Fractionally Integrated Models With ARCH Errors

Michael A. Hauser[*] *and Robert M. Kunst*[**]

Forschungsbericht/
Research Memorandum No. 322

May 1993

* University of Economics and Business Administration, Augasse 2-6, A-1090 Vienna, Austria

** Institute for Advanced Studies, Stumpergasse 56, A-1060 Vienna, Austria

ABSTRACT

We introduce ARFIMA-ARCH models which simultaneously incorporate fractional differencing and conditional heteroskedasticity. We develop the likelihood function and a numerical estimation procedure for this model class. Two ARCH models - Engle- and Weiss-type - are explicitly treated and stationarity conditions are derived. Finite-sample properties of the estimation procedure are explored by Monte Carlo simulation. An application to the Standard & Poor 500 Index indicates existence of intermediate memory ($d<0$) for the 1980's and no fractional differencing ($d=0$) but strong conditional heteroskedastic effects for the 1960's. For the latter time period, contrary to the suggestion of long memory by Mandelbrot, we only found evidence for a positive first-order autoregressive parameter.

ZUSAMMENFASSUNG

Wir stellen ARFIMA-ARCH-Modelle vor, welche gleichzeitig fraktionales Differenzieren und bedingte Heteroskedastie inkorporieren. Wir entwickeln die Likelihood-Funktion und eine numerische Schätzprozedur für diese Modellklasse. Zwei ARCH-Modelle - Engle- und Weiss-Typus - werden explizit behandelt und die Stationaritätsbedingungen abgeleitet. Endlich-Stichproben-Eigenschaften der Schätzprozedur werden durch Monte Carlo-Simulation erforscht. Eine Anwendung auf den Standard & Poor 500 Index zeigt Existenz von "intermediate memory" ($d<0$) für die 1980er-Jahre und kein fraktionales Differenzieren für die 1960er-Jahre an. Für letztere Zeitperiode - im Gegensatz zu Mandelbrot's Vermutung - fanden wir nur Evidenz für einen positiven autoregressiven Parameter 1. Ordnung.

ABSTRACT

We introduce ARFIMA-ARCH models which simultaneously incorporate fractional differencing and conditional heteroskedasticity. We develop the likelihood function and a numerical estimation procedure for this model class. Two ARCH models - Engle- and Weiss-type - are explicitly treated and stationarity conditions are derived. Finite-sample properties of the estimation procedure are explored by Monte Carlo simulation. An application to the Standard & Poor 500 Index indicates existence of intermediate memory ($d<0$) for the 1980's and no fractional differencing ($d=0$) but strong conditional heteroskedastic effects for the 1960's. For the latter time period, contrary to the suggestion of long memory by Mandelbrot, we only found evidence for a positive first-order autoregressive parameter.

KEY WORDS: Conditional heteroskedasticity; Autoregressive moving average model; Fractional differencing; Long memory; Maximum likelihood estimation.

1. INTRODUCTION

Models involving fractional differences have recently drawn much attention, both in the fields of financial data (exchange rates (Cheung 1993; Diebold, Husted, and Rush 1992), interest rates (Backus and Zin 1993), stock prices (Lo 1991)) and of macroeconomics and business cycle analysis (Diebold and Rudebusch 1989; Sowell 1992b). In the latter area, fractional differences are often seen as a sort of compromise between the concepts of trend stationarity and of integrated series. Viewing e.g. industrial output as being fractionally integrated with a differencing parameter $d<1$ explains the concentration of spectral mass at frequency 0 and permits to retain the more classical picture of trend reversion and transitory shocks.

On the other hand, financial series typically are rather close to random walks and little dynamic structure can be found. Repeatedly, researchers have taken up the difficult challenge to detect as much structure as possible. Whereas short-run dependence of increments or returns is low, dependence in volatility can be found in high-frequency series which has incited research on ARCH models of various types (in particular, see Engle (1982) and Weiss (1984)). Other researchers claim that long-run dependence or long memory is present in some financial series which would allow systematic gains from speculation over sufficiently long time intervals. Such long memory can be well represented by fractional models of the ARFIMA(p,d,q) type (see Granger and Joyeux (1980) and Hosking (1981)). Empirical evidence on this phenomenon remains mixed.

Because both non-standard features (ARCH and ARFIMA) play such a big role in the literature on financial time series, it seems worth while to study both at the same time and to gauge possible cross-effects between estimation of heteroskedasticity and of modeling long-run dependence (compare e.g. Cheung (1993, p.95) or Lo (1991, p.1283)). Consequently, we suggest the use of ARFIMA-ARCH models and develop a maximum-likelihood estimation procedure. We also apply the procedure to an empirical time series, viz. the Standard & Poor 500 Index.

This paper is organized as follows. Section 2 presents the ARFIMA-ARCH model. We allow for two different forms of modeling heteroskedasticity that have been suggested in the literature. In addition, we give stationarity conditions for the model. Section 3 develops the maximum-likelihood estimation procedure and presents some evaluation of its small-sample performance by Monte Carlo simulation. Section 4 applies the ML procedure to the Standard & Poor 500 Index for different time periods. Section 5 concludes.

2. STATIONARITY CONDITIONS

The ARFIMA model with heteroskedastic errors in its most general form can be written as follows:

$$\begin{aligned} &\Phi(B)(1-B)^d(Y_t-\mu)=\Theta(B)\varepsilon_t \\ &E(\varepsilon_t^2|I_{t-1})=h_t \end{aligned} \qquad (2.1)$$

Here, $\Phi(B) = 1-\varphi_1 B-\ldots$ is the autoregressive (AR) polynomial of order p and $\Theta(B) = 1+\theta_1 B+\ldots$ is the moving average (MA) polynomial of order q. I_t denotes the information set generated by $\{Y_t, Y_{t-1}, Y_{t-2},\ldots\}$ and h_t is some function of t to be specified in more detail. To facilitate notation, the constant μ will be suppressed in the following. Two versions of the heteroskedastic ARFIMA model are treated in this paper that are distinct due to the way heteroskedasticity is modeled. The first model is based on Engle's (1982) ARCH model which relates conditional heteroskedasticity to squares of past innovations:

$$h_t = E(\varepsilon_t^2|I_{t-1}) = \alpha_0 + \alpha_1\,\varepsilon_{t-1}^2 + \ldots + \alpha_s\,\varepsilon_{t-s}^2 \qquad (2.2)$$

The model defined by (2.1)&(2.2) will be denoted ARFIMA-ARCH($p,d,q/s$) in the following. It is seen that stability properties of the

ARCH process in (2.2) do not depend on the linear ARFIMA model in (2.1). Engle (1982) proved that a weakly stationary solution for the ε_t process exists iff $\alpha_1+\ldots+\alpha_s < 1$, under the assumption of finite second moments in the conditional error distribution. Nelson (1990) showed that strongly stationary solutions may still exist if this condition fails. Such solutions necessarily display infinite variances. Even if a weakly stationary solution exists and a conditional Gaussian law is assumed, the stationary distribution of ε_t is non-Gaussian and leptokurtic.

From the literature on long-memory processes it is known that the linear part has a weakly stationary solution as long as $d < ½$ and all roots of $\Phi(.)$ are outside the unit disk. It follows that the ARFIMA-ARCH model (2.1)&(2.2) is weakly stationary as long as these conditions and the restriction on the sum of the ARCH coefficients are both satisfied. We have summarized this result as a theorem.

THEOREM 1.

(a) The ARFIMA-ARCH($p,d,q/s$) model given by (2.1)&(2.2) has a strictly and covariance-stationary solution as long as

(i) all p roots of the AR polynomial $\Phi(.)$ have modulus greater than 1,

(ii) $d < ½$,

(iii) $\sum_{i=1}^{s} \alpha_i < 1$, $\alpha_i \geq 0$ for all $i \geq 1$ and $\alpha_0 > 0$.

(b) The variance of the resulting ARFIMA-ARCH process can be expressed as

$$\frac{\alpha_0}{1-\sum_{i=1}^{s}\alpha_i}\sum_{i=0}^{\infty}\psi_i^2 \text{ where } \Psi(B)=\Theta(B)\Phi^{-1}(B)(1-B)^{-d}$$

In particular, for pure fractional processes, the variance is given by

$$\frac{\alpha_0}{1-\sum_{i=1}^{s}\alpha_i}\,\frac{\Gamma(1-2d)}{\Gamma^2(1-d)}\ .$$

Proof: The ARCH conditions only concern the innovations process. Hence, if they hold, the innovations process is strictly and covariance-stationary. On the other hand, if an uncorrelated and stationary innovations sequence with finite variance exists, then the variance of the ARFIMA process with $d < ½$ is finite. In consequence, as it is defined from a time-constant law, the ARFIMA-ARCH process is stationary.

The formula in (b) evolves from the variance formula for white noise ARCH given in Engle (1982), the variance of an ARMA process, and the variance of a fractionally integrated process (see e.g. Brockwell and Davis (1991)). It can also be calculated directly in a somewhat tedious way making repeated use of Mertens' Theorem for the convergence of product sums. ■

Note that the stationarity conditions for the ARFIMA-ARCH process are not stronger than for each process separately. Nothing is said, however, about the existence of higher moments or of other properties of the stationary distribution. Of course, such distribution depends on assumptions on the conditional distribution of the ARCH innovations but, at the moment, we do not even know this distribution in the case of a conditional Gaussian law.

A different model arises from assuming that conditional error variances depend on squared past observations of the process Y_t itself. Such model is a special case of the approach suggested by Weiss (1984):

$$h_t = \alpha_0 + \alpha_1 Y_{t-1}^2 + \cdots + \alpha_s Y_{t-s}^2 \tag{2.3}$$

(2.3) is, in fact, a simplification of Weiss' original model. Weiss' model allows for amalgams of Engle-type dependence on past errors and (2.3) as well as for explicit dependence on the squared linear predictor for Y_t but these extensions will not be treated here. To distinguish it from the Engle-type model, a process defined by (2.1) & (2.3) will be denoted by ARFIMA-ARCH$_W(p,d,q/s)$ in the following.

Weiss (1984) has shown that the existence of a weakly stationary solution for (2.3) depends on joint conditions for the linear and the ARCH parameters. These conditions can be interpreted in such a way that stronger ARCH dependence (i.e. larger α_i) can be "traded" against weaker linear dependence in order to retain stability. To explicitly state the conditions, again denote the Wold infinite-order moving average representation of the linear process Y_t by $\Psi(B) = (1-B)^{-d}\Phi^{-1}(B)\Theta(B)$. Building on a result by Weiss (1984) it is easy to derive the following theorem.

THEOREM 2.

(a) The ARFIMA-ARCH$_W(p,d,q/s)$ model given by (2.1)&(2.3) has a strictly and covariance-stationary solution if the following conditions are satisfied

(i) all p roots of the AR polynomial $\Phi(.)$ have modulus greater than 1.

(ii) $d < ½$

(iii) $\alpha_i \geq 0$ for all $i \geq 1$ and $\alpha_0 > 0$.

(iv) $\sum_{i=1}^{s} \alpha_i \sum_{j=0}^{\infty} \psi_j^2 < 1$

(b) The variance of the resulting ARFIMA-ARCH$_W$ process can be expressed as

$$\frac{\alpha_0 \sum_{j=0}^{\infty} \psi_j^2}{(1 - \sum_{i=1}^{s} \alpha_i \sum_{j=0}^{\infty} \psi_j^2)}$$

Some numerical examples for the condition (b) for the simple case $\Phi \equiv \Theta \equiv 1$ and hence $\Psi(B) = (1-B)^{-d}$ are given in Table 1. Existence of the unconditional second moment depends on the joint values of d and the α_i. In particular, if d approaches 0.5 then the $\Sigma\alpha_i$ has to decrease quickly to 0 in order to preserve stationarity. In contrast, for larger negative values of d, the bound decreases only slightly.

TABLE 1. Stationarity conditions for ARFIMA-ARCH$_W$(0,d,0/s) processes.

d	$\max \sum_{i=1}^{s} \alpha_i$
0.49	0.06
0.4	0.48
0.2	0.91
0	1.00
-0.2	0.95
-0.4	0.84
-0.49	0.79

Within the limits of this paper, we restrict ourselves to the conditional Gaussian specification in the ARCH models (2.2) and (2.3). This means that, e.g. in place of the general formulation (2.2), we explicitly assume

$$\varepsilon_t \sim N(0, \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_s \varepsilon_{t-s}^2) \qquad (2.4)$$

The importance of this assumption will be seen in the next section.

3. ESTIMATION

Estimation can be based on the maximum likelihood algorithm for ARFIMA models by Hosking (1984). This method starts from the full likelihood for T successive observations $Y_T = (y_1,\ldots,y_T)$ from an ARFIMA model

$$\ell(Y_T, \Sigma) = (2\pi)^{-T/2} |\Sigma|^{-\frac{1}{2}} exp(-\tfrac{1}{2} Y_T' \Sigma^{-1} Y_T) \qquad (3.1)$$

where Σ is a symmetric Toeplitz matrix that depends on the autoregressive parameters Φ as well as on the moving average parameters Θ and on the fractional differencing parameter d. The numerical evaluation of Σ in (3.1) for given Φ and Θ is derived in Sowell (1992a) using the hypergeometric function and the roots of the AR polynomial.

The ARFIMA-ARCH model can be viewed as a model of the above type where the condition that Σ be Toeplitz is relaxed. More specifically, if Y_T is

ARFIMA-ARCH then its linear innovations vector $E_T = (\varepsilon_1,\ldots,\varepsilon_T)$ is ARCH and

$$H^{-½}E_T = \text{diag}(h_1^{-½},\ldots,h_T^{-½})E_T \tag{3.2}$$

is distributed standard normal $N(0,I_T)$. H may be chosen to be generated by Engle- or Weiss-type ARCH. Let us denote the Cholesky unit-diagonal (Banachiewicz) factorization of Σ as LDL'. D is a non-constant diagonal matrix with its (T,T)-element coming closest to the true ε_t variance. $L^{-1}Y_T$ defines a vector W_T of serially uncorrelated but (unconditionally) heteroskedastic errors w_t. These errors have second-order properties distinct from the ε_t and the ARCH models (2.2) and (2.3) are not related to them.

Writing the covariance matrix of W_T as $D=\sigma^2P^2$ where σ^2 is the variance of ε_t, one can also consider the transformations vector $P^{-1}L^{-1}Y_T = U_T = (u_1,\ldots,u_T)'$. The u_t are uncorrelated and homoskedastic with constant variance σ^2. The u_t can be viewed as process innovations under the assumption of appropriate pre-sample observations Y_0, Y_{-1},... Hence, unlike the w_t, the u_t approximate the theoretical innovations ε_t and share their second-moment properties. Whereas the w_t assume zero values during the pre-sampling period, the u_t rather assume that the process before time point 1 behaved similarly to after that time.

For the ARFIMA-ARCH model, the covariance matrix of Y_T is (formally, though the unconditional covariance matrix remains $\Sigma = LDL'$ notwithstanding any ARCH features) $\Sigma_H = LPHPL'$ where H reduces to $\sigma^2 I$ in the homoskedastic case. The likelihood then becomes

$$\ell(Y_T,\Sigma_H) = (2\pi)^{-T/2}|\Sigma_H|^{-1/2}\exp\{-½Y_T'\Sigma_H^{-1}Y_T\} \tag{3.3}$$

Because the covariance matrix now depends on Y_T, (3.3) is no more the likelihood of a Gaussian distribution problem. The numerical evaluation of the likelihood can be conducted via the Levinson algorithm which is applied to Σ. This efficiently generates the Banachiewicz decomposition of its inverse as $\Sigma^{-1} = MD^{-1}M'$. In consequence,

$$\Sigma_H^{-1} = MP^{-2}H^{-1}M' = M_c H^{-1} M_c' \tag{3.4}$$

where M_c denotes the left factor in the classical Cholesky decomposition $\Sigma^{-1} = M_c M_c'$.

Engle (1982) has shown that, for all combinations of linear and ARCH models which obey certain regularity conditions, the information matrix is block diagonal and estimation of ARCH and linear parameters can be conducted by alternating iterations. The ARFIMA-ARCH model is essentially regular in the sense of Engle, with the only caveat being moment conditions which, as shown in Section 2, are not stronger in the ARFIMA-ARCH case than for the pure ARCH model as long as the Engle-type model is adopted. Hence, one suggested procedure could be the following:

1. Estimate all ARFIMA parameters by the method of Sowell.
2. Calculate E_T as the U_T as outlined above.
3. Estimate the ARCH parameters via least squares.
4. Maximize (3.3) by a modified method of Sowell, conditional on a fixed H_T.

5. Re-calculate U_T and re-estimate the ARCH parameters, now based on weighted least squares with the weights provided by the previous estimation.
6. Re-maximize (3.3) and iterate until convergence.

The possible gains achieved by this simplified procedure are uncertain. Step 4 also uses matrices of dimension T which limitation is the most stringent one for the direct ML estimation in practice. The above algorithm, however, avoids estimation of ARCH parameters by optimization which can become inconvenient due to the constraints on these parameters.

In the rest of the paper, we proceed with direct maximization of the likelihood (3.3). To calculate h_i for $i \leq s$ in (2.2), starting values for ε_i^2 with $i \leq 0$ are required which we specified to equal the unconditional variance. In contrast, Diebold and Schuermann (1992) propose a rather time-consuming procedure to approximate exact ML for ARCH situations. However, they report substantial relative gains over our approximate method only for very small samples. Table 2 gives the results of a Monte Carlo simulation based on 500 replications of ARFIMA-ARCH(0,d,0/1) processes (Engle-type) of length $T=100$. The processes were simulated by generating ARCH innovations and calculating the Cholesky decomposition of Σ.

For the estimated parameters, standard deviations range from 0.0033 to 0.0055 for the d estimates, from 0.0030 to 0.0085 for the α_1 estimate.

TABLE 2. Small sample properties of the ML estimator for the ARFIMA-ARCH (0,d,0/1) model. (T=100, 500 replications)

(a) Mean estimates of the FI parameter d

	d						
α_1	-.4	-.2	-.1	0	.1	.2	.4
0	-.424	-.241	-.145	-.049	.050	.145	.331
.25	-.423	-.237	-.141	-.045	.051	.150	.334
.50	-.413	-.226	-.130	-.032	.065	.163	.348
.75	-.399	-.213	-.118	-.021	.076	.173	.360
.90	-.393	-.209	-.111	-.015	.082	.179	.366
.95	-.384	-.200	-.106	-.009	.089	.186	.371
.99	-.379	-.195	-.102	-.006	.090	.186	.371
1.00	-.378	-.195	-.101	-.004	.090	.187	.372

(b) Mean estimates of the ARCH parameter α_1

	d						
α_1	-.4	-.2	-.1	0	.1	.2	.4
0	.043	.043	.043	.044	.044	.044	.037
.25	.224	.227	.228	.229	.229	.226	.197
.50	.445	.446	.446	.446	.445	.440	.383
.75	.638	.640	.640	.640	.639	.633	.566
.90	.732	.734	.734	.734	.732	.726	.662
.95	.767	.769	.770	.769	.768	.762	.693
.99	.761	.763	.763	.763	.761	.755	.688
1.00	.766	.768	.768	.768	.766	.760	.693

(c) Rejection frequency at empirical 10 % significance bounds (for $d<0$) and acceptance frequency at same significance bounds (for $d>0$)

	P($d \le c_-$)			c_-	c_+	1-P($d \le c_+$)		
α_1	d=-.4	d=-.2	d=-.1	d=0		d=.1	d=.2	d=.4
0	1.00	.75	.33	-.178	.061	.47	.84	1.00
.25	.99	.75	.38	-.169	.068	.47	.84	.99
.50	1.00	.81	.41	-.148	.074	.47	.86	1.00
.75	.99	.86	.41	-.127	.084	.49	.87	.99
.90	.99	.87	.47	-.116	.083	.52	.89	.99
.95	.99	.87	.43	-.115	.091	.50	.88	.99
.99	.96	.85	.47	-.118	.109	.38	.83	.99
1.00	.96	.84	.47	-.117	.110	.38	.83	.99

NOTE: c_-, c_+ ... 10 % one-sided critical value under d=0 ("left-sided test, right-sided test")

The ML bias of the parameter d is clearly recognizable. The size of this bias corresponds to the one known from Cheung and Diebold (1993) and is robust against spurious estimation of an additional ARCH parameter. In contrast, if the true process is ARFIMA-ARCH, the d bias decreases with an increase of α_1.

Whereas d is biased to the negative in most cases, α_1 is positively biased for very small true values due to the non-negativity constraints. For larger values of true α_1, a bias toward zero becomes sizable. Except for the very long-memory case d=0.4, performance of the α_1 estimate is unaffected by d. Correctly detected heteroskedasticity appears to improve the accuracy of ARFIMA estimation.

Table 2c shows that the power of tests for d=0 tends to increase with higher α_1, a feature which works for $d<0$ monotonously and for $d>0$ as long as $\alpha_1<0.9$.

4. APPLICATION TO THE STANDARD & POOR 500 INDEX SERIES

We use logarithmic changes (approximate returns) of the Standard & Poor 500 Index (S&P) from July 1962 to December 1990, 7167 daily observations in total. The logarithmic index and its first difference are depicted in Figure 1 and 2. In particular, the sample contains the year 1987 when a crash occurred which came to be known as the "Black Monday". Maximization of the likelihood over the entire sampling interval would involve inverting symmetric Toeplitz matrices of dimension 7167^2 which exceeded the capacity of accessible computers. Hence, we resorted to splitting the data into subsamples. This also has the advantage that we can evaluate possible structural changes in certain parameters which may not be obvious from analyzing the entire sample.

4.1 Single Years

First we split the sample into calendar years (one calendar year has approximately 250 observations), starting with 1963 and ending with 1990. At first, simple ARFIMA-ARCH(0,d,0/1) and (0,d,0/0) models (Engle-type ARCH) were estimated. In general, including first-order ARCH effects has little effects on the d estimates. Exceptions are those years where outliers are most pronounced such as 1963, 1970, and of course 1987. These outliers and the concomitant increased volatility in the series generate large ARCH parameters. In the crash year 1987, even estimates beyond the kurtosis boundary are obtained, i.e. the implied stationary distribution does not have finite fourth moments. (Engle (1982) gives the boundary for finite kurtosis in a first-order ARCH model as $1/\sqrt{3}$. This means that for $\alpha_1 > 1/\sqrt{3}$ fourth moments do not exist. For higher-order ARCH, this condition becomes more complicated.) Interestingly, in these years with strong ARCH effects, d estimates can be both higher and lower as compared to the pure FI model. On the whole, d estimates appear to have decreased over the sampling interval, with significantly positive values encountered in the 1960's and near-zero values in the 1980's.

Introduction of a short-run MA parameter θ_1 changes d estimates dramatically. This effect is similar to that seen from comparing Table 3a with 3b. Estimated θ_1 was rather stable up to the mid-1970's but also showed a downward trend after that time. (Results on this preliminary experiment with first-order ARCH are not shown as they are essentially similar to those in Table 3. All detailed results can be obtained on request from the authors.)

Visual inspection of the sample autocorrelation function (ACF) of the S&P return series for the whole sample (compare Figure 3) indicates a first-order MA with a positive θ_1. Inspection of the ACF of squared values shown as Figure 4 - as suggested e.g. by Weiss (1984) - gives evidence on noteworthy correlations at the orders 2, 5 (the weekly lag), 8, and at some larger values such as around 60 whose interpretation remains unclear. This evidence would suggest a fifth-order or maybe even an eighth-order ARCH model for the errors and hence an ARFIMA-ARCH process of order $(0,d,1/5)$ for the S&P series. Table 3a reports the results from fitting an

ARFIMA-ARCH(0,d,0/5) process to the data whereas Table 3b results from fitting an ARFIMA-ARCH(0,d,1/5) model.

The tables again point to a decrease of the systematic linear part (the "memory") since around 1975. Unlike the data variance, which has increased during the last decades, ARCH effects appear to follow an irregular downward tendency. This downward tendency, however, is much less pronounced if the sum of all coefficients instead of only α_1 is considered. This sum - which determines stationarity (see our Theorem 1) - even reached 0.886 in the crash year 1987. In particular Table 3b demonstrates that in the presence of a short-run MA parameter long-memory characteristics become very small and, in a wider and somewhat casual sense, insignificant.

Estimation based on the Weiss-type ARCH models has also been conducted, both for ARFIMA-ARCH$_W$(0,d,0/5) and (0,d,1/5). For 20 out of 28 years, the Engle-type model produced a slightly better fit. Due to the small explanatory power of the linear part of these ARFIMA-ARCH models, results were close to those for the Engle-type model shown in Table 3.

TABLE 3a. Fitting an ARFIMA-ARCH (0,d,0/5) process to the S&P series.

Year	d	α_1	α_5	$\Sigma\alpha_i$
1963	0.103	0.332	0.000	0.588
1964	0.147	0.215	0.060	0.380
1965	0.177	0.386	0.021	0.602
1966	0.192	0.000	0.115	0.756
1967	0.172	0.197	0.000	0.385
1968	0.201	0.050	0.047	0.438
1969	0.217	0.099	0.000	0.403
1970	0.211	0.056	0.191	0.597
1971	0.223	0.059	0.083	0.495
1972	0.205	0.099	0.000	0.102
1973	0.066	0.087	0.191	0.590
1974	0.174	0.000	0.158	0.457
1975	0.163	0.039	0.023	0.174
1976	0.086	0.039	0.143	0.224
1977	0.132	0.008	0.016	0.072
1978	0.165	0.245	0.000	0.317
1979	0.098	0.096	0.000	0.104
1980	0.098	0.000	0.056	0.287
1981	0.029	0.005	0.168	0.173
1982	0.114	0.000	0.153	0.451
1983	0.013	0.081	0.072	0.293
1984	-0.012	0.074	0.000	0.142
1985	0.031	0.000	0.050	0.091
1986	0.049	0.000	0.134	0.134
1987	0.076	0.491	0.057	0.886
1988	-0.045	0.000	0.031	0.334
1989	0.010	0.034	0.000	0.048
1990	0.063	0.087	0.077	0.460

TABLE 3b. Fitting an ARFIMA-ARCH (0,d,1/5) process to the S&P series.

Year	d	θ_1	α_1	α_5	$\Sigma\alpha_i$
1963	0.087	0.024	0.334	0.003	0.589
1964	0.090	0.103	0.214	0.060	0.374
1965	0.115	0.184	0.397	0.017	0.617
1966	-0.057	0.330	0.000	0.124	0.787
1967	0.084	0.175	0.213	0.000	0.401
1968	0.032	0.319	0.082	0.086	0.499
1969	-0.004	0.361	0.060	0.000	0.427
1970	-0.002	0.418	0.000	0.134	0.555
1971	0.032	0.306	0.035	0.173	0.558
1972	0.035	0.304	0.166	0.000	0.188
1973	-0.137	0.351	0.000	0.132	0.568
1974	-0.059	0.372	0.000	0.163	0.521
1975	-0.066	0.444	0.013	0.129	0.346
1976	-0.046	0.211	0.040	0.140	0.277
1977	-0.009	0.203	0.002	0.015	0.045
1978	0.066	0.176	0.229	0.000	0.338
1979	-0.036	0.226	0.072	0.000	0.104
1980	0.060	0.063	0.000	0.054	0.276
1981	-0.086	0.172	0.011	0.166	0.177
1982	0.070	0.058	0.000	0.140	0.443
1983	-0.004	0.026	0.079	0.072	0.289
1984	-0.042	0.108	0.058	0.000	0.116
1985	0.036	0.071	0.000	0.064	0.126
1986	-0.040	0.140	0.000	0.088	0.088
1987	0.070	0.011	0.490	0.057	0.885
1988	-0.150	0.160	0.009	0.021	0.346
1989	0.021	-0.019	0.033	0.000	0.047
1990	-0.024	0.132	0.058	0.097	0.405

4.2 Pre- and Post-Crash 1987 Periods

Whereas Table 3b points to an overall small or insignificant amount of long memory in the series, the evidence on negative d is corroborated if longer time segments are used (see Table 4). We looked for the best model according to the Akaike's Information Criterion (AIC)

$$AIC = 2\log \ell(Y_T, \Sigma) - 2(p+\delta+q+1+s) \qquad (4.1)$$

with $\delta=0$ in the pure ARMA and $\delta=1$ in the ARFIMA case. In general, the maximum order was set at 2 for p and q. For s, the values 0 and 5 were considered. First, estimation was conducted over the "post-crash" years 1988-1990. The selected specification turned out to be ARFIMA-ARCH(1,1,0/0). Taking the sample size into account, the d estimate can be regarded as being significantly negative. Such series are said to have "intermediate memory".

For the three pre-crash years 1984-1986 (T=757), AIC selected a pure MA(1) model with an estimated θ_1 of 0.107. According to this selection, there was no evidence on ARCH effects and no evidence on a non-zero differencing parameter. This result may change, however, if ARCH models with explicit zero restrictions on certain parameters were included in the search as α_1 and α_5 appear to be greater than the intermediate coefficients which were estimated as 0 in a variety of model specifications.

If 1987 is included in the sample and the time range is extended to cover 1984-1990, AIC selects the ARFIMA-ARCH(0,1,1/5) model which, as pointed out above, corresponds to the model suggested by a visual analysis of some first- and second-order characteristics based on the full sample 1963-1990. The sum of the ARCH coefficients increases to 0.54 obviously due to the crash in 1987 and its aftermath. Due to the large sample size, the bias in the d estimate is relatively small and cannot explain the negative value. Again, a process with intermediate memory is suggested.

TABLE 4: Fitting ARFIMA-ARCH models to subsamples of the S&P series. Parameter estimates are given for best model according to AIC.

	Time range					
	1962-67	1969-78	1979-90	1984-90	1984-86	1988-90
T	1384	2522	3033	1768	757	757
d		-0.034	-0.067	-0.053		-0.105
φ_1	0.177		0.152			0.109
θ_1		0.321		0.119	0.107	
α_1	0.229	0.093	0.155	0.225		
α_2	0.166	0.103	0.028	0.022		
α_3	0.131	0.141	0.091	0.085		
α_4	0.130	0.148	0.129	0.161		
α_5	0.091	0.174	0.078	0.051		
$\Sigma\alpha_i$	0.747	0.659	0.425	0.544		

4.3 Other Periods

Extending the sample further backward to 1979-1990 results in a further decrease of the d estimate. AIC suggests an ARFIMA(1,1,0/5) model but the difference to the ARFIMA-ARCH(0,1,1/5) model is small.

A data portion of a similar size are the earlier years 1969-1978. Again, an ARFIMA-ARCH(0,1,1/5) is selected. The sum of the α_i reaches the high value of 0.66 and θ_1 increases to 0.321, corroborating our finding from the analysis of single years reported in section 4.1 that S&P used to show more systematic features in earlier years. Again, the d estimate turned out to be negative. Due to the large sample size, a value of -0.034 can be regarded as significant in spite of the known bias of the ML d estimate.

For the earliest years in the sample, i.e. 1962-1967, AIC selects a pure AR(1) model with strong ARCH effects - $\Sigma\alpha_i$ reaches 0.75 - but without fractional differencing. Note that the lag pattern in the α_i for the early years is different from the later time periods as it reflects a monotonous decay without explicit evidence on a weekly lag. This result is interesting insofar as Mandelbrot (1971) claimed findings of long memory in asset returns of the 1960's but possibly misinterpreted the autocorrelation structure because of pronounced conditional heteroskedasticity.

5. SUMMARY AND CONCLUSION

We have introduced the ARFIMA-ARCH model which incorporates conditional heteroskedasticity and fractional differencing simultaneously.

The corresponding likelihood function was developed which a maximum likelihood estimation algorithm can be based on. It is an exact ML conditional on starting values for the ARCH part of the model. Alternatively, we outlined an iterative scoring algorithm but did not explore its performance yet as we suppose that its merits over straightforward optimization may be small.

For the sample size that we used in our simulation ($T=100$), the ML estimation procedure yields a bias for d that corresponds to the values reported from ML estimation in the pure ARFIMA model whereas the ARCH effects are small. If the ARCH effects become stronger, this bias in d decreases. As in the case of estimation of pure ARCH models, estimates of ARCH parameters are biased upward for very small ARCH effects but are biased downward for larger effects.

Our empirical example shows that our suggested ARFIMA-ARCH model brings a better fit for most subperiods of the entire sample 1963-1990 than pure ARIMA or pure ARCH models, the exception being the earliest years in the 1960's. The explanatory power of these models, however, remains small and coefficient estimates change a lot over time, pointing to some basic instability of the model. Hence, we do not purport that an ARFIMA-ARCH model with constant coefficients is the "correct" specification for the Standard & Poor 500 Index as such but, as AIC selects mixed models for most periods, the ARFIMA-ARCH appears to be a valuable model relative to the previously used simpler models.

The specification of the functional form for conditional heteroskedasticity in the ARFIMA-ARCH specification may deserve further research. As an alternative to the popular Engle model, we also used the Weiss specification which, however, produced a slightly less satisfactory fit for most subperiods. Similarly, other models, including the GARCH model suggested by Bollerslev (1986) and the AARCH model by Bera, Higgins, and Lee (1992) could be incorporated into the presented framework.

ACKNOWLEDGMENTS

The authors wish to thank Manfred Deistler and Benedikt Poetscher for helpful comments on a previous version of the paper. We also thank Fallaw

Sowell for providing us with a FORTRAN code for maximum likelihood estimation for ARFIMA models.

REFERENCES

Backus, D. K., and Zin, S. E. (1993), "Long-Memory Inflation Uncertainty: Evidence From the Term Structure of Interest Rates," Technical Working Paper No. 133, National Bureau of Economic Research.

Bera, A. K., Higgins, M. L., and Lee, S. (1992), "Interaction Between Autocorrelation and Conditional Heteroscedasticity: A Random-Coefficient Approach," *Journal of Business & Economic Statistics*, 10,133-142.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31, 307-327.

Brockwell, P. J., and Davis, R. A. (1991), *Time Series: Theory and Methods* (2nd ed.), New York: Springer-Verlag.

Cheung, Y. (1993), "Long Memory in Foreign-Exchange Rates," *Journal of Business & Economic Statistics*, 11, 93-102.

——— and Diebold, F. X. (1993), "On Maximum-Likelihood Estimation of the Differencing Parameter of Fractionally Integrated Noise with Unknown Mean," *Journal of Econometrics*, forthcoming.

Diebold, F. X., Husted, St., and Rush, M. (1991), "Real Exchange Rates Under the Gold Standard," *Journal of Political Economics*, 99, 1252-1271.

Diebold, F. X., and Rudebusch, G. (1989), "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, 24, 189-209.

Diebold, F. X., and Schuermann, T. (1992), "Exact Maximum Likelihood Estimation for ARCH Models," Mimeo, University of Pennsylvania.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Granger, C. W. J. and Joyeux, R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15-29.

Hosking, J. R. M. (1981), "Fractional Differencing," *Biometrika*, 68, 165-176.

——— (1984), "Modeling Persistence In Hydrological Time Series Using Fractional Differencing," *Water Resources Research*, 20, 1898-1908.

Lo, A. W. (1991), "Long-Term Memory in Stock Market Prices," *Econometrica*, 59, 1279-1314.

Mandelbrot, B. B. (1971), "When Can Price Be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models," *Review of Economics and Statistics*, 53, 225-236.

Nelson, D. B. (1990), "Stationarity and Persistence in the GARCH(1,1) Model," *Econometric Theory*, 6, 318-334.

Sowell, F. (1992a), "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics*, 53, 165—188.

——— (1992b), "Modeling Long-Run Behavior With the Fractional ARIMA Model," *Journal of Monetary Economics*, 29, 277-302.

Weiss, A. A. (1984), "ARMA Models With ARCH Errors," *Journal of Time Series Analysis*, 5, 129-143.

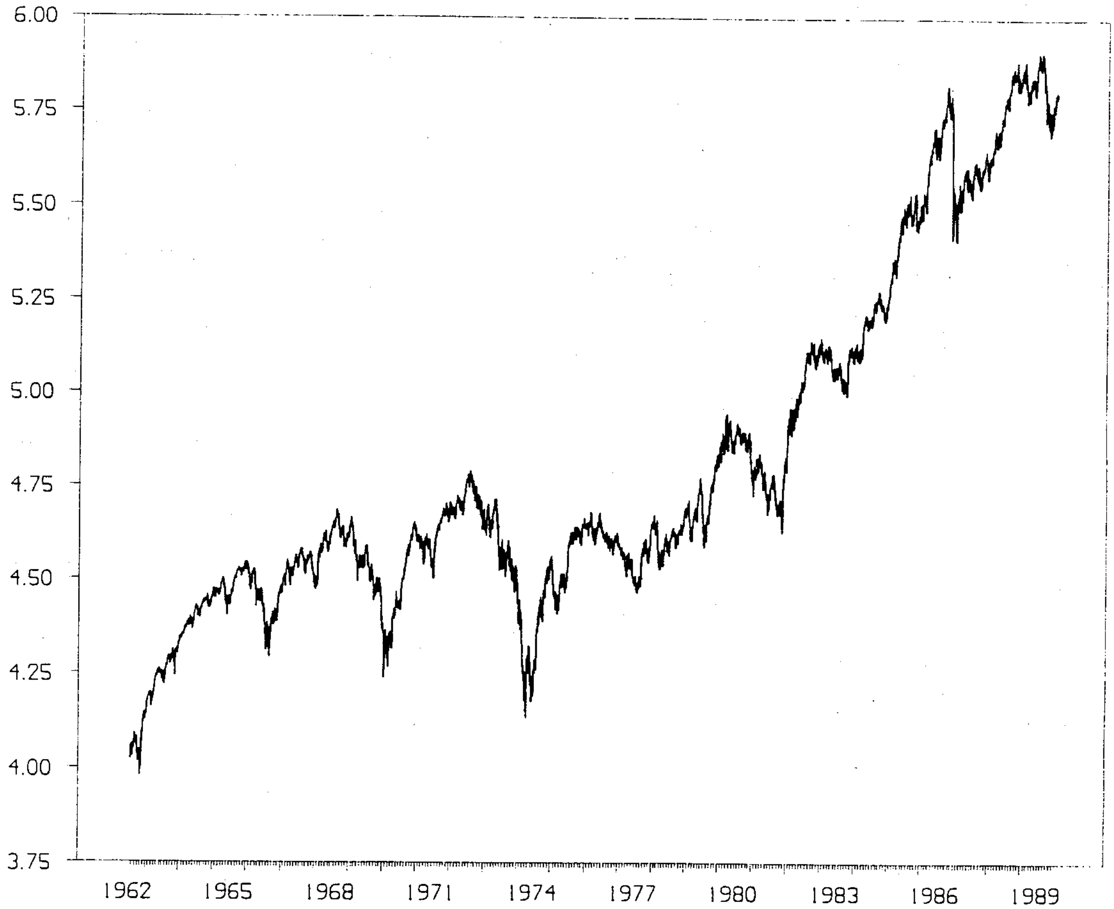

Figure 1: Logarithm of Standard & Poor 500 Index series. The sample runs from July 1, 1962 to December 31, 1990 and thus contains 7167 observations. The horizontal time scale is not exact, particularly toward its right end, due to missing values caused by holidays.

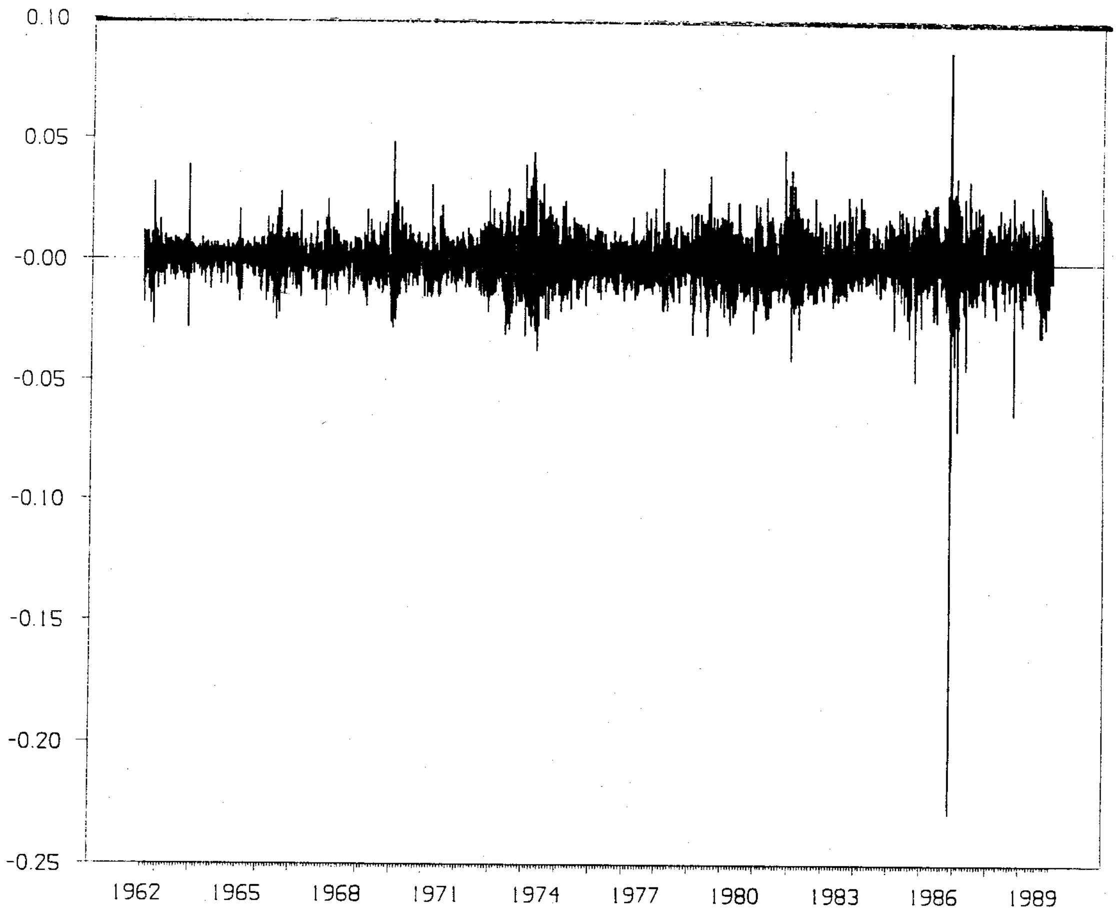

Figure 2: Log changes of Standard & Poor Index 500 series. The sample runs from July 2, 1962 to December 31, 1990 and thus contains 7166 observations. The remark for Figure 1 again applies.

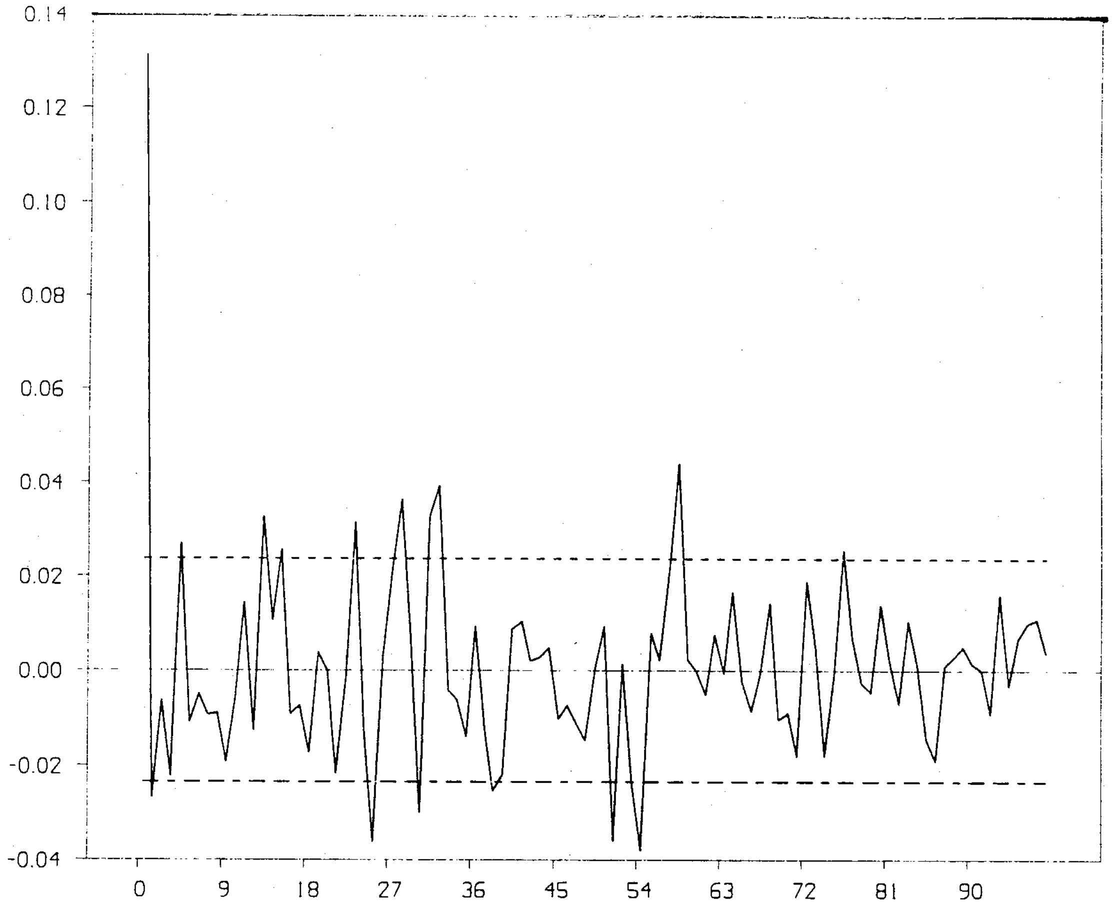

Figure 3: Sample ACF of Standard & Poor returns (log changes). Dashed lines correspond to conventional $2/\sqrt{T}$ guidelines for significance of correlation functions.

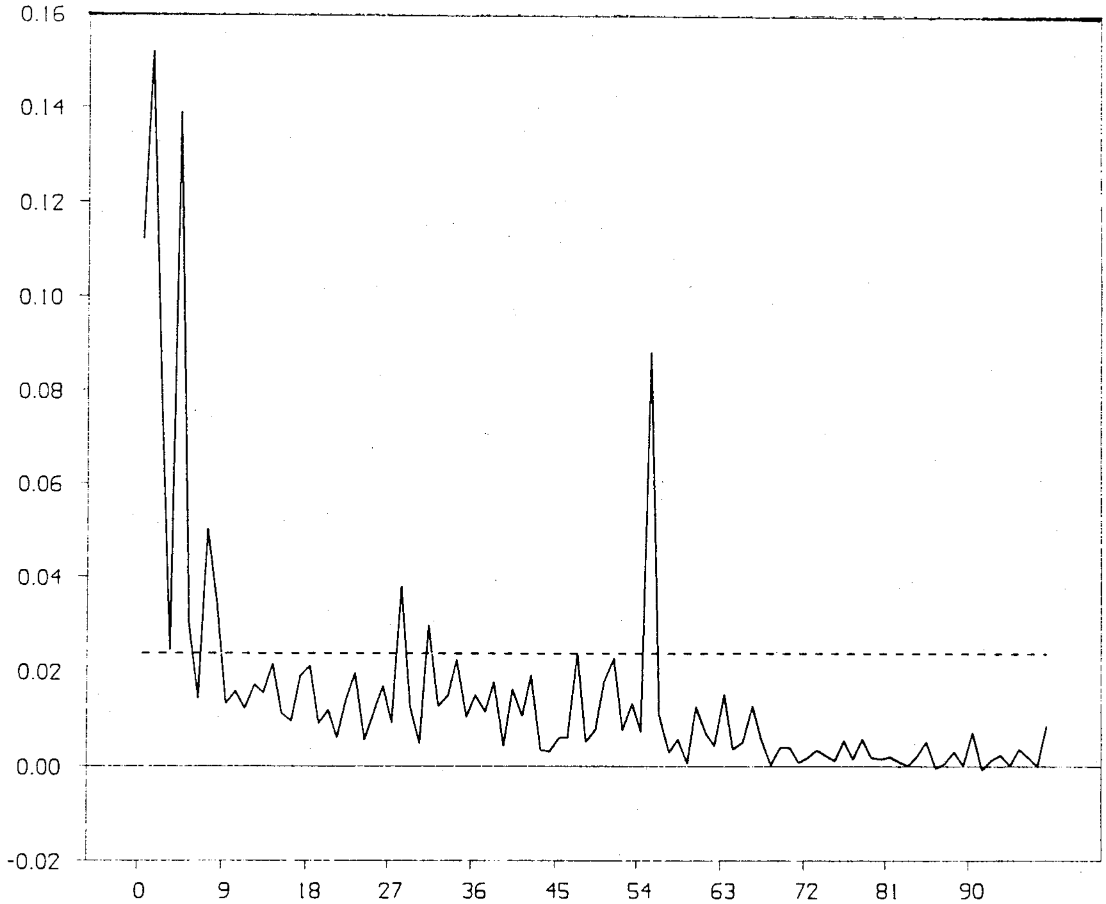

Figure 4: Sample ACF of squared Standard & Poor returns (log changes). Dashed lines correspond to conventional $2/\sqrt{T}$ guidelines for significance of correlation functions.