Is it really long memory we see in financial returns?[1]

Thomas Mikosch
University of Groningen

and

Cătălin Stărică
The Wharton School, Philadelphia
Chalmers University of Technology, Gothenburg

Abstract

Our study supports the hypothesis of global non-stationarity of the return time series. We bring forth both theoretical and empirical evidence that the long range dependence (LRD) type behavior of the *sample ACF* and *the periodogram* of absolute return series and the IGARCH effect documented in the econometrics literature could be due to the impact of non-stationarity on statistical instruments and estimation procedures. In particular, contrary to the common-hold belief that the LRD characteristic and the IGARCH phenomena carry meaningful information about the price generating process, these so-called stylized facts could be just artifacts due to structural changes in the data. The effect that the switch to a different regime has on the sample ACF and the periodogram is theoretically explained and empirically documented using time series that were the object of LRD modeling efforts (S&P500, DEM/USD FX) in various publications.

AMS 1991 Subject Classification: Primary: 62P20 Secondary: 60G10 60F17 60B12 60G15 62M15 62P20
Key Words and Phrases. sample autocorrelation, change point, GARCH process, long range dependence.

[1]This research supported in part by a grant of the Dutch Science Foundation (NWO).

1 Introduction

1.1 Preliminaries on ARCH and GARCH processes

Among the models for log-returns

$$X_t = \log(P_t/P_{t-1}), \quad t = 1, 2, \ldots,$$

of stock indices, foreign exchange rates, share prices, etc., P_t, $t = 0, 1, \ldots$, the ARCH (*autoregressive conditionally heteroscedastic*) processes have gained particular popularity. Besides the *stochastic volatility models* (see for example Ghysels et al. [34] for a recent survey paper) they have become *the* standard models in the financial econometrics literature. In particular, they appear in many recent textbooks and monographs on time series analysis (see for example Brockwell and Davis [19] or Embrechts et al. [30]) or econometrics (see Campbell et al. [21]), thus every student of statistics or econometrics might have heard of them. The (not inexpensive) GARCH modules of various software packages have certainly contributed to the increasing popularity of this kind of econometric time series model as well.

The success story of the ARCH family started in 1982 when Engle introduced the ARCH(p) processes (ARCH of order p) by requiring that

$$X_t = \sigma_t \, Z_t, \tag{1.1}$$

where σ_t (the so-called *stochastic volatility*) obeys the recurrence equation

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \cdots + \alpha_p X_{t-p}^2. \tag{1.2}$$

Here (Z_t) is a white noise process with variance 1, usually supposed to be iid or a strictly stationary martingale difference sequence, and the α_i's are non-negative parameters. In what follows we will always assume that (Z_t) is iid. This implies that, conditionally upon $X_{t-1}, \ldots, X_{t-p}$ (the p past observations of the time series), X_t has variance σ_t^2.

The basic idea behind the construction (1.1) is quite intuitive: for a "forecast" of the distribution of X_t we only have to know two ingredients: σ_t and the distribution of Z_t. For example, if Z_t is normal $N(0,1)$, then $X_t \sim N(0, \sigma_t^2)$, given the past observations of the time series. Hence, conditionally upon $X_{t-1}, \ldots, X_{t-p}$, the present value X_t may assume values in $[-1.96\sigma_t, 1.96\sigma_t]$ with 95% probability. Similarly, there is a 5% chance for the log-return X_t to fall below the threshold $-1.64\sigma_t$. The 5%–quantile of the log-return distribution is considered as a measure of risk for the underlying asset. In the financial area, this quantile is known under the name of *Value at Risk* or *VaR*; see RiskMetrics [**?**].

These simple calculations show why models of type (1.1) have become so popular; in the presence of non-Gaussian distributions for log-returns (this is a fact no specialist would doubt!) mixture models such as (1.1) allow one to get updated (i.e. conditional) probability "forecasts" without too much sophistication.

However, it turned out that the simple ARCH(p) given by (1.1) and (1.2) has a reasonable fit to real-life data *only if the number of the parameters α_i is rather large.* Since the rationale for the definition (1.2) is to take a time-changing weighted average of the past squared observations as an approximation to the conditional variance σ_t^2 (an "updated estimate of the variance", if you like), it is quite natural to define σ_t^2 not only as a weighted average of past X_j^2's but also of past σ_j^2's. This new idea resulted in Bollerslev's [13] and Taylor's [?] *generalised ARCH process of order* (p, q) (GARCH(p, q)): the process (X_t) is again given by (1.1), but now the squared stochastic volatility satisfies

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j \, X_{t-j}^2 + \sum_{k=1}^{p} \beta_k \, \sigma_{t-k}^2 := \alpha_0 + \alpha(L) X_t + \beta(L) \sigma_t^2 , \tag{1.3}$$

where the α_j's and β_k's are non-negative parameters and L is the lag operator. Clearly, σ_t^2 could have been defined in many other reasonable ways, and therefore it is perhaps not totally surprising that a high wave of different ARCH-type models has flooded the econometrics journals. Each of these models was introduced in order to improve upon (1.2) or (1.3) in some sense. Some of them have gained popularity such as Nelson's [53] EGARCH (exponential GARCH) model. They have been known to the specialists for some years. Not all of them are directly comparable with the GARCH processes, from a mathematical point of view, but we do not have here the space to discuss these modifications in detail; see for example Bollerslev et al. [14] or Shephard [?] for review papers.

1.2 Can heavy tails save the GARCH(1, 1) model?

In what follows, we mostly concentrate on the GARCH(1, 1) model. We do so for several reasons. First of all it is theoretically well understood (Bollerslev [13], Nelson [52], Bougerol and Picard [16]). Secondly, it can be easily fitted to log-returns

$$X_t = \log P_t - \log P_{t-1} \, , \quad t = 1, 2, \ldots \, ,$$

of speculative prices P_t. Moreover, it is commonly believed that, despite its simplicity (3 parameters only), it captures some of the basic features of log-returns. In particular, this model adequately describes the *heavy-tailedness of the marginal distribution.* Indeed, under very general conditions on the noise sequence (Z_t), the GARCH(1, 1) in particular, and GARCH(p, q) processes in general, have Pareto-like marginal distributions, i.e.

$$P(X > x) \sim c_0 \, x^{-\kappa} \quad \text{as } x \to \infty \text{ for some } c_0, \kappa > 0. \tag{1.4}$$

The GARCH(1, 1) model can also explain to some extent another empirical feature related to the extremal behavior of log-return data, i.e. *the exceedances of high and low thresholds do not occur* *separated over time but are heavily clustered.* This fact is usually referred to as *dependence in the tails.* For some recent research in this area we refer to Mikosch and Stărică [50] for the GARCH$(1,1)$ case and Davis et al. [23] in the general GARCH(p,q) case.

However, the real data displays a series of features that seem to disagree with the theoretical properties of the GARCH$(1,1)$ process, pointing out the limitations of the model. These discrepancies become obvious mainly when long series of log-returns are analysed. In the sequel we will discus two of these so called *stylized facts* that characterize the log-returns of foreign exchange rates, stock indices, share prices, bond yields, etc., P_t. The first one can be summarized as follows.

- Although the *sample autocorrelations* of the data are tiny (uncorrelated log-returns), the *sample autocorrelations* of the absolute and squared values are significantly different from zero even for large lags.

This empirical finding (which we will refer to as *long memory type of behavior of the sample ACF of absolute and squared log-returns*) is usually interpreted as evidence for *long memory in the volatility* of financial log-returns and seems to provide strong evidence against the GARCH models. Indeed, GARCH models have *short memory*, i.e. if $\rho_X(h) := corr(X_0 X_h)$, with $h \in \mathbb{Z}$ is the theoretical autocorrelation functions (ACFs) of a GARCH process (X_t), ρ_X decays to zero at an exponential rate. More generally, the theoretical ACF of the process $(f(X_t))$ for any measurable function f, decays to zero at an exponential rate. This fact implies *short memory* for the GARCH process and the corresponding processes of its absolute values and squares. One might expect then that the *short memory property* that characterizes the theoretical ACF should be reflected in the behavior of the *sample ACF*

$$\widehat{\rho}_X(h) = \widehat{\gamma}_X(h)/\widehat{\gamma}_X(0) \quad \text{where} \quad \widehat{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n-h} X_t X_{t+h}\,, \quad h \in \mathbb{Z},$$

the natural estimator of the theoretical ACF.

A hasty identification of the behaviors of the theoretical ACF and the sample ACF would imply that the sample ACF of a GARCH$(1,1)$ process should decay to zero at an exponential rate and vanish rather quickly. This, however, is in contradiction with the stylized fact mentioned above.

We want to emphasize that all GARCH models, regardless their order, have the same basic properties. In particular, the extremal behavior encompassing heavy tails and dependence in the tails, the asymptotic behavior of the sample ACF, etc., are similar in nature, independent of the order of the process. For example, independently of the order, all GARCH(p,q) processes have heavy tails and display short memory. The only difference is the degree of analytical tractability of these properties. Therefore the conclusions we draw from examining the GARCH$(1,1)$ case are also valid for the wider class of GARCH(p,q) models.

Before concluding that the GARCH(p, q) processes fail to capture the mentioned stylized feature of slowly decaying sample ACF, an attempt should be made to reconcile the empirical findings with the theoretical facts. A possible explanation (which we had considered as plausible for some time) for the mentioned contradiction could be the neglection of the statistical uncertainty present in the *sample ACF*. This uncertainty could render problematic the identification between the the behaviour of the theoretical and sample ACF. It is conceivable that the sample ACF at large lags could be non-zero and at the same time statistically insignificant. This argument seems especially plausible in the light of the mentioned heaviness of the tails of log-returns which implies a considerable variability in the estimated ACF.

Since little was known about the behavior of the *sample ACF* of GARCH processes when the assumption of finite 4th moment for the marginal distribution is violated, investigations, aiming at assessing the truth behind the mentioned explanation, were conducted in Davis and Mikosch (1998) for the ARCH(1), Mikosch and Stărică [50] for the GARCH$(1, 1)$ and Davis et al. [23] in the general GARCH(p, q) case. The motivation for this work can be summarized as follows.

Log-returns could have infinite 3rd or 4th moments (see Embrechts et al. [30] for the statistical theory of tail and high quantile estimation, in particular Chapter 6 for methods to detect how heavy the tails of real-life data are). Therefore one expects that the rate of convergence of the sample autocorrelations to their theoretical counterparts is much slower than $\sqrt{n}$-rate encountered in classical time series theory and that the asymptotic normal limiting distributions in the classical large sample theory are replaced with much heavier tailed stable distributions. This results in extremely wide confidence bands that would render the non-zero values of the sample ACF of the absolute values and squares at large lags statistically unsignificant, even for huge sample sizes. In particular, the long memory type of behavior observed in the sample ACF of absolute values and squares would not be in contradiction with the short memory property of the GARCH processes.

The investigations conducted in the above mentioned papers show that the sample ACF is indeed a poor estimator of the theoretical ACF in the situations described. In particular, if the variance of the data is believed to be finite, but the 4th moment is not, it does not make sense to look at the sample ACF of the squared log-returns because the ACF is not defined and, even worse, the sample autocorrelations (which can be defined for any time series, independent of the existence of the moments) have non-degenerate limit distributions.

In the light of these theoretical findings, the heavy-tailedness of the GARCH model *could* be one possible explanation for the slow decay of the sample ACF of absolute and squared log-returns. If this was the case the long memory in the volatility would be spurious.

This theoretically plausible reconciliation between the GARCH$(1, 1)$ model and the mentioned stylized fact is however infirmed by the data. In Mikosch and Stărică [50] this issue is investigated empirically by means of an analysis of a long high frequency time series of log-returns on the JPY/USD spot rates. Even when accounting for the mentioned larger-than-usual statistical uncertainty, by applying the wide confidence bands for the sample ACFs which are imposed by the estimated parameters α_0, α_1 and β_1 of a GARCH$(1,1)$ process, we did not find sufficient evidence that a GARCH process could explain the effect of almost constant sample ACF at large lags for absolute values and squares of long time series of log-returns.

1.3 How forecastable is the volatility?

Let us now concentrate on another less known anomaly that affects the modeling of long log-return series by one GARCH$(1,1)$ process. As we have already mentioned, it is known that GARCH processes have Pareto-like tails (1.4). However, only in a very few cases the value of κ, the tail index occurring in (1.4), can be obtained as closed form solution to an equation involving both the coefficients and the innovations in the model, allowing for a fully parametric estimation of the tail index. Hence, under most of the GARCH(p,q) models, one has to rely on semi-parametric statistical estimation procedures for the tail index. These methods can be applied whenever the tails of the marginal distribution have a behavior of the type (1.4); see Section 6.4 in Embrechts et al. [30]. *The* exception is the GARCH$(1,1)$ for which one can calculate κ as a function of the parameters α_1, β_1 and the distribution of the innovations Z. For this model we can compare the semi-parametric estimates of κ based only on the assumption (1.4) with the fully parametric ones based on the estimated parameters, α_1 and β_1, and the fitted innovations. The result of this comparison can be summarized as follows:

- The semi-parametric estimation techniques suggest values of κ which are significantly higher than the parametric estimates which are solutions to the equation for κ. In other words, the tails implied by the GARCH$(1,1)$ model fitted to long log-returns time series are significantly heavier than the tails of the data.

This is mainly due to the fact that the sum of the estimated coefficients of a GARCH$(1,1)$ model fit to longer time series, $\hat{\varphi}_1 = \hat{\alpha}_1 + \hat{\beta}_1$ is usually close to 1. Indeed, the theory for the GARCH$(1,1)$ in Bollerslev [13], cf. Mikosch and Stărică [50], explains that, in this case, one is close to the situation when the second moment of the theoretical model is infinite.

In the light of this fact and since there is little statistical evidence that log-returns have infinite variance (see Embrechts et al. [30], Section 6.4, for an extensive discussion and various examples) another natural question appears:

How does one explain that $\hat{\varphi}_1 \approx 1$?

Another aim of this paper is to show by theoretical means and empirical examples that the $\hat{\varphi}_1 \approx 1$ effect is spurious and might be due to non-stationarity in the time series.

We mention at this point that the occurence of almost integrated GARCH(1, 1) in the practice of fitting ARCH type models to log-returns, i.e. $\hat{\varphi}_1 \approx 1$, is another *stylized fact* of the empirical research in financial time series analysis. Besides its impact on the tail behavior of the estimated GARCH(1, 1) model, this empirical finding, if taken at face value, makes a strong statement about how forecastable the volatility is. This can be phrased as:

- The persistence in variance, i.e. the degree to which past volatility explains current volatility, as measured by ARCH models fitted to the data, is substantial.

Apparently, both the long memory type of behavior of the sample ACF of the absolute values and squares of the log-returns and the persistence in variance, seem to establish a strong connection between volatilities separated by large intervals of time. The first finding seems to say that significant correlation exists between the present volatility and remote past volatilities while an interpretation of the second fact can be given in terms of the forecast of the variance. The GARCH(1, 1) model describes the behavior of the m-period-ahead forecast of the variance

$$\sigma^2_{t+m|t} := E(X^2_{t+m} \mid X_t, X_{t-1}, \ldots)$$

through a first-order difference equation in the forecast horizon m:

$$\sigma^2_{t+m|t} = \alpha_0 + \varphi_1 \, \sigma^2_{t+m-1|t}. \tag{1.5}$$

Since estimation produces values of φ_1 close to 1, the equation (1.5) implies a strong persistence of shocks to volatility. For example, a value of $\varphi_1 = 0.96$ estimated from weekly log-returns on the New York Stock Exchange *July 1962–December 1987* (Hamilton and Susmel [40]), would imply that any change in the stock market this week will continue to have non-negligible consequences a full year later: $0.96^{52} = 0.12$.

However, given the importance of measuring the degree to which past volatilities determine and explain the current volatility, no hasty conclusions should be drawn. The modeling of the prices of contingent claims, such as options, relies on the perception of how permanent shocks to variance are: a shock that is expected to vanish fast will have a smaller impact on the price of an option far from its expiration than a shock that is mostly permanent. The correct assessment of the relationship between past and present volatility is a key aspect of understanding such issues. Hence, a careful investigation of various possible explanations for the mentioned empirical facts should be carried out, with emphasis on the understanding of the statistical subtleties of this issue.

1.4 Tell me what data thou use, and I'll tell thee what your volatility doest.

Let us say a few words about the data which have been used to uncover the mentioned *stylized facts*, bearing in mind that the kind of data one analyzes can strongly determine the results of the analysis. One fact immediately draws even the attention of a not-so-careful reader of the econometrics literature on the long memory property of the volatility: an overwhelming proportion of the time series used to document and model long memory cover *extremely long time spans*, usually decades of economic activity. The latter are inevitably marked by turbulent years of crises and, possibly, structural shifts. Ding et al. [28] discuss the slow decay of the sample ACF of powers of absolute daily log-returns in the S&P *1928–1990*. Ding and Granger [27] found the same type of behavior for the sample ACF of the daily log-returns of the Nikkei index *1970–1992*, of foreign exchange (FX) rate log-returns Deutsche Mark versus U.S. Dollar *1971–1992* and of Chevron stock *1962–1991*. Bollerslev and Mikkelsen [15] use the S&P500 daily log-returns *1953–1990* to fit their FIEGARCH model while Breidt et al. [17] fit a stochastic volatility model with long memory in the volatility to the daily log-returns of the value-weighted market index *1962–1987*.

There is also plenty of empirical evidence for integrated GARCH(1,1) behavior, i.e. evidence for the parameter $\varphi_1 = \alpha_1 + \beta_1$ being close to 1; see Bollerslev, Chou and Kroner [14] and the references therein. The results reported in the two last mentioned studies are exemplary. For example, Bollerslev and Mikkelsen [15] fit a GARCH(1,1) to the S&P500 daily log-returns *1953–1990* and get an estimate of $\hat{\varphi}_1 = 0.995$. For the same model, Breidt et al. [17] obtain an estimate of $\hat{\varphi}_1 = 0.999$ for the daily log-returns of the value-weighted market index *1962–1987*.

It is worth mentioning that, while studies of daily asset log-returns have frequently found integrated GARCH behavior, studies with higher-frequency data *over shorter time spans* have often uncovered weaker persistence. For example, Baillie and Bollerslev [5] report on the estimation of GARCH(1,1) models for hourly log-returns on FX rates of British Pound (GBP), Deutsche Mark (DM), Swiss Franc (CHF) and Japanese Yen (JPY) versus U.S. Dollar (USD) *January 1986–July 1986* (this is only a 6.5 months period!). The estimated parameters $\hat{\varphi}_1^{GBP} = 0.606$, $\hat{\varphi}_1^{DM} = 0.568$, $\hat{\varphi}_1^{CHF} = 0.341$, $\hat{\varphi}_1^{JPY} = 0.717$ are in sharp contrast to the almost integrated GARCH(1,1) models fitted to daily log-returns.

Much lower persistence in variance than suggested by almost integrated GARCH(1,1) models can be detected if one allows for changes in the level of unconditional variance. Such models were considered, for example, by Hamilton and Susmel [40] and Cai [20]. Hamilton and Susmel fitted their SWARCH model to weekly log-returns from the New York Stock Exchange *July 1962–December 1987*. They reported a measure of persistence (related to $\hat{\varphi}_1$ of the GARCH(1,1)) of 0.4. This implies that shocks to volatility die out almost completely after a month ($0.4^4 = 0.05$).

1.5 IGARCH, long memory: just symptoms of non-stationarity?

The message of this article can be formulated in one sentence: *it might be misleading to take the empirical evidence of long memory and strong persistence of the volatility in log-returns at face value, especially when it comes from the analysis of time series that cover long periods.*

In particular, we challenge the following two implications:

1. if *the sample autocorrelations of the absolute and squared values are significantly different from zero for large lags* then *there is long range dependence (LRD) or long memory in the data*,

2. if *the persistence in variance as measured by ARCH type processes is high* then *past volatility explains current volatility.*

In doing so we follow on the steps of Boes and Salas [12], Potter [57], Bhattacharya et al. [9], Anderson and Turkman [1], Teverovsky and Taqqu [63] and others with respect to the first statement (although these references are not directly related to finance) and Diebold [26], Lamoureux and Lastrapes [47], Hamilton and Susmel [40], Cai [20] with respect to the second one.

In Mikosch and Stărică [51] we showed that the type of behavior described by the mentioned stylized facts can be due simply to a very plausible type of *non-stationarity*: *shifts in the unconditional variance* of the model underlying the log-returns. Given the fast rate at which new technological and financial tools have been introduced in the financial markets, the case for the existence of structural changes (and thus for lack of stationarity) seems quite strong. A detailed analysis of the S&P500 log-return series identified the recession periods as being structurally different, i.e. characterized by higher variance. The major structural change is detected between 1973 and 1975 and corresponds to the oil crises.

A conclusion relevant to the methodology of GARCH modeling also follows: *long log-return series should not be modeled just by one GARCH process.* The parameters of the model must be periodically updated.

Since the GARCH$(1,1)$ does not seem to provide a good description of long series of log-returns even when treated with the right amount of statistical care, and if one does not want to exclude GARCH processes as realistic models for log-returns, the modelling effort should be focused on shorter time series. We have noticed in our empirical work with log-returns that, at least when the sample ACF behavior is concerned, GARCH models fitted to daily time series covering one or two years of data are in better agreement with the theoretical results. More concretely, the sample ACFs of the data, their absolute values and squares behave very much in line with the theory on the sample ACF of GARCH processes: they vasnish at all lags or they decay exponentially to zero, respectively. Thus a possible strategy (popular among practitioners) is to fit the model to short time series and change it as soon as one realizes that something goes wrong with it (for example when one sees long memory developing in the sample ACF of the absolute values of the log-returns). This approach is clearly related to the question whether or not the real-life time series can be modeled as a sequence of stationary GARCH processes with varying parameters.

1.6 Misused statistical tools behave awfully

Our perspective on the time series analysis of long log-return data sets can be summarized as follows: statistical tools and procedures (such as the sample autocorrelations, parameter estimators, periodogram) gather meaningful information and perform the tasks for which they were designed only under certain assumptions (stationarity, light tails, ergodicity, etc.). Hence, if one or several of these assumptions are violated by the data at hand, the reading of these tools as well as the output of our procedures are rendered meaningless. More colorfully, when used inappropriately, the statistical tools and procedures could "see things that are not there".

Let us briefly illustrate this point in relation to the firstly mentioned *stylized fact* that we reformulate in a more specific form: *power law decay of the sample ACF of the absolute values of log-returns is an indication for long memory in the volatility of log-return process.* We simulated realizations $(X_t)_{t=1,\dots,1000}$ and $(Y_t)_{t=1,\dots,1000}$ from two independent GARCH$(1,1)$ processes with corresponding parameters

$$\alpha_0 = 0.13 \times 10^{-6}, \quad \alpha_1 = 0.11, \quad \beta_1 = 0.52\,, \tag{1.6}$$

$$\alpha_0 = 0.17 \times 10^{-6}, \quad \alpha_1 = 0.20, \quad \beta_1 = 0.65. \tag{1.7}$$

The innovations are standard normal. The top left graph in Figure 1.1 displays the two time series concatenated. The difference in unconditional variance of the two pieces is clearly noticeable. The top right and bottom left graphs display the sample ACFs of the time series $|X_1|, \dots, |X_{1000}|$ and $|Y_1|, \dots, |Y_{1000}|$, respectively. For the time series $X_1, \dots, X_n$, the *theoretical autocorrelation (ACF)* function is defined as

$$\rho_X(h) = \mathrm{corr}(X_0 X_h), \quad h \in \mathbb{Z}$$

while the *sample ACF* is defined as

$$\widehat{\rho}_X(h) = \widehat{\gamma}_X(h)/\widehat{\gamma}_X(0) \quad \text{where} \quad \widehat{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n-h} X_t X_{t+h}\,, \quad h \in \mathbb{Z},.$$

GARCH$(1,1)$ processes are strongly mixing with geometric rate and, hence, the theoretical ACF of the absolute values is well defined (for our parameter choices) and decays to zero exponentially fast. Under the choice of parameters (1.6) and (1.7), the 4th moments of X and Y are finite. Therefore the sample ACFs converge to the theoretical ones at $\sqrt{n}$-rate and the asymptotic limits are normal; see Mikosch and Stărică [50]. Here we are using the sample ACF tool in a proper way and the readings are meaningful: the sample ACF decays to zero quickly, as expected.

The bottom right graph in Figure 1.1 displays the sample ACF for the juxtaposition of the two pieces, i.e. for $|X_1|, \dots, |X_{1000}|, |Y_1|, \dots, |Y_{1000}|$. Long range dependence type behavior of the sample ACF develops even though we know there is no long memory in the data. The explanation

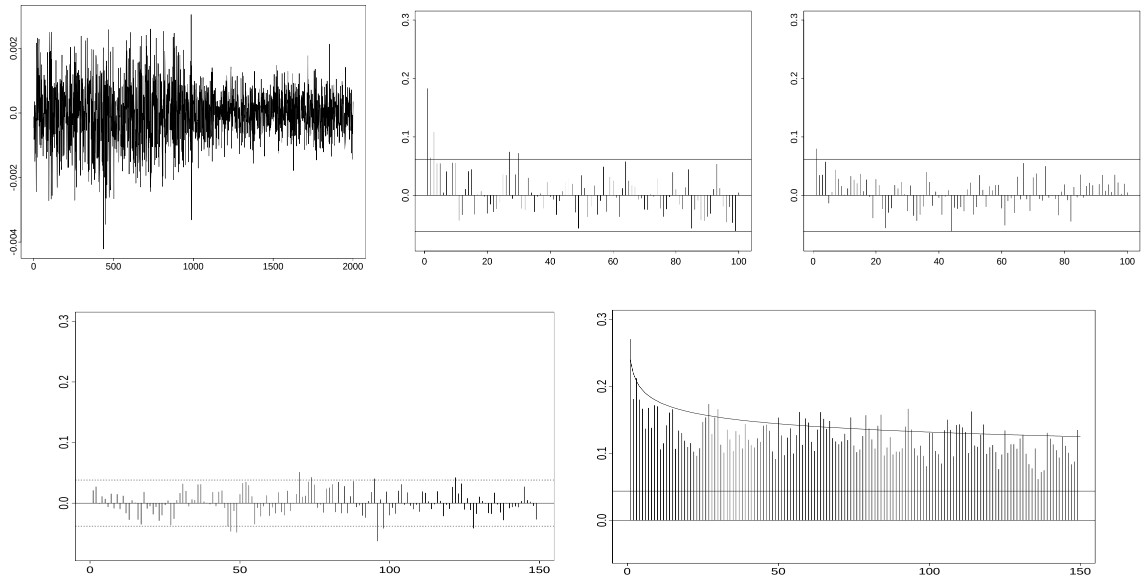

Figure 1.1 Top left: *Concatenation of two* GARCH(1, 1) *time series with parameters* (1.6) *and* (1.7). *The unconditional variances are different.* Top center: *Sample ACF of the absolute values of the first 1000 values.* top left: *Sample ACF of the absolute values of the second 1000 values.* Bottom left: *Sample ACF for the whole original time series. The returns are uncorrelated.* Bottom right: *Sample ACF for the absolute values of the whole time series. The hyperbolic line is given by the function* $\gamma(h) = 0.24h^{-0.13}$. *The horizontal lines in the ACF plots indicate the usual* 95% *confidence bands for the sample ACF of iid Gaussian noise.*

lies in the change of the unconditional variance of the time series as we rigorously proved in Mikosch and Stărică [51]. Inferring the presence of long memory from this sample ACF would be an instance where, through misuse, statistical tools could "see things that are not there". The sample ACF reading of LRD is rendered meaningless by the non-stationarity in the data. This example shows that LRD type behavior of the sample ACF can be caused either by stationary long memory time series or, equally well, by non-stationarity in the time series.

1.7 The everlasting spell of long memory

The field of long memory detection and estimation is particularly (in)famous for the numerous statistical instruments that behave similarly under the assumptions of long range dependence and stationarity or under weak dependence affected by some type of non-stationarity. The following three examples of statistics that are frequently used in the detection and estimation of long memory and that are mired by the mentioned lack of power to discriminate between possible scenarios will certainly raise reader's awareness of the difficulties specific to the this area of statistics.

The first one is probably the most famous, as it is related to the very phenomenon that brought the issue of long range dependence to the forefront of statistical research, i.e the *Hurst effect*. This effect is defined in terms of the asymptotic behavior of the so-called R/S (range over standard deviation) statistic. This statistic was proved to have the same kind of asymptotic behavior when applied to a stationary long memory time series or to a short memory time series perturbed by a small monotonic trend that even converges to 0 as time goes to infinity (Bhattacharya et al. [9]). See Anderson and Turkman [1] for another instance where a long memory type behavior for the R/S statistic of a stationary 1-dependent sequence is proved.

The second tool deals with one of the traditional methods used in the detection of long memory. It examines the sample variance of the time series at various levels of aggregation. Under stationarity and long memory, the variance of the aggregated series

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \quad m \geq 1, \quad k = 1, \ldots, [n/m]$$

behaves asymptotically like a power of the aggregation level m:

$$\mathrm{var}(X^{(m)}) \sim cm^{\beta},$$

where $\beta = 2H - 2$ and H is the so-called *Hurst exponent.* This behavior suggests to take the slope of the regression line of the plot of the logarithm of the sample variance versus $\log m$ as an estimator of β. However, Teverovsky and Taqqu [63] showed that this estimator performs similarly when applied to a long memory stationary time series or to a stationary short memory one that was perturbed by shifts in the mean or small trends. This could make one believe that there is long range dependence in the data when in reality a jump or a trend is added to a series with no long range dependence.

In the more specialized context of long memory detection in the volatility of log-returns, the results of a simulation study in de Lima and Crato [25] give clear evidence of the unreliability of the methods for detecting long memory which are commonly used in the econometrics literature. In the first step of the study, the Geweke–Porter–Hudak procedure and the R/S statistic (adjusted in the spirit of Lo [48]) were used to document the presence of long memory in the volatility of five log-return time series. In the second step, GARCH$(1,1)$ models were fitted to these time series and a thousand simulated log-return series were generated for each set of parameters. The simulated time series clearly had short memory. The two mentioned methods were then used to test the short memory null hypothesis. *Both tests for long memory clearly reject the null of short memory in the data generated by the* GARCH$(1,1)$ *models.* (It is, however, surprising to see how determined the authors are to "find" long memory in the volatility. At odds with the evidence they provide against the methods under discussion, their conclusion reads: "Using both semi-parametric and non-parametric statistical tests, we found compelling evidence of persistent long run dependence in the squared returns series". However, this fact does not diminish the value of their evidence.)

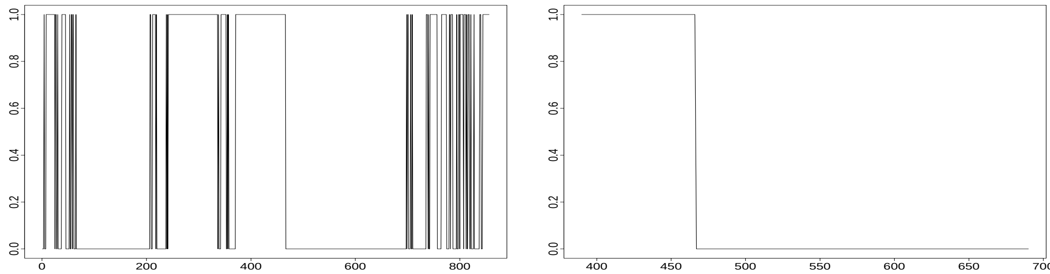

Figure 1.2 Left: X_t, $t = 1, \ldots, 856$. *The long memory stationary assumption is plausible.* Right: X_t, $t = 390, \ldots, 690$. *The hypothesis of structural change is plausible if only a part of the original time series is available.*

In Mikosch and Stărică [51] we showed, theoretically and through examples, that the basic statistical tools for gauging the long memory phenomenon: the sample ACF and the periodogram do also suffer from the mentioned lack of power to discriminate between long memory and non-stationarity. We prove there that the sample ACF and the periodogram for a time series with deterministically changing mean exhibit behavior similar to that implied by the assumptions of stationarity and LRD. Our study shows how a change in the unconditional variance of a time series with short memory causes not only slow decay, but even almost constancy of the sample ACF of absolute and squared log-returns. This behavior translates into the periodogram of the absolute (squared) values: the periodogram has a significant peak at zero.

1.8 One cannot prove God exists. One can only believe in her.

The task of uncovering the phenomena behind the empirical findings of slowly decaying sample ACFs and high persistence of the volatility is rendered more difficult by the fine distinction (many times just a matter of belief) between non-stationarity and LRD stationarity. To illustrate this consider the time series in Figure 1.2. Here the observations X_t are 0 or 1 and the lengths of ON periods (spells of 1s) and OFF periods (spells of 0s) are iid Pareto distributed with tail index 1. It is easy to see that the behavior of the theoretical ACF is: $\text{corr}(X_t, X_{t+h}) \approx ch^{-1}$, hence the time series exhibits long memory. When the time series in the left graph of Figure 1.2 is analyzed, the assumption of stationarity (and long memory) is plausible, while, when only in possession of the observations displayed in the second graph, i.e. observation 390 up to 690, the hypothesis of a structural change is more plausible.

This fine distinction between long memory stationarity and non-stationarity is the source for various competing explanations for the empirical findings under discussion. Our paper postulates *non-stationarity of the unconditional variance* as a possible source of both the slow decay of the sample ACF and the high persistence of the volatility in long log-return time series as measured by ARCH type models. Hence we claim that both these findings in long time series could be spurious. Other studies (Baillie et al. [6]) have argued that the slow decay of the sample ACF correctly reflects the presence of a stationary long memory time series while the apparent integrated GARCH behavior is an artifact of a long memory process investigated through the estimation of a GARCH$(1,1)$ model. While the "correct" explanation behind the empirical facts is still elusive, it is worth mentioning that, while Baillie et al. [6] base their explanation on simulations, we prove our results analytically.

2 A closer look at real-life data

2.1 The LRD effect

Long log-return series (X_t) of foreign exchange rates, stock indices and share prices have the following properties in common:

- The sample ACF $\widehat{\gamma}_X$ of the data is tiny for all lags, save possibly the first ones; the sample mean is not significantly different from zero. This indicates that (X_t) is a white noise process.
- The sample ACFs $\widehat{\gamma}_{|X|}$ and $\widehat{\gamma}_{X^2}$ of the absolute values and their squares
  - are all positive,
  - decay fast for the first few lags,
  - remain "almost constant" for larger lags.

  This is what we call the *LRD effect*.

The first mentioned empirical property of the sample ACF for the data fits nicely with the fact that the X_t's from a GARCH(p,q) process are uncorrelated (provided their second moment exists). Recalling that the GARCH(p,q) process actually has exponentially decaying autocorrelations, we may doubt that a GARCH process can capture the particular behaviour of the sample ACFs of the real-life $|X_t|$'s and X_t^2's described as the *LRD effect*.

In order to illustrate the mentioned "stylized sample ACF facts" we consider the daily log-returns of the Standard & Poor's 500 composite stock index from January 2, 1953, to December 31, 1990. The sample ACF of the log-returns and their absolute values (called *absolute log-returns* in what follows) are displayed in Figure 2.2. The same data set will be used in the sequel to substantiate most of our statements.

2.2 The IGARCH effect

The estimation of GARCH processes on log-return data produces with regularity the following results:

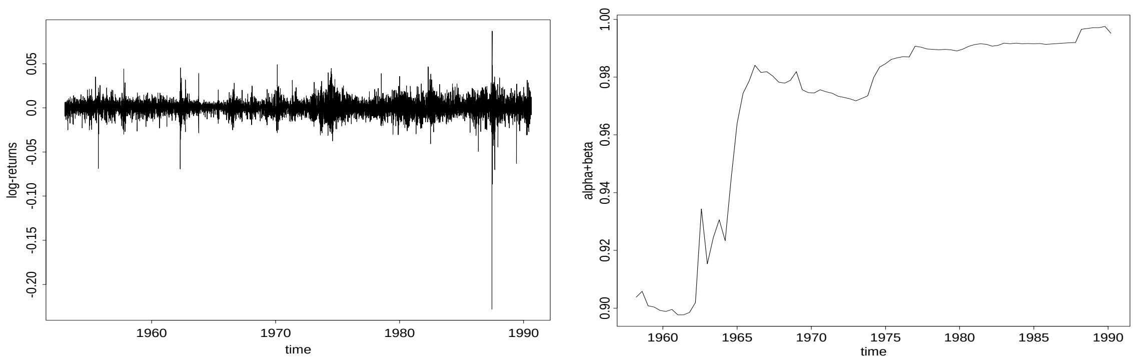

Figure 2.1 Left: *Plot of* 9558 S&P500 *log-returns. The year marks indicate the beginning of the calendar year.* Right: *The estimated values of* $\alpha_1 + \beta_1$ *for an increasing sample of* S&P500 *log-returns. An initial* GARCH(1,1) *model was estimated on the first* 1500 *observations. Then* α_1 *and* β_1 *were re-estimated on increasing samples of size* $1500 + k * 100$, $k > 0$. *The labels on the x-axis indicate the date of the latest observation used for the estimation procedure. The graph shows how the* IGARCH effect *builds up when the sample size increases.*

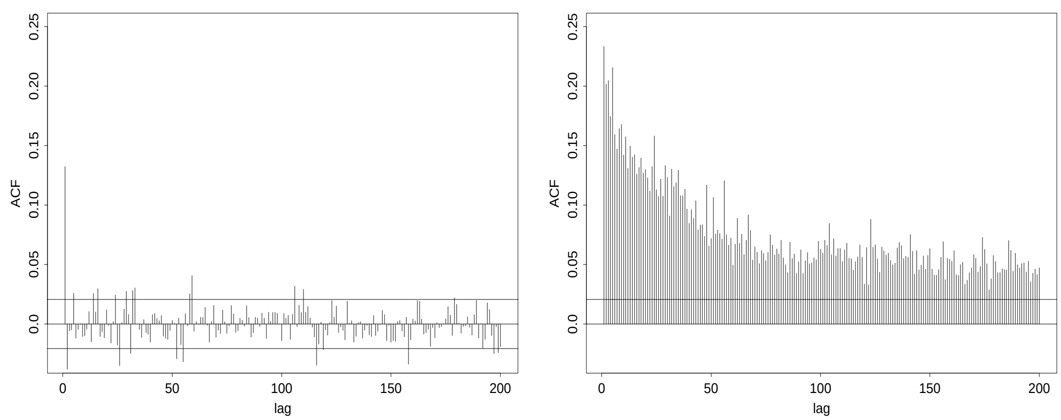

Figure 2.2 Left: *Sample ACF for the* S&P500 *log-returns. Here and in what follows, the horizontal lines in graphs displaying sample ACFs are set as the* 95% *confidence bands* ($\pm 1.96/\sqrt{n}$) *corresponding to the ACF of iid Gaussian white noise.* Right: *Sample ACF for the* S&P500 *absolute log-returns.*

- For longer samples, the estimated parameters $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ of the model (1.1), (1.3) sum up to values *close* to one.

- When shorter subsamples are used for estimation, the sum of the coefficients, although not small, *stays away* from 1.

We will refer to these two regularities as the *IGARCH effect.* Figure 2.1 illustrates the IGARCH(1, 1) *effect* with the S&P500 data.

This stylized fact motivated the introduction of the *integrated* GARCH(p, q) (IGARCH(p, q)) process by Engle and Bollerslev [32] as a possible generating process for log-returns

$$\alpha_0 > 0 \quad \text{and} \quad \sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k = 1\,. \tag{2.1}$$

Under the assumptions given above, in particular $EZ^2 = 1$, the IGARCH model has a strictly stationary solution, but the X_t's do not have a finite second moment. To see this take expectations in the defining equation (2.1) and note that $E\sigma^2 = EX^2$:

$$E\sigma^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \; EX^2 + \sum_{j=1}^{q} \beta_k \; E\sigma^2 = \alpha_0 + E\sigma^2\,.$$

Since $\alpha_0 > 0$ is necessary for strict stationarity, $E\sigma^2 = \infty$ follows. For an IGARCH(1,1) process, if the distribution of Z satisfies some mild assumptions (such as the existence of a density with infinite support), it follows from a classical result of Kesten [43] (see also Goldie [36] for an alternative proof) that

$$P(X > x) \sim c x^{-2}\,, \quad x \to \infty, .$$

We refer to Mikosch and Stărică [50] for details and further references.

At this point it is important to notice that

the IGARCH model and the LRD notion are incompatible.

Indeed, our definition of *LRD* in terms of the ACF is not applicable since the ACF is not well defined. Thus, *if* the IGARCH model was correct, in particular the variance infinite, the sample ACFs of (X_t), $(|X_t|)$ and (X_t^2) would estimate nothing meaningful. A plausible explanation of the empirically observed *LRD effect* would then be:

> If *the IGARCH model is the generating process of the log-returns, the* LRD effect *has nothing to do with* LRD*; it is simply an artifact since the sample ACFs do not measure anything.*

The fact that both the LRD and the IGARCH effects *only* become apparent in long time series raises the question we believe to be central to the understanding of the issue at hand:

Is it possible that both, the LRD *and the* IGARCH effects*, are caused by the same simple reason:*

non-stationarity of the data?

A possible answer is given in the next sections.

3 How long it takes until your GARCH(1,1) model fails you

Another title for this section could very well be *In what way are recessions different from the normal economic tempo?* and its purpouse is two fold. First we want to prove statistically that one GARCH$(1,1)$ process cannot describe well a return time series for long intervals of time. The parameters of the model must be updated periodically. The second one is to investigate the reason for this, i.e. to identify the main cause that induces a GARCH$(1,1)$ model that fit the data for a certain period to stop describing it appropriately. Our findings seem to indicate that the culprits are *the changes in the unconditional variance* of the time series and that these changes are associated mainly with *economic recessions.*

In order to verify in which period of time a GARCH$(1,1)$ model gives a good fit to real-life data we constructed the following statistic:

$$S_n := \sqrt{n} \sup_{\lambda \in [0,\pi]} \left| \sum_{h=1}^{n-1} \frac{\widehat{\gamma}_X(h)}{[\mathrm{var}(X_0X_h)]^{1/2}} \frac{\sin(\lambda h)}{h} \right| \tag{3.1}$$

where

$$\widehat{\gamma}_X(h) = \frac{1}{n} \sum_{t=1}^{n-h} X_t X_{t+h}\,, \quad h = 0,1,2,\ldots$$

is the *sample ACF* function.

Under the assumption that (X_t) comes from a GARCH$(1,1)$ model with given parameters α_0, α_1 and β_1, Mikosch and Stărică [51] have calculated the approximate distribution of the statistic S_n. This approximate distribution can be used to decide if the data is generated by a certain GARCH$(1,1)$ model or not in the following way. Assume you want to verify that your favorite time series $X_1,\ldots,X_n$ is well described by your favorite GARCH$(1,1)$ model that has coefficients, say, $\alpha_0 = 0.13 \times 10^{-6}$, $\alpha_1 = 0.11$, $\beta_1 = 0.72$ and $EZ^4 = 3.8$. The theoretical quantities $\mathrm{var}(X_0X_h)$ are functions of α_0, α_1, β_1 and EZ^4. Hence they can be explicitly calculated. For your favorite GARCH$(1,1)$ model for example, $\mathrm{var}(X_0X_1) = 0.54$, $\mathrm{var}(X_0X_2) = 0.45$, $\mathrm{var}(X_0X_3) = 0.37$, etc. The quantities $\widehat{\gamma}_X$ are directly calculated from your favorite time series. Hence a numerical value of the statistic is produced. If the data comes indeed from your favourite GARCH$(1,1)$ process, the value of the statistic should be in the range described by the approximate distribution calculated in Mikosch and Stărică [51]. If in fact the data is generated by some other process, then the value of the statistic will be much bigger (smaller) than what the approximate distribution predicts as a possible range of values.

We want to emphasize that the statistic S_n (3.1) is designed to be sensitive to changes in the variance of the time series. Hence deviations of the statistic from the range prescribed by the theory signal, most often, that the log-returns are more volatile than data that your favourite model would produce. Doing this type of analysis on a window that moves through the data allows for revealing the time periods where the volatility of the returns is higher than usual.

In Figure **??** we show how one can apply S_n in order to detect changes in the GARCH structure of the S&P500 log-return series. A GARCH(1,1) model was fitted using quasi-maximum likelihood estimation (see for example Gourieroux [37]) to the first 3 years of the data (750 observations), yielding the following parameters by:

$$\alpha_0 = 8.58 \times 10^{-6}, \quad \alpha_1 = 0.072, \quad \beta_1 = 0.759\,, \tag{3.2}$$

and an estimated 4th moment for the residuals of 3.72. These quantities are used for calculating $v_X(h) = \mathrm{var}(X_0 X_h)$.

A closer look at the left-hand graph of Figure 2.1 together with the left-hand graph in Figure **??** reveals an almost one-to-one correspondence between the periods of larger absolute log-returns (larger volatility) and the periods when the goodness of fit test statistic S_{125} falls outside the confidence region. This observation has theoretical grounding since the statistics S_n is sensitive to changes in the model mainly through changes in the variance σ_X^2 of the data. Therefore, one can identify the excursions of the statistics S_{125} above the 99% quantile threshold with periods of higher data volatility then that of the fitted GARCH(1,1) model.

It is then interesting to verify whether a periodically updated GARCH(1,1) can describe the pattern of changing unconditional variance in the data. Towards that goal we calculate and display in the right-hand graph of Figure **??** the implied unconditional GARCH(1,1) variance for a periodically re-estimated GARCH(1,1) model. Recall that the variance of a GARCH(1,1) process is given by

$$\sigma_X^2 = \alpha_0/(1-(\alpha_1+\beta_1)).$$

The calculations of the values in Figure **??** are based on the estimated parameters α_1 and β_1.

More concretely, we fitted a GARCH(1,1) model every 6 months, i.e. every 125 days, based on a moving window of 508 past observations, equivalent to roughly two years of daily log-returns. We then plotted the implied variance σ_X^2 corresponding to every 6 months period. One notices that the pattern of increased implied variance is quite similar to the pattern of the excursions of the statistic S_{125} above the 99% quantile threshold. For us, this similarity seems to imply that one can capture the changing patterns of volatility present in the data by periodically updating the GARCH(1,1) model.

As a conclusion, the graphs in Figure **??** show quite convincingly that:

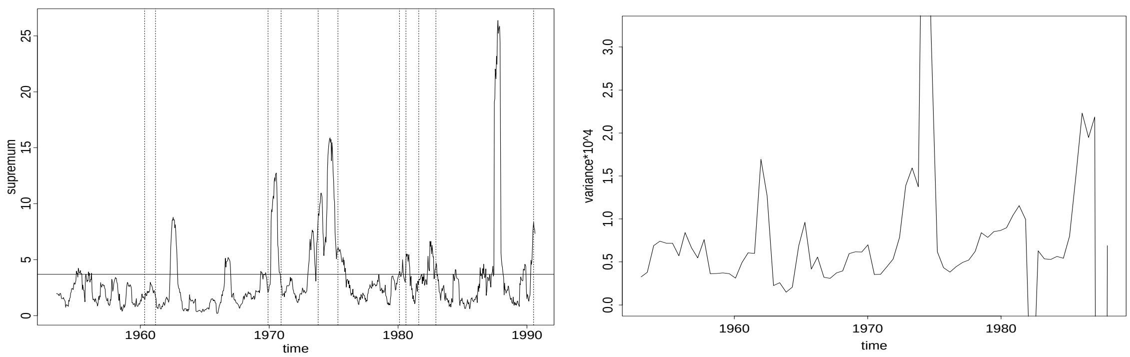

Figure 3.1 Left: *The statistic* S_{125} *calculated weekly based on previous* 125 *observations (approximately* 6 *months) of the* S&P500. *The horizontal line, set at* 3.6*, corresponds to the* 99% *quantile of the approximate distribution of* S_n *(i.e. 99 out of 100 values of the statistic* S_n *calculated on samples that come from the* GARCH(1, 1) *model specified above are smaller than* 3.6*). Values above* 3.6 *correspond to* 6 *months periods when the hypothesised model is inappropriate. Recall that high values of the statistic* S_n *correspond to higher unconditional variance than that of the hypothesized model. The dotted vertical lines correspond to the beginning and the end of economic recessions as determined by the NBER. They nicely show the coincidence between the recession periods and the intervals of higher unconditional variance detected by our tool.*
Right: *The implied* GARCH(1,1) *unconditional variance of the* S&P500 *data. A* GARCH(1,1) *model is estimated every* 6 *months using the previous* 2 *years of data. The graph displays the variances* $\sigma_X^2 = \alpha_0/(1-\alpha_1-\beta_1)$*. The similarities between the two graphs seem to show that a frequently re-estimated* GARCH(1,1) *model captures to a certain extent the changing unconditional variance of the log-returns.*

- *One particular GARCH process is a good model for the log-return time series only for a relatively short period of time.*

- *Periodically re-estimating a* GARCH(1, 1) *process produces models that seem to capture the changes in variance.*

- *The unconditional variance of the data changes through time.* In particular, the periods of recession are characterized by a higher variance of the returns, i.e. markets are more volatile during economic recessions.

It is the last conclusion that motivates our next step: understanding the impact of the type of non-stationarity evident in the log-return data on statistical instruments and procedures. Statistical tools such as the *sample ACF* and the *periodogram* together with the results of *parametric model estimation* are understood and can be interpreted meaningfully *only under the assumption that the underlying data is a strictly stationary process.* Violations of this assumption, of the type we see in the log-return data, render questionable the interpretation of the behavior of these tools and procedures, i.e. the *LRD effect* and the *IGARCH effect.*

4 Statistical instruments and procedures under non-stationarity

Before discussing in detail the effects of non-stationarity on various statistical tools and procedures let us introduce the *spectral density* and its sample version, *the periodogram.* Under general conditions, γ_X, the theoretical ACF of the time series $X_1, \ldots, X_n$ can be writen as

$$\gamma_X(h) = \int_{(-\pi,\pi]} e^{ih\lambda} f_X(\lambda)\, d\lambda, \quad h \in \mathbb{Z}.$$

The function $f_X(\lambda)$ is called the *spectral density* at frequency λ. It is a well known fact that power law decay of the ACF translates into power law behavior of the spectral density for small frequencies (for a precise formulation of this statement, the ACF has to satisfy some subtle conditions; we refrain from discussing them here.) The sample version of the spectral density is the *periodogram*

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

which is the natural (method of moment) estimator of the spectral density f_X of the stationary sequence (X_t) at frequency λ; see Brockwell and Davis [18] or Priestley [58].

We now turn to discussing the effects changes in variance might have on statistical tools like the sample ACF and the periodogram and on statistical estimation procedures. In [51] we considered a time series

$$Y_1^{(1)}, \ldots, Y_{[np]}^{(1)}, Y_{[np]+1}^{(2)}, \ldots, Y_n^{(2)},$$

where $p \in (0,1)$ is fixed. The two pieces of this time series come from distinct stationary ergodic models. One can easily show that as $n \to \infty$, the sample ACF at lag h converges:

$$\widehat{\gamma}_Y(h) \xrightarrow{P} p\, \gamma_{Y^{(1)}}(h) + (1-p)\, \gamma_{Y^{(2)}}(h) + p(1-p)\, (EY^{(1)} - EY^{(2)})^2 \,. \tag{4.1}$$

If the two subsamples are also uncorrelated and $\lambda_j = 2\pi j/n$, $j = 1, 2, \ldots$, then the periodogram $I_Y(\lambda_j)$ satisfies as $n \to \infty$

$$EI_Y(\lambda_j) \sim p\, 2\pi f_{Y^{(1)}}(\lambda_j) + (1-p)\, 2\pi f_{Y^{(2)}}(\lambda_j) + \frac{2}{n\lambda_j^2}(EY^{(1)} - EY^{(2)})^2(1 - \cos(2\pi j p)). \tag{4.2}$$

The sample ACF. More concretly let us now take a look at what happens when dealing with a sample that consists of subsamples $X_t^{(1)}, t = 1, \ldots, [np]$, and $X_t^{(2)}, t = [np]+1, \ldots, n$ from two GARCH$(1,1)$ processes with *different* unconditional variance (ths also implies $E|X^{(1)}| \neq E|X^{(2)}|$). Applying (**??**) to first the data itself and second the absolute values of the data reveals very different behaviors of the sample ACF. For the data itself, since all the variables have mean zero, (**??**) implies that

$$\widehat{\gamma}_X \quad \xrightarrow{P} \quad p\, \gamma_{X^{(1)}}(h) + (1-p)\, \gamma_{X^{(2)}}(h) = 0\,,$$

Hence *the sample ACF estimates zero at all lags;* see Figure 1.1, bottom-left graph. For the absolute values and squares of the time series one gets

$$\widehat{\gamma}_{|X|} \quad \xrightarrow{P} \quad p\, \gamma_{|X^{(1)}|}(h) + (1-p)\, \gamma_{|X^{(2)}|}(h) + p(1-p)(E|X^{(1)}| - E|X^{(2)}|)^2 \,.$$

Since the theoretical ACF of the absolute values of a GARCH process decays to zero exponentially the terms

$$p\, \gamma_{|X^{(1)}|}(h) + (1-p)\, \gamma_{|X^{(2)}|}(h)$$

decay to zero at an exponential rate. Hence the sample ACF will decay fastly at the first few lags. This is indeed in agreement with the behavior of the sample ACF of the absolute values of real log-return series. The typical shape of the sample ACF at large lags of such a time series is however determined by the constant term

$$p(1-p)\, (E|X^{(1)}| - E|X^{(2)}|)^2 \,.$$

which forces the sample ACF to stay positive and almost constant for a large number of lags. This term that reflects the difference in the unconditional variances in the two subsamples is responsible for the *LRD effect* in the absolute log-returns; see Figure 1.1, bottom-left graph.

The Periodogram. The discussion on the periodogram follows a similar trajectory. Applying (**??**) to the sample $X_1^{(1)}, \ldots, X_{[np]}^{(1)}, X_{[np]+1}^{(2)}, \ldots, X_n^{(2)}$ yields

$$EI(\lambda_j) \quad \sim \quad p\ 2\pi f_{X^{(1)}}(\lambda_j) + (1-p)\ 2\pi f_{X^{(2)}}(\lambda_j) = p\ \mathrm{var}(X^{(1)}) + (1-p)\ \mathrm{var}(X^{(2)})\,.$$

Thus the periodogram of the data estimates a constant; see Figure **??**.

Now consider the expected periodogram at small Fourier frequencies for the absolute values of the sample. From (**??**),

$$EI_{|X|}(\lambda_j) \quad \sim \quad p\ 2\pi f_{|X^{(1)}|}(\lambda_j) + (1-p)\ 2\pi f_{|X^{(2)}|}(\lambda_j) + \frac{2}{n\lambda_j^2}(E|X^{(1)}| - E|X^{(2)}|)^2(1-\cos(2\pi jp))\,.$$

The situation is similar to that we encountered in the case of the sample ACF. The first two terms are continuous functions in λ and hence bounded in a neighbourhood of 0.

$$p\ 2\pi f_{|X^{(1)}|}(\lambda_j) + (1-p)\ 2\pi f_{|X^{(2)}|}(\lambda_j) \to p\ 2\pi f_{|X^{(1)}|}(0) + (1-p)\ 2\pi f_{|X^{(2)}|}(0) = \text{constant}.$$

It is the third term (whose presence is due to the difference between the variances of the two subsamples) that determines the shape of the periodogram. Let us take a closer look at its behavior in a neighborhood of the origin. As $n\lambda_j^2 \to 0$ (p is assumed to be r_1/r_2, a ratio of two relatively prime integers)

$$\frac{2}{n\lambda_j^2}(E|X^{(1)}| - E|X^{(2)}|)^2(1-\cos(2\pi j\ r_1/r_2))$$

takes large values for all frequencies for which $1-\cos(2\pi j\ r_1/r_2) \neq= 0$. This will create the impression of a spectral density with a singularity at zero, which is also the mark of stationary LRD processes; see Figure **??**.

The IGARCH effect. Finally, it is also possible to show that the *IGARCH effect* might be due to non-stationarity as well. In [51] we proved that if, as above, we assume that the sample consists of pieces from different GARCH(1,1) models, the Whittle estimate $\hat{\varphi}_1^W$ of $\varphi_1 = \alpha_1 + \beta_1$ behaves as

$$\hat{\varphi}_1^W \sim 1 - \frac{c_1}{c_2 + [\mathrm{var}(X^{(1)}) - \mathrm{var}(X^{(2)})]^2}\,,$$

where c_1 and c_2 are positive constants depending on the coefficients of both GARCH models. This implies that the larger the difference between the variances of the two models, the closer to one the estimate of φ_1. This might explain the *IGARCH effect* since the longer the time series the larger the chance that strong non-stationarity will affect it and hence the closer to one the estimated value of φ_1 (see Figure 2.1 for an example of the way the *IGARCH effect* builds up in longer time series).

The impact of the change of regimes in the simulated data set on the estimation of α_1 and β_1 is ilustrated in Figure **??**. We used quasi-maximum likelihood as, for example, proposed in Gourieroux [37]. GARCH(1,1) models have been fitted to the increasing samples

$X_1, \ldots, X_{150+t*50}$, $t = 1, \ldots, 38$. For sample sizes less then 1000, the sum of the theoretical parameters is 0.85. The estimated sum varies for these sample sizes between 0.75 and 0.90. The graph in Figure **??** clearly shows how the switch of regimes (which happens at $t = 1000$) makes the sum increase to 1. This is in agreement with the theory in Mikosch and Stărică [51] since we expect the quasi-maximum likelihood estimator of $\alpha_1 + \beta_1$ to behave similar to the Whittle estimate.

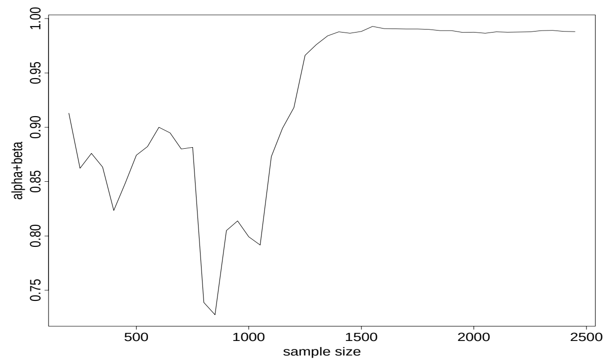

Figure 4.1 *The estimated values $\alpha_1 + \beta_1$ for a* GARCH(1, 1) *model fitted to an increasing sample from the simulated data in Figure 2.2. The labels on the x-axis indicate the size of the sample used in the quasi-maximum likelihood estimation.*

5 Estimation of the long memory parameter

Let us now consider the issue of statistical estimation of the so called *Hurst exponent*. This quantity, commonly denoted by H has been proposed in the literature as a measure of LRD; see Beran [8] for details on the definition, properties and statistical estimation of H. In the sequel we discuss two statistical estimation procedures for H.

If one assumes that the theoretical ACF $\rho(h)$ of the time series has a hyperbolic decay rate, i.e. $\rho(h) \approx ch^{-\beta}$ for some positive β and c, the Hurst coefficient is usually determined as $H = 1 - \beta/2$. In particular, the presence of LRD in the time series is signalled if $H \in (0.5, 1)$. In this case, the sequence $(\rho(h))$ is not absolutely summable. The closer H to 1, the further the dependency reaches. An estimation procedure for H is then suggested by the following argument. Since

$$\log \rho(h) \approx \log c - \beta \log h$$

and the sample ACF estimates the theoretical one, a log-log plot of the lags versus the sample ACF should be roughly linear, the slope of the regression line yielding an estimate of the quantity β, hence of H.

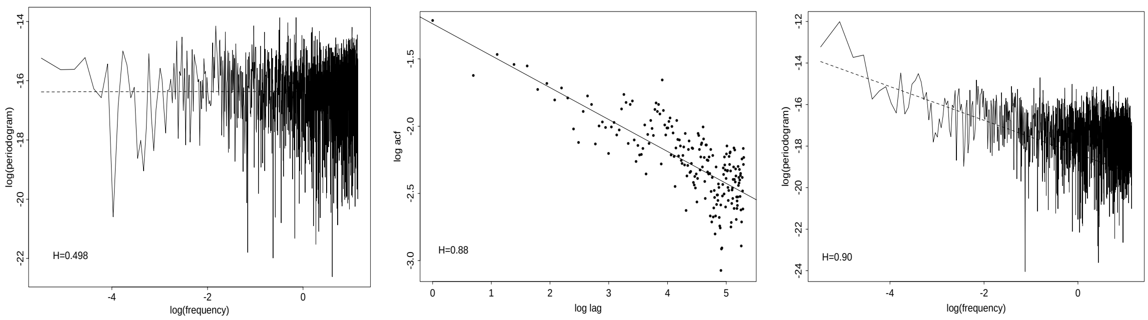

Figure 5.1 Left: *Log-log fit of the periodogram for the simulated data displayed in Figure 1.1.* Center: *Log-log fit of the sample ACF for the absolute values of the simulated data displayed in Figure 1.1. Estimated Hurst coefficient $H = 0.88$.* Left: *Log-log fit of the periodogram for the absolute values of the same sample. Estimated $H = 0.90$.*

Figure **??** shows in more detail the misleading effect the non-stationarity paradigm can have on statistical estimation of the *Hurst parameter*. We applied the procedure described above to the simulated data displayed in Figure 1.1 and to its absolute values. The central graph in Figure **??** displays the fit of a regression line through the plot of the log lags versus the log sample ACF. The slope is -0.23, the intercept -1.3. Hence $\hat{\rho}(h) = 0.27h^{-0.23}$. The resulting Hurst coefficent is $H = 1 - 0.23/2 = 0.88$. This would imply a strong LRD effect if the data came from a stationary sequence. The actual fit of the hyperbolically decaying function to the sample ACF is illustrated in the right-hand side graph of the second row of Figure 1.1 and is particularly noteworthy. A good fit of a hyperbolic decay line to the sample ACF is often used in the econometrics literature as an argument for the presence of LRD in the volatility process (see for example Anderson and Bollerslev [2]). Our example shows how misleading this approach could be. It says that a good hyperbolic fit might have nothing to do with the presence or absence of LRD in the volatility and it can be just a byproduct of changing unconditional variance.

Another popular estimation procedure uses the periodogram and it is based on the well known fact that power law decay of the ACF translates into power law behavior of the spectral density) for small frequencies. Hence, frequently used statistical tests for detecting LRD and measuring its strength are based, one way or the other, on the following behavior of the spectral density:

$$f_{|X|}(\lambda) \approx c\lambda^{2H-1},$$

for small $\lambda > 0$. Equivalently, $\log f(\lambda)$ is linear in λ,

$$\log f_{|X|}(\lambda) \approx \log c + (2H - 1)\log \lambda.$$

Extrapolating this relationship to the sample, on a log-log plot, the periodogram should roughly exhibit a linear behavior, the slope of the line yielding an estimate of H. (For example, using a regression to capture the linear dependency yields the ubiquitous Geweke–Porter–Hudak estimator [33].)

We apply this procedure to our non-stationary simulated sequence. The graphs on the sides of Figure **??** show the plot of the log frequencies versus the log periodogram with a regression line fit to the first 10% of the frequencies (discarding the first $3000^{0.2} = 5$ lowest ones) for the data (left-hand graph) and for absolute values of the data (the right-hand graph). The regressions yield an estimate $H = 0.498$ for the data and $H = 0.90$ for the absolute values of the data. Interestingly enough, both methods give similar values for the Hurst coefficient of the absolute values of the data and suggest that there is a strong LRD effect in the data.

This concordance in the values yielded by the time and the frequency domain methods is sometimes used as an additional argument supporting the existence of LRD (see, for example Anderson and Bollerslev [2]). Our example shows that this concordance might have nothing to say about the presence of LRD. It serves as a warning sign and shows that linear behavior in the log periodogram for low frequences can develop for other reasons than LRD.

The estimates for H are clearly subject to statistical uncertainty. *It is not our primary purpose to give confidence bands for the estimation of H; all we intend to show here is that standard estimation procedures for stationary time series, when applied to a non-stationary sequence, may give misleading answers as to whether there is LRD in the data.*

The above discussion shows:

Non-stationarity of a time series could be responsible for the spurious LRD effect *in the behavior of the* sample ACF *and the* periodogram *of absolute log-returns.*

6 A study of the Standard & Poors 500 series

Now we proceed to analyse a time series that has been previously used to exemplify the presence of LRD in financial log-return series: the Standard 90 and Standard and Poor's 500 composite stock index. This series, covering the period between January 3, 1928, to August 30, 1991, was used in Ding et al. [28], Granger et al. [38], Ding and Granger [27] for an analysis of its autocorrelation structure. It led the authors to the conclusion that the powers of the absolute values of the log-returns are positively correlated over more than 2500 lags, i.e. 10 years. Hardly any proof is needed to convince one that this time series is likely to be non-stationary. It covers the Great Depression, a world war together with the most recent period, marked by major structural changes in the world's economy. In addition, there was a compositional change in the S&P composite index that happened in January 1953 when the Standard 90 was replaced by the broader Standard and Poor's 500 index. Despite all these, Ding et al. [28] conclude the section which describes the data as follows (page 85): "During the Great Depression of 1929 and early 1930s, volatilities are much higher than any other period. There is a sudden drop in prices on Black Monday's stock market crash of 1987, but unlike the Great Depression, the high market volatility did not last very long. *Otherwise, the market is relatively stable.*"

Bollerslev and Mikkelsen [15] used the daily returns on the Standard and Poor's 500 composite stock index from January 2, 1953, to December 31, 1990 (a total of 9559 observations) to fit a FIGARCH model under the assumptions of stationarity and LRD.

In Mikosch and Stărică [51] we performed a detailed analysis of the same data set covering the time span from January 2, 1953, to December 31, 1990. Contrary to the belief that the LRD characteristic carries meaningful information about the price generating process, we show that the LRD behavior could be just an artifact due to structural changes in the data. We have already used the statistic (3.1) to detect the moments in time when a GARCH$(1,1)$ model estimated on past data stops describing the behavior of the time series; see Figure **??**. Next we document the effect which the switch to a different regime of variance has on the sample ACF. We find that the aspect of the sample ACF changes drastically after episodes of increased variance that cannot be properly described by the estimated model.

In Figure **??** the sample ACFs for the absolute values of the log-returns for the first 9-year and 11-year periods are compared. While the first period's autocorrelations seem to be insignificant after 50 lags, the autocorrelations for the 11-year period are still significant at lag 100. We also note that the size of the significant autocorrelations increases together with the proportion of positively correlated lags. This is indeed consistent with the explanation of this phenomenon provided in Section **??**.

However, from the view point of the behavior of the sample ACF, the most interesting part is the period beginning in 1973 and lasting for almost 4 years. The values of our statistic S_n (3.1) are quite extreme, strongly indicating that the period between 1973 and 1977 is a long interval when the model estimated on the first 1500 observations does not describe the data; see Figure **??**. Let us analyze the changes in the sample ACF caused by this long period of different behavior. Figure **??** displays the sample ACF of the absolute values $|X_t|$ up to the moment when the change is detected, next to the sample ACF including the 4-year period that followed. We see that the sample ACF up to 1973 does not differ significantly from the one based only on the data up to 1964; see Figure **??**. However, the impact of the change in regime between 1973 and 1977 on the form of the sample ACF is extremely strong as one sees in the second graph of Figure **??**. The graph clearly displays the LRD features given by the theoretical explanation of Section **??**: exponential decay at small lags followed by almost constant plateau for larger lags together with strictly positive correlations.

In the end of this section, it is interesting to take a closer look at the behavior of our statistic around the Black Monday crash in October 1987. Following on the observation that a frequently re-estimated GARCH$(1,1)$ model seems to follow better the changing patterns of volatility in the

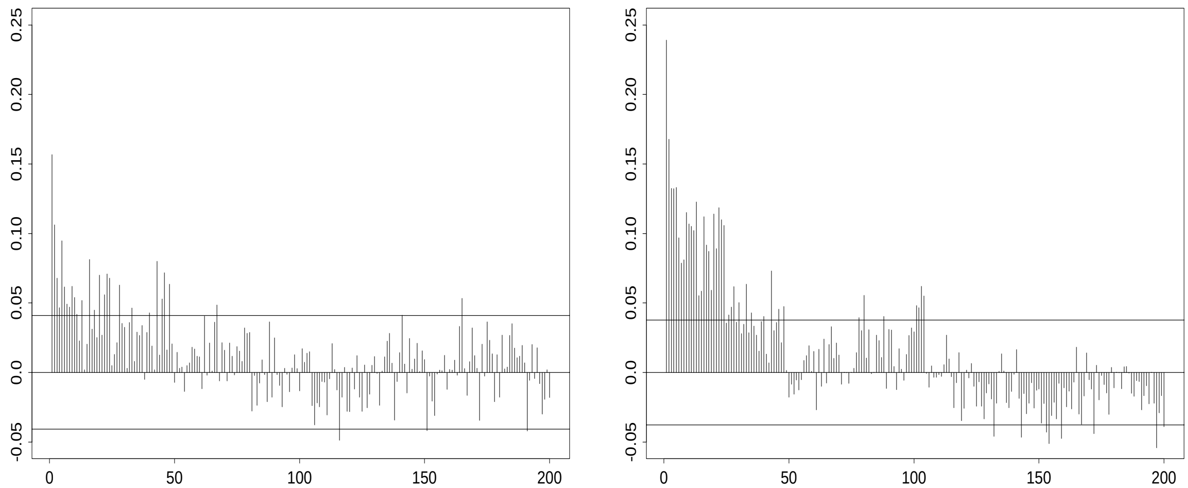

Figure 6.1 *The sample ACF for the absolute values of the log-returns for the first* 9 *years* (left) *and the first* 11 *years of the* S&P *data.*

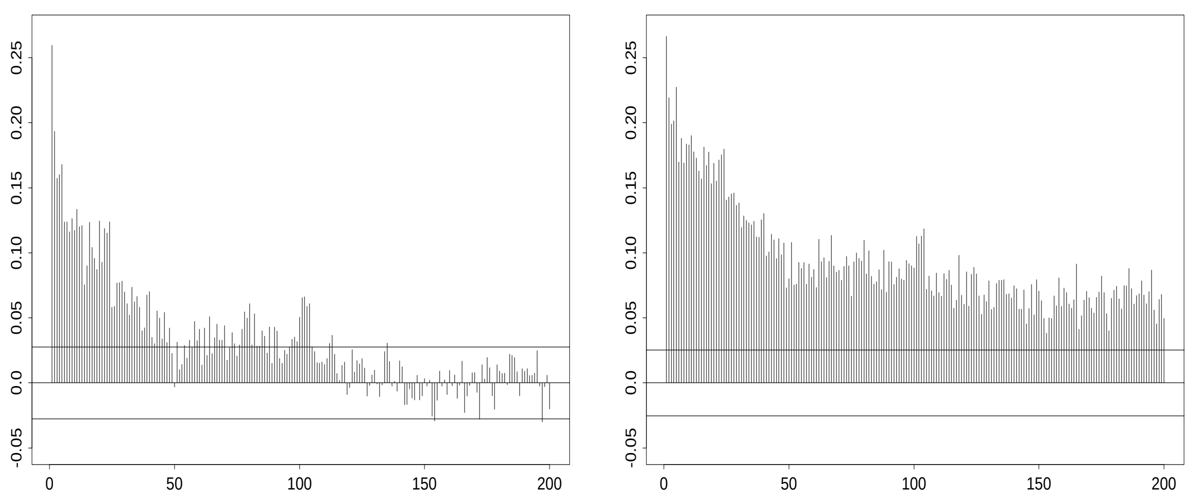

Figure 6.2 *The sample ACF for the absolute values of the log-returns of the first* 20 *years* (left) *and* 24 *years* (right) *of the* S&P *data.*

time series (see Figure **??**), a GARCH(1,1) model is estimated in the beginning of June 1987 using the observations between January 1986 and June 1987 (375 observations). The estimated coefficients

$$\alpha_0 = 16 \times 10^{-6}, \quad \alpha_1 = 0.013, \quad \beta_1 = 0.812 \tag{6.1}$$

together with the 4th moment of the estimated residuals, $E\hat{Z}^4 = 3.4$ are used to build the S_{125} statistic. Figure **??** shows the behavior of the statistic during the 100 days preceding and the 25 following the crash. To allow for a better analysis, the log of the statistic is displayed. One sees

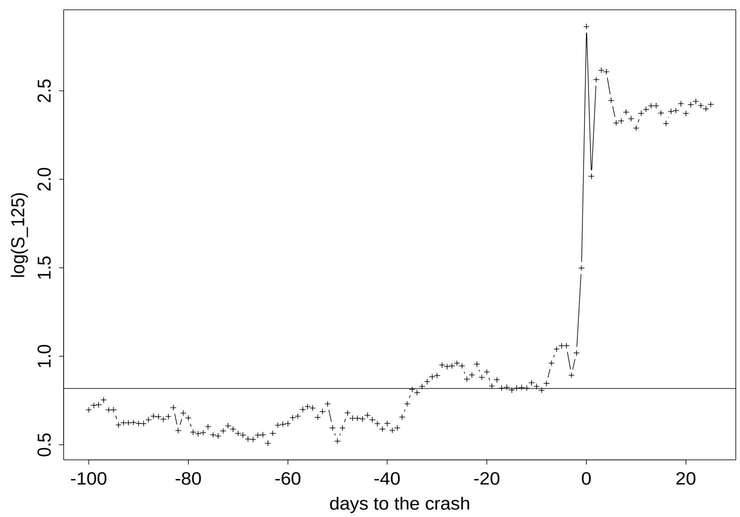

Figure 6.3 *The* $\log(S_{125})$ *statistic before and after the Black Monday 1987 crash. The horizontal line is the log of the* 75% *quantile* ($q_{0.75} = 2.26$) *of the limiting distribution of* S_{125}.

that the statistic reaches above the 75% quantile (2.26) of the limiting distribution 6 weeks before the crash and never falls below. The days just before the crash mark an increase in the statistic. The days -6 up to -2 are above the 90% percentile of the limit distribution ($q_{0.90} = 2.71$), while day -1 is well above the 99% quantile ($q_{0.99} = 3.6$).

7 A study of exchange rates

In this section we focus on log-returns of foreign exchange (FX) data. We consider 2004 daily log-returns from the DEM/USD exchange rate between January 1975 and December 1982. We chose this period for a couple of reasons. This period is marked by two important events in the recent history of monetary policy. In Europe, several central banks improved upon the coordination of their monetary policy around March 1979. In the U.S., the Federal Reserve changed its instrument for controlling the money stock. It switched from using the federal funds rate to targeting the non-borrowed reserve. This change of policy lasted from *October 1979* until *October 1982* and, according to the economics literature (Spindt and Tarhan (1987), p. 107), it was "one of the more dramatic events in the recent history of monetary policy". Since monetary policy changes undoubtly have great impact on exchange rate dynamics, we expect to see the traces of these events in the data set.

Another reason is that data sets from this period which is marked by significant structural changes, have been included in studies conducted by various authors, both in the contexts of GARCH$(1,1)$ modeling and long memory analysis. Baillie and Bollerslev (1989) used daily spot exchange rates for the currencies of France, Italy, Japan, Switzerland, the United Kingdom and Germany against the USD between *March 1, 1980* and *Janury 28, 1985* to fit GARCH$(1,1)$ models with conditional t-distribution. They report estimated values of $\alpha_1 + \beta_1$ close to unity in all cases. Baillie et al. (1996) found evidence of long memory and fitted their FIGARCH model to a time series of daily DEM/USD spot exchange rates from *March 13, 1979* through *December 30, 1992.* Note that these periods contain (at least) the major event of the return of the Federal Reserve to the use of the federal funds rate to control the money stock.

Our analysis shows that both events seem to have produced sensible changes in the dynamics of the exchange rates during this period. In the light of this analysis and of the theoretical developments of the paper, both findings of almost integrated GARCH$(1,1)$ and long memory in the mentioned studies are rendered questionable.

Further evidence for the fact that the mentioned structural shifts could be responsible for the spurious persistence in volatility documented in Baillie and Bollerslev (1989), comes from Cai (1994). This author finds that fitting a GARCH$(1,1)$ to the monthly excess returns of the three-month T-bill from *August 1964* to *November 1991* implies highly persistent volatility ($\hat{\phi}_1 = 0.98$). In contrast to that, fitting a model which allows for shifts in the unconditional variance yields significantly reduced ARCH parameters and, hence, a model with much less persistency in the volatility. The model clearly associates the periods of regime shifts with the oil shock and the Federal Reserve policy change.

Let us now commence our analysis. The left-hand graph of Figure **??** displays the data. Visible changes in the appearence of the data can be detected during 1978, begining of 1979 and end of 1980, 1981. A GARCH$(1,1)$ fit to the first 2 years yields the following parameters:

$$\alpha_0 = 8.3 \times 10^{-6}\,, \quad \alpha_1 = 0.18\,, \quad \beta_1 = 0.77\,, \tag{7.2}$$

and an estimated 4th moment for the residuals of 4.6.

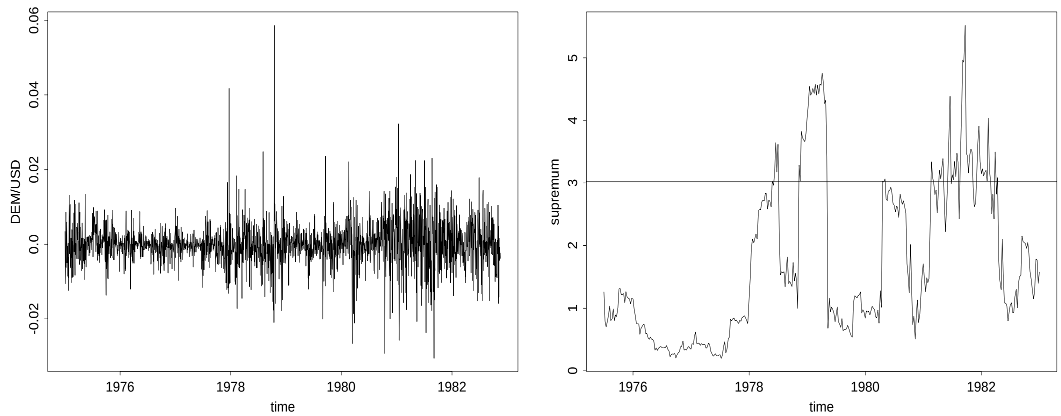

Figure 7.1 Left: *Plot of the* 2032 *DEM/USD log-returns. The year marks indicate the beginning of the calendar year.* Right: *The corresponding goodness of fit test statistics* S_{125}.

7.1 Goodness of fit test and the sample ACF

The right-hand graph of Figure **??** displays the values of the statistic S_n calculated on a weekly basis from the previous $n = 125$ observations which amount to roughly six months. The horizontal line is set at the asymptotic 95% quantile. The graph shows the presence of two intervals where the estimated model clearly does not fit the data: a shorter period of less than a year, covering the beginning of 1979, and a longer period of about one year and a half, covering 1981 and the beginning of 1982. A look at the sample ACF of the absolute values $|X_t|$ before and after the period in the discussion reveals a minor change in the behavior of this statistical instrument; see Figure **??**. Here the first 100 lags of the sample ACF before and after the first episode are displayed.

In contrast to the first episode that could possibly be associated with the mentioned increase in the policy coordination among several European countries, the second period during which the goodness of fit test statistic exceeds the threshold lasts longer. A visual inspection of the graphs in Figure **??** gives the impression that the structure of the time series in the period between spring 1980 and the end of 1981 is different from the remaining observations. The statistic S_{125} confirms this fact by frequently switching sides of the threshold line during this period. The dramatic changes in the behavior of the sample ACF are illustrated in Figure **??** where we can see the first 100 lags of the sample ACF before and after the second episode. Again, in accordance with our explanation, the sample ACF displays exponential decay at small lags followed by almost constant plateau for larger lags together with strictly positive correlations.

Acknowledgment: Cătălin Stărică would like to thank the Department of Mathematics of the University of Groningen and the Dutch Science Foundation (NWO) for financial support. Thomas

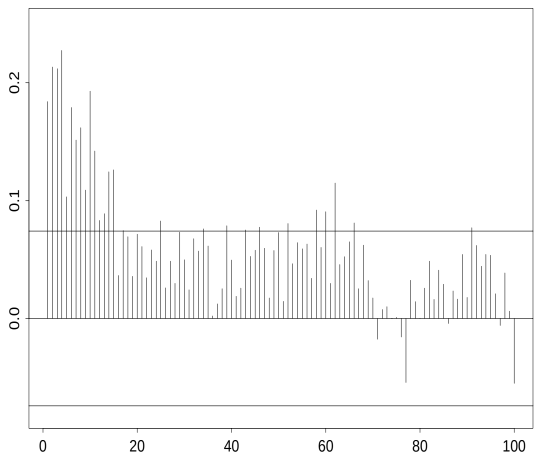

Figure 7.2 *The sample ACF for the absolute values of the log-returns of the DEM/USD FX data up to the beginning of* 1978 (left) *and up to June* 1979 (right)*.*

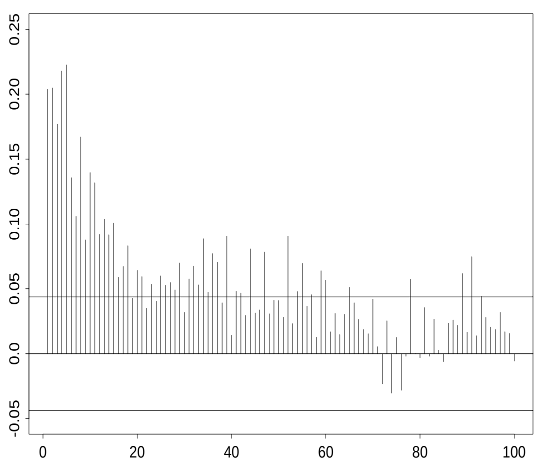

Figure 7.3 *The sample ACF for the absolute values of the log-returns of the DEM/USD FX data up to the beginning of* 1981 (left) *and up to June* 1982 (right)*.*

Mikosch worked on this paper in January 1999 during a visit to the University of Aarhus. He would like to thank his colleagues at the Institute of Statistics, in particular Ole Barndorff-Nielsen, for their warm hospitality and generous support.

References

[1] ANDERSON, C.W. AND TURKMAN, K.F. (1995) Sums and maxima of stationary sequences with heavy tailed distributions. *Sankhya* **57** , 1–10.

[2] ANDERSEN, T. AND BOLLERSLEV T. (1998) Deutsche Mark–Dollar volatility: intraday activity patterns, macroeconomic announcements and longer run dependencies. *J. Finance* **LIII**, 219–262.

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

[4] BAILLIE, R.T. AND BOLLERSLEV, T. (1989) The message in daily exchange rates: a conditional-variance tale. *J. Business and Economic Statist.* **7**, 297–305.

[5] BAILLIE, R.T. AND BOLLERSLEV, T. (1990) Intra-day and inter-market volatility in foreign exchange rates. *Review of Economic Studies* **58**, 565–585

[6] BAILLIE, R.T., BOLLERSLEV, T. AND MIKKELSEN, H.O. (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **74**, 3–30.

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

[8] BERAN, J. (1994) *Statistics for Long–Memory Processes.* Monographs on Statistics and Applied Probability, No. 61. Chapman and Hall, New York.

[9] BHATTACHARYA, R.N., GUPTA, V.K. AND WAYMIRE, E. (1983) The Hurst effect under trends. *J. Appl. Probab.* **20**, 649–662.

[10] BICKEL, P.J. AND WICHURA, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, 1656–1670.

[11] BILLINGSLEY, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

[12] BOES, D. C. AND SALAS-LA CRUZ, J. D. (1978) Non stationarity of the mean and the Hurst phenomenon. *Water Resour. Res.* **14**, 135–143.

[13] BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.

[14] BOLLERSLEV, T., CHOU, R.Y. AND KRONER, K.F. (1992) ARCH modeling in finance: a review of the theory and empirical evidence. *J. Econometrics* **52**, 5–59.

[15] BOLLERSLEV, T. AND MIKKELSEN, H.O. (1996) Modeling and pricing long memory in stock market volatility. *J. Econometrics* **73**, 151–184.

[16] BOUGEROL, P. AND PICARD, N. (1992) Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**, 115–127.

[17] BREIDT, F.J., CRATO, N. AND DE LIMA, P. (1996) The detection and estimation of long memory in stochastic volatility. *J. Econometrics* **83**, 325–348.

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

[19] Brockwell, P.J. and Davis, R.A. (1996) *Introduction to Time Series and Forecasting.* Springer, New York.

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

[21] Campbell, J.Y., Lo, A.W. and MacKinley, A.C. (1997) *The Econometrics of Financial Markets.* Princeton University Press, Princeton.

[22] Davis, R.A. and Mikosch, T. (1999) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26**, 2049–2080.

[23] Davis, R.A., Mikosch, T. and Basrak, B. (1998) Sample ACF of multivariate stochastic recurrence equations with applications to GARCH. Technical Report, University of Groningen.

[24] de la Peña, V.H. and Montgomery–Smith, S.J. (1995) Decoupling inequalities for the tail probabilities of multivariate U–statistics. *Ann. Probab.* **23**, 806–816.

[25] de Lima, P. and Crato, N. (1994) Long range dependence in the conditional variance of stock returns. *Economic Letters* **45**, 281–285.

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

[27] Ding, Z. and Granger, C.W.J. (1996) Modeling volatility persistence of speculative returns: A new approach. *J. Econometrics* **73**, 185–215.

[28] Ding, Z., Granger, C.W.J. and Engle, R. (1993) A long memory property of stock market returns and a new model. *J. Empirical Finance* **1**, 83–106.

[29] Doukhan, P. (1994) *Mixing. Properties and Examples.* Lecture Notes in Statistics **85**. Springer Verlag, New York.

[30] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[31] Engle, R.F. (ed.) (1995) *ARCH Selected Readings.* Oxford University Press, Oxford (U.K.).

[32] Engle, R.F. and Bollerslev, T. (1986) Modelling the persistence of conditional variances. With comments and a reply by the authors. *Econometric Rev.* **5**, 1–87.

[33] Geweke, J. and Porter–Hudak, S. (1983) The estimation and application of long memory time series models. *J. Time Series Analysis* **4**, 221–238.

[34] Ghysels, E., Harvey, A. and Renault, E. (1997) Stochastic volatility. In: Madala, G.S. and Rao, C.R. (Eds.) *Statistical Methods of Finance. Handbook of Statistics,* vol. 14, pp. 119–191.

[35] 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.

[36] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.

[37] Gourieroux, C. (1997) *ARCH Models and Financial Applications.* Springer Series in Statistics. Springer, New York.

[38] Granger, C.W.J. and Ding, Z. (1996) Varieties of long memory models. *J. Econometrics* **73**, 61–77.

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

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

[41] Hida, T. (1980) *Brownian Motion.* Springer, New York.

[42] Jacod, J. and Shiryaev, A.N. (1987) *Limit Theorems for Stochastic Processes.* Springer, Berlin, New York.

[43] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

[44] Klüppelberg, C. and Mikosch, T. (1996) The integrated periodogram for stable processes. *Ann. Statist.* **24**, 1855–1879.

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

[46] Kokoszka, P. and Mikosch, T. (1997) The integrated periodogram for long–memory processes with finite or infinite variance. *Stoch. Proc. Appl.* **66**, 55–78.

[47] 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.

[48] Lo, A. (1991) Long memory in stock market prices. *Econometrica* **59**, 1279–1313.

[49] 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.

[50] Mikosch, T. and Stărică, C. (1998) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. To appear, *Ann. Stat.* Available at *www.math.rug.nl/˜mikosch*

[51] Mikosch, T. and Stărică, C. (1999) Change of structure in financial time series, long–range dependence and the GARCH model. Technical Report. University of Groningen. Available at *www.math.rug.nl/˜mikosch*

[52] Nelson, D.B. (1990) Stationarity and persistence in the GARCH$(1,1)$ model. *Econometric Theory* **6**, 318–334.

[53] Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns. A new approach. *Econometrica* **59**, 347–370.

[54] Oodaira, H. and Yoshihara, K. (1972) Functional central limit theorems for strictly stationary processes satisfying the strong mixing condition. *Kōdai Math. Sem. Rep.* **24**, 259–269.

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

[56] Pollard, D. (1984) *Convergence of Stochastic Processes.* Springer, Berlin.

[57] Potter, K. (1976) Evidence for non stationarity as physical explanation of the Hurst phenomenon. *Water Resour. Res.* **12**, 1047–1052.

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

[59] Resnick, S.I. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Probab.* **18**, 66–138.

[60] Rosiński, J. and Woyczyński, W.A. (1987) Multilinear forms in Pareto–like random variables and product random measures. *Coll. Math.* **51**, 303–313.

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

[62] Spindt, P.A. and Tarhan, V. (1987) The Federal Reserve's new operating procedure: a postmortem. *J. Monetary Econom.* **19**, 107–123.

[63] Teverovsky, V. and Taqqu, M. (1997) Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator. *J. Time Ser. Anal.* **18**, 279–304.

[64] Whittle, P. (1951) *Hypothesis Testing in Time Series Analysis.* Almqvist och Wicksel, Uppsala.

[65] Zygmund, A. (1988) *Trigonometric Series.* First paperback edition. Cambridge University Press, Cambridge (UK).

THOMAS MIKOSCH
Department of Mathematics
P.O. Box 800
University of Groningen
NL-9700 AV Groningen
THE NETHERLANDS
mikosch@math.rug.nl

CĂTĂLIN STĂRICĂ
The Wharton School, Philadelphia, and
Department of Statistics
Chalmers University of Technology
S–412 96 Gothenburg
SWEDEN
starica@math.chalmers.se