ANNALS OF ECONOMICS AND FINANCE **14-2(B)**, 721–746 (2013)

Occasional Structural Breaks and Long Memory

Clive W.J. Granger

Department of Economics University of California, San Diego 9500 Gilman Drive La Jolla, CA 92093-0508

and

Namwon Hyung

Department of Economics University of California, San Diego 9500 Gilman Drive La Jolla, CA 92093-0508

This paper shows that a linear process with breaks can mimic autocorrelations and other properties of $I(d)$ processes, where d can be a fraction. Simulation results show that S&P 500 absolute stock returns are more likely to show the "long memory" property because of the presence of breaks in the series rather than an I(d) process.

1. INTRODUCTION

There have been several works analyzing the long-run properties of stock returns. Granger and Ding (1995a,b) considered long return series, using the well-known Standard and Poor's (S&P) 500 index of about 17,000 daily observations, and established a set of temporal and distributional properties for such series. They suggested that the absolute returns are well characterized by long memory process, but the parameter estimates of the long-memory model sometimes vary considerably from one subseries to the next as shown by Granger and Ding (1996). There are several attempts to explain these findings, such as Rydén, Teräsvirta, and Åsbrink (1998). They suggested that the temporal higher-order dependence observed in return series may be well described by a hidden Markov model. Such a model is estimated for the series of the S&P 500 from 1928 to 1991 which is the series considered by Granger and Ding (1996). However they failed to explain the one stylized fact that is the very slowly decaying autocorrelation function for the absolute returns. One possible alternative explanation for this time-varying long-memory property in the stock market might be de-

1529-7373/2013

rived from a recent paper of Granger and Teräsvirta (1999) using a simple nonlinear model.

The purpose of this paper is to explain the long memory property in the stock market by extending the idea of Granger and Teräsvirta (1999). Occasional structural breaks might cause the "long memory" property of absolute returns since Granger and Ding (1996) examined the series from 1928 to 1991. We suspect that there were structural changes in the absolute returns or in the volatility of returns during this long period. If such changes exist, a stationary process that encounters occasional regime switches will have some properties that are similar to those of a long-memory process.

The plan of the paper is as follows. Section 2 contains considerations of temporal properties of a linear model with occasional breaks. Section 3 introduces an estimation method of multiple breaks at unknown dates (Bai (1997) and Bai and Perron (1998)) or changes in variance (Inclan and Tiao, 1994). Section 4 contains our analysis of spurious long memory properties of a simple linear model with occasional breaks in mean. In Section 5 we discuss spurious breaks in I(d) process and overdifference caused by removal of estimated breaks. Section 6 is devoted to an application using the S&P 500 stock returns and finally Section 7 presents some extensions of this paper and conclusions.

2. OCCASIONAL BREAKS AND AUTOCORRELATION FUNCTION

We shall consider a simple linear model with occasional breaks in mean (Chen and Tiao (1990) and Engle and Smith (1999)).

$$y_t = m_t + \varepsilon_t \tag{1}$$

$$m_t = m_{t-1} + q_t\eta_t = m_0 + \sum_{i=1}^{t} q_i\eta_i \tag{2}$$

where $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$, $\eta_t \sim iid(0, \sigma_\eta^2)$ for $t = 1, \ldots, T$. Assume that q_t follows an $i.i.d$ binomial distribution as below[1]:

$$q_t = \begin{cases} 0 & \text{with probility } 1-p \\ 1 & \text{with probility } p \end{cases} \tag{3}$$

Then combining (1) and (2) yields

$$y_t = \varepsilon_t + m_0 + q_t\eta_1 + q_t\eta_2 + \cdots + q_t\eta_t \tag{4}$$

[1]Engle and Smith (1999) use an endogenous smooth transition function, $q_t(\gamma) = \varepsilon_t^2/(\gamma + \varepsilon_t^2)$, $\gamma > 0$, to capture different regime, i.e., random walk and stationary period.

One problem with the binomial model in (3) is that this model implies sudden change only. Structural change may occur gradually. One, therefore, can use a simple extension of (3), with q_t following a regime switching model:

$$q_t = \begin{cases} 0 & \text{when } s_t = 1 \\ 1 & \text{when } s_t = 2 \end{cases} \tag{3'}$$

Let s_t be a latent random variable with two discrete values: 1,2. Each value of s_t represents a different state in the length of memory of shock. s_t is assumed to be governed by the following Markov probability law: $p_{ij} = Pr(s_t = j|s_{t-1} = i)$. Then (1) can be solved exactly to yield (4). In this specification, the state of s_t will determine if the shock of η_t is permanent or not. A regime with $s_t = 2$ represents a period of structural change.

Now focus on the binomial case first. Assume that ε_t, η_τ and q_s are independent for all t, τ and s. For the initial conditions, let $m_0 = 0$, $q_t = 0$, $\varepsilon_t = 0$ and $\eta_t = 0$ for all $t \leq 0$, then the mean of y_t is $E(y_t) = 0$, and its variance is

$$\mathrm{var}(y_t) = tp\sigma_\eta^2 + \sigma_\varepsilon^2 \tag{5}$$

Similarly, one finds that the covariance between y_t and y_{t+k} is

$$\mathrm{cov}(y_t, y_{t+k}) = tp\sigma_\eta^2 \tag{6}$$

We can get the following k-th autocorrelation equation of this process,

$$\mathrm{corr}(y_t, y_{t+k}) = \frac{tp\sigma_\eta^2}{\sqrt{tp\sigma_\eta^2 + \sigma_\varepsilon^2}\sqrt{(t+k)p\sigma_\eta^2 + \sigma_\varepsilon^2}} \tag{7}$$

We can also derive properties of a sample autocorrelation function of this process. A k-th sample autocorrelation equation is,

$$\hat{\rho}_{T,k} = \frac{\sum_{t=1}^{T-k}(y_t - \overline{y})(y_{t+k} - \overline{y})}{\sum_{t=1}^{T}(y_t - \overline{y})^2} \approx \frac{\frac{\langle Tp\rangle\sigma_\eta^2}{6}\left(1 - \frac{k}{T}\right)\left(1 - 2\frac{k}{T} + 4\left(\frac{k}{T}\right)^2\right)}{\frac{\langle Tp\rangle\sigma_\eta^2}{6} + \sigma_\varepsilon^2} \tag{8}$$

for large T and k where $\overline{y} = \sum_{t=1}^{T} y_t/T$

$$\hat{\rho}_{T,k} \xrightarrow{p} \left(1 + \frac{6\sigma_\varepsilon^2}{\langle Tp\rangle\sigma_\eta^2}\right)^{-1}, \text{as } T \to \infty \text{ for fixed } k \text{ and } 0 < Tp < \infty. \tag{9}$$

$$\hat{\rho}_{T,k} \xrightarrow{p} 0, \text{ as } T \to \infty \text{ and } k/T \to 1. \tag{10}$$

In this process, Tp is an expected number of structural breaks within the sample period T and σ_η^2 will be related to the size of breaks. As explained in the following proposition, these parameters are closely linked to the basic properties of this process. Three qualitative conclusions can be drawn and are stated in Proposition 1.

Proposition 1. *A linear process with occasional breaks as described in (1) and (2) has the following properties.*

(a) Suppose that the probability of breaks converges to zero slowly as the sample size increases (i.e., $p \to 0$ as $T \to \infty$, yet $\lim Tp \to c < \infty$ where c is non-zero constant)[2]*, then a k-th sample autocorrelation in (8), $\hat{\rho}_{T,k}$, converges to nonzero value for fixed k as $T \to \infty$.*

(b) If $p > 0$, $\hat{\rho}_{T,k}$ appears to approximate the autocorrelations of an $I(1)$ process. In fact, since Tp increases to infinity as T increases to infinity, this process is an $I(1)$ process.

(c) When $p = 0$, this process is an $I(0)$ process.

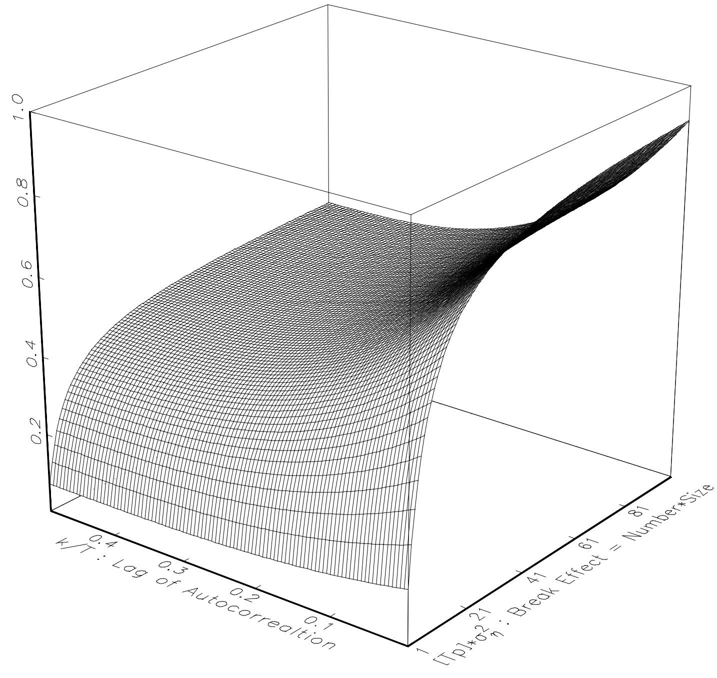
Proof. (a) As $T \to \infty$, $\hat{\rho}_{T,k} \xrightarrow{p} (1+\Delta)^{-1}$, where Δ is determined by the values of $T, p, \sigma_\varepsilon^2$, and σ_η^2 (see (9) and the 1st rows of Table 1). (b) As $T \to \infty$, and $Tp \to \infty$, we get $\hat{\rho}_{T,k} \to 1$ for any fixed k. (c) If $p = 0$, then $\hat{\rho}_{T,k} = 0$ for all k. ▮

Let's focus on the property of proposition 1(a) [spurious long-memory process] in this following paper. Proposition 1(a) reflects the finite sample property of autocorrelation, i.e., for T large (but finite) and Tp small but not zero, $\mathrm{corr}(y_t, y_{t+k})$ appears to approximate the autocorrelations of an $I(d)$ process which have a slow hyperbolic decay after the initial dropoff from $k = 0$ to $k = 1$. The autocorrelations in (8) will not decline exponentially (even if ε_t of equation (1) has serial correlation), but decay very slowly as in Table 1. Figure 1 exhibits slow hyperbolic decay of the autocorrelations as k increases. However, it will converge to zero as usual for the linear, stationary model or the stationary long memory model ($d < 0.5$) as k increases. For a given k, the sample autocorrelation approaches a nonzero constant as $T \to \infty$ as shown (9).

If $[Tp]$ is a small positive integer, a plot of y_t against t shows a few breaks in level since the value $[Tp]$ is in some sense the expected number of breaks within the sample period. As $[Tp]$ increases there will be more breaks and a higher value of the sample autocorrelation. The increase of σ_η^2, which means larger magnitude of breaks, has similar effects on the

[2]For example, let R be a number of breaks for the given sample size T, then $p = R/T$.

FIG. 1. Sample Autocorrelation



autocorrelations. The intuition is that an increase of Tp or σ_η^2 make this process closer to a random walk.

In the stock market, the correlogram of absolute stock return declines steadily but not exponentially. It may start with $\hat{\rho}_1 = 0.4$, say, and then declines only slowly from this value. Another stylized fact of financial data is that the correlogram is low but remains positive for many lags[3]. Granger and Ding (1996) suggest a fractionally integrated model amongst the models known to generate series having such properties. In this paper, we are suggesting a new class of model, a linear model with occasional breaks, as possible generating mechanism, instead of the $I(d)$ model with a fraction d.

[3]Granger and Marmol (1997) explain this property by using a process which consists of a stationary, long memory component plus a white noise component of much larger variance.

TABLE 1.

Theoretical Values of Autocorrelation

p, $[Tp]$	0.0025,	[5]	0.005,	[10]	0.01,	[20]	0.05,	[100]
$\sigma_\eta^2 = 0.1$	0.077	0.043	0.143	0.080	0.250	0.141	0.624	0.352
	0.070	0.004	0.126	0.008	0.217	0.015	0.543	0.025
	0.352	(3.22)	0.464	(4.23)	0.587	(5.47)	0.825	(8.09)
$\sigma_\eta^2 = 0.5$	0.294	0.165	0.454	0.256	0.624	0.352	0.892	0.502
	0.236	0.013	0.369	0.018	0.528	0.024	0.832	0.024
	0.585	(5.43)	0.707	(6.71)	0.815	(8.02)	0.951	(9.75)
$\sigma_\eta^2 = 1$	0.454	0.256	0.624	0.352	0.768	0.433	0.942	0.531
	0.352	0.017	0.510	0.020	0.670	0.023	0.903	0.022
	0.683	(6.46)	0.795	(7.78)	0.884	(8.90)	0.975	(10.11)
$\sigma_\eta^2 = 5$	0.806	0.454	0.892	0.502	0.942	0.531	0.987	0.556
	0.654	0.021	0.799	0.018	0.894	0.017	0.975	0.020
	0.858	(8.59)	0.929	(9.60)	0.970	(10.16)	0.997	(10.48)

Note: The number of series is $T = 2,000$ and $\sigma_\varepsilon^2 = 1$. $[Tp]$ means the number of structural breaks within sample, and σ_η^2 related the size of breaks. The first numbers are theoretical values of the autocorrelation at $k = 1$ and 500, by using (8). The numbers in the second rows are average values of the autocorrelations at $k = 1$ and 500, which are simulated from 1,000 replications with 2,000 sample size. The numbers in the third rows are the estimated d by the GPH method and t-values in the parenthesis.

3. ESTIMATION OF BREAK POINTS AND NUMBER OF BREAKS

As discussed before, the aim of this work is to investigate the properties of time series with multiple unknown structural breaks. This may include the case when there are a few known breaks in a finite sample. One could deal with structural breaks in variance in the same way. Our goal is to estimate break points, the number of breaks and their sizes. Bai (1997) and Bai and Perron (1998) suggests sequential estimation of multiple breaks in mean, and show that this method can estimate break points consistently with order T even when the number of breaks in practice is unknown. To detect discrete changes in the variance of an observed time series, we use the ICSS (iterated cumulative sum of squares) algorithm presented by Inclan and Tiao (1994).

Bai (1997)'s procedure works as follows. When the first break point is identified at k, the whole sample is divided into two subsamples with the first subsample consisting of k observations and the second subsample consisting of the rest of the observations $(T-k)$. One then estimates a break point for the subsample where a hypothesis test of parameter constancy is rejected. Divide the corresponding subsample further into subsamples at the newly estimated break point, and perform parameter constancy test for the hierarchically obtained subsamples. This procedure is repeated until the parameter constancy test is not rejected for all subsamples. The number of break points is equal to the number of subsamples minus 1. He shows how the sequential procedure coupled with hypothesis testing can yield a consistent estimate for the true number of breaks.

Although asymptotic theory implies that the sequential procedure will not underestimate the number of breaks, Bai (1997) shows by Monte Carlo simulations that the procedure has a tendency to underestimate. This problem can be overcome by using a two-step procedure as suggested by Bai (1997). In the first step, the goal is to obtain a consistent (or less biased) estimate for the error variance. This can be achieved by allowing more breaks (R), solely for the purpose of constructing the error variance. It is evident that as long as $R \geq R_0$ (the true number of breaks) the error variance will be consistently estimated. Obviously one does not know whether $R \geq R_0$, but the specification of R in this stage is not as important as in the final model estimation. When R is fixed, the R break points can be selected either by the Schwarz-Bayesian criterion simultaneously or by the "one additional break" sequential procedure described in Bai and Perron (1998). In the second step, the number of breaks is determined by the sequential procedure coupled with hypothesis testing. The test statistics use the error variance estimator obtained in the first step. The details of the test statistic and methods were given in Bai (1997).

4. SPURIOUS LONG-MEMORY PROPERTY OF OCCASIONAL STRUCTURAL BREAKS PROCESS

Suppose that m_t is the series from (2) that, when differenced once, gives the series

$$\Delta m_t = (1 - L)m_t = q_t \eta_t$$

which has an ARMA representation (here we assume a white noise process in (2)). m_t will then be called an integrated series, and denoted $m_t \sim I(1)$. If Δm_t has spectrum $f(\omega)$, then m_t does not strictly posses a spectrum, but from filtering considerations the spectrum of m_t can be thought of as

$$f_m(\omega) = |1 - z|^{-2} f(\omega), \quad \omega \neq 0$$

where $z = e^{-i\omega}$. Since Δm_t is strictly ARMA with $f(\omega) = \frac{1}{2\pi} p\sigma_\eta^2$, then

$$\lim_{\omega \to 0} f(\omega) = \frac{1}{2\pi} p\sigma_\eta^2$$

For small ω,

$$f_m(\omega) = c\omega^{-2}$$

where $c = p\sigma_\eta^2/2\pi$.

Since m_t is independent of ε_s for all t and s the spectrum of y_t in (1) is then

$$f_y(\omega) = f_m(\omega) + f_\varepsilon(\omega) = |1-z|^{-2}\frac{1}{2\pi}p\sigma_\eta^2 + \frac{1}{2\pi}\sigma_\varepsilon^2, \quad \omega \neq 0 \tag{11}$$

It follows that

$$f_y(\omega) = c\omega^{-2} + c', \text{ for } \omega \text{ small.} \tag{12}$$

where $c = p\sigma_\eta^2/2\pi$, $c' = \sigma_\varepsilon^2/2\pi$.

Geweke and Porter-Hudak (1983) (henceforth GPH) show that, when attention is confined to frequencies near zero, the differencing parameter can be estimated consistently from the least square regression since

$$\frac{\partial \ln(f_z(\omega))}{\partial \ln(\omega)} = -2d, \text{ with some } z_t \sim I(d) \text{ series.}$$

If we apply this to (12), then

$$-\frac{1}{2}\frac{\partial \ln(f_y(\omega))}{\partial \ln(\omega)} = \left(1 + \omega^2 \frac{\sigma_\varepsilon^2}{p\sigma_\eta^2}\right)^{-1}$$

If we evaluate at $\omega_0 = -2\pi T^{-1/2}$ using $g(T) = T^{1/2}$, which is a popular choice in GPH estimation, then

$$\tilde{d} = -\frac{1}{2}\left.\frac{\partial \ln(f_y(\omega))}{\partial \ln(\omega)}\right|_{\omega=\omega_0} = \left(1 + 4\pi^2 \frac{\sigma_\varepsilon^2}{Tp\sigma_\eta^2}\right)^{-1} \tag{13}$$

where $0 < \tilde{d} < 1$. Of course it is true $\tilde{d} \to 1$ as $\omega \to 0$. However, we would get an estimated d less than 1 only because $\sigma_\varepsilon^2/(Tp\sigma_\eta^2)$ is small enough to counter-balance small w and an estimated d is calculated from $0 < \omega < \omega_0$ in a finite sample estimation.

In support, we conduct a Monte Carlo analysis of GPH estimation and Lobato and Robinson's LM test (Lobato and Savin, 1998) using the series with occasional breaks. The purpose of our Monte Carlo simulation is to investigate whether the apparent long-run dependence could be due to occasional structural breaks in a finite sample. Instead of inspecting the autocorrelation, an easy way to summarize these correlograms is to present the value of the estimated d by the GPH method or Lobato and Robinson's LM statistics.

$$\log\{I(\lambda_j)\} = c - d\log\{4\sin^2(\lambda_j/2)\} + u_j, \; j = 1, \ldots, g(T). \tag{14}$$

where $I(\lambda_j) = \frac{1}{2\pi T}|\sum_{i=1}^{T} y_t \exp(i\lambda_j t)|^2$ is the periodogram at frequency $\lambda_j = 2\pi j/T$ which will depend on the sample size T. The following (15) is Lobato and Robinson's LM test statistic to test $H_0 : d = 0$ against $H_a : d \neq 0$,

$$LM = m\left(\sum_{j=1}^{m} v_j I(\lambda_j)/\sum_{j=1}^{m} I(\lambda_j)\right)^2 \tag{15}$$

where $v_j = \log j - \frac{1}{m}\sum_{j=1}^{m}\log j$ and $m = g(T)$.

The following two artificial data sets illustrate the long memory property of a linear model with occasional breaks with different parameters.

FIG. 2-1. A Series with Occasional Breaks

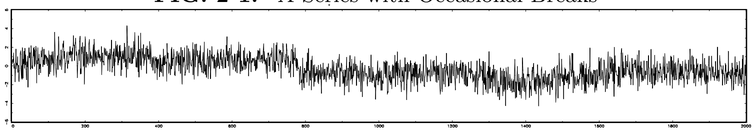


FIG. 2-2. Break Points (Binomial Distribution)

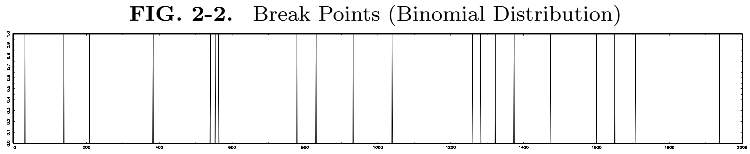


FIG. 2-3. Autocorrelations

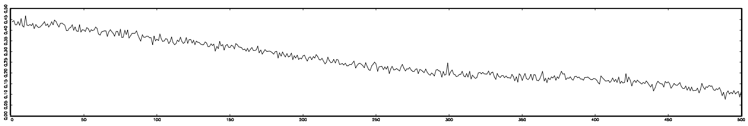
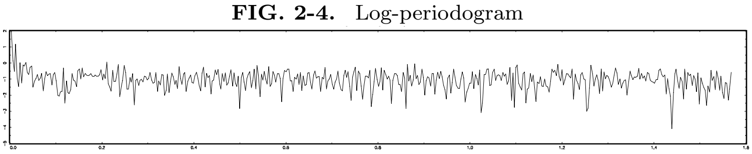


FIG. 2-4. Log-periodogram



Example 1: q_t has a binomial distribution of (3). Length of series = 2,000, $\sigma_\varepsilon^2 = 1$, $\sigma_\eta^2 = 0.25$, $p = 0.01$. The second graph of Figure 2 plots sample autocorrelations of this series up to lag 500. Sample autocorrelations start around 0.44 (quite close to 0.45 calculated by (8) with $k = 1$),

but decrease very slowly with increasing lags and the corresponding periodogram in Figure 2 clearly has a pole at the origin. These figures might suggest that this series has long memory. By the GPH method, $\hat{d}$ is 0.747 with t-value = 6.17, a clear long memory.

FIG. 3-1. A Series with Occasional Breaks

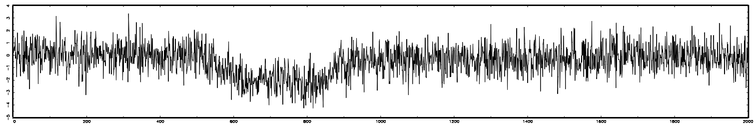


FIG. 3-2. Break Periods (Markov Switching Process)

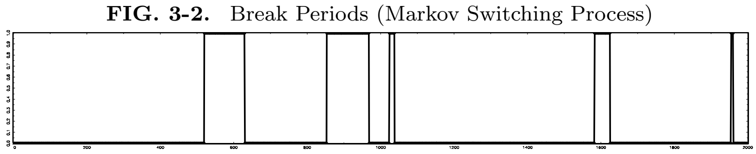


FIG. 3-3. Autocorrelations

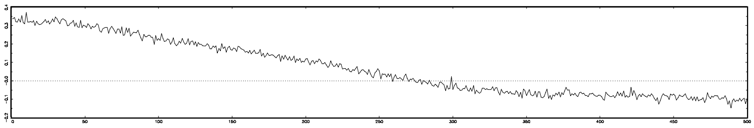
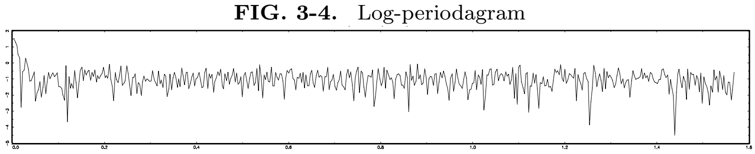


FIG. 3-4. Log-periodagram



Example 2: q_t has a regime switching process (3'). Length of series = 2,000, $\sigma_\varepsilon^2 = 1$, $\sigma_\eta^2 = 0.01$, $p_{11} = 0.998$, $p_{22} = 0.99$, implying unconditional probability of state 2 (break) is 0.167. A sample autocorrelation at $k = 1$ is 0.337 close to 0.357 calculated by (8). The second graph of Figure 3 plots the value of q_t which is determined by the state variable, and $\hat{d}$ for this series is 0.860 with t-value = 5.38. We could observe gradual changes when $s_t = 2$.

Estimated values of d by GPH in the simulation are presented in the 3rd rows of Table 1 using a sample size of 2,000 and $\sigma_\varepsilon^2 = 1$ with different values of p and σ_η^2. Simliarly to Granger and Teräsvirta (1999)'s examples, this process has long-memory rather than short-memory, if we just consider linear properties of the data. As Tp increases there will be more breaks

TABLE 2.

Estimated d for 20 Sub-periods

	Case I		Case II		Case III		Case IV	
Period	d	t-stat	d	t-stat	d	t-stat	d	t-stat
1	0.405	2.754	0.337	1.923	0.840	5.121	0.668	5.831
2	0.031	0.151	0.204	1.117	0.274	1.649	0.595	3.836
3	0.079	0.634	0.493	4.946	0.079	0.634	0.798	8.035
4	0.155	1.456	0.461	2.729	0.155	1.456	0.802	4.247
5	−0.105	−0.884	0.313	2.109	−0.105	−0.884	0.923	5.095
6	−0.167	−1.197	0.347	2.446	−0.166	−1.439	0.710	6.945
7	−0.087	−0.891	0.238	2.317	−0.148	−1.437	0.568	4.441
8	−0.145	−1.496	0.292	3.019	−0.115	−1.142	0.708	5.250
9	−0.067	−0.532	0.226	1.883	−0.067	−0.532	0.638	4.811
10	−0.003	−0.020	0.419	2.850	0.210	1.267	0.918	6.434
11	0.142	1.563	0.517	5.346	0.142	1.563	0.927	8.025
12	0.003	0.034	0.362	3.289	0.003	0.034	0.765	4.564
13	0.252	1.842	0.723	5.744	0.252	1.842	1.005	10.343
14	−0.058	−0.548	0.309	3.341	−0.058	−0.548	0.722	3.602
15	0.262	2.128	0.445	3.971	0.374	2.067	0.868	6.767
16	−0.063	−0.592	0.386	2.653	−0.063	−0.592	0.802	4.800
17	0.083	0.632	0.284	2.428	0.083	0.632	0.621	6.322
18	−0.239	−1.975	0.269	2.515	−0.239	−1.975	0.728	5.492
19	0.123	1.127	0.288	2.147	0.336	2.843	0.849	4.833
20	0.068	0.554	0.711	3.947	0.068	0.554	0.978	7.317

Note: The number of total sample is 20,000 and the size of each sub-sample is 1,000. d is estimated by the GPH method and $g(T) = 32 \approx 1,000^{0.5}$. (Tp, σ_η^2) for Case I$= (20, 0.1)$, Case II $= (200, 0.1)$, Case III $= (20, 0.5)$, and Case IV $= (200, 0.5)$ where Tp means the expected number of structural breaks within sample.

and a higher value of $\hat{d}$ is obtained from the regression (14). An increase of σ_η^2 will have similar effects on $\hat{d}$ since it is more likely that breaks are detected. Table 2 shows estimated values of d from 20 subsamples of size 1,000. The $\hat{d}$'s based on the total sample are 0.376 for Case I, 0.782 for Case II, 0.687 for Case III and 0.888 for Case IV. For Case I and III, only a few subsamples have significant changes, so most of them look like white noise. For Case II and IV, d seems to be time-varying, for example, going from 0.337 in period 1 to 0.204 in period 2 and upto 0.723 in period 13 in Case II. As Tp or σ_η^2 increase (i.e., as the number of breaks or the size of breaks increase), a higher value of estimated d is obtained from the regression. The value of $\hat{d}$ will depend on the values of Tp or σ_η^2 of each subsample.

In Table 3, Lobato and Robinson's LM test was conducted for the various values of parameters, $T = 200$, $p = 0.025, 0.05, 0.1$, $\sigma_\varepsilon^2 = 1$, and $\sigma_\eta^2 = 0.001$,

TABLE 3.

Rejection Rates of the Null of Stationarity against $I(d)$ Process

p (Tp)	0.025 (5)	0.05 (10)	0.1 (20)
$\sigma_\eta^2 = 0.001$	6.6	6.4	8.5
	1.0	1.6	1.9
0.005	7.2	9.9	12.9
	1.7	3.3	7.2
0.01	10.4	12.7	20.0
	2.5	7.0	14.0
0.05	23.6	36.5	55.0
	17.7	33.6	52.8
0.1	33.6	52.6	72.7
	30.9	54.4	70.6
0.5	69.2	86.1	93.1
	68.9	85.9	93.9
1	81.6	91.1	95.7
	80.2	90.0	94.8

Note: The first numbers equals the % of $t-$value > 1.645 in the GPH. The number in the second rows is the % of p-value of the LM test which is less than 0.05. The results are based on 1,000 replications with 200 sample size.

0.005, 0.01, 0.05, 0.1, 0.5, 1. As the value of Tp or σ_η^2 is getting larger, the rejection rate of the LM test is close to 100% spuriously. The GPH shows similar results. For the same value of $Tp\sigma_\eta^2$, the rejection rates are similar. For example, when $p = 0.1$ (20 breaks on average within sample) and $\sigma_\eta^2 = 0.05$, the rejection rate of the LM test is 52.8% which is reasonably close to the value, 54.4% with $p = 0.05$ (10 breaks) and $\sigma_\eta^2 = 0.1$.

5. SPURIOUS NUMBER OF BREAKS AND $I(D)$ PROCESS

Nonstationarity of the DGP causes many breaks to be detected spuriously by standard estimation methods. Unlike stationary processes, $I(d)$ or $I(1)$ processes would have different effects on the estimated number of breaks. Suppose y_t is generated from $(1-B)^d y_t = \varepsilon_t$, when d is not an integer, $0 < d < 1$, the fractional difference operator $(1-B)^d$ defined by its Maclaurin series is,

$$(1-B)^d = \sum_{j=0}^{\infty} \pi_j B^j, \ \pi_j = \frac{j-1-d}{j}\pi_{j-1}, \ \pi_0 = 1 \qquad (16)$$

FIG. 4-1. $I(d)$ process ($d = 0.2$) and Estimated Level Shifts

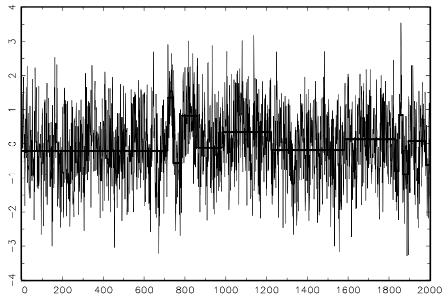


FIG. 4-2. $I(d)$ process ($d = 0.4$) and Estimated Level Shifts

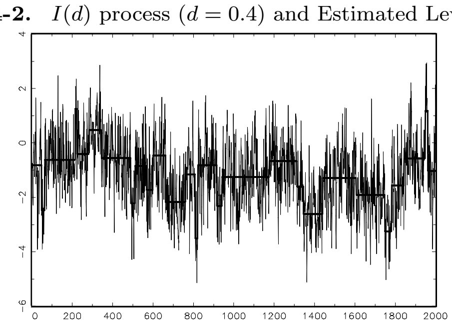
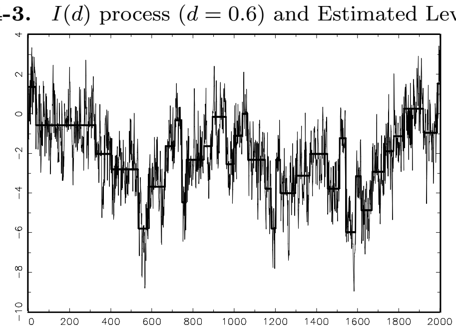
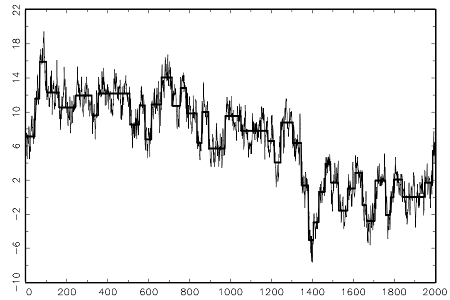


FIG. 4-3. $I(d)$ process ($d = 0.6$) and Estimated Level Shifts



In Table 4, the number of breaks is estimated by Bai's method for various values of d in the DGP. To simulate the series it is assumed that $\pi_j = 0$ for $j > 1,000$, and the first 2,000 observations are discarded. It is initially assumed that the maximum number of breaks is 50 and we do not estimate a break if the size of a subsample is less than 50 observations. Figures 4-1 - 4-4 show examples of plots of $I(d)$ processes and estimated means of each regime: (4-1) $\hat{d} = 0.092$ ($t-$value $= 1.04$), 12 breaks, (4-2) $\hat{d} = 0.500$ (4.18), 26 breaks, (4-3) $\hat{d} = 0.519$ (5.51), 32 breaks, and (4-4) $\hat{d} = 0.808$ (7.80),

FIG. 4-4. $I(d)$ process ($d = 0.8$) and Estimated Level Shifts



and 42 breaks. These graphs clearly suggest a positive relation between the number of breaks and the value of d as in the Table 4. After removing the break component from the original series, $y_t - m_t$, we find some evidences of overdifferencing, i.e., the estimated value of d is less than zero. This might suggest that estimated breaks in $I(d)$ may be spurious.

TABLE 4.

Estimated d for Decomposed $I(d)$ Process into Break Process and Residual

DGP of $I(d)$	Number of Breaks	$y_t : d$ (t-value)	$m_t : d$ (t-value)	$y_t - m_t : d$ (t-value)
$d = 0.2$	8.19	0.202	0.630	−0.168
		(1.90)	(6.05)	(−1.53)
$d = 0.4$	19.76	0.409	0.657	−0.453
		(3.82)	(6.04)	(−4.07)
$d = 0.6$	28.36	0.618	0.797	−0.623
		(5.78)	(7.43)	(−5.64)
$d = 0.8$	34.77	0.827	0.953	−0.708
		(7.84)	(9.04)	(−6.41)

Note: Numbers of breaks are estimated by Bai's method. d is estimated by the GPH method. Results are based on 1,000 replications with 2,000 sample.

We could further investigate if the estimated number of breaks has something to do with the underlying DGP. Since the Schwarz-Bayesian criterion (SBC) can be considered as an alternative for which might be inferred from visual inspection of a time series, we can show that spuriously[4] many breaks

[4]They might be interpreted as real breaks in a finite sample approximation.

(but with different numbers depending on value of d) will be inferred when the process is $I(1)$[5] or $I(d)$ without breaks in the parameters of the DGP.

Consider the following simple model of R structural breaks:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim iidN(0, \sigma_\varepsilon^2),$$

and $\mu_t = \mu^i$, for $k_{i-1} < t \leq k_i$, where $i = 1, \ldots, R+1$, $k_0 = 0, k_{R+1} = T$ and $0 < k_1 < \cdots < k_R < T$.

Then the maximum likelihood estimator of σ^2 is

$$\hat{\sigma}_T(R)^2 = \min_{0<k_1<\cdots<k_R<T} \frac{1}{T}\sum_{t=1}^{T} y_t^2 - \frac{1}{T}\sum_{r=1}^{R+1} \frac{1}{k_r - k_{r-1}} \left(\sum_{t=k_{r-1}+1}^{k_r} y_t \right)^2 \tag{17}$$

Using SBC, the estimated number of break points R is found by

$$\hat{R} = \arg\min_R SBC(R) = \log(\hat{\sigma}_T(R)^2) + (1+2R)\frac{\log T}{T} \tag{18}$$

subject to $R < R_U$, with R_U a given fixed upper bound for R. Since the second term in (18) goes to 0 as T increases, for any fixed R, only the first term in (18) matters asymptotically. For the given large T, we would have different value of SBC with different number of R because of the second term in (17).

In this section, we focus on following data generating processes without break, where ε_t is white noise:

(I) $y_t = \varepsilon_t$,
(II) $(1-B)^d y_t = \varepsilon_t$ with $0 < d < 1/2$,
(III) $(1-B)^d y_t = \varepsilon_t$ with $1/2 < d < 1$, and
(IV) $(1-B)y_t = \varepsilon_t$.

The following asymptotic facts will be useful to obtain the results of Proposition 2:

(II') $T^{-1/2-d} S_{[T\gamma]} = O_p(1)$, with $0 < d < 1/2$,
(III') $T^{-1/2-d} S_{[T\gamma]} = O_p(1)$, with $1/2 < d < 1$, and

PROPOSITION 2. *Assume y_t follows four different DGP of (I), (II), (III) and (IV) with no break, then the estimated numbers, $\hat{R}$, of breaks by the SBC method as $T \to \infty$ are*

(a) $\hat{R} \to 0$, *with* $d = 0$,
(b) $\hat{R} \to \infty$, *with* $0 < d < 1/2$
(c) $\hat{R} \to \infty$, *with* $1/2 < d < 1$

[5] Nunes, Kuan and Newbold (1995) showed that the SBC failed to estimate the true number of breaks, but estimated the maximum permitted number of breaks when the DGP was a random walk.

(d) $\hat{R} \to \infty$, *with* $d = 1$.

Proof. The proofs of part (a) and (d) are in Yao (1988) and Bai (1998). The proof of part (b) is as follows: Define

$$M_T^*(k) = T^{-1-2d}\left\{\frac{1}{k}\left(\sum_{t=1}^{k}\varepsilon_t\right)^2 + \frac{1}{T-k}\left(\sum_{t=k+1}^{T}\varepsilon_t\right)^2\right\}$$

For an arbitrary projection matrix P, we have $z'Pz \le z'z$. Then $\sup_{\lambda\in(0,1)} M^*(\lambda) = O_p(1)$ from Theorem 1 of Bai (1998), since $M_T^*(k) \le T^{-1-2d}\sum_{t=1}^{k}\varepsilon_t^2 \le T^{-1}\sum_{t=1}^{k}\varepsilon_t^2 = O_p(1)$ for $0 < d < 1/2$. By applying Theorem 2 of Bai (1998), we have

$M^*(0) = M^*(1) < M^*(\lambda)$, for every $0 < \lambda < 1$ with probability 1, since $T^{-1/2-d}S_{[T\gamma]} = T^{-1/2-d}\sum_{t=1}^{[T\gamma]}\varepsilon_t \Rightarrow G(r)$ where $G(.)$ has a continuous distribution for each r. The proof of part (c) is obvious since

$$M_T^*(k) \le T^{-1-2d}\sum_{t=1}^{k}\varepsilon_t^2 \le T^{-2d}\sum_{t=1}^{k}\varepsilon_t^2 = O_p(1) \text{ for } 1/2 < d < 1.$$

∎

In Table 5, we present simulation results of proposition 2 in a small sample. Six different DGPs are considered and show a positive relation between the number of breaks and the value of d in a finite sample. When the DGP is an $I(d)$ or a random walk, the SBC selects the maximum permitted number of breaks on the majority of occasions as d close to 1. Clearly the $I(d)$ process is an intermediate process between $I(0)$ and $I(1)$ in terms of the estimated number of breaks in finite sample.

6. OCCASIONAL BREAK IN THE STOCK MARKET AND LONG MEMORY

In this section we investigate the "long memory" property in the stock market. As a preliminary analysis, we plot stock returns, absolute returns and the 220 days moving average of absolute returns from January 4, 1928 to August 30, 1991 with 17054 daily observations in Figure 5. By the GPH, $\hat{d}$ is 0.475 with t-value = 8.22. However, it can be seen from Figure 5-3 that large absolute returns are more likely to be followed by large absolute return than small returns. For example, the October 1987 crash (16,076-th observation) significantly increased volatility for a while.

It is possible to suggest two conjectures. The first one is that the absolute returns follow a linear process with occasional breaks (If Tp is a constant

TABLE 5.

Number of Breaks Selected by the SBC

R_U	R	$I(0)$	$d=.2$	$d=.4$	$d=.6$	$d=.8$	$I(1)$
1	0	92.3	62.6	30.7	12.3	5.8	1.4
	1	7.7	37.4	69.3	87.7	94.2	98.6
2	0	90.9	54.9	19.9	5.5	2.0	0.3
	1	7.3	28.4	37.1	25.5	13.4	4.5
	2	1.8	16.7	43.0	69.0	84.6	95.2
3	0	90.9	54.2	19.1	4.7	1.5	0.3
	1	7.3	27.2	31.7	17.6	6.4	1.3
	2	1.8	14.4	30.3	32.0	24.8	12.2
	3	0.0	4.2	18.9	45.7	67.3	86.2
4	0	90.9	54.2	19.1	4.7	1.5	0.2
	1	7.3	27.2	31.3	16.7	5.5	1.3
	2	1.8	14.1	28.2	28.1	18.5	8.5
	3	0.0	4.2	15.2	30.5	29.2	24.7
	4	0.0	0.3	6.2	20.0	45.3	65.3

Note: Results based on 1,000 replications with 50 observations, We consider 6 different DGP's, i.e., white noise, $I(d)$ and random walk without breaks.

FIG. 5-1. S&P 500 Daily Retures 1/4/28-8/30/91

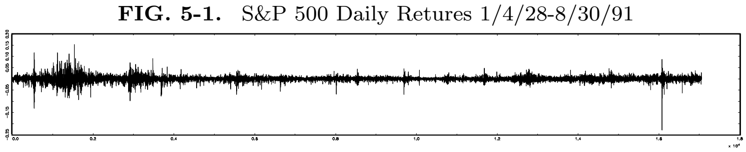


FIG. 5-2. Absolute Returns

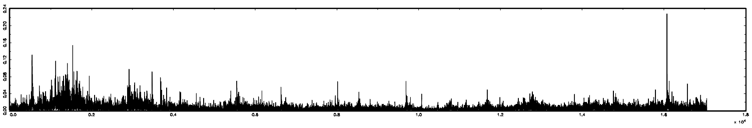
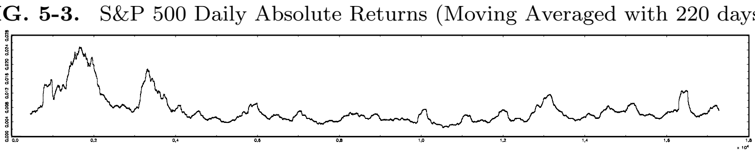
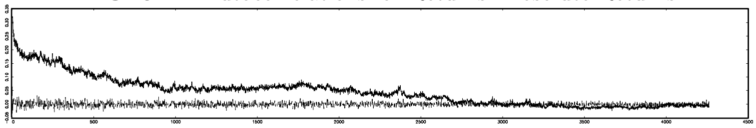


FIG. 5-3. S&P 500 Daily Absolute Returns (Moving Averaged with 220 days)



positive integer as T increases to infinite, this process is still not explosive. That means the variance of this series is bounded). The second one is that the breaks may not be determined by the size of shock as suggested by the

FIG. 5-4. Autocorrelations for Returns. Absolute Returns



specification of Engle and Smith (1999). Sometimes a big shock could be a transitory, not permanent effect on the volatility of returns in Figure 5-2.

There would be several competing theories to explain the long memory properties, and also time-varying d in the stock market volatility. We suggest that this series would be well characterized by a linear model with occasional breaks. An $I(d)$ process could be a possible alternative but it has its own drawbacks, for example,

PROPOSITION 3. *Let $\varepsilon_t > 0$, $y_t > 0$ for all t, and $(1-B)^d y_t = \varepsilon_t$, then y_t is a process with upward trend component.*

Proof. Even for $d < 1/2$, $y_t = \sum_{j=0}^{\infty} d_j \varepsilon_{t-j}$ such that $\sum_{j=0}^{\infty} d_j^2 < \infty$ but $\sum_{j=0}^{\infty} |d_j| \to +\infty$.

Since $\varepsilon_t > 0$ for all t, $y_t = \sum_{j=0}^{t} d_j \varepsilon_{t-j} \to +\infty$ as $t \to \infty$. ▮

That is, if we characterize stock return volatility by an $I(d)$ process, it implies explosive volatility. It might be a good approximation in finite sample, but can not be a true DGP of stock return volatility.

Lobato and Savin (1998) point out that the evidence in favor of long memory in stock return might be due to structural breaks, and they investigate if the observed evidence of long memory is, in fact, due to nonstationarity during long period. They split their sample (1962-1994) into two periods, taking January 1973 as the break point. But they do not find any evidence that long memory was caused by the structural break of 1973. A different, but related work by Aggarwal, Inclan and Leal (1999) indicates how to investigate the effect of breaks on the long memory property. An iterated cumulative sums of squares (ICSS) algorithm is used to identify the points of breaks in the variance of returns in the stock market and how long the shift lasts. Dummy variables are then introduced into the variance equation of a GARCH model to account for the sudden changes in variance, to yield a combined model with GARCH and dummy variables.

Instead of Lobato and Savin's approach (using pre-determined break), we estimate unknown structural breaks in the stock market by adapting the method of Aggarwal, Inclan and Leal, and check whether these breaks contain long memory component. In our analysis, the structural breaks in volatility detected by either Bais method or the ICSS algorithm can be used directly to decompose stock returns into a break component and residuals.

TABLE 6.
Estimated d and LM statistics of the Absolute Stock Returns, Break Process and Residuals

		Abs. return ($\|y_t\|$)			Break Process (m_t)		Residual ($\|y_t\|/m_t$)	
	time	d (t-stat)	LM (p-val)	Break	d (t-stat)	LM (p-val)	d (t-stat)	LM (p-val)
1	1928-1934	0.352	49.35	10	0.520	88.61	0.034	0.163
		(2.72)	(0.00)		(3.75)	(0.00)	(0.34)	(0.69)
2	1934-1940	0.405	52.08	9	0.583	85.01	0.104	1.001
		(3.10)	(0.00)		(6.00)	(0.00)	(0.84)	(0.31)
3	1941-1947	0.438	18.34	5	0.519	31.10	0.082	1.308
		(4.51)	(0.00)		(5.36)	(0.00)	(0.65)	(0.25)
4	1947-1953	0.347	14.97	8	0.492	40.06	0.137	0.827
		(2.31)	(0.00)		(5.81)	(0.00)	(1.12)	(0.36)
5	1954-1960	0.154	15.44	9	0.567	82.80	0.091	0.859
		(1.49)	(0.00)		(5.71)	(0.00)	(0.77)	(0.35)
6	1960-1966	0.451	14.87	5	0.686	41.23	0.082	0.737
		(4.11)	(0.00)		(7.90)	(0.00)	(0.70)	(0.39)
7	1967-1973	0.517	38.38	11	0.681	57.29	0.151	0.039
		(5.96)	(0.00)		(5.99)	(0.00)	(1.06)	(0.84)
8	1973-1979	0.715	129.67	17	0.746	131.52	0.109	0.517
		(6.80)	(0.00)		(6.74)	(0.00)	(1.16)	(0.47)
9	1980-1986	0.418	34.63	17	0.499	42.88	0.231	0.002
		(3.80)	(0.00)		(4.64)	(0.00)	(1.60)	(0.97)
10	1986-1991	0.352	20.73	9	0.350	26.66	0.117	0.640
		(5.00)	(0.00)		(7.99)	(0.00)	(1.07)	(0.42)

Note: 1705 daily absolute stock returns, $|y_t|$, for each subperiod. After identifying breaks in mean of the absolute returns by Bai's method, decompose the absolute stock returns into break component (m_t) and break-free component ($|y_t|/m_t$). m_t is a sample mean of each regime.

We found clear evidence of a positive relation between the time-varying property of d and the number of breaks (see Tables 6 - 8).

For the analysis of variance, we introduce a simple version of occasional breaks in variance. Let $\{y_t\}$ be independent observations from a normal distribution with zero mean and variance σ_t^2. When the total number of variance changes in T observations is R, and $1 < k_1 < k_2 < \cdots < k_R < T$ are the set of change points, the variance between two consecutive break points is denoted by τ_j^2, $j = 0, 1, \ldots, R$. One could use the following GARCH process as a general case of a process with breaks in variance[6]:

$$y_t = \sigma_t \nu_t, \nu_t \sim iid(0,1), \ \sigma_t^2 = E(y_t^2 | I_{t-1}) \qquad (19)$$

[6]Recently Mikosch and Starica (1999) show long range dependence of financial time series by using shifts in the variance.

TABLE 7.

Estimate Breaks in Variance by the ICSS, and Apply to the Absolute Stock Returns

			Break Process (m_t)		Residual ($\|y_t\|/m_t$)	
	time	Breaks	d (t-stat)	LM (p-val)	d (t-stat)	LM (p-val)
1	1928-1934	9	.669 (9.58)	129.57 (0.00)	−.041(−0.31)	0.693 (0.41)
2	1934-1940	7	.691 (7.16)	105.36 (0.00)	−.102(−0.80)	1.149 (0.28)
3	1941-1947	17	.458 (4.61)	17.70 (0.00)	−.363(−2.59)	0.580 (0.45)
4	1947-1953	14	.397 (4.48)	27.31 (0.00)	−.304(−2.87)	3.253 (0.07)
5	1954-1960	8	.608 (6.77)	83.40 (0.00)	−.132(−1.16)	0.964 (0.35)
6	1960-1966	11	.431 (4.67)	21.99 (0.00)	.084 (1.09)	0.052 (0.82)
7	1967-1973	17	.567 (6.58)	40.52 (0.00)	−.519(−4.89)	3.782 (0.05)
8	1973-1979	9	.704 (5.22)	158.85 (0.00)	−.204(−1.99)	0.196 (0.66)
9	1980-1986	4	.986 (11.06)	120.24 (0.00)	.125 (1.11)	1.628 (0.20)
10	1986-1991	8	.278 (6.375)	18.33 (0.00)	.080 (0.89)	0.697 (0.40)

Note: Estimate breaks in variance of stock returns by the ICSS algorithm for 10 sub-samples with 1,705 observations. Not estimate break if the size of observation is less than 50 observations. After identifying break points, decompose the absolute stock returns into break component (m_t) and break-free component ($|y_t|/m_t$). m_t is a sample mean of $|y_t|$ of each regime. $|y_t|$ is the absolute stock returns.

$$\sigma_t^2 = \omega_t + \beta\sigma_{t-1}^2 + \alpha y_{t-1}^2 \tag{20}$$
$$\omega_t = \omega_{t-1} + q_t\eta_t,\ \eta_t \sim iid(0, \sigma_\eta^2) \tag{21}$$

where ν_t and η_s are independent for all t and s. In the GARCH equation (20), the parameters α, β are nonnegative and ω_t are positive for all t. It might be better to use EGARCH instead of GARCH because no restriction on α, β and ω_t are necessary. Let $\alpha = \beta = 0$, for the simplest case, i.e., this is a process with occasional breaks in variance.

After identifying unknown structural breaks, we could obtain a break process and a series, y_t^*, without any break component. If we estimate break points in level, we could get y_t^*, by dummy variables regression:

$$y_t^* = y_t - (a_0 + a_1 D_1 + \cdots + a_{R-1} D_{R+1}) \tag{22}$$

where $D_i = \begin{cases} 1 & \text{if } t \in \text{regime } i \\ 0 & \text{otherwise} \end{cases}$, for $i = 1, \ldots, R+1$.

Although there is a similarity between a model for breaks in mean and in variance, the method of removing the break component in (19) - (21) is completely different from (22). Instead divide series y_t by $\tau_i = \omega_t$, for each different regime i as below:

$$y_t^* = y_t/\tau_i,\ \text{if } y_t \text{ belongs to regime } i = 1, \ldots, R+1. \tag{23}$$

TABLE 8.

Estimate Breaks in Variance by the ICSS, and Estimate d and LM statistics of the Squared Stock Return, Break Process and Squared Residuals

		Squ. Return (y_t^2)			Break Process (m_t)		Residual $(y_t/m_t)^2$	
	time	d	LM	Break	d	LM	d	LM
		(t-stat)	(p-val)		(t-stat)	(p-val)	(t-stat)	(p-val)
1	1928-1934	0.218	15.14	9	0.668	133.93	−0.077	0.709
		(1.85)	(.000)		(9.09)	(0.00)	(−0.64)	(0.40)
2	1934-1940	0.457	35.85	7	0.707	105.39	−0.110	0.890
		(4.68)	(.000)		(8.03)	(0.00)	(−0.73)	(0.35)
3	1941-1947	0.445	8.23	17	0.483	19.48	−0.449	1.962
		(2.37)	(.004)		(4.15)	(0.00)	(−3.78)	(0.16)
4	1947-1953	0.200	7.69	14	0.441	31.89	−0.295	3.672
		(2.04)	(.006)		(5.02)	(0.00)	(−2.87)	(0.06)
5	1954-1960	0.106	5.61	8	0.564	62.41	−0.145	0.658
		(0.87)	(.018)		(6.75)	(0.00)	(−1.16)	(0.42)
6	1960-1966	0.286	5.63	11	0.351	17.90	0.112	0.046
		(3.37)	(.018)		(3.21)	(0.00)	(1.15)	(0.83)
7	1967-1973	0.485	23.33	17	0.619	47.85	−0.463	4.734
		(4.95)	(.000)		(7.34)	(0.00)	(−3.76)	(0.03)
8	1973-1979	0.925	108.75	9	0.834	158.83	−0.334	0.401
		(6.67)	(.000)		(7.82)	(0.00)	(−2.77)	(0.53)
9	1980-1986	0.341	22.51	4	0.985	115.74	−0.138	0.271
		(3.17)	(.000)		(10.53)	(0.00)	(−1.01)	(0.60)
10	1986-1991	0.101	1.29	8	0.283	20.03	−0.013	0.024
		(5.77)	(.256)		(6.77)	(0.00)	(−0.13)	(0.88)

Note: Estimate breaks in variance of the stock returns, y_t, by the ICSS algorithm for 10 subsamples with sample size 1,705. After identifying break points, decompose the stock returns into break component (m_t) and break-free component (y_t/m_t). m_t is a sample standard deviation of each regime. The GPH and LM are applied to the squared returns and squared residuals.

For 10 sub-periods of S&P 500 daily return, Tables 6 - 8 provide the estimated d with t-statistics, number of breaks, and LM statistics with p-value. The 1st panel of Table 6 presents the estimated values of d by GPH in the absolute returns and LM statistics. $\hat{d}$'s are changing from 0.352 in period 1 to 0.154 in period 5 and upto 0.715 in period 8. All of the sub-periods have strong evidence of long memory in the absolute stock return. The estimated number of breaks in the level of absolute returns by Bai's method (in the 2nd panel) obviously has positive relation with $\hat{d}$. For example, period 8 has 17 breaks and has the highest value of $d = 0.715$ amongst all sub-periods. But they do not show an exact relation since d is affected by the magnitude of the break too. We assume the maximum number of breaks is 50 in the first step, and restrict not to estimate a break if the size of the sample is less than 50 observations[7]. The 4th panel shows the results of fitting structural breaks that correspond to the points of level shifts in the absolute returns. Although all of the sub-periods have strong evidence of long memory in the absolute returns, none of its residuals, $\{y_t^*\}$[8], has long memory.

However, Bai's procedure might be problematic when applied to this data since it is based on the assumption of no breaks in variance. It is a well known fact that absolute stock returns have an exponential distribution, which implies mean and variance are determined by one parameter. To avoid this problem, we use the ICSS method to identify breaks in variances of stock returns by using the model (19) - (21). Table 7 reports the number of sudden changes in variance as identified by the ICSS algorithm for stock returns. Periods 3 and 7 have 17 break points and period 9 has only 4 change points and so on. Figures 6-1 - 6-10 present plots of absolute stock returns for each sub-period. The solid line shows the sudden changes detected by plotting sample means of absolute returns, where the sample means calculated for the observation between the change points. The significant changes in variance are a little bit more than those in level of absolute returns. The 3rd panel of Table 7 shows the results of fitting breaks that correspond to the points of breaks in variance to the level of absolute stock returns. When breaks in variance are introduced, the evidence is somewhat mixed. In some sub-series, negative estimates of d in the residuals are obtained, so there is some possibility of overdifference as pointed out in section 5.

As an additional analysis, we also examined long memory in the squared stock returns in Table 8. As occasional breaks are incorporated directly into return series, the existence of long memory in volatility is mixed, too.

In the empirical analysis the evidence is somewhat mixed as to whether volatility has $I(d)$ process or structural breaks. However, structural changes in the absolute stock return or squared returns are evident and result in spurious "long memory" of these series in a certain degree. Either model by itself may not capture all of the persistence in the volatility, i.e., there may be residual $I(d)$ effects when a model is fitted that includes only structural breaks, and there may still be sudden changes in the volatility of residuals after fitting an $I(d)$ model. Therefore, a more complete analysis would allow for both kinds of effects.

---

[7]As we increase the maximum permitted number of breaks or decrease the minimum number of observations to detect a break within that sample, we get a little bit more breaks.

[8]In Table 6, we use (23) instead of (22) to get break 'free' series since the absolute stock returns show some co-break in mean and variance. Because we estimate breaks by Bai's method, (22) is suitable one. However, the application of (22) to the absolute returns shows similar long memory property as Table 6, since spurious long memory depends on changes in mean.

FIG. 6. Absolute Stock Return and Breaks of 10 Sub-Periods in 1/4/1928-8/30/1991

FIG. 6-1. Period 1 (1928-1934)

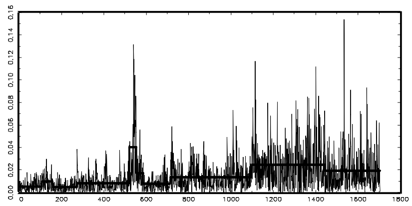


FIG. 6-2. Period 2 (1934-1940)

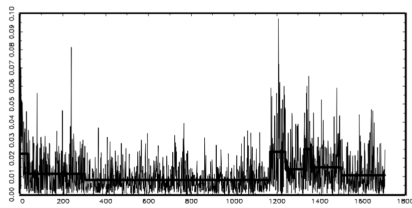


FIG. 6-3. Period 3 (1941-1947)

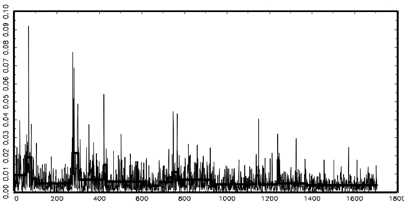


FIG. 6-4. Period 4 (1947-1953)

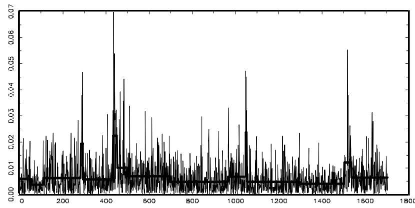


FIG. 6-5. Period 5 (1954-1960)

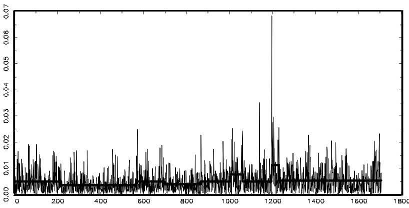


FIG. 6-6. Period 6 (1960-1966)

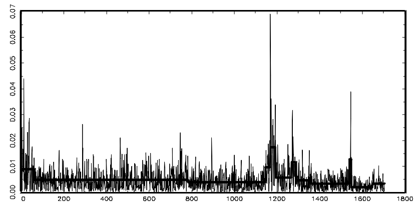


FIG. 6-7. Period 7 (1967-1973)

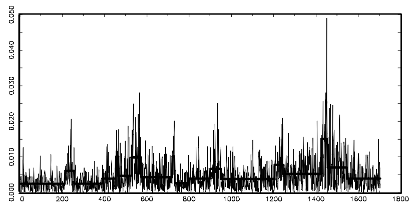


FIG. 6-8. Period 8 (1973-1979)

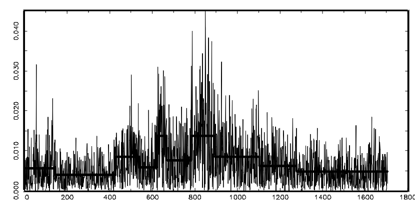


FIG. 6-9. Period 9 (1980-1986)

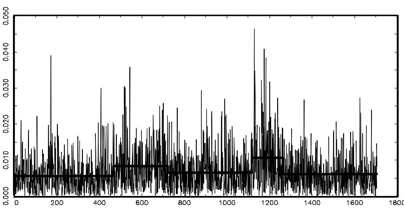
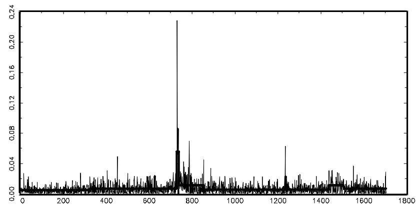


FIG. 6-10. Period 10 (1986-1990)



7. EXTENSION: COMMON OCCASIONAL BREAKS PROCESS IN MULTIVARIATE SERIES

Ray and Tsay (1997) find that common long-range dependent components occur as frequently in a randomly chosen group of S&P 500 companies as in companies of similar market capitalization, but that common long-range dependent components are more likely for companies in the same business sector. These results suggest that there might exist common breaks in stock markets since stock market volatility can be characterized by a linear model with occasional breaks instead of an $I(d)$ process. One could have alternative explanations of common long-range dependence by applying the common break concept which might capture any common long memory component of these multivariate series. Structural breaks may be related across variables, in an analogous way to common long-range dependence or cointegration.

The following is a model for multivariate series with a common break process in mean. This model could easily be altered for common breaks in variance. If there are N individual series which have one common break component,

$$y_{it} = m_{it} + \varepsilon_{it}, i = 1, \ldots, N, t = 1, \ldots, T. \tag{24}$$

$$m_{it} = m_{it-1} + q_t \eta_{it} = m_{i0} + \sum_{s=1}^{t} q_s \eta_{is} \tag{25}$$

The process, m_{it}, will capture coincident breaks among several series. Individual series might have different levels of mean within each regime. The break component 'free' series here is $y^*_{it} = y_{it} - m_{it}$, where m_{it} is the mean of each regime of individual i. This y^* is a series without any break components, where all of the series have the same structure of breaks, q_t, but allowing different size of shocks to each series by the amount, η_{it}.

8. CONCLUSIONS

It has been shown that a series with breaks can mimic some of the properties of $I(d)$ processes, particularly the autocorrelations, where d can be a fraction, its value depending on the number of breaks for a particular sample size. From simulation results it is also shown that absolute returns for the S&P 500 stock index are more likely to show the "long memory" property because of the presence of breaks in the series rather than being an $I(d)$ process. This has relevance for the forecastability of absolute returns, which are potentially useful for value at risk estimates, especially if the timing and the size of breaks can be shown to be forecastable. This is potentially possible if the breaks are endogenous and needs to be explored further.

REFERENCES

Aggarwal, R., C. Inclan, and R. Leal, 1999. Volatility in Emerging Stock Markets. *Journal of Financial and Quantitative Analysis*, Forthcoming.

Bai, J., 1997. Estimating Multiple Breaks One at a Time. *Econometric Theory* **13**, 315-352.

Bai, J., 1998. A Note on Spurious Breaks. *Econometric Theory* **14**, 663-669.

Bai, J. and P. Perron, 1998. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* **66**, 47-78.

Chen, C. and G. C. Tiao, 1990. Random Level-Shift Time Series Models, ARIMA Approximations, and Level-Shift Detection. *Journal of Business & Economic Statistics* **8**, 83-97.

Ding, Z., C. W. J. Granger, and R. F. Engle, 1993. A Long Memory Properties of Stock Returns and a New Model. *Journal of Empirical Finance* **1**, 83-106.

Engle, R. F. and A. D. Smith, 1999. Stochastic Permanent Breaks. forthcoming, *Review of Economics and Statistics.*

Geweke, J. and S. Porter-Hudak, 1983. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis* **4**, 15-39.

Granger, C. W. J. and Z. Ding, 1995. Some Properties of Absolute Returns. An Alternative Measure of Risk. *Annales d'economie et de statistique* **40**, 67-91.

Granger, C. W. J. and Z. Ding, 1995. Stylized Facts on the Temporal and Distributional Properties of Daily Data from Speculative Markets, UCSD Working Paper.

Granger, C. W. J. and Z. Ding, 1996. Varieties of Long Memory Models. *Journal of Econometrics* **73**, 61-78.

Granger, C. W. J. and F. Marmol, 1997. The Correlogram of a Long Memory Process Plus a Simple Noise, UCSD Working Paper.

Granger, C. W. J. and T. Teräsvirta, 1999. A Simple Nonlinear Time Series Model with Misleading Linear Properties. *Economics Letters* **62**, 161-165.

Kim, In-Moo, 1997. Detecting the Number of Structural Breaks. *Economics Letters* **57**, 145-148.

Inclan, C. and G. C. Tiao, 1994. Use of Cumulative Sum of Squares for Retrospective Detection of Change of Variance. *Journal of the American Statistical Association* **89**, 913-923.

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

Mikosch, T. and C. Starica, 1999. Change of Structure in Financial Time Series, Long Range Dependence and the GARCH Model, Working paper, The Wharton School, Chalmers University of Technology.

Nunes, L. C., C. M. Kuan, and P. Newbold, 1995. Spurious Break. *Econometric Theory* **11**, 736-749.

Ray, K. B. and R. S. Tsay, 1997. Identifying Common Long-Range Dependence in a Vector Time Series, Working paper, Statistics Research Center, Graduate School of Business, University of Chicago.

Rydén, T., T. Teräsvirta, and S. Åsbrink, 1998. Stylized Facts of Daily Returns Series and Hidden Markov Model. *Journal of Applied Econometrics* **13**, 217-244.

Yao, Y. C., 1988. Estimating the Number of Change-Points via Schwarz' Criterion. *Statistics & Probabilities Letters* **6**, 181-189.