Non-stationarities in stock returns

Cătălin Stărică
Department of Mathematical Statistics,
Chalmers University of Technology,
S–412 96 Gothenburg, Sweden [1]

Clive Granger
Department of Economics,
University of California, San Diego,
9500 Gilman Drive, La Jolla, CA 92093, USA [2]

Draft October 14, 2001

Abstract

Modeling financial returns on longer time intervals under the assumption of stationarity is, at least intuitively, given the pace of change in world's economy, a choice hard to defend. Relinquishing the global stationarity hypothesis, this paper conducts a data analysis focused on the size of the returns, i.e. the absolute values of returns, under the assumptions that, at least locally, the $S\&P500$ daily return series can be modeled by stationary processes. The challenging task when working under the assumption of local stationarity is to define the intervals on which stationary processes provide a good approximation. This task is accomplished by using a goodness of fit test based on the integrated periodogram (Picard ([21]), Klüppelberg and Mikosch ([15])). The conclusion of the paper is that almost all the dynamics of return time series seem to be concentrated in the shifts of the variance. More concretely, the $S\&P500$ absolute returns, $|r_t|$ can be modeled as

$$|r_t| = h(t)exp(\epsilon_t), \quad t = 0, 1, \ldots$$

where (ϵ_t) is white noise, $E\epsilon = 0$, $E\epsilon^2 = \sigma^2$ and $h(t)$ a function of t which can be well approximated by a step function, yielding a model with piecewise constant variance.

[1] starica@math.chalmers.se
[2] cgranger@weber.ucsd.edu

1 Introduction

Standard time series asset price models assume time homogeneous unconditional dynamics of returns, even when long periods are considered. This class includes, among others most of the autoregressive conditionally heteroschedastic (ARCH) and stochastic volatility models which describe asset returns as stationary non-linear processes. While these models are designed to capture a time-varying conditional second moment, they fundamentally assume stationarity, in particular constant *unconditional* variance, even when decades of economic activity are modeled. The consistency of an econometric model specification over time, however, is questionable. Given the pace at which new technological tools and financial instruments have been introduced on the financial markets, the case for lack of stationarity seems quite strong. The recent unmatched and unexpectedly long period of economic expansion only highlights the need for reassessing our answer to the question: Does history really repeats itself? The standard stationary time series models answers: Always.

Rethinking our modeling approaches is even more actual in the light of an increasing amount of scientific evidence that questions the relevance of the stationarity assumption. The most important type of non-stationarity encountered in the return time series seems to affect the unconditional variance of the returns. Modeling stock returns as a non-stationary process with discrete shifts in the unconditional variance can be traced back to Hsu, Miller and Wichern (1974). More recently a number of studies have shown that some of the statistical features of return time series that puzzled researchers through their omnipresence (the so called "stylized facts"), like the ARCH effects, the slowly decaying sample ACF for absolute returns or the IGARCH effect receive an alternative simple explanation under the hypothesis of non-stationary changes in the unconditional variance.

ARCH-type models that have become the dominant time series models for stock returns postulates the existence of a constant pattern of changing conditional variance, the so called ARCH effects. Simonato ([23]) and Cai ([4]) among others brought empirical and theoretical evidence that presence of level shifts in the unconditional variance can produce spurious ARCH effects. Statistical estimation of this class of models on data produces extremely persistent processes of conditional variance. Diebold ([6]), Lamoureux and Lastrapes ([17]), Mikosch and Stărică([19]) among others suggested that shifts in the unconditional variance could explain this common finding of persistence in the conditional variance. Lobato and Savin ([16]) and Mikosch and Stărică([19]) argue also that the slowly decaying sample ACF for absolute returns could be caused by level changes in the mean of the absolute returns series.

Consequently, classes of models that allow for patterns of changing unconditional variance have recently been proposed (Hamilton and Susmel ([11]), Cai ([4])). In these models the unconditional variance changes according to the states of a Markov chain. However, the assumption of global stationarity is preserved. These models are hence based on the hypothesis that the pattern of change we observed in the 70's was similar with the one in the 80's and 90's and it will remain unchanged in the future. This is another hypothesis we find difficult to defend, particularly when modeling such a dynamic environment as the financial markets.

In our view a more realistic approach to modeling financial returns is based on the hypothesis of continuous change. Under the assumption that the returns follow a non-stationary process, an interesting class of models is those with time depending parameters. The price one pays for relinquishing the hypothesis of global stationarity is a more complicated approach to the statistical estimation (of the changing parameters) of these models (see Dahlhaus ([5])). One possible approach is to approximate locally the non-stationary process by stationary models. One is then interested in identifying intervals of homogeneity, i.e. intervals where a certain estimated stationary model describes well the reality of the data. On an interval of homogeneity the parameters of the return process do not vary much relative to the estimation error of the parameters of the stationary model used as an approximation on that particular time interval (see Härdle et al.[12]).

For example, if one thinks of the sequence of returns as of an independent sequence of random variables with changing variance, an interval of homogeneity is a period of time where one has reasons to believe that the variance was almost constant (more precisely, that the change in variance cannot be distinguished from estimation error). On the intervals of homogeneity, one approximates the changing variance of returns with a constant. Hence, in the end, the changing pattern of variance will be approximated by a step function, yielding a model with piecewise constant variance.

In this paper we assume that the return generating process is locally stationary and we use ARMA processes with changing coefficients to approximate it locally. We apply a statistical methodology based on a goodness of fit test in the spectral domain (similar to the ones proposed in Picard ([21]) and Klüppelberg and Mikosch ([15])) to identify the homogeneity periods. Our approach leads to modeling the returns as a sequence of uncorrelated (possibly independent) variables with a piecewise constant variance function. We show that even our rough approximation of the variance dynamics by a step function suffices to explains most of the dependency structure present in the sample ACF of long absolute return series. Our results seem to point out that most of the dynamics of stock returns can be explained through a relatively simple pattern of changing unconditional variance.

A commonly hold belief in the econometric community is that taking the slow decay of the sample ACF at face value (even though it might be caused by shifts in the unconditional variance) is a meaningful way of making use of the past in forecasting the future. In other words, estimating long memory *stationary* models (based on the slow decay of the sample ACF) and using them in forecasting exploits in a meaningful way the patterns of change observed in the past. Faced with a modeling choice between a stationary long-memory model and a non-stationary model (informed by the paradigm of changing unconditional variance described above), we compare in the second half of the paper the forecasting performance of the two alternatives. Our results argue for the superiority of the non-stationary modeling approach and support the hypothesis that the non-stationarity of the unconditional variance is the main source of long memory in absolute stock returns.

2 Delimitation of intervals of homogeneity using the integrated periodogram

In this section we describe a statistical methodology based on a goodness of fit test in the spectral domain (see Picard ([21]) and Klüppelberg and Mikosch ([15])) for identifying the homogeneity intervals in series of stock returns.

Our analysis is based on the spectral properties of the underlying time series. Recall that the *periodogram*

$$I_{n,X}(\lambda) = \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \mathrm{e}^{-i\lambda t} X_t \right|^2 , \quad \lambda \in [0,\pi] ,$$

is the natural (method of moment) estimator of the spectral density f_X of a stationary sequence (X_t); see Brockwell and Davis [3] or Priestley [22]. Under general conditions, the *integrated periodogram* or *empirical spectral distribution function*

$$\frac{1}{2\pi} J_{n,X}(\lambda) = \frac{1}{2\pi} \int_0^\lambda I_{n,X}(x)\, dx , \quad \lambda \in [0,\pi] , \tag{2.1}$$

is a consistent estimator of the *spectral distribution function* given by

$$F_X(\lambda) = \int_0^\lambda f_X(x)\, dx , \quad \lambda \in [0,\pi] ,$$

provided the density f_X is well defined. Given a finite 4th moment for X and supposing that X_t is a linear process, the limit of $\sqrt{n}(J_{n,X} - F_X)$ in $C\,[0,\pi]$, the space of continuous functions on $[0,\pi]$ endowed with the uniform topology, is an unfamiliar Gaussian process, i.e. its covariance structure depends on the spectral density f_X; see for example Anderson [1] or Mikosch [18]. Since one wants to use the distributional limit of $\sqrt{n}(J_{n,X} - F_X)$ for the construction of goodness of fit tests of the spectral distribution function, as proposed by Grenander and Rosenblatt [10], one needs to modify the integrated periodogram to get a more familiar Gaussian process, if possible a bridge type process. Bartlett [2] (cf. Priestley [22]) had the idea to use a weighted form of the integrated periodogram:

$$J_{n,X,f_X}(x,\lambda) = \int_0^\lambda \frac{I_{n,[nx],X}(y)}{f_X(y)}\, dy , \quad x \in [0,1] , \quad \lambda \in [0,\pi] ,$$

where

$$I_{n,k,X}(\lambda) = \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{k} \mathrm{e}^{-i\lambda t} X_t \right|^2 , \quad k = 0,\ldots,n , \quad \lambda \in [0,\pi] .$$

Dividing the periodogram by the spectral density makes the limit process independent of the spectral density. The process $J_{n,X}(x, \lambda)$, properly centered and scaled, converges in distribution in the Skorokhod space $D([0,1] \times [0,\pi])$ to a two-parameter Gaussian process which, for fixed λ, is a Brownian motion and, for fixed x, a Brownian bridge. Such a process is known as *Kiefer–Müller process*; see Shorack and Wellner [24].

In the sequel the integrated periodogram will be used to determine the intervals of homogeneity by monitoring the changes in the spectral distribution function of a time series. The method we discuss is related to the one proposed in Picard [21] for detecting changes in the spectral distribution function of a time seriesand further developed for various linear processes under mild assumptions on the moments of X and the coefficients of the process by Giraitis and Leipus [7] and Klüppelberg and Mikosch [15].

Functionals of $J_{n,X}(x, \lambda)$ of Kolmogorov–Smirnov type will be used to define intervals of homogeneity as follows. Assume we know that the subsample $X_m, X_{m+1}, \ldots, X_{m_1}$ is well described by a linear parametric model with mean μ, variance of the noise σ^2 and spectral density f, i.e. the interval of homogeneity contains at least the observations between the m-th one and the m_1-th one, and we try to decide whether the following p observations, $X_{m_1+1}, \ldots, X_{m_1+p}$ do also belong to the interval or not. A functional of $J_{n,X}(x, \lambda)$ based on the subsample $X_{m_1-l}, \ldots, X_{m_1+p}$ $(m_1 - l > m)$ is calculated and compared with the asymptotic distribution of the functional under the null hypothesis that the subsample $X_{m_1-l}, \ldots, X_{m_1+p}$ is part of a stationary sequence, with the prescribed mean μ, variance σ^2 and spectrum f . If the value of the statistic falls within the asymptotic confidence interval, the homogeneity interval is extended to include observations, $X_{m_1+1}, \ldots, X_{m_1+p}$. Otherwise a new homogeneity interval is started with the observations $X_{m_1+1}, \ldots, X_{m_1+p}$.

Before recalling the main result that rests the theoretical foundation of our procedure we need to introduce a few notations. We consider the linear process

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \quad t \in Z \tag{2.2}$$

where the innovations (Z_t) are a sequence of iid random variables with mean 0 and finite variance σ^2. The assumption

$$\sum_{j=-\infty}^{\infty} |\psi_j| j < \infty$$

ensures that X_t is properly defined as an a.s. absolutely converging series. The function

$$|\psi(e^{-i\lambda})|^2 = |\sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j}|^2, \quad -\pi \le \lambda \le \pi,$$

is called the power transfer function of the linear filter (ψ_j) and

$$f(\lambda) = |\psi(e^{-i\lambda})|^2 \sigma^2 / (2\pi) \tag{2.3}$$

is the spectral density of the linear process (X_t). The following result describes the asymptotic behavior of the integrated periodogram for linear processea and provides the theoretical basis for the analysis to follow.

Theorem 2.1 (Klüppelberg and Mikosch ([15])) *Assume that $EZ = 0$ and that $EZ^4 < \infty$ and denote $var(Z) = \sigma^2$.*

a. If X_t is the linear processes (2.2) with $EZ^4 < \infty$, the following holds

$$\sqrt{n}\left(\int_{-\pi}^{\lambda}\left(\frac{I_{n,[nx],X}(y)}{\psi(e^{-iy})} - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{I_{n,[nx],X}(z)}{\psi(e^{-iz})}dz\right)dy\right) \tag{2.4}$$

$$\xrightarrow{d} 2\sigma^2(K(x,\lambda))_{x\in[0,1],\lambda\in[0,\pi]} = 2\sigma^2\left(\sum_{h=1}^{\infty} W_h(x)\frac{\sin(\lambda h)}{h}\right)_{x\in[0,1],\lambda\in[0,\pi]}, \tag{2.5}$$

in $D([0,1]\times[0,\pi])$ where $(W_h(\cdot))_{h=1,\dots}$ is a sequence of iid standard Brownian motions on $[0,1]$. The infinite series on the right-hand side converges with probability 1 and represents a Kiefer–Müller process, i.e., a two-parameter Gaussian field with covariance structure

$$E\left(K\left(x_1,\lambda_1\right)\,K\left(x_2,\lambda_2\right)\right) = \min\left(x_1,x_2\right)\sum_{t=1}^{\infty}\frac{\sin\left(\lambda_1 t\right)\sin\left(\lambda_2 t\right)}{t^2}$$

$$= 2^{-1}\pi^2\min\left(x_1,x_2\right)\left(\min\left(\frac{\lambda_1}{\pi},\frac{\lambda_2}{\pi}\right) - \frac{\lambda_1}{\pi}\frac{\lambda_2}{\pi}\right). \tag{2.6}$$

b. In the case $(X_t) = (Z_t)$ the following holds

$$\sqrt{n}\int_{-\pi}^{\lambda}\left(I_{n,[nx],X}(y) - n^{-1}\sum_{t=1}^{[nx]}Z_t^2\right)dy \tag{2.7}$$

$$\xrightarrow{d} 2\sigma^2(K(x,\lambda))_{x\in[0,1],\lambda\in[0,\pi]}, \tag{2.8}$$

where $K(x,\lambda))_{x\in[0,1],\lambda\in[0,\pi]}$ is the Kiefer–Müller process (2.6).

3 Data analysis

The data we use for our analysis are the daily returns $r_t := \log P_t - \log P_{t-1}$, where P_t is the daily closing level of *S&P500* between the 3rd of January 1928 until the 25th of May 2000. From a traditional time series point of view, the information contained in the time series of daily returns can be split in two components: the sign of the return and its size. Empirical evidence (Granger et al. ([9]) shows that the sign of daily returns is not predictable. In what follows we concentrate on studying the time series of absolute

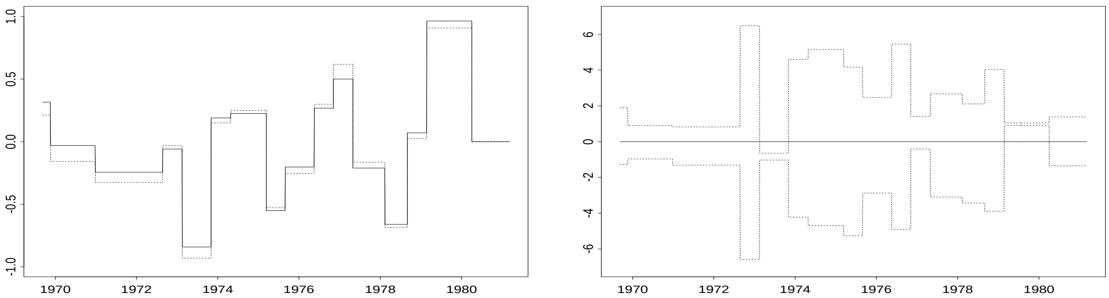

Figure 3.1 Left: The AR (*continuous line*) and the MA (*dotted line*) estimated coefficients for some intervals of homogeneity of the log-absolute returns of the *S&P500*. Right: The 95% confidence intervals for the AR coefficient displayed on the left. Zero is most of the time covered by the interval.

returns. More precisely, due to the presence of heavy tails in the absolute returns, we chose to analyze the series of the logarithm of the absolute values of daily returns.

We assume that the *logarithm of the absolute values of daily returns*, $X_t := \log(|r_t|)$ follows a locally stationary process. For simplicity one can think of an ARMA process with time varying parameters

$$\phi(t, B)(X_t - \mu(t)) = \theta(t, B)Z_t, \quad Z_t = \sigma(t)\epsilon_t, \tag{3.1}$$

where ϕ, θ are polynomials of degree p, q respectively, B is the back shift operator, ϵ_t are iid with $E\epsilon_t = 0$, $E\epsilon_t^2$=1. In words, the model displays local linear dependence built with innovations with changing mean and variance. We intend to approximate the functions $\phi(t, B)$, $\theta(t, B)$, $\mu(t)$, $\sigma(t)$ with step functions on appropriately defined homogeneity intervals. In other words, we will approximate the process $\log(|r_t|)$ by stationary ARMA models whose parameters change every once in a while. We start hence with a very general model and we will let the data render it more specific by means of testing and deciding on the relevance of different parameters.

In order to construct this approximation we proceed as follows. Assume we know that the subsample $X_m, X_{m+1}, \ldots, X_{m_1}$ is well described by a linear parametric model with mean μ, power transfer function ψ and variance of the innovations σ^2 (these values were estimated on the observations in the beginning of the homogeneity interval), i.e. the interval of homogeneity contains at least the observations between the m-th one and the m_1-th one. Next we try to decide whether the following 20 observations (or the next month), $X_{m_1+1}, \ldots, X_{m_1+20}$ belong to the same homogeneity interval. Towards this goal we calculate the following statistic (whose asymptotic distribution is a Brownian bridge as it follows from Theorem 2.1 by making $x = 1$ and taking a supremum over λ)

$$(3.2) \qquad \sqrt{n} \sup_{\lambda \in [0,\pi]} \left| \int_{-\pi}^{\lambda} \left(\frac{I_{n,n,X}(y)}{\psi(e^{-iy})} - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_{n,n,X}(z)}{\psi(e^{-iz})} dz \right) dy \right| / 2\sigma^2$$

using the subsample $X_{m_1-180}, \ldots, X_{m_1+20}$. If the statistics value is well explained by the distribution of the supremum of a Brownian bridge, the interval of homogeneity is extended to include the subsample $X_{m_1+1}, \ldots, X_{m_1+20}$. Otherwise a new homogeneity interval commences with the observation X_{m_1+1} and the parameters μ, ψ and σ^2 are reestimated on the block $X_{m_1+1}, \ldots, X_{m_1+200}$. Then the procedure is reiterated.

In this procedure we choose to neglect the error committed by taking the estimated values of the model's parameters for the real ones. The reason is two fold. We try to keep things as simple as possible but more important our empirical experiece showed that allowing for large deviations of the asymptotic variance of the goodness of fit statistic from the value prescribed by the Brownian bridge does not change at all the results of the analysis.

We begin by performing this analysis using an ARMA(1,1) process as a local approximation of the series of logarithm of the absolute values of daily returns. Figure 3.1 displays the estimated AR and MA coefficients for the intervals of homogeneity corresponding to roughly 10 years of data, the 70's. (We chose to show only a part of the data for the sake of visual clarity. The rest of the data displays the same traits). Figure 3.1 presents a remarkable finding of our analysis: in the framework of the general model of local linear dependency with innovations with possibly changing mean and variance, the data rejects the significance of the linear dependency (retaining as significant the changes in the mean and the variance of the innovations). First, that the AR and MA estimated coefficients are always almost equal although taking a wide range of values. Second, the 95% confidence interval almost always covers 0. As anyone working with time series models knows, this situation is typical of fitting ARMA(1,1) models to independent data. Other linear models were used to locally approximate the data and the results were consistent with the findings for the case ARMA(1,1). *Figure 3.1 suggests that piecewise, on the intervals of homogeneity, the data is in fact white noise.*

Hence the following simple model for $X_t := \log(|r_t|)$, where r_t are the daily returns on the $S\&P500$, could be considered

$$(3.3) \qquad X_t = \mu(t) + \sigma(t)\epsilon_t,$$

where (ϵ_t) is a white noise with $E(\epsilon_t) = 0$, $E\epsilon_t^2$=1 and $\mu(t)$ and $\sigma(t)$ are functions of t. For the absolute returns, this model translates into

$$(3.4) \qquad |r_t| = exp(\mu(t))exp(\sigma(t)\epsilon_t).$$

From this discussion it follows that a simpler statistic corresponding to an i.i.d. model can

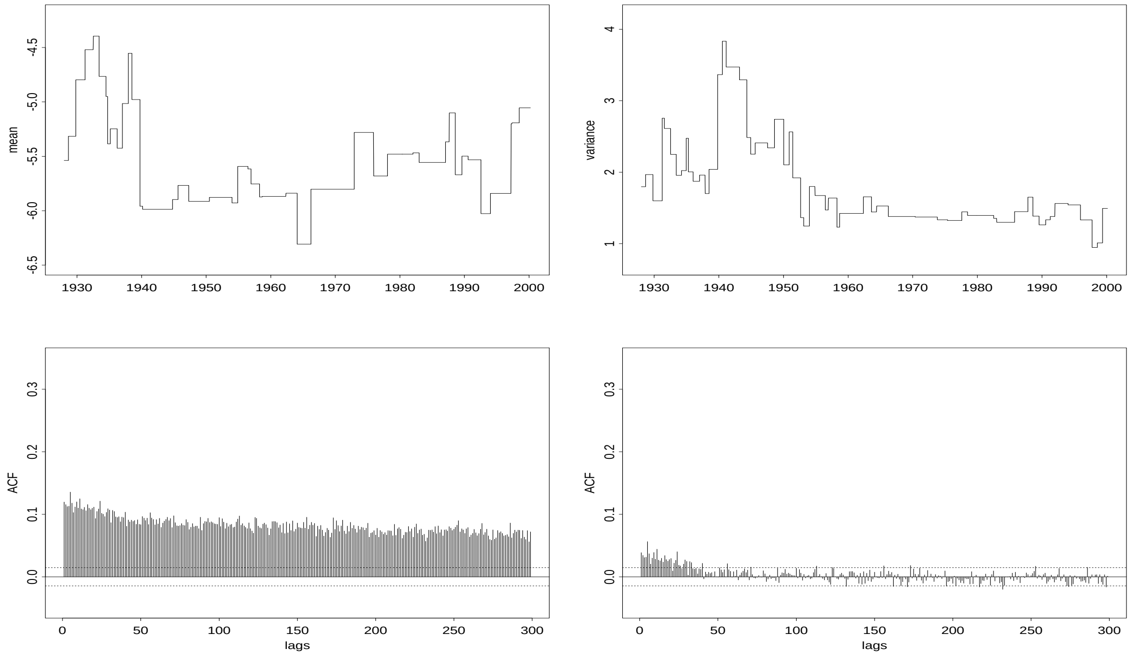

Figure 3.2 Top: The mean *(left)* and the variance *(right)* of the logarithm of the absolute returns of the *S&P500* . Bottom: Sample ACFs of the logarithm of the absolute returns of the *S&P500* before *(left)* and after *(right)* subtracting the mean in Top *(left)*.

be used in place of (3.2) for determining the homogeneity intervals, i.e.

$$\sqrt{n} \sup_{\lambda \in [0,\pi]} \left| \int_{-\pi}^{\lambda} \left(I_{n,n,X}(y) - n^{-1} \sum_{t=1}^{n} Z_t^2 \right) dy \right| /2\sigma^2 \tag{3.5}$$

(see Theorem 2.1). Applying the methodology described above (with statistic (3.2) replaced by (3.5)), the homogeneity intervals associated with logarithm of the absolute values of daily returns are produced. Figure 3.2 displays the mean and the standard deviation of the absolute values of logarithm of daily returns estimated within the homogeneity intervals.

The top-left graph in Figure 3.2 shows a very volatile decade between 1928-1938 and a general upwards trend for the period from 1938 up to the present. One can possibly see a certain connection between the higher levels of the mean of log-absolute returns (hence variance of returns), the 1973 oil crisis and the economic recessions in the beginning of the 80's and 90's. The recent period of economic expansion covering the past decade is also characterized by higher levels of the mean of log-absolute returns, i.e. variance of returns.

Figure 3.2 bottom displays the sample ACF for the logarithm of absolute values of daily returns before and after the data was centered by the mean estimated by our methodology. The two graphs show a strong reduction of the dependency present in the sample ACF

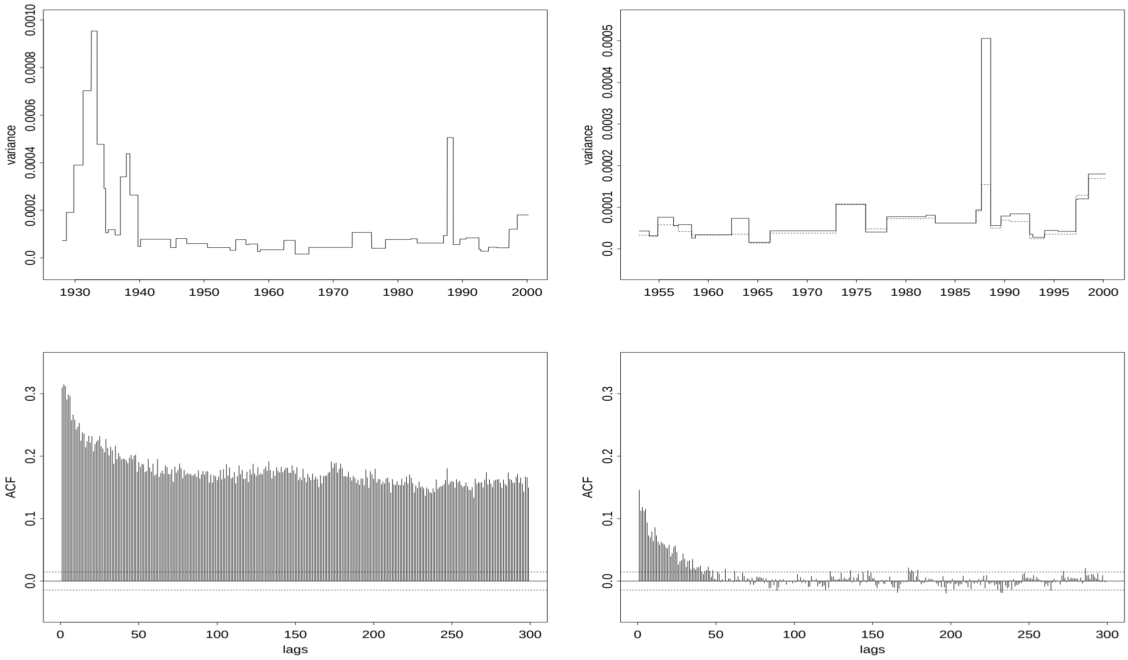

Figure 3.3 Top: *(left)* Variance of the absolute returns of the *S&P500*. *(right)* $E|r_t|^2$ (continuous line) and $0.23\ exp(2\mu(t))$ (dotted line) as evidence that $\sigma(t) = \sigma$ in (3.6). Bottom: Sample ACFs of the absolute returns of the *S&P500* before *(left)* and after *(right)* scaling with the standard deviation estimated on the homogeneity intervals constructed using the log-absolute returns.

(the dependency is reduced to the first circa 40 lags). Work in progress shows that finding a better approximation of the changing mean (than our present rather rough step function approximation) removes even the dependence still present in the sample ACF of the residuals in Figure 3.2. Before discussing the possible implications of the top-right graph of this figure, let us say a few things about modeling the absolute returns.

The intervals of homogeneity for the logarithms of absolute returns define intervals of homogeneity for the absolute returns. (A direct application of our methodology although theoretically feasible does not produce meaningful results. Due to the presence of heavy tails a close tracking of the variance is practically impossible.) Figure 3.3 top-left displays the variance of the absolute returns of the *S&P500*. Figure 3.3 bottom displays the sample ACF for the absolute values of daily returns 1928-2000 before and after the data was scaled by the standard deviation estimated on every interval of homogeneity constructed by our method. The two graphs show a strong reduction of the dependency present in the sample ACF (the dependency is reduced to the first circa 50 lags). The still unexplained dependency present in the sample ACF of the residuals is very likely due (as work in progress shows) to the imperfection of our tracking of the changes in variance.

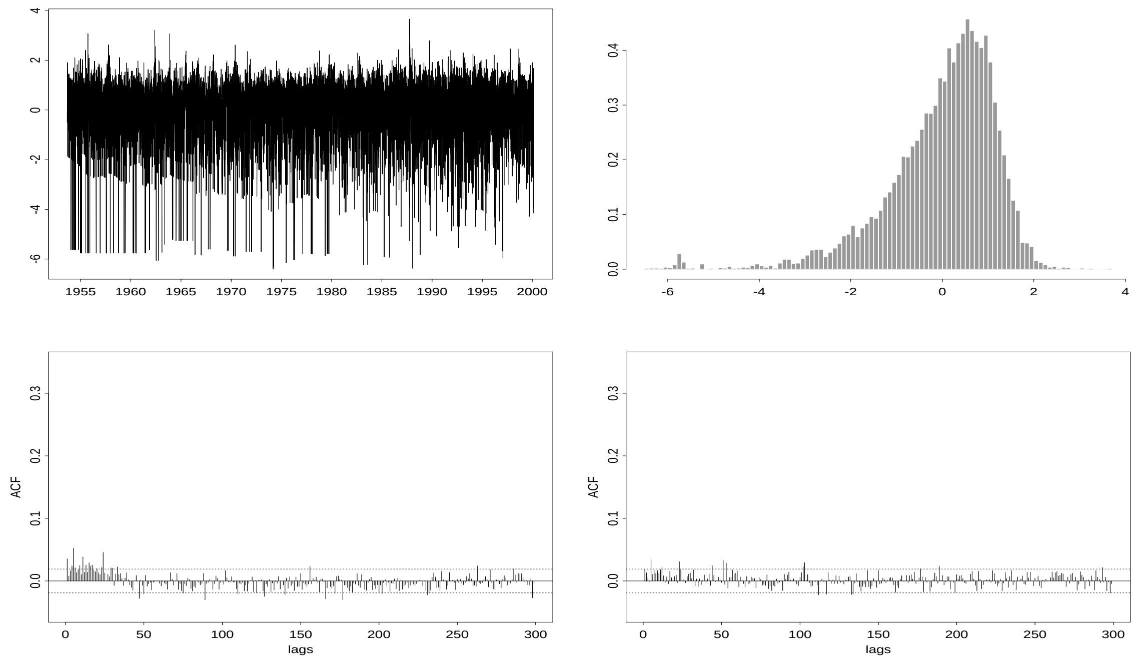

Figure 3.4 Top: Plot (*left*) and histogram *(right)* of the estimated residuals ϵ_t based on the model (3.6) corresponding to the period 1953-present. Bottom: *(right)* Sample ACF for the estimated residuals ϵ_t. Possibly due to the rough approximation of the mean of the time series by a step function, the first circa 25 lags are slightly significant. *(left)* Sample ACF of absolute estimated residuals $|\epsilon_t|$. The graph shows no properties in the variance.

We return now at Figure 3.2. The top-right graph in Figure 3.2 shows a change in the role played by the variance $\sigma^2(t)$ around 1953 (the year the modern and current structure of the $S\&P500$ was defined). Before 1953 the variance seems to have been a meaningful parameter carrying a certain amount of information. After 1953 the value of σ^2 stayed roughly constant around the value of 1.4. For the time period 1953-present the model for log-absolute returns can be then simplified to

$$X_t = \mu(t) + \sigma\epsilon_t, \tag{3.6}$$

where ϵ_t are white noise with $E\epsilon_t = 0$ and $E\epsilon_t^2$=1. Figure 3.4 displays a brief analysis of the estimated residuals ϵ_t based on the model (3.6). Both the sample ACF of the residuals ϵ_t as well as $|\epsilon_t|$ are close to being statistically insignificant. Almost all the dynamics of the log-absolute returns (hence absolute returns) time series seem to be concentrated in the shifts of the mean (variance).

More evidence that the assumption of a constant $\sigma(t)$ in the model (3.6) for the period 1953-2000 is plausible is displayed in the top-right graph of Figure 3.3. If $\sigma(t) = \sigma$, a

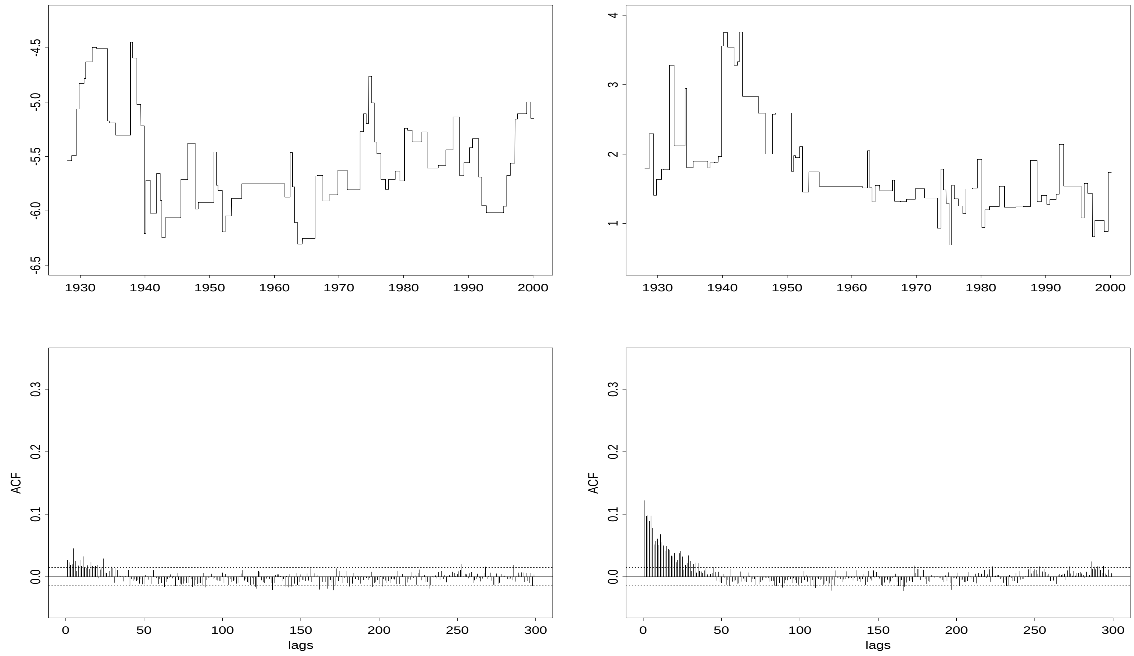

Figure 3.5 Top: The mean *(left)* and the variance *(right)* of the logarithm of the absolute returns of the *S&P500* (intervals of homogeneity constructed using the t-test). Bottom: Sample ACFs of the absolute log-absolute *(left)* and absolute *(right)* returns of the *S&P500* after subtracting the mean (dividing with the standard deviation, respectively) estimated on intervals of homogeneity constructed using the t-test.

constant, (3.4) can be rewritten as

$$|r_t| = exp(\mu(t))exp(\sigma\epsilon_t) = h(t)\varepsilon_t, \tag{3.7}$$

with (ε_t) a white noise sequence. Hence $E|r_t|^2/h(t)^2$ would be constant and equal to $E(\varepsilon_t^2)$. A variable $\sigma(t)$ would imply that the the ratio $E|r_t|^2/h(t)^2$ changes from an interval of homogeneity to another as $E(\varepsilon_t^2)$ changes. The top-right graph in 3.3 displays both $E|r_t|^2$ and 0.23 $h(t)^2$ and shows a very close concordance between the two series in all the intervals but the one containing the October 1987 crash.

We end this section with a discussion on the performance of our procedure when replacing the goodness of fit test based on the integrated periodogram with a very simple goodness of fit based on the central limit theorem. The null hypothesis is that on a certain interval of length n, the data is independent with the mean μ and variance σ^2 and the test statistic is simply $\sqrt{n}(\bar{X} - \mu)/\sigma$. The results of this approach are displayed in Figure 3.5. The overall pattern of change is the same. However the t-statistic finds more changes then the integrated periodogram. Although the overall goodness-of-fit (as measured by the shape of the residual correlation in sample ACF) is a bit higher for the t-statistic method, the integrated periodogram approach offers a simpler overall picture of the pattern of changes in the long time series we analyzed (at the price of a worse fit) and serves to motivate the assumption of locally independent log-absolute returns on which the t-statistic based method rests.

4 Forecasting comparison

In the stationary framework, a sample ACF behavior as shown in Figure 3.2 and Figure 3.3 will be interpreted as evidence of long memory. Hence we are facing a modeling choice for $X_t := \log(|r_t|)$, the *logarithm of the absolute values of daily returns.* The choice is between a stationary long memory model and a non-stationary model with its dynamics mainly concentrated in the changes of the mean. One possible way of solving this dilemma is to compare the forecasting behavior of two paradigms on the data at hand. Since our approach is to describe the volatility directly by analyzing the sequence of absolute returns, a natural choice for a long memory stationary model is the fractionally ARIMA class introduced by Granger and Joyeux ([8]) and Hosking ([13]).

The process $\{X_t, t = 0, 1, \ldots\}$ is said to be a FARIMA(p,d,q) with $d \in (0, 0.5)$ if $\{X_t\}$ is stationary and satisfies the difference equation

$$\phi(B) \bigtriangledown^d X_t = \theta(B) Z_t, \tag{4.8}$$

where $\{Z_t\}$ is white noise and ϕ, θ are polynomials of degree p, q respectively. The operator $\bigtriangledown^d$ is defined by

$$\bigtriangledown^d := (1 - B)^d = \sum_{j=0}^{\infty} \pi_j B^j, \tag{4.9}$$

where

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(1-d)} = \Pi_{0<k\leq j} \frac{k-1-d}{k}, \quad j = 1, 2, \ldots.$$

The data used in the comparison are the logarithm of the absolute values of daily returns in the interval 1957-2000. A FARIMA(1,d,1) model (LM) was estimated on the first 1000 observations and reestimated every 2 years (i.e. every 500 observations) using the last 1000 observations. With the estimated long memory model, predictions $\{f^{LM}\}$ for the future values of $X_t = \log(|r_t|)$ were made every month (i.e. every 20 observations). The maximal forecasting horizon was 200 days ahead. The other model used which will be referred as the shifts-in-the-mean model (SM) is described by equation (3.3). The observations anterior to the date when a forecast was made were used for determining the (then) current interval of homogeneity. The forecasts for the future values of $X_t = \log(|r_t|)$ (independent of the horizon), $\{f^{SM}\}$ were simply the estimated mean on this homogeneity interval.

One way of comparing the two forecasts would be by assesing the orthogonality of one forecast error (at horrizon h) to the other forecast. Concretly one can test something less

Horizon (days)	P-value of the Wald statistic for $H_0 : \alpha = 0, \beta_1 = 0, \beta_2 = 1$ i.e. f_t^{SM} orthogonal to LM forecast error	$H_0 : \alpha = 0, \beta_1 = 1, \beta_2 = 2$ i.e. f_t^{LM} orthogonal to SM forecast error
10	0.00	0.21
20	0.03	0.27
30	0.00	0.31
40	0.06	0.67
50	0.01	0.50
60	0.09	0.23
70	0.01	0.32
80	0.09	0.39
90	0.01	0.44
100	0.01	0.12
110	0.03	0.22
120	0.06	0.27
130	0.00	0.06
140	0.06	0.40
150	0.01	0.27
160	0.03	0.08
170	0.00	0.48
180	0.06	0.16
190	0.00	0.58
200	0.06	0.23

Table 1: Comparison of forecasting performance between the LM model and SM model. The LM process is reestimated every 2 years using the observations of the previous 4 years. The Wald statistic of the F-ratio test is calculated under the two alternatives and the p p-values are reported. A small p-value is a signal of the failure of the null. Overall the table seems to show a better performance of the SM model in forecasting.

general, i.e whether one forecast error is uncorrelated with the other forecast. This would be accomplieshed by means of a regression. For example, to test if the SM forecasts and the LM forecast errors are uncorrelated, one would test whether $\alpha = 0$, $\beta_1 = 0$ in the regression

$$X_{t+h} - f_t^{LM} = \alpha + \beta_1 f_t^{SM} + \varepsilon_t. \tag{4.10}$$

However, given the possibly non-stationary nature of the time series the assumption of ergodic stationarity of the regressor and dependent variables needed for the well-functioning of the GMM machinery, is likely to be violated (f_t^{SM} is close to a piecewise constant function).

To address this possible problem we reformulate our test. Testing whether $\alpha = 0$, $\beta_1 = 0$ in (4.10) is equivalent to testing whether

$$\alpha = 0, \quad \beta_1 = 0, \quad \beta_2 = 1 \tag{4.11}$$

in the following regression

$$X_{t+h} - X_t = \alpha + \beta_1 (f_t^{SM} - X_t) + \beta_2 (f_t^{LM} - X_t) + \varepsilon_t, \tag{4.12}$$

Notice that testing for

$$\alpha = 0, \quad \beta_1 = 1, \quad \beta_2 = 0 \tag{4.13}$$

in the same regression would be equivalent to verifying that the SM forecast error $X_{t+h} - f_t^{SM}$ is uncorrelated with the FARIMA forecast f_t^{LM}. For this regression the violations of the assumption of ergodic stationarity of the regressor and dependent variables are likely to be less severe. Indeed, the vector $(X_{t+h} - X_t, f_t^{SM} - X_t, f_t^{LM} - X_t)$ is stationary and ergodic on every interval on which the volatility process $\sigma(t)$ is constant, i.e. on any interval of homogeneity. Because under the null hypothesis (4.13) the error term ε_t in regression (4.12) equals the forecast error $X_{t+h} - f_t^{SM}$ orthogonal to anything known at date t including $f_t^{(i)} - X_t$, $i = 1, 2$, the regressors are quaranteed to be orthogonal to the error term (a similar statement holds under the null hypothesis (4.11)). Even more, Figure 4.1 supports the hypothesis of uncorrelated forecasting errors (the forecast errors $X_{t+h} - f_t^{LM}$ have a similar behavior), and hence it appears that an OLS estimate would suffice. Note that the regression (4.12) is closely related to the so-called *forecast encompassing equation*

$$X_{t+h} = \alpha + \beta_1 f_t^{SM} + \beta_2 f_t^{LM} + \varepsilon_t \tag{4.14}$$

which cannot be employed due to the possibly non-stationary nature both of the regressors and dependent variables. The p-values of the F-test Wald statistic corrresponding to $H_0 : \alpha = 0$, $\beta_1 = 0$ and $\beta_2 = 1$ and $H_0 : \alpha = 0$, $\beta_1 = 1$ and $\beta_2 = 0$, respectively are reported in Table 4. For most of the forecast horizons the hypothesis of orthogonality of the SM forecast on the LM forecast errors is rejected while the hypothesis of orthogonality of the LM forecast on the SM forecast errors remains unchallanged. Hence these tests seem to support the conclusion that the SM model overperforms the LM model in forecasting.

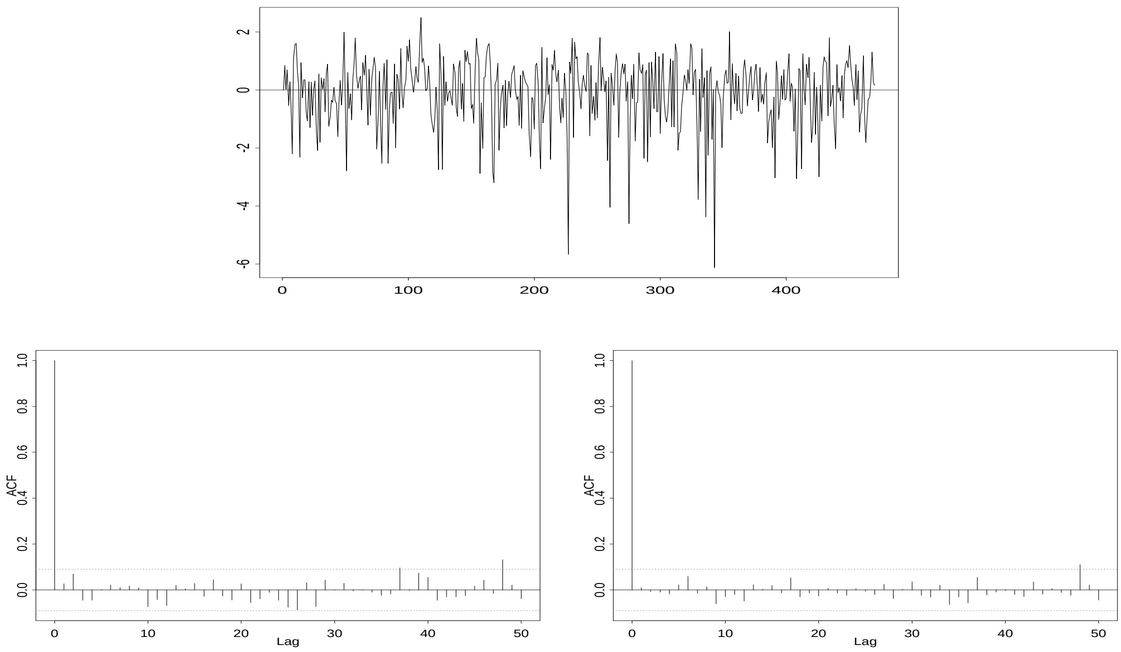

Figure 4.1 Top: The forecast errors $X_{t+h} - f_t^{SM}$ based on the model (3.6) corresponding to the period 1957-present (the period 1953-1957 is used for the preliminary estimation of the LM model. Bottom: *(left)* Sample ACF for the forecast errors $X_{t+h} - f_t^{SM}$. *(right)* Sample ACF of absolute forecast errors $|X_{t+h} - f_t^{SM}|$. The graphs suggest that the forecast errors are uncorrelated and homoskedastic.

5 Conclusions

Understanding volatility of stock returns lies at the core of modern econometric research. From a time series point of view, it means understanding and modeling the series of absolute returns. In this paper an analysis of the $S\&P500$ absolute returns is conducted giving up the usual assumption of global stationarity. Supposing that the returns follow a non-stationary process, an interesting class of models is those with time depending parameters. We approximate locally the non-stationary process by stationary models and identify the intervals on which stationary processes provide a good approximation by using a goodness of fit test based on the integrated periodogram (Picard ([21]), Klüppelberg and Mikosch ([15])). Our approach leads to modeling the returns as a sequence of independent variables with a piecewise constant variance function. More concretely, the $S\&P500$ absolute returns, $|r_t|$ can be described by the following

$$|r_t| = h(t)exp(\epsilon_t), \quad t = 0, 1, \ldots$$

where (ϵ_t) is a white noise sequence, $E(\epsilon) = 0$, $E(\epsilon)^2 = \sigma^2$ and $h(t)$ a function of t which can be well approximated by a step function, yielding a model with piecewise constant variance. We show that even a rough approximation of the variance dynamics by a step function is enough to explains most of the dependency structure present in the sample ACF of long absolute return series, providing an explanation for the so called "long memory in volatility" phenomenon. In other words, we find that all the dynamics of return time series seem to be concentrated in the shifts of the variance.

Acknowledgment. Part of the research presented in this paper was carried out while the first author visited the Department of Economics of the University of California at San Diego. He would like to thank his colleagues there, in particular Rob Engle, for their warm hospitality and generous support that made this work possible. The idea of trying the goodness of fit test based on the central limit theorem came from Olivier Perrin. His patience, curiosity and statistical intuition made a difference. The paper gained in clarity at different stages due to the careful readings of Holger Drees. His constructive skepticism is highly appreciated.

References

[1] Anderson, T.W. (1993) Goodness of fit tests for spectral distributions. *Ann. Statist.* **21**, 830–847.

[2] Bartlett, M.S. (1954). Problemes de l'analyse spectrale des séries temporelles stationnaires. *Publ. Inst. Statist. Univ. Paris.* **III-3**, 119–134.

[3] Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods,* 2nd edition. Springer, New York.

[4] Cai, J. (1994) A Markov model of unconditional variance in ARCH. *J. Business and Economic Statist.* **12**, 309–316.

[5] Dahlhaus, R. (1997) Fitting time series models to nonstationary processes. *Ann. Statist.* **25**, 1–37

[6] Diebold, F.X. (1986) Modeling the persistence of the conditional variances: a comment. *Econometric Reviews* **5**, 51–56.

[7] Giraitis, L. and Leipus, R. (1992) Testing and estimating in the change-point problem of the spectral function. *Lith. Math. Trans. (Lit. Mat. Sb.)* **32**, 20–38.

[8] Granger, C., W. and Joyeux, R. (1980) An introduction to long-memory time series models and fractional differencing. *J. Time Series Analysis*, **1**, 15–29.

[9] Granger, C., W., Spear, S., Ding, Z-X. (2000) Stylized facts on the temporal and distributional properties of absolute returns: an update. In Chan, W.-S., Li, W., K. and Tong, H., eds. *Statistics and Finance* Imperial College Press.

[10] Grenander, U. and Rosenblatt, M. (1984) *Statistical Analysis of Stationary Time Series,* 2nd edition. Chelsea Publishing Co., New York.

[11] Hamilton, J. and Susmel, R. (1994) Autoregresive conditional heteroskedasticity and changes in regime. *J. Econometrics* **64**, 307–333.

[12] Härdle, W., Spokoiny, V., Teyssière, G. Adaptative estimation for a time inhomogeneous stochastic volatility model. *Preprint.*

[13] Hosking, J. R. (1981) Fractional differencing . *Biometrika* **68**, 165–176.

[14] Jorion, p. (1996) Risk and turnover in the foreign exchange market. In Frankel, J.,A., Galli, G. and Giovannini, A. eds *The Microstructure of Foreign Exchange Markets* The University of Chicago Press.

[15] Klüppelberg, C. and Mikosch, T. (1996) Gaussian limit fields for the integrated periodogram. *Ann. Appl. Probab.* **6**, 969–991.

[16] Lobato, I. N. and Savin, N. E (1998). Real and Spurious Long-Memory Properties of Stock-Market Data *Journal of Business & Economic Statistics*, 261–268.

[17] Lamoureux, C.G. and Lastrapes, W.D. (1990) Persistence in variance, structural change and the GARCH model. *J. Business and Economic Statist.* **8**, 225–234.

[18] Mikosch, T. (1998) Periodogram estimates from heavy–tailed data. In: R. Adler, R. Feldman and M.S. Taqqu (eds.) *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy–Tailed Distributions,* pp. 241–258. Birkhäuser, Boston.

[19] Mikosch, T. and Stărică, C. (2000) Change of structure in financial time series, long range dependence and the GARCH model. Preprint. Available at www.math.chalmers.se/ starica

[20] Mikosch, T. and Stărică, C. (2000) Long memory and the ARCH models. To appear . Available at www.math.chalmers.se/˜starica

[21] Picard, D. (1985) Testing and estimating change-points in time series. *Adv. Appl. Probab.* **17**, 841–867.

[22] Priestley, M.B. (1981) *Spectral Analysis and Time Series, vols. I and II.* Academic Press, New York.

[23] Simonato, J., G. (1992) Estimation of GARCH processes in the presence of structural change. *Economic Letters* **40**, 155–158.

[24] Shorack, G.R. and Wellner, J.A. (1986) *Empirical Processes with Applications to Statistics.* Wiley, New York.