-diversifiable uncertainty of the stock market. The model is characterized as a three dimensional diffusion process that evolves in market time with occasional periods of extremely high market activity. When viewed in calendar time, the market activity shows spikes during such periods. With two initial values and four economically meaningful parameters the proposed model explains naturally the occurrence of spikes in volatility and rare flash crashes, as well as, the leverage effect and other stylized empirical index properties.

The new model class leads beyond classical no-arbitrage theory and is derived un-
der the more general benchmark approach. The proposed model reveals perfect negative correlation between index and volatility fluctuations when observed in market time. Through the extraction of the hidden model volatility the leverage effect puzzle has been resolved, where the correlation between model volatility and index seemed inaccessible.

Due to occasional periods of extremely high market activity, higher-order Wagner-Platen expansions for increments of model components turn out to be necessary for fitting and simulating the model. Finally, the paper demonstrates how to extract from the given discretely observed trajectory of the index the hidden path of the postulated driving Brownian motion, where the latter cannot be easily rejected as that of a true Brownian motion. In forthcoming work this methodology will be generalized into a new powerful inference method for multi-dimensional stochastic differential equations with hidden components.

References


