

Sattarhoff, Cristina

Conference Paper

GMM estimation of Multifractal Random Walks using an efficient algorithm for HAC covariance matrix estimation

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session:
Asset Price Dynamics, No. A19-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Sattarhoff, Cristina (2010) : GMM estimation of Multifractal Random Walks using an efficient algorithm for HAC covariance matrix estimation, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Asset Price Dynamics, No. A19-V3, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37308>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

GMM estimation of Multifractal Random Walks using an efficient algorithm for HAC covariance matrix estimation

Cristina Sattarhoff

University of Hamburg

Institute of Statistics and Econometrics

Faculty of Economics and Social Sciences

Von-Melle-Park 5, 20146 Hamburg, Germany

E-mail: cristina.sattarhoff@uni-hamburg.de

Abstract

This paper improves the estimation procedure of the Multifractal Random Walk model by means of an optimal iterated Generalized Method of Moments (GMM) estimator using an enhanced moments function. We report good estimation results within the scope of a Monte Carlo simulation study, with normally distributed estimates for the intermittency coefficient λ^2 . This allows us to construct statistical hypothesis tests about λ^2 . Moreover, the GMM estimator proves to be robust to variations in the parameter starting values. In a financial application we estimate the Multifractal Random Walk model from the daily values of the German DAX stock market index. Throughout our study, computing time is considerably reduced by means of an efficient algorithm for Heteroscedasticity and Autocorrelation Consistent (HAC) covariance matrix estimation. This algorithm outperforms the classical HAC estimation methods developed for GAUSS or R due to a fast Toeplitz matrix-vector multiplication procedure.

Keywords: GMM, multifractal processes, Multifractal Random Walk, HAC, Toeplitz matrices

JEL classification: C13, C63, G12

1 Introduction

The current financial crisis, which has been affecting international financial markets since 2007, has set a challenge to improve risk management methods and to rigorously assure their quality. This challenge is addressed in this paper with respect to a multifractal model of asset returns, the Multifractal Random Walk.

It is well known that the dynamics of financial prices varies with the employed sampling rate, from high-frequency data to daily, quarterly or annual data. Multifractal processes offer a reach and parsimonious tool for modeling price fluctuations at different time scales within one unifying framework. They account for the relation between the sampling period and the probability distribution of financial returns, which constitutes a valuable piece of information for risk forecasting. Correspondingly, multifractal models offer a fruitful approach to volatility modeling outperforming standard volatility forecasts with models of the GARCH/FIGARCH family (for example, see Calvet and Fisher 2004, Bacry et al. 2008a, Lux 2008).

The first multifractal model of asset returns was introduced in 1997 by the pioneering work of Calvet, Fisher and Mandelbrot.¹ Today, financial multifractality can be described in the context of two models: the Markov-Switching Multifractal (MSM) by Calvet and Fisher (2001) and the Multifractal Random Walk (MRW) by Muzy and Bacry (2002). Recently, statistical inference issues for multifractal processes were brought into focus following two directions of research: Monte Carlo simulation studies, to assess the performance of estimation and forecast methods in small samples² and analytical derivation of probability distributions, a particularly exigent research area due to the nonlinearity of multifractal models (Bacry et al. 2008b).

This paper pursues these recent developments, in respect of the MRW model. In this connection, we introduce a new Generalized Method of Moments (GMM) estimator for the MRW with enhanced performance in finite samples. The second contribution of this paper is in the field of computational econometrics. The resource efficient computer-based manipulation of large data sets is a typical challenge in finance. We propose a new algorithm for the computation of Heteroscedasticity and Autocorrelation Consistent covariance matrix estimators that can cope with large data sets due to a fast Toeplitz matrix-vector multiplication procedure.

After a short overview of the MRW model in the next section, a new optimal iterated GMM estimator for the MRW is introduced in Section 3. Due to the nonlinearity of the model, the GMM estimation is performed numerically by means of a constrained nonlinear optimization procedure, which can be initialized effectively using preliminary estimates. Section 4 is dedicated to the finite sample behaviour

¹See Calvet et al. 1997, Fisher et al. 1997, and Mandelbrot et al. 1997.

²For example, see Calvet and Fisher (2004) on the performance of Maximum Likelihood estimation for the MSM, Lux (2008) on the performance of Generalized Method of Moments estimation and linear forecasts of volatility with the MSM, Bacry et al. (2008b) on the performance of Generalized Method of Moments estimation for the MRW.

of the MRW estimator within the scope of a Monte Carlo study. Moreover, we discuss the robustness of our estimator to variations of the parameter starting values. Section 5 contains an application to the evolution of the German stock market over the past 50 years using the DAX index. The algorithm for the Heteroscedasticity and Autocorrelation Consistent covariance matrix estimation introduced in Section 6 outperforms standard estimation routines for GAUSS and R significantly in terms of computational efficiency. Concluding remarks and prospects are presented in Section 7.

2 The Multifractal Random Walk model

Today, financial data is available at various sampling rates. Electronic stock markets allow the collection of tick by tick (transaction by transaction) data sampled every 0.01 seconds, also called high-frequency data. The comparison of differently sampled time series can give a deeper insight into the dynamics of the corresponding financial price process. The following figure displays the evolution of the German DAX stock index with sampling intervals (time scales) of 1 day, half an hour and 2.5 minutes. Going over the plot from top to bottom, one can zoom in on the evolution of the DAX from its activity over a period of 19 years to its activity in the space of 1 month. Is there a statistical correspondence between these time series? Is it possible to gain useful information about the activity of a price process over a wider period of time from its evolution over a shorter period? And if so, can this knowledge constitute a solid basis for price forecasts?

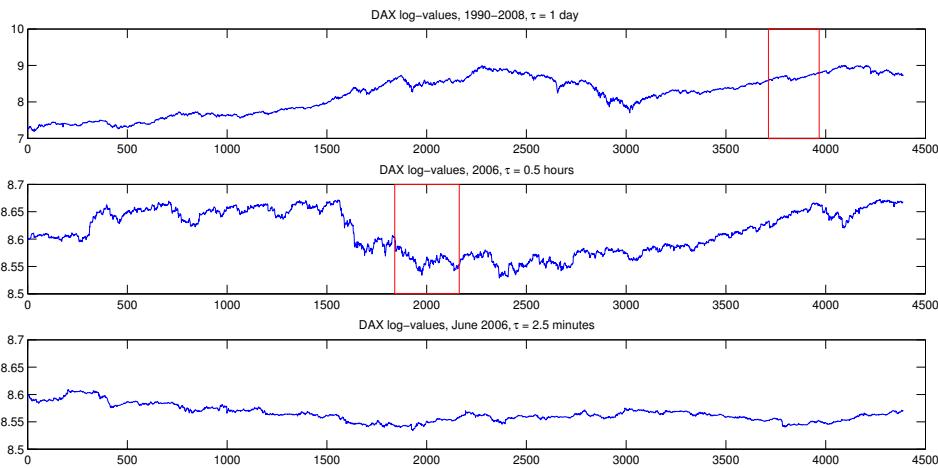


Figure 1: The log-values of the German DAX stock market index with the sampling interval $\tau = 1$ day, $\tau = 30$ minutes and $\tau = 2.5$ minutes respectively. The red lines mark the data section displayed in the subsequent plot.

These questions can be explored within the framework of *scale-invariant processes*,

which can capture the relation between the sampling period and the probability distribution of financial returns (Calvet and Fisher 2002, pp. 385-386). In the following we will denote the increments of a real-valued process $X(t)$ between times t and $t - \tau$ by $\delta_\tau X(t) := X(t) - X(t - \tau)$.

Definition 1 Let $X(t)$ be a stochastic process with stationary increments $\delta_\tau X(t)$. Suppose that the absolute moments $E[|\delta_\tau X(t)|^q]$ exist and are finite for all $q \in Q$, where $Q \subset \mathbb{R}$ is a suitable set containing the interval $[0, 1]$. Consider c_q the q -th absolute moment of the increments $\delta_\tau X(t)$ with scale $\tau = 1$: $c_q := E[|\delta_1 X(t)|^q]$. Further, let $\zeta : Q \rightarrow \mathbb{R}$ denote a differentiable function and T a real positive number. The process $X(t)$ is scale-invariant with multifractal exponent $\zeta(q)$ and integral scale T if we have

$$E[|\delta_\tau X(t)|^q] = c_q \tau^{\zeta(q)} \quad (1)$$

for all $\tau \in [0, T]$, for all $t \in \mathbb{R}_+$ and for all moment orders $q \in Q$.

According to (1), the statistical moments of the increments of a scale-invariant process are exact power laws as a function of the sampling interval τ . Generally one can distinguish between two types of scale-invariant processes subject to their multifractal exponents $\zeta(q)$: *fractal processes* (also called monofractal, unifractal or uniscaling) when $\zeta(q)$ is linear³ and *multifractal processes* (also called multiscaling or intermittent) when $\zeta(q)$ is non-linear (Muzy and Bacry 2002). Non-linear exponents $\zeta(q)$ are necessarily concave and have finite T (Bacry and Muzy 2003, p. 450). The stronger the concavity of the multifractal exponent $\zeta(q)$, the stronger the multifractality of the scale-invariant process. A common measure of the degree of multifractality is the *intervallency coefficient*

$$\lambda^2 := -\zeta''(0) \geq 0.$$

λ^2 increases with the bending of the multifractal exponent, i.e. with the degree of multifractality, and it is 0 in the case of fractal processes.

The Multifractal Random Walk is a continuous nonlinear multifractal process which was first propounded in Bacry et al. (2001). The model was generalized by Bacry and Muzy (Muzy and Bacry 2002, Bacry and Muzy 2003), who introduce a continuous random cascade model on the upper half-plane by means of an independently scattered infinitely divisible 2D random measure. This cascade model generates the log-infinitely divisible Multifractal Random Walk in the limit of small scales $\ell \rightarrow 0$.

In this paper we employ a (log-normal) Multifractal Random Walk, which can be constructed by means of a Gaussian process $\omega_\ell(t)$.

Definition 2 Consider the stationary Gaussian process $\omega_\ell(t)$ with mean

$$E[\omega_\ell(t)] = -\lambda^2 \left(\ln \left(\frac{T}{\ell} \right) + 1 \right)$$

³In this case, the scale-invariance property holds for all scales τ ($T = \infty$). Self-similar processes with stationary increments (e.g. the fractional Brownian motion) are fractal.

and autocovariance function $\gamma_{\omega_\ell}(h) := \text{Cov}[\omega_\ell(t), \omega_\ell(t+h)]$,

$$\gamma_{\omega_\ell}(h) = \begin{cases} \lambda^2 \left(\ln\left(\frac{T}{\ell}\right) + 1 - \frac{h}{\ell} \right), & 0 \leq h < \ell \\ \lambda^2 \ln\left(\frac{T}{h}\right), & \ell \leq h < T \\ 0, & h \geq T \end{cases},$$

also called a Gaussian magnitude.

The Multifractal Random Walk (MRW) is the weak limit process $X(t)$

$$X(t) = \lim_{\ell \rightarrow 0^+} \int_0^t e^{\omega_\ell(u)} dB(u), \quad (2)$$

where $\omega_\ell(u)$ is a Gaussian magnitude with $\lambda^2 < \frac{1}{2}$ and $dB(u)$ is a Gaussian White Noise with mean 0 and variance σ^2 , independent of $\omega_\ell(u)$. The MRW $X(t)$ has the increments

$$\delta_\tau X(t) = \lim_{\ell \rightarrow 0^+} \int_{t-\tau}^t e^{\omega_\ell(u)} dB(u). \quad (3)$$

The structure of the Gaussian magnitude $\omega_\ell(t)$ plays a decisive role for the properties of the MRW. Let us remark that the magnitude $\omega_\ell(t)$ is uncorrelated and hence is independent for lags $h \geq T$. For this reason, the integral scale T is also called the *decorrelation scale*. Analogously, the MRW increments (3) are independent for lags $h \geq T + \tau$. According to this, the increments process is ergodic. This property will facilitate the estimation of the model in the following section.⁴

Note that the (limit) probability distribution of the MRW is unknown except for some statistical moments:

$$E[X(t)^2] = \sigma^2 t$$

and accordingly

$$E[\delta_\tau X(t)^2] = \sigma^2 \tau. \quad (4)$$

Additionally, Bacry et al. (2008a, 2008b) derived an approximate distribution of the MRW increments in the limit of small intermittency $\lambda^2 \ll 1$ given by the rescaled product of a Gaussian White Noise and a log-normal process. This facilitated the formulation of approximate expressions for the mean and the autocovariance function (ACF) of the logarithmic absolute increments $Z_\tau(t) := \ln(|\delta_\tau X(t)|)$. In the following we reproduce these moment approximations for increments $Z(t) := Z_1(t)$

⁴Bacry and Muzy (2003, pp. 456-460) prove the existence and non-degeneracy of the MRW. The authors show that the model is stochastically self-similar with the random self-similarity factor $W_{\ell/L} = \sqrt{\ell/L} e^{\Omega_{\ell/L}}$, $\Omega_{\ell/L} \sim \mathcal{N}(\lambda^2 \ln(\ell/L), -\lambda^2 \ln(\ell/L))$, $\ell < L$ and that it has stationary increments. See also Bacry et al. (2008b, Section II) on a detailed exposition of the model.

with $\tau = 1$. We denote the ACF $Cov[Z(t), Z(t+h)]$ by $\gamma(h)$ and the mean $E[Z(t)]$ by μ . We then have for $h \geq 1$ for all t

$$\gamma(h) = \begin{cases} \lambda^2 g(h, T) + o(\lambda^2), & h \leq T-1 \\ o(\lambda^2), & h \geq T+1 \end{cases} \quad (5)$$

with

$$g(h, T) = \begin{cases} \ln(T) + 1.5 - 2\ln(2), & h = 1, T \geq 2 \\ \ln\left(\frac{T}{h}\right) - \frac{(h+1)^2}{2} \ln\left(1 + \frac{1}{h}\right) - \frac{(h-1)^2}{2} \ln\left(1 - \frac{1}{h}\right) + 1.5, & 2 \leq h \leq T-1 \end{cases}$$

where $o(\lambda^2)$ satisfies $\lim_{\lambda^2 \rightarrow 0^+} \frac{o(\lambda^2)}{\lambda^2} = 0$. The mean of the process $Z(t)$ is

$$\mu = \ln(\sigma) - \frac{\gamma + \ln(2)}{2} - \lambda^2 (1.5 + \ln(T)) + o(\lambda^2) \quad (6)$$

with γ the Euler constant.

The MRW features three parameters: the intermittency coefficient λ^2 , the logarithmic decorrelation scale $\ln(T)$ and the logarithmic standard deviation $\ln(\sigma)$. The following plot exhibits the first 300 lags of the ACF in (5) with $\lambda^2 = 0.02$ and $T = 200$. The value of T is given by the point where the ACF becomes zero. One can see that the ACF passes very smoothly on to zero. On account of this, the identification of T (or $\ln(T)$) proves to be a demanding task, even within the deterministic framework. The estimation of $\ln(T)$ is consequently all the more ambitious, and will be considered together with the estimation of λ^2 and $\ln(\sigma)$ in the following two sections.

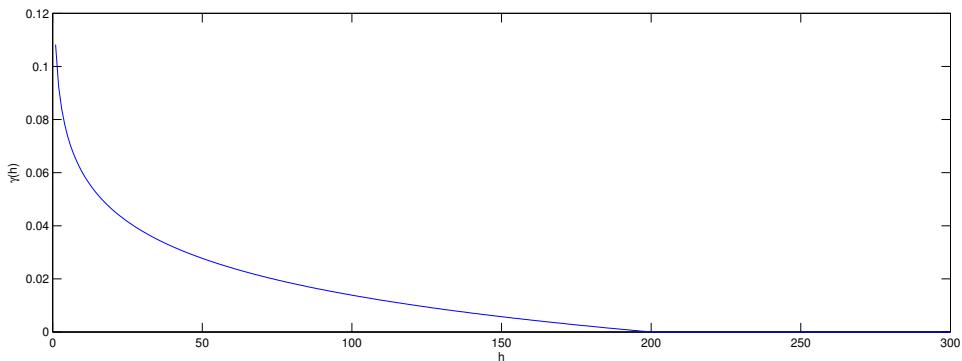


Figure 2: The ACF $\gamma(h)$ of the process $Z(t)$ with $\lambda^2 = 0.02$ and $T = 200$.

3 The estimation procedure

This section addresses the estimation of the MRW parameters: the intermittency coefficient λ^2 , the logarithmic decorrelation scale $\ln(T)$ and the logarithmic stan-

dard deviation $\ln(\sigma)$. Bacry et al. (2008b) supplied the 2nd moment of the MRW increments as well as approximate expressions for the mean and the autocovariance function of the logarithmic increments $Z(t)$ in the limit of small intermittency $\lambda^2 \ll 1$. These moment conditions can be used to estimate the MRW parameters with the Generalized Method of Moments (GMM). In this paper we employ a new moments configuration which improves the performance of the GMM estimation in finite samples. Moreover, we introduce an effective initialization method of the GMM estimation for the variance parameter using a preliminary estimate.

3.1 The GMM estimator

Let us denote with index 0 the unknown MRW parameters λ_0^2 , $\ln(T_0)$ and $\ln(\sigma_0)$, which we want to estimate from a time series $Z(k), k \in \mathbb{N}, 1 \leq k \leq N$ with N the time series length. Accordingly

$$\theta_0 = (\lambda_0^2 \quad \ln(T_0) \quad \ln(\sigma_0))'$$

will identify the unknown parameter vector. This is assumed to lie in a given set $\Theta \subset [0, 0.5) \times \mathbb{R}^2$. Without loss of generality we choose Θ compact. The moment conditions (4), (5) and (6) with lags $h_1 \leq h \leq h_H$ for the ACF can be written compactly as follows:

$$E[f(Z(t), \theta_0)] = 0 \quad (7)$$

with the moments function f given by

$$f(Z(t), \theta) = \begin{pmatrix} \exp(2Z(t)) \\ Z(t) \\ (Z(t) - \mu(\theta))(Z(t + h_1) - \mu(\theta)) \\ \vdots \\ (Z(t) - \mu(\theta))(Z(t + h_H) - \mu(\theta)) \end{pmatrix} - \begin{pmatrix} \sigma^2 \\ \mu(\theta) \\ \gamma(h_1, \theta) \\ \vdots \\ \gamma(h_H, \theta) \end{pmatrix} \quad \forall \theta \in \Theta. \quad (8)$$

By contrast Bacry et al. (2008a, 2008b) use the moments function

$$f^*(Z(t), \theta) = \begin{pmatrix} \exp(2Z(t)) \\ (Z(t) - \mu(\theta))(Z(t + h_1) - \mu(\theta)) \\ \vdots \\ (Z(t) - \mu(\theta))(Z(t + h_H) - \mu(\theta)) \end{pmatrix} - \begin{pmatrix} \sigma^2 \\ \gamma(h_1, \theta) \\ \vdots \\ \gamma(h_H, \theta) \end{pmatrix} \quad \forall \theta \in \Theta, \quad (9)$$

which doesn't exploit the mean condition directly but only by means of the ACF.

It is reasonable to investigate to what extent the moment conditions (7) hold in the sample $Z(k)$. The GMM estimator $\hat{\theta}_N$ is the value of θ , which minimizes the deviation of the sample moments from their theoretical counterparts with respect to the quadratic norm:

$$\hat{\theta}_N = \arg \min_{\theta \in \Theta} Q_N(\theta) \quad (10)$$

with

$$Q_N(\theta) = \left(\frac{1}{N} \sum_{k=1}^N f(Z(k), \theta)' \right) W_N \left(\frac{1}{N} \sum_{k=1}^N f(Z(k), \theta) \right) \quad (11)$$

where the weighting matrix W_N is positive semi-definite and is assumed to converge in probability to a positive definite matrix of constants, W (Greene 2003, pp. 536-538; Hall 2005, p. 14).

As discussed in Section 2, the MRW increments and accordingly the process $Z(t)$ are stationary and ergodic. These properties together with some regularity conditions⁵ ensure the weak consistency of the GMM estimator

$$\underset{N \rightarrow \infty}{\text{plim}} \widehat{\theta}_N = \theta_0.$$

Additionally, $\widehat{\theta}_N$ is asymptotically normally distributed

$$\sqrt{N} (\widehat{\theta}_N - \theta_0) \xrightarrow{d} \mathcal{N}(0, MSM') \quad (12)$$

with the large sample distribution

$$\widehat{\theta}_N \stackrel{\text{d}}{\sim} \mathcal{N}\left(\theta_0, \frac{1}{N} MSM'\right),$$

where

$$S = \lim_{N \rightarrow \infty} N \cdot \text{Var} \left[\frac{1}{N} \sum_{k=1}^N f(Z(k), \theta_0) \right]$$

is the asymptotic covariance matrix of the normalized sample mean, and

$$M = (J'WJ)^{-1} J'W$$

with $J = E[\partial f(Z(t), \theta_0) / \partial \theta']$ the expected value of the Jacobian matrix of f at θ_0 (Hall 2005, pp. 66-72).

The GMM estimator has minimum asymptotic variance $MSM' = (J'S^{-1}J)^{-1}$ with respect to the choice of the limit weighting matrix W , when using $W = S^{-1}$ (Davidson and MacKinnon 1993, pp. 597-598). On account of this $\widehat{\theta}_N$ with the limit weighting matrix $W = S^{-1}$ is called the *optimal GMM estimator*. Its determination requires a weakly consistent estimator of the unknown asymptotic covariance matrix S and can be accomplished as follows: first, one constructs a suboptimal estimator $\widehat{\theta}_N(1)$ of θ_0 with the weighting matrix $W_N(1) = I$, where I denotes the identity matrix. The estimator $\widehat{\theta}_N(1)$ is used to construct the weakly consistent estimator $\widehat{S}_N(1)$ for S . As a second step, $W_N(2) = \widehat{S}_N^{-1}(1)$ is employed to estimate $\widehat{\theta}_N(2)$. The estimator $\widehat{\theta}_N(2)$ is then optimal for θ_0 .

One can iterate this estimation algorithm and re-estimate $\widehat{S}_N(2)$ by means of $\widehat{\theta}_N(2)$ and consecutively $\widehat{\theta}_N(3)$ with $W_N(3) = \widehat{S}_N^{-1}(2)$. $\widehat{\theta}_N(3)$ is also optimal.

⁵See Hall (2005, pp 50-71) for a detailed discussion of the regularity conditions.

Moreover $\widehat{\theta}_N(3)$ tends to fluctuate less than $\widehat{\theta}_N(2)$ around θ_0 in finite samples. Generally, the larger the number of estimation steps, the more efficient the optimal estimator in finite samples (Hall 2005, Chapter 6). This estimation procedure can be repeated until the estimates converge to a solution $\widehat{\theta}_N(i)$ called *optimal iterated GMM estimator*.⁶

The MRW estimation in this paper employs an optimal iterated GMM estimator with the moments function (8). For this purpose the asymptotic covariance matrix S will be estimated by means of an Heteroscedasticity and Autocorrelation Consistent (HAC) covariance matrix estimator.

3.2 The HAC estimator

Consider the following alternative expression for S

$$S = \Gamma_0 + \sum_{i=1}^{\infty} (\Gamma_i + \Gamma'_i) \quad (13)$$

with Γ_i the i^{th} autocovariance function matrix of f

$$\Gamma_i = E [(f(Z(k), \theta_0) - E[f(Z(k), \theta_0)]) (f(Z(k-i), \theta_0) - E[f(Z(k), \theta_0)])']$$

(Hamilton 1994, pp. 261-262, 279-281). In the GMM context, the autocovariance function matrix Γ_i can be simplified to

$$\Gamma_i = E [f(Z(k), \theta_0) f(Z(k-i), \theta_0)'],$$

since the moment conditions (7) are satisfied. The HAC estimator \widehat{S}_N is a nonparametric kernel estimator originated from the estimation of spectral density, introduced to econometrics literature by Newey and West (1987):

$$\widehat{S}_N = \widehat{\Gamma}_0 + \sum_{i=1}^{N-1} \omega_{i,N} (\widehat{\Gamma}_i + \widehat{\Gamma}'_i) \quad (14)$$

where $\widehat{\Gamma}_i$ is the i^{th} empirical autocovariance matrix of f

$$\widehat{\Gamma}_i = \frac{1}{N} \sum_{k=i+1}^N f(Z(k), \widehat{\theta}_N) f(Z(k-i), \widehat{\theta}_N)' \quad (15)$$

(Hall 2005, pp. 79-86). Under weak assumptions on f the HAC estimator is weakly consistent

$$\text{plim}_{N \rightarrow \infty} \widehat{S}_N = S,$$

⁶We say that the estimation procedure converged to a solution $\widehat{\theta}_N(i)$ in iteration step i when the difference between the estimates $\widehat{\theta}_N(i)$ and $\widehat{\theta}_N(i+1)$ becomes negligible, i.e. $\|\widehat{\theta}_N(i+1) - \widehat{\theta}_N(i)\| < \varepsilon$ with ε an arbitrarily small positive constant. In this study the maximum number of iteration steps was restricted to 50.

which ensures the weak consistency of the optimal GMM estimator.

We can write $\widehat{\Gamma}_i$ compactly as follows

$$\widehat{\Gamma}_i = \frac{1}{N} (F((i+1) : N))' F(1 : (N-i)) \quad (16)$$

by means of the matrix $F \in \mathbb{R}^{N \times q}$ containing the sample values of the moments function f

$$F = \begin{pmatrix} f(Z(1), \widehat{\theta}_N)' \\ f(Z(2), \widehat{\theta}_N)' \\ \vdots \\ f(Z(N), \widehat{\theta}_N)' \end{pmatrix} \quad (17)$$

and where the expression $F(j : k)$ denotes the rows j till k of F for any $j, k \in \mathbb{N}$, $1 \leq j, k \leq N$

$$F(j : k) = \begin{pmatrix} f(Z(j), \widehat{\theta}_N)' \\ f(Z(j+1), \widehat{\theta}_N)' \\ \vdots \\ f(Z(k), \widehat{\theta}_N)' \end{pmatrix}.$$

Based on a sample of size N , one can compute N empirical autocovariance matrices $\widehat{\Gamma}_i$ with lags i between 0 and $N - 1$. These empirical estimates are less efficient with i . Hence it is pointless to use the estimates for large lags i . Instead these are excluded from the HAC estimator (14) by multiplying them with weights $\omega_{i,N} = 0$. In this way, the weights function $\omega_{i,N}$ controls the number of estimates $\widehat{\Gamma}_i$ to be considered in (14). Furthermore, $\omega_{i,N}$ controls the relative contribution of $\widehat{\Gamma}_i$ to \widehat{S}_N .

In this paper, we use Bartlett weights

$$\omega_{i,N} = \begin{cases} 1 - \frac{i}{b_N + 1}, & 1 \leq i \leq b_N + 1 \\ 0, & i > b_N + 1 \end{cases}$$

with b_N the optimal bandwidth selected according to Newey and West (Hall 2005, pp. 81-82). The bandwidth b_N gives the number of estimates $\widehat{\Gamma}_i, i \geq 1$ to be considered in (14) (weighted with $\omega_{i,N} \neq 0$) so that we can also write

$$\widehat{S}_N = \widehat{\Gamma}_0 + \sum_{i=1}^{b_N+1} \omega_{i,N} (\widehat{\Gamma}_i + \widehat{\Gamma}'_i). \quad (18)$$

3.3 The numerical minimization procedure

The GMM estimator $\widehat{\theta}_N$ is the global solution of the minimization problem (10) with bound constraints $\theta \in \Theta$. For the purpose of the MRW estimation in this paper the parameter space was restricted to $\Theta = [0, 0.5] \times [0, \infty) \times (-\infty, \infty)$. This choice

of Θ is justified by the moment conditions (5) and (6), which are valid only in the small intermittency regime $\lambda_0^2 \ll 1$. Furthermore, we consider $\ln(T_0) \geq 0$ ($T_0 \geq 1$) for the sake of simplicity. In this case the MRW increments exert influence at least on the increments during the following two periods. This is satisfied in financial applications, where the MRW increments model logarithmic asset returns.

Due to the nonlinearity of the MRW, the minimization problem (10) cannot be solved to produce a closed form solution for $\hat{\theta}_N$. Instead we determine $\hat{\theta}_N$ numerically by means of an iterative search using a *trust region algorithm for constrained nonlinear optimization*. Trust region algorithms build a distinct class of relatively new optimization techniques with very strong convergence properties. Unlike line search algorithms, in a trust region algorithm, the minimization problem is approximated at each iteration step only within a neighbourhood centered at the current candidate value for $\hat{\theta}_N$, called the *trust region*. This approximate model is then solved for search direction and trial step simultaneously (Geiger and Kanzow 1999, Chapter 14).

In this study we implemented the trust region optimization algorithm with the computer program MATLAB (The MathWorks 2009a).⁷ We assess convergence of the algorithm to a solution $\hat{\theta}_N := \bar{\theta}(i)$ at iteration step i if one of the following three stopping criteria is met:

- stopping criterion with respect to the parameters: the new candidate $\bar{\theta}(i+1)$ is located within a negligible distance from $\bar{\theta}(i)$

$$\|\bar{\theta}(i+1) - \bar{\theta}(i)\| < \varepsilon_1$$

- stopping criterion with respect to the minimand: the difference between the minimands on step i and $i+1$ is insignificant

$$|Q_N(\bar{\theta}(i+1)) - Q_N(\bar{\theta}(i))| < \varepsilon_2$$

- stopping criterion with respect to the first-order conditions: the first-order conditions for constrained optimization are given by the Karush-Kuhn-Tucker conditions. MATLAB provides a *first-order optimality measure* based on the Karush-Kuhn-Tucker conditions. Convergence is assessed if the optimality measure is less than ε_2 .

The tolerances ε_1 and ε_2 were set to 10^{-6} . Alternatively we exit the procedure if the maximum number of iteration steps, here set to 400, is reached.⁸

The choice of the parameter starting value for θ plays a decisive role in the performance of the minimization algorithm. If the starting value lies far from the

⁷We use the MATLAB function *fmincon* in the Optimization Toolbox.

⁸See also The MathWorks (2009b, Chapter 2, pp. 35-36 and Chapter 3, pp.11-13) for a detailed discussion of the optimization settings.

global solution of the minimization problem or it lies in a complex region of the parameter space, the search method may get stuck in a parameter area of local optimality and fail to reach the global solution (Hall 2005, pp. 59-60). In this case, the GMM estimator $\widehat{\theta}_N$ is set automatically to a local minimum and performs poorly.

The moment condition (4) facilitates the construction of a preliminary estimate $\overline{\ln(\sigma)}$ for $\ln(\sigma)$ which can be used to initialize the minimization algorithm:

$$\overline{\ln(\sigma)} = \frac{1}{2} \ln\left(\widehat{\sigma}^2\right). \quad (19)$$

This pre-estimate employs the *trivial variance estimator*

$$\widehat{\sigma}^2 = \frac{1}{N} \sum_{k=1}^N \delta_1 X(k)^2$$

based on (4) with $\tau = 1$. In the next section we will address the performance of this initialization method for $\ln(\sigma)$ within the scope of a Monte Carlo simulation study.

4 Finite sample behaviour of the estimation procedure

The GMM estimator is weakly consistent and asymptotically normally distributed. However for the purpose of empirical studies it is important to understand how it behaves in finite samples. In this section we investigate the performance of the GMM estimator for the MRW in finite samples by means of an extensive Monte Carlo simulation study. We also examine the performance of the initialization method for $\ln(\sigma)$ as well as the influence of the starting values for λ^2 and $\ln(T)$ on the estimation.

4.1 The Monte Carlo simulation study

In the last section the GMM estimator $\widehat{\theta}_N$ for the true MRW parameter vector θ_0 was considered. For simplicity reasons we will further omit the subscripts for the true values and estimators of the individual MRW parameters. These will be denoted simply by λ^2 , $\ln(T)$, $\ln(\sigma)$ and $\widehat{\lambda}^2$, $\widehat{\ln(T)}$, $\widehat{\ln(\sigma)}$ respectively.

We consider three parameter specifications, which are depicted in Table (1),⁹ and five sample sizes $N \in \{1897, 3945, 8041, 16233, 31849\}$. For each model configuration and sample size we simulate the MRW process at discrete time points $n \in \mathbb{N}, 1 \leq n \leq N_{sim}$ 10,000 times, with $N_{sim} = N + 151$, and estimate the true model parameters λ^2 , $\ln(T)$ and $\ln(\sigma)$ via GMM. The simulated sample sizes N_{sim}

⁹The variance parameter σ^2 is a simple multiplicative factor for the MRW, which doesn't affect the estimation performance (Bacry et al. 2008b, p. 12). On account of this we keep the parameter $\ln(\sigma)$ constant throughout the entire simulation study.

differ from the sample sizes N employed in the estimation for two reasons:¹⁰ First, the estimation is based on the simulated series $Z(n) = \ln(|\delta_1 X(n)|)$ of length $N_{sim} - 1$. Second, the maximum lag of the ACF considered in the GMM estimation is $h = 150$. Accordingly, we can employ only $N = N_{sim} - 1 - 150$ datapoints in the estimation in order to compute the empirical ACF with lag 150.

Model configuration	λ^2	$\ln(T)$	$\ln(\sigma)$
MRW1	0.02	5.3	0
MRW2	0.04	5.3	0
MRW3	0.02	9.7	0

Table 1: Model configurations. The values of the parameter $\ln(T)$ correspond to the decorrelation scales $T = 200$ and $T = 16384$ respectively.

The simulation procedure is based on a discrete time representation of equation (2) which can be accomplished in two steps according to Bacry et al. (2008b, p. 6). First we construct the process $X_\ell(m\ell)$ on a fine grid $m\ell, m \in \mathbb{N}$ with $\ell > 0, \ell \ll 1$ and ℓ^{-1} integer

$$X_\ell(m\ell) = \sum_{k=1}^m e^{\omega_\ell(k\ell)} \varepsilon(k\ell) \sqrt{\ell}$$

where $\omega_\ell(k\ell)$ is a stationary Gaussian process with the mean and the ACF in Definition 2 and $\varepsilon(k\ell)$ is a Gaussian White Noise with mean 0 and variance σ^2 , independent of $\omega_\ell(k\ell)$. In this paper we employ sampling intervals of length $\ell = 2^{-7}$. The process $X_\ell(m\ell)$ corresponds to the right-hand side of equation (2) when we omit the limit for $\ell \rightarrow 0^+$. The limit is then replicated in the second step by extracting only the representations of $X_\ell(m\ell)$ at integer time points $m\ell$

$$X(n) := X_\ell(m\ell) \text{ with } n = m\ell \in \mathbb{N}$$

with $X(n)$ the simulated discrete-time MRW.

The simulation study in this paper comprises 150,000 series $X(n)$. All the simulation parameters are concordant with the parameters in Bacry et al. (2008b) and ensure the comparability of our results. By contrast we employ an optimal iterated GMM estimator of the MRW, with an enhanced moments configuration whose performance will be evaluated below.

4.2 The estimation performance

The MRW parameters were estimated from the simulated data by means of the optimal iterated GMM estimator introduced in Section 3. We used the moments function (8) with 22 lags $1 \leq h \leq 150$ for the ACF $\gamma(h)$. For each simulated series

¹⁰The simulated sample sizes are $N_{sim} \in \{2048, 4096, 8192, 16384, 32000\}$.

the minimization procedure was initialized with the corresponding true parameter values λ , $\ln(T)$ and $\ln(\sigma)$ employed in the simulation. This estimation procedure was implemented in MATLAB using a revised and reprogrammed GMM toolbox as created by Kyriakoulis (2007).

The estimation performance can be evaluated using the empirical counterparts of the loss functions Bias and MSE. In the following we denote the empirical Bias with $eBias$ and the empirical MSE with $eMSE$. These are given by the following expressions when using the example of the GMM estimator $\widehat{\lambda}^2$:

$$eBias\left(\widehat{\lambda}^2\right) = \frac{1}{10000} \sum_{k=1}^{10000} \left(\widehat{\lambda}^2_k\right) - \lambda^2,$$

$$eMSE\left(\widehat{\lambda}^2\right) = \frac{1}{10000} \sum_{k=1}^{10000} \left[\left(\widehat{\lambda}^2_k - \lambda^2\right)^2\right],$$

where $\widehat{\lambda}^2_k$ denotes the GMM estimate for λ^2 using the k^{th} simulated MRW series. Tables (2), (3) and (4) comprise the values of the eBias and eMSE for the MRW estimates for each model configuration and sample size in our simulation study. Further, we investigate the normality of the MRW estimates by means of QQ plots which can be found in the Appendix.

	$\widehat{\lambda}^2$		$\widehat{\ln(T)}$		$\widehat{\ln(\sigma)}$		
N	eBias	eMSE	eBias	eMSE	eBias	eMSE	
1897	-0.0025	$3 \cdot 10^{-5}$	0.0082	2.0242	-0.0102	0.0050	
3945	-0.0014	10^{-5}	-0.0257	0.6568	-0.0056	0.0024	
8041	-0.0007	$6 \cdot 10^{-6}$	-0.0222	0.2767	-0.0028	0.0012	
16233	-0.0004	$3 \cdot 10^{-6}$	-0.0087	0.1269	-0.0012	0.0006	
31849	-0.0002	10^{-6}	-0.0074	0.0614	-0.0007	0.0003	

Table 2: The eBias and eMSE values for the model configuration MRW1

	$\widehat{\lambda}^2$		$\widehat{\ln(T)}$		$\widehat{\ln(\sigma)}$		
N	eBias	eMSE	eBias	eMSE	eBias	eMSE	
1897	-0.0035	$6 \cdot 10^{-5}$	-0.1872	0.8199	-0.0236	0.0111	
3945	-0.0018	$3 \cdot 10^{-5}$	-0.1013	0.3658	-0.0139	0.0054	
8041	-0.0009	10^{-5}	-0.0547	0.1702	-0.0077	0.0027	
16233	-0.0003	$6 \cdot 10^{-6}$	-0.0338	0.0796	-0.0035	0.0013	
31849	0.00003	$3 \cdot 10^{-6}$	-0.0224	0.0375	-0.0021	0.0007	

Table 3: The eBias and eMSE values for the model configuration MRW2

N	$\widehat{\lambda^2}$		$\widehat{\ln(T)}$		$\widehat{\ln(\sigma)}$		
	eBias	eMSE	eBias	eMSE	eBias	eMSE	
1897	-0.0028	$4 \cdot 10^{-5}$	-3.0798	17.5060	-0.0832	0.0824	
3945	-0.0016	$2 \cdot 10^{-5}$	-2.6151	11.2747	-0.0655	0.0653	
8041	-0.0009	$8 \cdot 10^{-6}$	-1.9928	7.3995	-0.0484	0.0483	
16233	-0.0005	$4 \cdot 10^{-6}$	-1.3603	4.9398	-0.0322	0.0326	
31849	-0.0003	$2 \cdot 10^{-6}$	-0.8093	3.2351	-0.0205	0.0193	

Table 4: The eBias and eMSE values for the model configuration MRW3

We report very good estimation results for the parameters λ^2 and $\ln(\sigma)$ for all three model configurations, featuring normally distributed estimates¹¹ with very small eBias and eMSE. By contrast, the estimator $\widehat{\ln(T)}$ for the parameter $\ln(T)$ performs poorly. This confirms our guess expressed in Section 2 based on the plot of the ACF $\gamma(h)$ which passes very slowly on to zero. According to this, the transition point of $\gamma(h)$ to zero, at T , is difficult to identify from deterministic data and a fortiori within a stochastic framework.

$\widehat{\ln(T)}$ exhibits small eBias and eMSE for the models MRW1 and MRW2, yet its estimates do not satisfy the normal distribution. The deviations of $\widehat{\ln(T)}$ from the normal distribution are small for large sample sizes N , however the normal distribution could not be reached for any of the sample sizes $N \leq 31849$ in this study.

The estimator $\widehat{\ln(T)}$ fails in the case of the model MRW3 due to its high eBias and eMSE as well as its considerable deviations from the normal distribution. This is no surprise since in most cases the MRW3 samples are shorter than the decorrelation scale T . Hence the MRW3 samples can capture only the first phase $h < T$ of the correlation structure of the model and miss any information about the transition point T . Even in the case of $N = 31849$ with N almost twice as large as T , the estimator $\widehat{\ln(T)}$ doesn't work satisfactorily. The sample size is still too short compared to the decorrelation scale T to allow for its proper identification.

On account of this, we recommend for the purpose of empirical studies the use of long time series combined with preliminary information about T derived from the economic theory, in order to ensure that N is much greater than T ($N \gg T$).

The GMM estimators in this paper clearly outperform the results in Bacry et al. (2008b) in respect of its considerably smaller eMSE. Moreover, the estimates for λ^2 are normally distributed in contrast to Bacry et al. (2008b). This allows the employment of statistical hypothesis tests for the MRW model with respect to λ^2 (see Section 5).

¹¹The λ^2 estimates deviate from the normal distribution in the case of MRW1 with $N = 1897$ and $N = 3945$ respectively.

4.3 The performance of the initialization method for $\ln(\sigma)$

The choice of the starting value for θ in the numerical minimization procedure plays a decisive role in its performance and consequently in the performance of the GMM estimation. In the best case, the parameter starting value should lie as close as possible to the global minimum in order to ensure the (fast) convergence of the minimization procedure towards the global solution (Hall 2005, pp. 59-60).

In the following we investigate the performance of the starting value $\overline{\ln(\sigma)}$ introduced in Section 3.3, in the estimation of the MRW, using the simulated MRW1 series with sample sizes $N \in \{1897, 3945, 8041, 16233\}$. For each sample size N we will evaluate how close the starting value $\overline{\ln(\sigma)}$ lies on average to the global minimum of the GMM problem, for the parameter $\ln(\sigma)$.

For this purpose we compute the starting values $\overline{\ln(\sigma)}_k$ with $k \in \{1, 2, \dots, 10000\}$ identifying the employed MRW1 realization. The global solutions of the GMM problem are given by the estimated values $\widehat{\ln(\sigma)}_k$ in Section 4.2 respectively. We compute the mean distance \bar{d} between $\overline{\ln(\sigma)}_k$ and $\widehat{\ln(\sigma)}_k$

$$\bar{d} = \frac{1}{10000} \sum_{k=1}^{10000} \left(\overline{\ln(\sigma)}_k - \widehat{\ln(\sigma)}_k \right),$$

and the mean squared distance \bar{d}^2

$$\bar{d}^2 = \frac{1}{10000} \sum_{k=1}^{10000} \left(\overline{\ln(\sigma)}_k - \widehat{\ln(\sigma)}_k \right)^2.$$

Our results in the following table show that $\overline{\ln(\sigma)}$ lies close to $\widehat{\ln(\sigma)}$ on average. Consequently $\overline{\ln(\sigma)}$ is a reliable starting value for $\ln(\sigma)$. Starting the search at $\overline{\ln(\sigma)}$ will contribute to the (fast) convergence of the search method towards the global solution.

N	\bar{d}	\bar{d}^2
1897	0.0055	$5 \cdot 10^{-4}$
3945	0.0030	$2 \cdot 10^{-4}$
8041	0.0017	$6 \cdot 10^{-5}$
16233	0.0009	$2 \cdot 10^{-5}$

Table 5: The mean distance \bar{d} and the mean squared distance \bar{d}^2

Currently corresponding preliminary estimates for the parameters λ^2 and $\ln(T)$ are not available. On account of this, the minimization procedure will be arbitrarily initialized for the purpose of empirical applications. In the following subsection we investigate the influence of the starting values for λ^2 and $\ln(T)$ on the estimation procedure.

4.4 The robustness of the estimation procedure to variations of the starting values for λ^2 and $\ln(T)$

Consider the GMM minimand Q_N in (11) in the first iteration step, as a function of the parameters λ^2 and $\ln(T)$, with $W_N = I$ and $\ln(\sigma)$ set to its starting value $\overline{\ln(\sigma)}$. Visual inspection of the simulated data for different model configurations and sample sizes points out that the objective function Q_N has a global minimum close to the true parameter value θ_0 and a local minimum, separated by a bump. This can be seen in the following plots when using the example of a simulated MRW1 series with $N = 1897$.

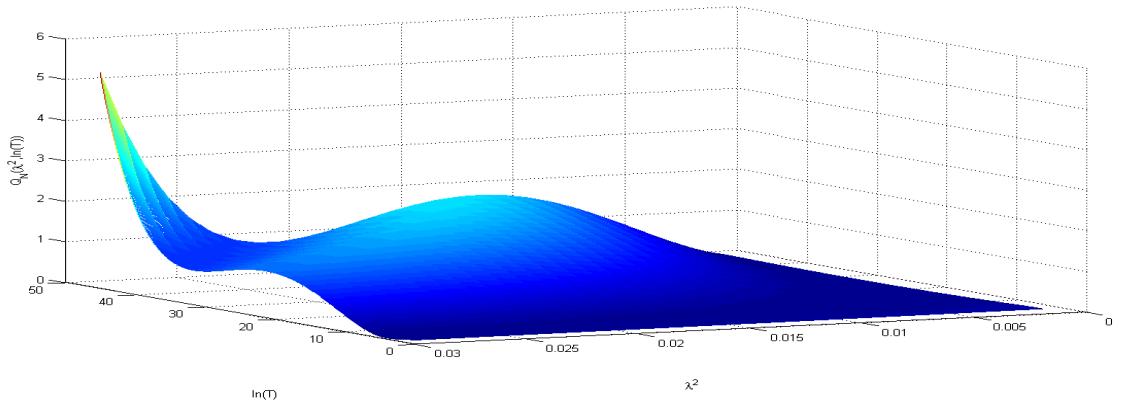


Figure 3: The GMM minimand $Q_N(\lambda^2, \ln(T))$ in the first iteration step, with $W_N = I$ and $\ln(\sigma) = \overline{\ln(\sigma)}$, for a simulated MRW1 series of size $N = 1897$. Q_N is based on the moments function f with 22 lags $1 \leq h \leq 150$ for the ACF $\gamma(h)$.

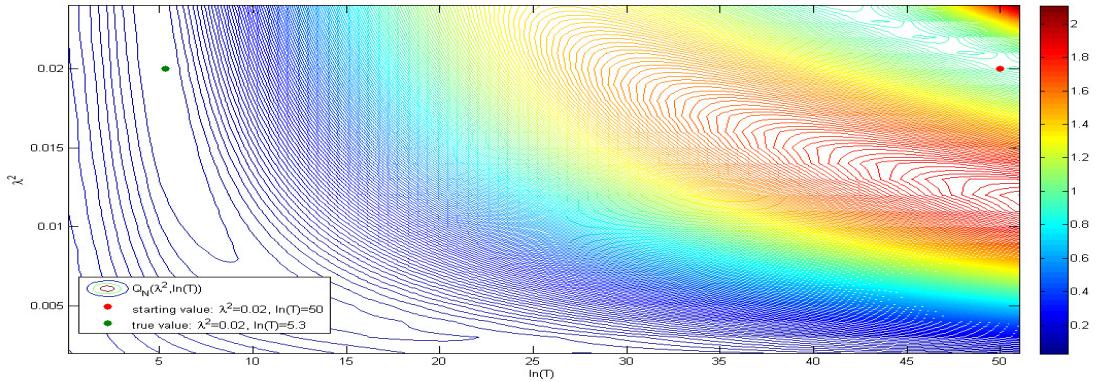


Figure 4: Contour plot of the GMM minimand $Q_N(\lambda^2, \ln(T))$ in the first iteration step, with $W_N = I$ and $\ln(\sigma) = \overline{\ln(\sigma)}$, for a simulated MRW1 series of size $N = 1897$. Q_N is based on the moments function f with 22 lags $1 \leq h \leq 150$ for the ACF $\gamma(h)$.

We want to evaluate the performance of the GMM estimation when starting the minimization procedure in the vicinity of the local minimum. Will the minimization algorithm in this case get stuck in the local minimum or will it still converge to the global solution? For this purpose we estimated the MRW parameters for 1000 simulated MRW1 series, with the starting value $\bar{\theta} = \begin{pmatrix} 0.02 & 50 & \ln(\sigma) \end{pmatrix}$, which will generally lie in the local minimum area. According to our results in Table (6) the GMM estimation is successful.

N	$\widehat{\lambda^2}$		$\widehat{\ln(T)}$		$\widehat{\ln(\sigma)}$		
	eBias	eMSE	eBias	eMSE	eBias	eMSE	
1897	-0.0026	$3 \cdot 10^{-5}$	0.0194	1.7823	-0.0109	0.0047	
3945	-0.0014	10^{-5}	-0.0281	0.6248	-0.0056	0.0024	
8041	-0.0008	$6 \cdot 10^{-6}$	-0.0078	0.2501	-0.0015	0.0012	
16233	-0.0004	$3 \cdot 10^{-6}$	-0.0034	0.1232	-0.0003	0.0006	

Table 6: The eBias and eMSE values for the model configuration MRW1, using the moments function f and the starting value $\bar{\theta}$

By contrast, the GMM estimation based on the moments function f^* in (9) employed by Bacry et al. (2008a, 2008b) is not robust against variations of the starting values. Figure (5) exhibits the minimand function Q_N using f^* , for the same MRW1 series as in the Figures (3) and (4). Its shape is similar to the shape of the minimand in Figure (3), however the minimum close to the true value θ_0 is only local. Table (7) comprises the estimation results with f^* , using the same 1000 simulated MRW1 series as for Table (6) and the starting value $\bar{\theta} = \begin{pmatrix} 0.02 & 50 & \ln(\sigma) \end{pmatrix}$. According to this, the GMM estimation performs well for the parameters λ^2 and $\ln(\sigma)$, yet the procedure fails dramatically to estimate $\ln(T)$.

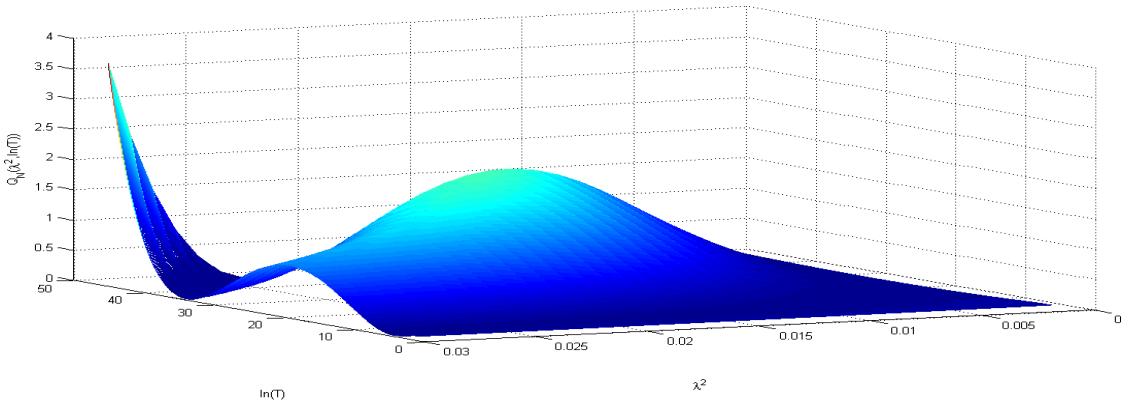


Figure 5: The GMM minimand $Q_N(\lambda^2, \ln(T))$ in the first iteration step, with $W_N = I$ and $\ln(\sigma) = \overline{\ln(\sigma)}$, for a simulated MRW1 series of size $N = 1897$. Q_N is based on the moments function f^* with 22 lags $1 \leq h \leq 150$ for the ACF $\gamma(h)$.

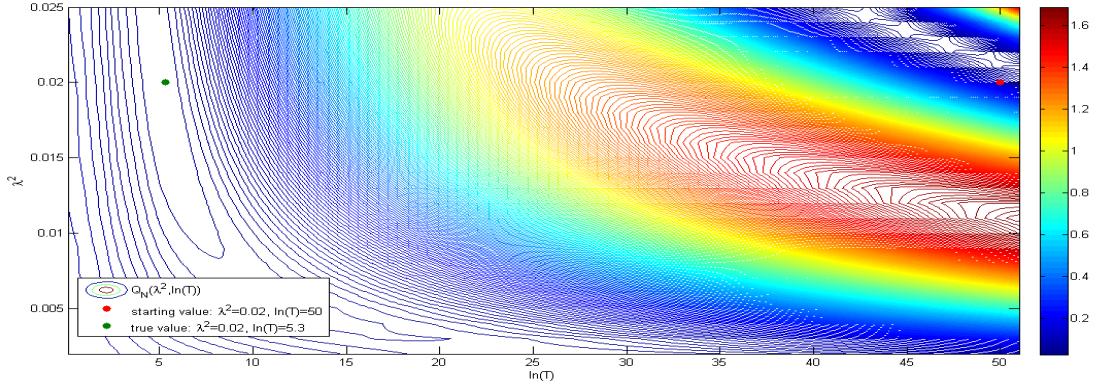


Figure 6: Contour plot of the GMM minimand $Q_N(\lambda^2, \ln(T))$ in the first iteration step, with $W_N = I$ and $\ln(\sigma) = \overline{\ln(\sigma)}$, for a simulated MRW1 series of size $N = 1897$. Q_N is based on the moments function f^* with 22 lags $1 \leq h \leq 150$ for the ACF $\gamma(h)$.

N	$\widehat{\lambda^2}$		$\widehat{\ln(T)}$		$\widehat{\ln(\sigma)}$	
	eBias	eMSE	eBias	eMSE	eBias	eMSE
1897	0.0029	$7 \cdot 10^{-5}$	48.4124	$29 \cdot 10^2$	-0.0110	0.0047
3945	0.0013	$3 \cdot 10^{-5}$	49.2939	$27 \cdot 10^2$	-0.0060	0.0024
8041	0.0008	$2 \cdot 10^{-5}$	49.3211	$26 \cdot 10^2$	-0.0019	0.0012
16233	0.0004	$8 \cdot 10^{-6}$	49.7714	$26 \cdot 10^2$	-0.0005	0.0006

Table 7: The eBias and eMSE values for the model configuration MRW1, using the moments function f^* and the starting value $\bar{\theta}$

5 Empirical study

In Section 3 we introduced a new optimal iterated GMM estimator for the MRW, with an enhanced moments function and an efficient initialization procedure for the variance parameter. This estimator is robust to variations in the starting values for λ^2 and $\ln(T)$, and allows for the estimation of the parameters intermittency coefficient λ^2 and logarithmic standard deviation $\ln(\sigma)$ with good performance in finite samples. Yet the logarithmic decorrelation scale $\ln(T)$ can only be reliably estimated from very large data sets with the sample size N much greater than T ($N \gg T$).

We now present an application of the MRW to the evolution of the German DAX stock market index over the past 50 years using the GMM estimation procedure described in Section 3. The DAX consists of the 30 most actively-traded German stocks. It was launched on July 1, 1988, whereas historical data for the index has

been retrospectively calculated back to 1959. Our data set comprises 11854 daily quotes between October 1, 1959 and April 30, 2009, which were kindly supplied by the German stock exchange *Deutsche Börse AG*.

The MRW models the logarithmic values of the DAX

$$X(t) := \ln(p(t)),$$

with $p(t)$ being the closing value at time t in days. The decorrelation scale T indicates the time span over which the DAX log-returns (logarithmic increments) $\delta_1 X(t) = \ln(p(t)) - \ln(p(t-1))$ are dependent. Accordingly, the log-returns are independent of the log-returns $T+1$ days later and the subsequent log-returns. It is sensible to assume that the log-returns are dependent only over a period of time shorter than 50 years ($N > T$), yet we can not say whether $N \gg T$ is satisfied. On account of this, the estimation of $\ln(T)$ should be handled circumspectly. Table (8) displays the estimation results for the DAX data.

	$\widehat{\lambda^2}$	$\widehat{\ln(T)}$	$\widehat{\ln(\sigma)}$
estimated value	0.0185	8.8941	-4.4786
s.e.	0.0029	1.0192	0.0406
c.i.	[0.0128, 0.0242]	[6.8964, 10.8918]	[-4.5583, -4.3990]

Table 8: Estimation results for the DAX data, via optimal iterated GMM based on the moments function f with 22 ACF lags ranging between 1 and 150 days, using the initialization method for $\ln(\sigma)$ introduced in Section 3.3. The estimation procedure converges invariably to the same estimates irrespective of the starting values for λ^2 and $\ln(T)$. s.e. denotes the standard error and c.i. denotes the 0.95 confidence interval.

Based on the asymptotic normality of the GMM estimators $\widehat{\theta}_{N,i}$ for the individual MRW parameters $\theta_{0,i}$ with $\theta_{0,1} = \lambda^2$, $\theta_{0,2} = \ln(T)$ and $\theta_{0,3} = \ln(\sigma)$, the quantities

$$\frac{\widehat{\theta}_{N,i} - \theta_{0,i}}{s.e.(\widehat{\theta}_{N,i})}$$

with $s.e.(\widehat{\theta}_{N,i})$ the standard error of $\widehat{\theta}_{N,i}$, are asymptotically standard normally distributed respectively. This facilitates the construction of approximate confidence intervals for the MRW parameters in large samples (Hall 2005, pp. 71, 161-164).

To simplify matters we will use in the following the example of $\widehat{\lambda^2}$ with

$$z = \frac{\widehat{\lambda^2} - \lambda^2}{s.e.(\widehat{\lambda^2})} \stackrel{d}{\sim} \mathcal{N}(0, 1).$$

The two-sided confidence interval for λ^2 at level α is then given by

$$\left[\widehat{\lambda^2} - z_{1-\frac{\alpha}{2}} \cdot s.e.(\widehat{\lambda^2}), \widehat{\lambda^2} + z_{1-\frac{\alpha}{2}} \cdot s.e.(\widehat{\lambda^2}) \right]$$

with $z_{1-\frac{\alpha}{2}}$ the $1 - \frac{\alpha}{2}$ quantile of the standard normal distribution.

The asymptotic covariance matrix of the GMM estimator $\widehat{\theta}_N$ can be estimated consistently by means of

$$\widehat{V} = \left(\left(J_N(\widehat{\theta}_N) \right)' \widehat{S}_N^{-1} J_N(\widehat{\theta}_N) \right)^{-1}$$

with \widehat{S}_N the HAC estimator and

$$J_N(\widehat{\theta}_N) = \frac{1}{N} \sum_{k=1}^N \partial f(Z(k), \widehat{\theta}_N) / \partial \theta' .$$

Accordingly, the standard error of $\widehat{\lambda}^2$ is given by

$$s.e.(\widehat{\lambda}^2) = \sqrt{\frac{\widehat{V}_{11}}{N}}$$

with \widehat{V}_{11} the first main diagonal element of \widehat{V} (Hall 2005, pp. 86, 88-89). The standard error and the 95% confidence interval for λ^2 from the DAX data can be found in Table (8) along with the corresponding results for $\ln(T)$ and $\ln(\sigma)$.

The most commonly reported value for λ^2 in financial applications is 0.02 (Bacry et al. 2008b, p. 6). We want to investigate whether the DAX data in this study is characterized by a different λ^2 value. For this purpose we test the hypothesis $H_0 : \lambda^2 = 0.02$ against the alternative $H_1 : \lambda^2 \neq 0.02$ using the Wald test. The 95% confidence interval for λ^2 reported in Table (8) comprises the value 0.02. On account of this, we fail to reject H_0 at the 5% significance level. The DAX data doesn't provide sufficient evidence for concluding that λ^2 is different from 0.02.

6 An efficient algorithm for the computation of the HAC estimator

The HAC estimator \widehat{S}_N is employed to estimate the asymptotic covariance matrix of the GMM estimator (see Section 5). Moreover, it is the basis for the optimal weighting methodology of the GMM estimator (see Section 3.1). Its practical computation proves to be both memory space- and computing time-intensive depending on the employed sample size N , the bandwidth b_N and the number of moment conditions to be denoted further by q . In this section we introduce a new algorithm for the computation of the HAC estimator which is more efficient than the classical HAC estimation procedures developed for the statistical computer programs GAUSS or R.

6.1 The algorithm

The HAC estimator (14) in Section 3.2 can be written compactly as follows

$$\widehat{S}_N = \frac{1}{N} F' T(\omega) F \quad (20)$$

with F given in (17). The expression $T(\omega)$ denotes the symmetric Toeplitz matrix, with the first column ω containing the weights of the empirical autocovariance matrices $\widehat{\Gamma}_i$ in (14)

$$\omega = \begin{pmatrix} 1 & \omega_{1,N} & \omega_{2,N} & \dots & \omega_{N-1,N} \end{pmatrix},$$

$T(\omega) \in \mathbb{R}^{N \times N}$ (Kyriakoulis 2007).

In this paper we employ a memory space-efficient computation of the matrix product $T(\omega)F$ in (20) using the properties of circulant matrices (Van Loan 1992, pp. 208-209). For this purpose we embed the Toeplitz matrix $T(\omega)$ in a symmetric circulant matrix $C(\omega^*) \in \mathbb{R}^{2N \times 2N}$ with the first column ω^*

$$\omega^* = \begin{pmatrix} 1 & \omega_{1,N} & \omega_{2,N} & \dots & \omega_{N-1,N} & 0 & \omega_{N-1,N} & \omega_{N-2,N} & \dots & \omega_{1,N} \end{pmatrix}.$$

We construct the $2N \times q$ matrix F^*

$$F^* = \begin{pmatrix} F \\ 0_{N \times q} \end{pmatrix}$$

by adding a $N \times q$ zero matrix at the end of F . The matrix product $T(\omega)F$ is then given by the upper $N \times q$ block of the $2N \times q$ matrix $C(\omega^*)F^*$. This is to our advantage since the matrix product $C(\omega^*)F^*$ can be computed quickly according to the following proposition (Golub and Van Loan 1996, pp. 201-202; Brockwell and Davis 1991, pp. 133-134):

Proposition 1 *Given a circulant matrix $C(c) \in \mathbb{R}^{n \times n}$ with c being the first column, consider its decomposition*

$$C(c) = V\Lambda V^*$$

with λ_i , $i \in \{1, 2, \dots, n\}$ the eigenvalues of $C(c)$ and v_i the associated eigenvectors, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $V = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$ and V^ the complex conjugate of V . The matrix $C(c)$ then has the following properties:*

- *The eigenvalues λ_i are the discrete Fourier transform (DFT) of the column vector c .*
- *For any vector $x \in \mathbb{R}^n$ the product V^*x is given by the DFT of x .*
- *For any vector $x \in \mathbb{R}^n$ the product Vx is given by the inverse discrete Fourier transform (IDFT) of x .*

Based on these preliminary considerations, we can now introduce the new algorithm for the computation of the HAC estimator. In the following we assume that the weights $\omega_{i,N}$ and the GMM estimator $\widehat{\theta}_N$ are known.

(Sattarhoff)

Step 1 Determine the eigenvalues λ_i , $i \in \{1, 2, \dots, 2N\}$ of the matrix $C(\omega^*)$ by means of the DFT of the vector ω^*

$$\omega^* = \begin{pmatrix} 1 & \omega_{1,N} & \omega_{2,N} & \dots & \omega_{N-1,N} & 0 & \omega_{N-1,N} & \omega_{N-2,N} & \dots & \omega_{1,N} \end{pmatrix}.$$

Step 2 Construct the $2N \times q$ matrix F^*

$$F^* = \begin{pmatrix} F \\ 0_{N \times q} \end{pmatrix}$$

with F defined in (17) and q the number of moment conditions.

Step 3 For j from 1 to q , compute the columns $C(\omega^*) F_j^* = V \Lambda V^* F_j^*$ of the matrix $C(\omega^*) F^*$, with F_j^* the j th column in F^* , in 3 steps respectively:

1. Determine $V^* F_j^*$ given by the DFT of F_j^* .
2. Multiply the rows $i \in \{1, 2, \dots, 2N\}$ in $V^* F_j^*$ with the eigenvalues λ_i respectively, in order to construct $\Lambda V^* F_j^*$.
3. Determine $C(\omega^*) F_j^*$ given by the IDFT of $\Lambda V^* F_j^*$.

Step 4 Select the matrix $Z = T(\omega) F$ given by the upper $N \times q$ block of $C(\omega^*) F^*$.

Step 5 Determine $\widehat{S}_N = \frac{1}{N} F' Z$.

6.2 Computational gains

In this section, we report the gains in computing times that are achieved by the new algorithm for the computation of the HAC estimator introduced above, in comparison with three alternative algorithms. For each alternative algorithm we assume that the weights $\omega_{i,N}$, the bandwidth b_N and the GMM estimator $\widehat{\theta}_N$ are known. Furthermore, q denotes the number of moment conditions, whereas the matrix F is given in (17).

The first algorithm that we consider was developed by Roncalli in 1996 and can be found in the time series library *TSM* (Time Series and Wavelets for Finance) for the statistical software GAUSS. Here the computation of the HAC estimator \widehat{S}_N is implemented by means of a *for*-loop according to expression (18) combined with an ingenious matrix product, which enables the fast computation of the autocovariance matrices $\widehat{\Gamma}_i$:

$$\widehat{\Gamma}_i = \frac{1}{N} F' \begin{pmatrix} 0_{i \times q} \\ F(1 : (N - i)) \end{pmatrix}. \quad (21)$$

(Roncalli)

Step 1 Determine $\widehat{\Gamma}_0 = \frac{1}{N}F'F$ and set $L = \widehat{\Gamma}_0$.

Step 2 For i from 1 to b_N determine $\widehat{\Gamma}_i$ according to (21) and $L = L + \omega_{i,N} (\widehat{\Gamma}_i + \widehat{\Gamma}'_i)$.

Step 3 Determine $\widehat{S}_N = L$.

The second algorithm by Zeileis (2004) is part of the *gmm* package for the statistical software R. This algorithm is similar to Roncalli's algorithm except for the calculation procedure for $\widehat{\Gamma}_i$ in Step 2. For the latter Zeileis employs expression (16). His procedure is less efficient than Roncalli's as it requires the sequential updating of two matrices instead of one.

(Zeileis)

Step 1 Determine $\widehat{\Gamma}_0 = \frac{1}{N}F'F$ and set $L = \widehat{\Gamma}_0$.

Step 2 For i from 1 to b_N determine $\widehat{\Gamma}_i$ according to (16) and $L = L + \omega_{i,N} (\widehat{\Gamma}_i + \widehat{\Gamma}'_i)$.

Step 3 Determine $\widehat{S}_N = L$.

Finally the algorithm by Kyriakoulis (2007) for MATLAB excels in terms of its elegance and simplicity. This algorithm avoids the resource-intensive recursive summations employed above by means of expression (20) and is the basis for the new algorithm introduced in the previous subsection. It consists of only two steps:

(Kyriakoulis)

Step 1 Construct the symmetric Toeplitz matrix $T(\omega)$ with the first column

$$\omega = (1 \ \omega_{1,N} \ \omega_{2,N} \ \dots \ \omega_{N-1,N}).$$

Step 2 Determine $\widehat{S}_N = \frac{1}{N}F'T(\omega)F$.

A big drawback of this algorithm is the memory space-inefficient handling of the $N \times N$ matrix $T(\omega)$. On account of this the program runs out of memory and fails to compute \widehat{S}_N for series longer than $N = 5000$ data points. This problem could be solved within the scope of our algorithm, which doesn't employ the matrix $T(\omega)$ explicitly.

In the following we investigate the computing performance of the four algorithms for the computation of the HAC estimator developed by Sattarhoff, Roncalli, Zeileis and Kyriakoulis respectively. For this purpose we employ an unique, randomly generated matrix F . For comparability reasons, all algorithms were programmed and

run under MATLAB.¹² The computations were performed on a 1.66 GHz Intel notebook computer with 1GB RAM. In the following discussion we ignore the algorithm by Kyriakoulis due to its poor performance.

Table (9) displays the computing times in seconds for various sample sizes N ranging from 5,000 to 1,000,000 data points. Throughout, the bandwidth and the number of moment conditions were held constant. One can see that the algorithm proposed in this paper performs best for all choices of N . This algorithm is at least twice as fast as the procedure by Roncalli, and at least three times as fast as the one by Zeileis.

N	5000	20000	100000	500000	1000000
Sattarhoff	0.0707	0.2271	1.5011	8.2581	41.5323
Roncalli	0.1550	0.6973	3.5636	17.9518	1040.5110
Zeileis	0.2765	1.0094	5.0762	24.7022	114.9194
Kyriakoulis	2.8100	-	-	-	-

Table 9: Computing times in seconds using $b_N = 30$ and $q = 24$. MATLAB runs out of memory in the case of the algorithm by Kyriakoulis with $N > 5000$.

Additionally, our algorithm has the advantage that the bandwidth value has no impact on its performance. By contrast, the number of iterations in Step 2 and accordingly computing time both increase with the bandwidth in the case of the algorithms by Roncalli and Zeileis. This can be seen in Table (10), which investigates the computing performance for different values of b_N and q , and a constant N .

q	10			24		
	b_N	30	60	100	30	60
Sattarhoff	3.5323	3.5658	3.7565	8.2581	8.1885	8.2767
Roncalli	9.7528	19.7211	31.7664	17.9518	35.2764	58.3308
Zeileis	11.2895	22.3449	38.1082	24.7022	48.9346	81.2929

Table 10: Computing times in seconds for $N = 500000$

The implications of these findings for the computation of the optimal iterated GMM estimator are considerable. The asymptotic covariance matrix S has to be estimated at each step of the iterated estimation procedure. This is necessary for the determination of the optimal weighting matrix in the following estimation step. Consequently, if we consider 50 estimation steps, our algorithm will save in the case of the parameter configurations in Table (10) between 5 and 42 minutes compared to Roncalli, and between 6 and 61 minutes compared to Zeileis.

¹²We used the programs created by Kyriakoulis (2007) and Borchers (2005).

7 Conclusion

Financial time series reveal a very rich and non-trivial statistical structure. Whereas log-return series are characterized by zero-autocorrelations, absolute and squared log-returns exhibit significant positive autocorrelations over long lags (Ding et al. 1993). Correspondingly, financial markets experience extended turbulent periods with large price changes followed by extended silent phases of small price activity. This phenomenon is known as volatility clustering. Another interesting property of financial time series is the variation of the log-return distribution as a function of the employed sampling interval, ranging from normal distributions for rather large sampling intervals (e.g. weekly or monthly log-returns) to fat tail distributions at fine scales (e.g. hourly and daily log-returns). From a statistical point of view, this property can be described within the more general framework of multifractality.¹³

The MRW is a recent volatility model which can capture the stylized facts of the financial time series depicted above in a parsimonious way, using only three parameters: the intermittency coefficient λ^2 , the logarithmic decorrelation scale $\ln(T)$ and the logarithmic standard deviation $\ln(\sigma)$. The model is employed by Bacry et al. (2008a, 2008b) to forecast volatility and Value at Risk with better results than for GARCH models. The estimation of the model parameters plays a crucial role for the subsequent prediction of risk.

In this paper we have introduced a new optimal iterated GMM estimator for the MRW. In addition to deploying the moment conditions of Bacry et al. (2008b), our estimation procedure also acquires valuable information from the mean condition, which is exploited explicitly. Consequently, we report better estimation results in finite samples within the scope of a Monte Carlo simulation study. The estimated intermittency coefficients are normally distributed. This allows us to construct statistical hypothesis tests about λ^2 using the Wald statistic. Based on the latter, we fail to reject within our empirical study on the DAX the hypothesis that $\lambda^2 = 0.02$ at the 5% significance level.

The performance of the GMM estimator using the moments configuration by Bacry et al. (2008b) depends on the parameter starting values when the minimization procedure employs an iterative search method, and accordingly on the parameter space in the case of grid search minimization. By contrast, the GMM estimator in this paper is robust to these factors. Moreover, we propose an efficient initialization method for the variance parameter. Throughout our study, the computational burden is considerably reduced by means of a computing-time-efficient algorithm for HAC covariance matrix estimation.

These results encourage further research in this field. In particular, the idea of constructing a simple test for multifractality using the Wald statistic is very appealing. Also, the prediction of volatility based on the optimal iterated GMM estimator introduced in this paper will be addressed in a forthcoming study. Based

¹³See Krämer (2002) for an overview of the stylized facts of financial time series.

on this sound estimation performance, we are confident that we will achieve further improvements in forecasting accuracy in the future.

Appendix

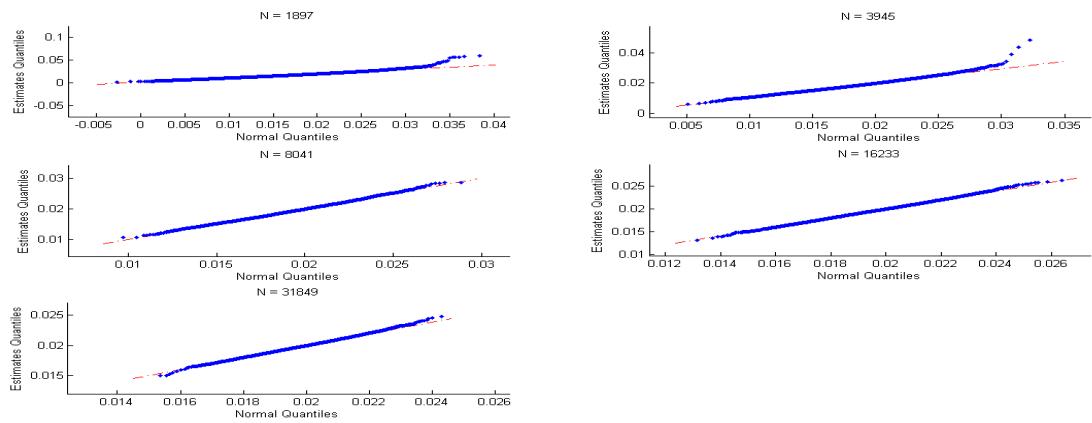


Figure 7: QQ plots of the estimates of λ^2 versus a normal distribution, for the model configuration MRW1.

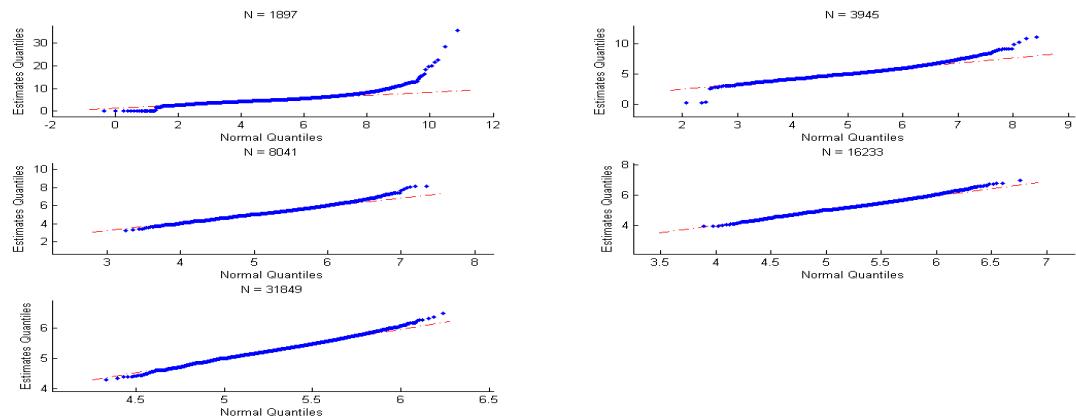


Figure 8: QQ plots of the estimates of $\ln(T)$ versus a normal distribution, for the model configuration MRW1.

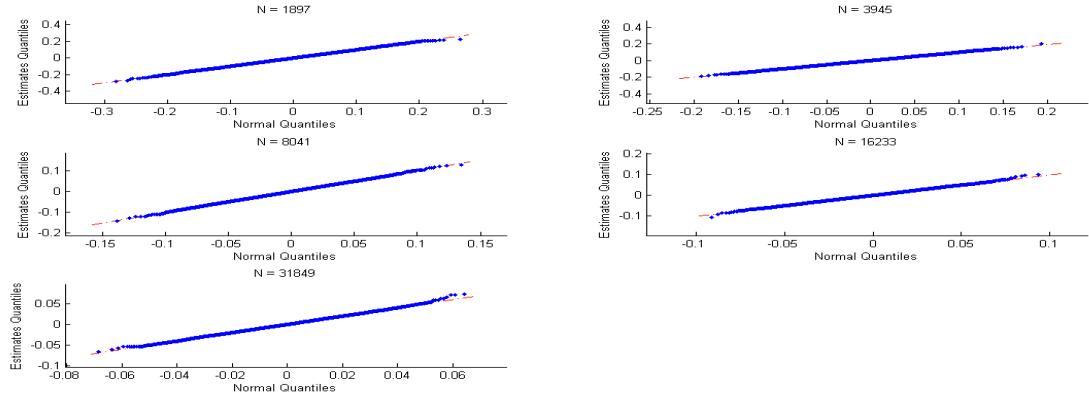


Figure 9: QQ plots of the estimates of $\ln(\sigma)$ versus a normal distribution, for the model configuration MRW1.

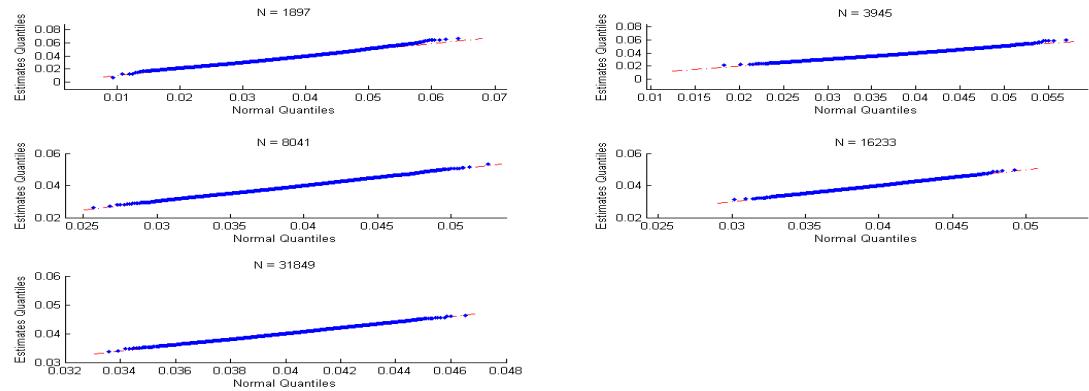


Figure 10: QQ plots of the estimates of λ^2 versus a normal distribution, for the model configuration MRW2.

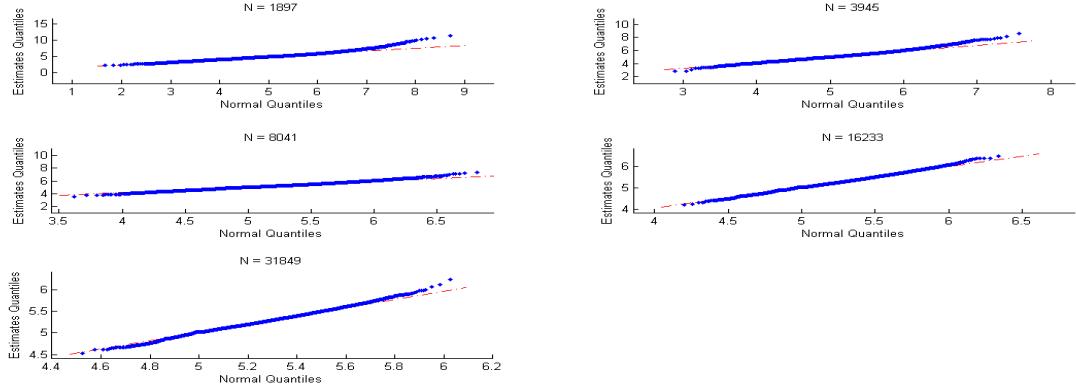


Figure 11: QQ plots of the estimates of $\ln(T)$ versus a normal distribution, for the model configuration MRW2.

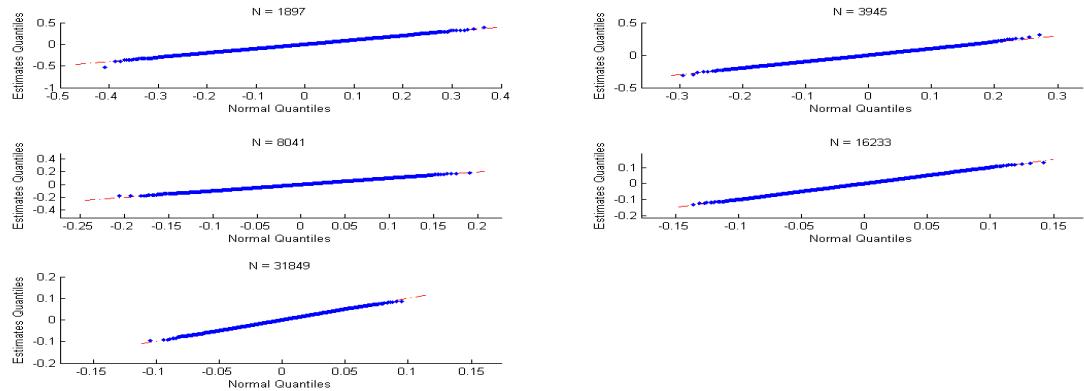


Figure 12: QQ plots of the estimates of $\ln(\sigma)$ versus a normal distribution, for the model configuration MRW2.

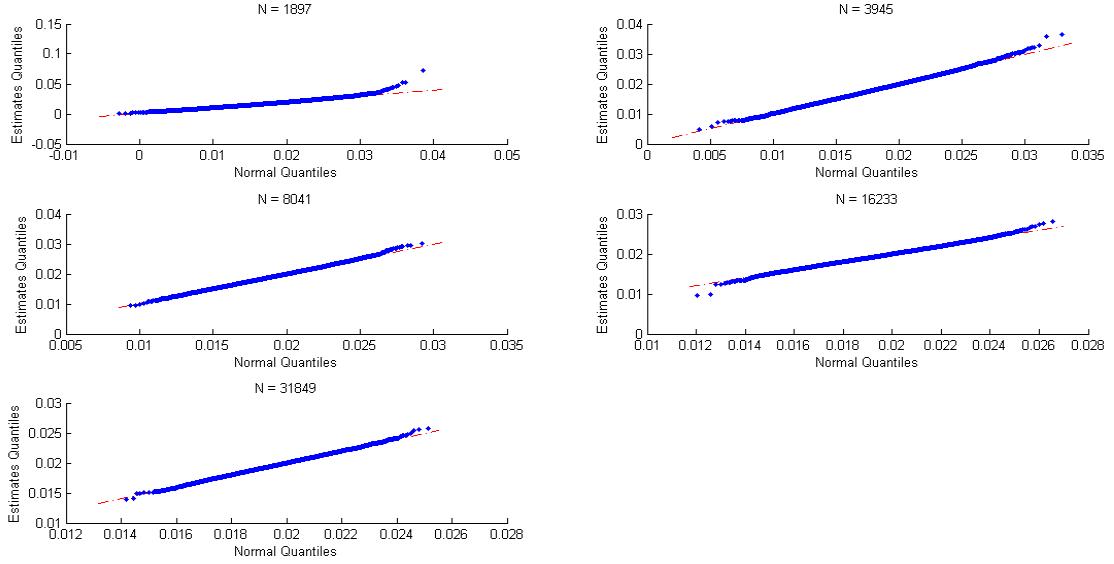


Figure 13: QQ plots of the estimates of λ^2 versus a normal distribution, for the model configuration MRW3.

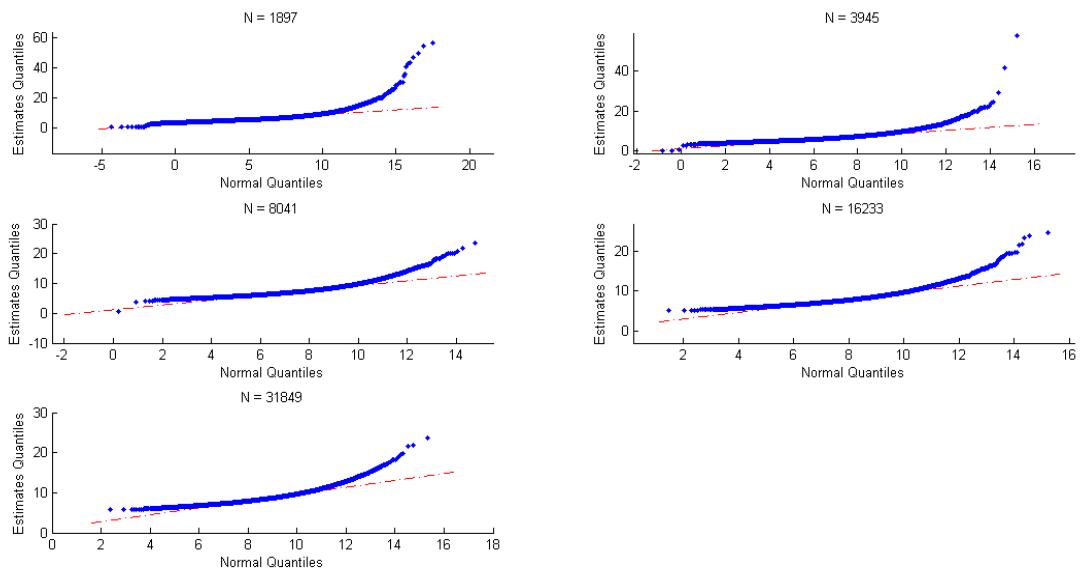


Figure 14: QQ plots of the estimates of $\ln(T)$ versus a normal distribution, for the model configuration MRW3.

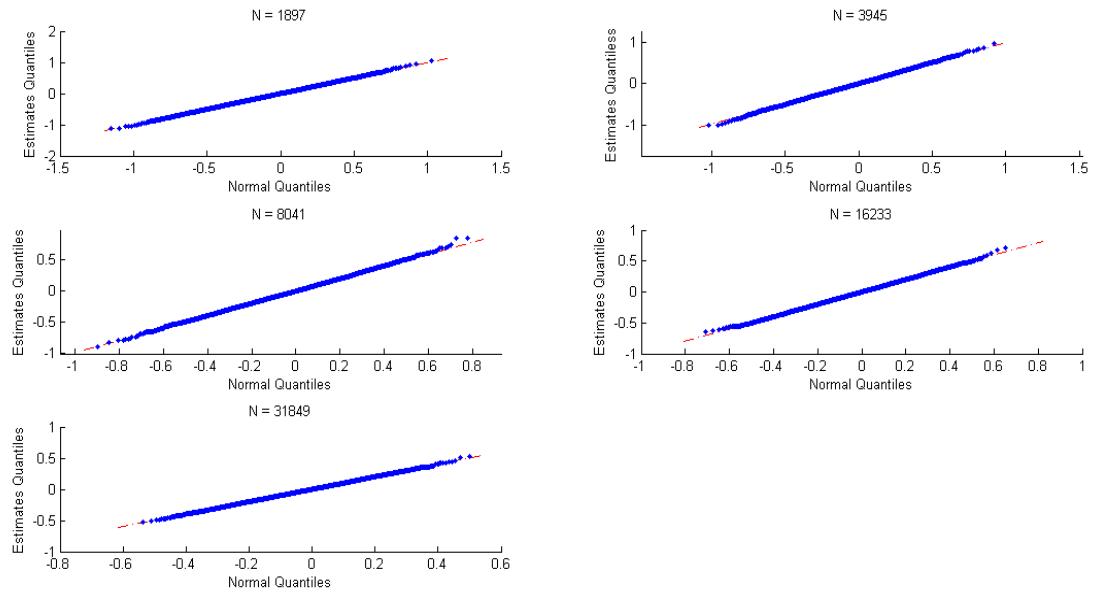


Figure 15: QQ plots of the estimates of $\ln(\sigma)$ versus a normal distribution, for the model configuration MRW3.

References

- Bacry, E., Kozhemyak, A. and Muzy, J. F. 2008a, *Continuous cascade models for asset returns*, Journal of Economic Dynamics and Control, 32(1), 156-199.
- Bacry, E., Kozhemyak, A. and Muzy, J. F. 2008b, *Log-Normal continuous cascades: aggregation properties and estimation. Application to financial time-series*, Unpublished results. Retrieved December 14, 2009 from
<http://www.cmap.polytechnique.fr/~bacry/ftpPapers1.php?paper=qf08.pdf>
- Bacry, E. and Muzy, J. F. 2003, *Log-infinitely divisible multifractal processes*, Communications in Mathematical Physics 236, 449-475.
- Bacry, E., Muzy, J. F. and Delour, J. 2001, *Multifractal Random Walks*, Physical Review E 64, 026103.
- Borchers, B. 2005, *Toeplitzmult*, computer program. Retrieved August 1, 2009 from
<http://www.mathworks.com/matlabcentral/fileexchange/8548>
- Brockwell, P. J. and Davis, R. A. 1991, *Time series: theory and methods*, second ed. Springer, New York.
- Calvet, L. and Fisher, A. 2001, *Forecasting multifractal volatility*, Journal of Econometrics, 105, 27-58.
- Calvet, L. and Fisher, A. 2002, *Multifractality in asset returns: theory and evidence*, The Review of Economics and Statistics, 84, 381-406.
- Calvet, L. and Fisher, A. 2004, *How to forecast long-run volatility: regime-switching and the estimation of multifractal processes*, Journal of Financial Econometrics, 2, 49-83.
- Calvet, L., Fisher, A. and Mandelbrot, B. 1997, *Large deviations and the distribution of price changes*, Cowles Foundation Discussion Paper No. 1165.
- Davidson, R. and MacKinnon, J. G. 1993, *Estimation and Inference in Econometrics*, Oxford University Press, New York.
- Ding, Z., Granger, C. W. J. and Engle, R. F. 1993, *A long memory property of stock market returns and a new model*, Journal of Empirical Finance, 1, 83-106.
- Fisher, A., Calvet, L. and Mandelbrot, B. 1997, *Multifractality of Deutschemark / US Dollar exchange rates*, Cowles Foundation Discussion Paper No. 1166.
- Geiger, C. and Kanzow, C. 1999, *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*, Springer, Berlin.
- Golub, G. H. and Van Loan, C. F. 1996, *Matrix computations*, third ed. Johns Hopkins University Press, Baltimore.

- Greene, W. H. 2003, *Econometric analysis*, fifth ed. Pearson Education, New Jersey.
- Hall, A. R. 2005, *Generalized Method of Moments*, Oxford University Press, New York.
- Hamilton, J. D. 1994, *Time series analysis*, Princeton University Press, Princeton.
- Krämer, W. 2002, *Statistische Besonderheiten von Finanzzeitreihen*, Jahrbücher für Nationalökonomie und Statistik, 222(2), 210-229.
- Kyriakoulis, K. 2007, *GMM toolbox for MATLAB*, computer program. Retrieved August 1, 2009 from <http://www.kostaskyriakoulis.com>
- Lux, T. 2008, *The Markov-Switching Multifractal model of asset returns: GMM estimation and linear forecasting of volatility*, Journal of Business and Economic Statistics, 26, 194–210.
- Mandelbrot, B., Fisher, A. and Calvet, L. 1997, *A Multifractal Model of Asset Returns*, Cowles Foundation Discussion Paper No. 1164.
- Muzy, J. F. and Bacry, E. 2002, *Multifractal stationary random measures and multifractal random walks with log-infinitely divisible scaling laws*, Physical Review E 66, 056121.
- Newey, W. K. and West, K. D. 1987, *A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix*, Econometrica, 55, 703–708.
- The MathWorks 2009a, *MATLAB* (R2009a), computer program, The MathWorks, Natick, Massachusetts.
- The MathWorks 2009b, *Optimization toolboxTM 4. User's guide*, The MathWorks, Natick, Massachusetts. Retrieved February 24, 2010 from
http://www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf
- Van Loan, C. F. 1992, *Computational frameworks for the fast Fourier transform*, SIAM Press, Philadelphia.
- Zeileis, A. 2004, *Econometric computing with HC and HAC covariance matrix estimators*, Journal of Statistical Software, 11(10), 1–17.