3-1996

Semiparametric Exploration of Long Memory in Stock Prices

David K. C. LEE
Singapore Management University, davidlee@smu.edu.sg

Peter M. ROBINSON
London School of Economics

Published in Journal of Statistical Planning and Inference, March 1996, Volume 50, Issue 2, Pages 155-174.
http://doi.org/10.1016/0378-3758(95)00051-8

Semiparametric exploration of long memory in stock prices

David K.C. Lee,[a,1] Peter M. Robinson[b,*,2]

[a] *Transearch Training Centre, 2 Finlayson Green, #09-03/04, Asia Insurance Building, Singapore 0104, Singapore*

[b] *Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK*

Received 12 March 1993; revised 8 February 1994

Abstract

New or modified methods for semiparametric analysis of fractional long memory in time series are described and applied to twenty-six stock prices and two stock indices. Evidence is found that some, but not all, of the stocks have long memory, while one of the indices exhibits mean reversion.

AMS Subject Classifications: 62M10, 60G18

Keywords: Long memory; Semiparametric model

1. Introduction

This paper applies new and modified methodology for investigating the possible presence of fractional long memory in financial time series. Let x_t, $t = 1, 2, \ldots$, be a covariance stationary time series having a spectral density

$$f_x(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_x(j) \cos j\lambda, \quad -\pi < \lambda \leqslant \pi, \tag{1.1}$$

where $\gamma_x(j) = \mathrm{Cov}(x_t, x_{t+j})$ is the lag-j autocovariance of x_t. Assume that

$$0 < f_x(0) < \infty. \tag{1.2}$$

A series with property (1.2) will be termed an $I(0)$ series. The simplest example of an $I(0)$ series is a series of uncorrelated, homoscedastic, random variables. Now consider

* Corresponding author.

[1] Research supported by grant RP900047 from National University of Singapore when the author was a faculty member there.

[2] Research supported by ESRC grant R000233609.

a series y_t, $t = 1, 2, \ldots,$ such that

$$(1 - L)^d y_t = x_t, \quad t = 1, 2, \ldots, \tag{1.3}$$

where L is the lag operator, so $Ly_t = y_{t-1}$, and d is a real number. A series y_t given by (1.2) and (1.3) will be termed an $I(d)$ series. We say that y_t exhibits long memory if $d > 0$. By far the most familiar such situation is when $d = 1$ is assumed, when y_t has unit root, and other integer values of d are sometimes assumed (see e.g. Box and Jenkins, 1970). We say that y_t exhibits fractional long memory when $d > 0$ but d is non-integer. In case $0 < d < \frac{1}{2}$ then y_t inherits the covariance stationarity of x_t. When $d \geqslant \frac{1}{2}$, y_t is non-stationary. When $d = 0$, that is y_t is an $I(0)$ process, y_t is said to have no long memory.

We employ some methods of point estimation and statistical inference on d which are semiparametric in that the parametric relation (1.3), involving the unknown parameter d and the known functional form $(1 - L)^d$, is complemented by a non-parametric $f_x(\lambda)$, satisfying (1.2) but no other assumptions apart perhaps, from mild regularity assumptions in a neighbourhood of zero frequency. In particular, $f_x(\lambda)$ might be infinite or zero at non-zero frequencies in $(-\pi, \pi]$. When $d < \frac{1}{2}$, (1.2) and (1.3) imply that y_t has spectral density $f_y(\lambda)$ (defined analogously to $f_x(\lambda)$ in (1.1)) satisfying

$$f_y(\lambda) \sim C\lambda^{-2d}, \quad \text{as } \lambda \to 0^+, \tag{1.4}$$

for $0 < C < \infty$. Thus $f_y(\lambda)$ tends to infinity as $\lambda \to 0^+$ if $0 < d < \frac{1}{2}$, but its behaviour away from zero frequency is unrestricted. For $d \geqslant \frac{1}{2}$, $f_y(\lambda)$ is not well-defined, but a suitable amount of integer differencing produces a covariance stationary time series in $I(d)$ for $d < \frac{1}{2}$. Notice that if y_t is in $I(d)$ for $\frac{1}{2} \leqslant d < 1$, the first differenced series $z_t = (1 - L)y_t$ is in $I(d)$ for $-\frac{1}{2} \leqslant d < 0$, so that z_t has spectral density $f_z(\lambda)$ satisfying

$$f_z(\lambda) \sim C\lambda^{-2d}, \quad \text{as } \lambda \to 0^+, \quad -\tfrac{1}{2} \leqslant d < 0, \tag{1.5}$$

and so $f_z(0) = 0$. Furthermore, overdifferencing of a stationary or non-stationary series also leads to a process with zero spectrum at zero frequency. Notice that our concept of long memory pertains only to zero frequency. The concept can be extended to apply to other frequencies but we will not discuss this here. The literature on the analysis of long memory time series has been reviewed by Robinson (1994a).

The following section reviews experience of modelling economic and financial time series, in relation to non-fractional and fractional $I(d)$ series. Section 3 describes methods of estimation and inference on d. Section 4 applies the methodology to financial series of stocks and stock indices. Section 5 contains some brief concluding remarks.

2. Long memory in financial and economic time series

The issue of whether macroeconomic variables and financial variables exhibit long memory is of great interest because it has important implications for economic planning and financial planning. The focus of macroeconomists is on time series that are related to the level of economic activity such as GNP, industrial production output and employment. For financial analysts, the focus is on stock prices and indices.

A general practice in macroeconomics is to model the series of interest as two separate components, namely a secular or growth component and a cyclical component. The secular component, assumed to be non-stationary, is associated with growth factors such as capital accumulation, population growth and technology improvement. The cyclical component, on the other hand, is associated with fundamental factors which are the primary cause of movement in these series, and is assumed to be covariance stationary, indeed typically $I(0)$, in that it contains no long memory.

Since the publication of the influential paper by Nelson and Plosser (1982), attention has focussed principally on modelling the non-stationary components of macro time series. It is common practice to attempt to take care of the non-stationary component either by regressing the series on time, or some function of time, or by first or second differencing of the series. In either case the residuals, if they appear stationary, are interpreted as approximating the cyclical component. There is a fundamental difference between these two approaches for approximating the secular or stationary component. The first approach of regressing on a function of time assumes a model of form

$$y_t = g(t;\theta) + x_t$$

for the observed time series, where $g(t;\theta)$ is a given function of t and an unknown parameter vector θ, while x_t is an $I(0)$ series. This model is consistent with a deterministic trend, and implies that the only information about the future is embodied in its mean. No past or future events will alter the long horizon expectation, and uncertainty is bounded, in the sense that the forecast error variance is finite. The second approach of differencing, on the other hand, assumes a model of form (1.3) with $d = 1$ or some other integer, where again x_t is an $I(0)$ series. This model is consistent with a stochastic trend, and implies that the long range forecast of the series will always depends on past events and that the variance of the forecast error will be unbounded.

To macroeconomists, any shock to the economic system will have a permanent effect on output if the series is thought to be generated by a process consistent with the second, differencing, approach. From a theoretical view point, this has important implications for modelling the business cycle. But, so far as practical economic planning is concerned, it implies that policy action is required to bring the variable back to its original long term projection if that is desired. In many developing countries, this is desired if one has targeted a certain level of GDP growth to be achieved by a certain time frame, for example, in 10 years. This is because a decline in output in the past will lower the forecasts of future output permanently. On the other hand, if the series is generated by a process which is consistent with the first, regression, approach, then there is not such a strong basis for policy action because the long term expectation of the series has not changed and the series will return to its trend, or potential output, sometime in the future.

However, the observed time series may not belong to one of the two classes described above, and in particular may satisfy (1.2) with non-integer d, so that y_t has fractional long memory. Indeed, Cochrane (1988) has provided evidence, using a variance-ratio test, that US GNP is not an $I(l)$ series but may have fractional long memory. One interesting interpretation of this observation given by Campbell and Mankiw (1987) is that the shock will appear to be persistent, but it will eventually move towards the original trend, but 'it does not get all the way there'. Other related theoretical and empirical papers include Haubrich and Lo (1989). Porter-Hudak (1990) and Sowell (1992a, b).

In the empirical analysis of financial time series, one of the most hotly debated topics is whether the holding period returns on a risky asset are serially independent. While earlier empirical evidence has supported the efficient market hypothesis by finding no evidence against it, recent studies by various authors have produced evidence that contradict earlier findings. Lo and MacKinlay (1988), Poterba and Summers (1988), using a variance-ratio test, have demonstrated that the lagged k variance ratio, which is defined to be k times the variance of the k-period return divided by the variance of the one-period return, is not unity as implied by the random walk hypothesis. In fact, stock returns exhibit mean reversion behaviour, that is there is a tendency for the returns to move away from the trend, then reverse direction, overshooting the trend before returning to it again. This is characterised by variance ratios below unity for lags longer than a year and above unity for shorter time periods. It is also consistent with the observations that there is negative autocorrelation for intervals longer than a year and positive autocorrelation for shorter periods. Using a generalized form of rescaled range (R/S) statistic, Lo (1991) has found no evidence against the random walk hypothesis for the stock indices, contradicting his earlier finding using variance-ratio tests. For financial analysis, the implication for investment planning and strategy is relatively straightforward in the case of transitory deviations from equilibrium which are strong and persistent as suggested by the mean reversion hypothesis. If one knows where the expected mean is for a stationary and long memory series, one can devise rules to buy and sell at the 'right time'.

Macro time series have been analyzed for long memory (e.g. Diebold and Rudebusch, 1989), but are apt to be short, almost certainly too short to justify the application of large sample inference rules based only on a semiparametric model for the data, while no finite sample theory yet exists even for rules of parametric inference on long memory. However, many financial time series can be very lengthy, seemingly sufficiently so to warrant semiparametric exploration of long memory. We examine individual stock returns and stock indices from the Singapore Stock Market. The methods we use have only recently been proposed, and their use on financial series is novel. One of the methods is Robinson's (1995) modified (trimmed and efficiency-improved) version of the log-periodogram estimate of d proposed by

Geweke and Porter-Hudak (1983), and an alternative estimate based on the (un-logged) periodogram due to Robinson (1994a, b).

The study of long range dependence in macroeconomic and financial variables has been on-going for at least fifteen years. A number of results, mainly based on the R/S statistic and variance ratio tests, have suggested the presence of long memory in financial series. These studies include Booth et al. (1982b) and Diebold et al. (1991) on the foreign exchange market; Booth et al. (1982a) on the gold market; Helms et al. (1984) on commodity futures; Greene and Fielitz (1977), Poterba and Summers (1988), Lo and MacKinlay (1988) and Fong (1992) on stock prices; Shea (1991) on interest rates. However, recently, Lo (1991) has produced results that contradict some of these findings, citing the low power of previous tests in the presence of short term dependence as the main reason for their rejection of the random walk hypothesis. The empirical results are now mixed. However, it appears that Lo and MacKinlay's (1988) results have suggested that individual stock returns exhibit long memory behaviour, but not the indices. One reason could be that individual stocks may not be traded as often. Lo (1988), however, has argued that one can form a portfolio of stocks to overcome the problem of the thin trading.

3. Semiparametric inference on long memory

We assume that (1.2) and (1.3) obtain for the observable, covariance stationary, time series y_t, $t = 1, 2, \ldots,$ so that the spectral density $f_y(\lambda)$ of y_t satisfies (1.4) with $d < \frac{1}{2}$. The series y_t is observed at $t = 1, 2, \ldots, n$. Given a correct, fully parametric model for $f_y(\lambda)$ at all frequencies, $\sqrt{n}$-consistent estimates of d and the other parameters are available; in particular, Gaussian maximum likelihood estimates are not only $\sqrt{n}$-consistent but asymptotically efficient in case y_t is Gaussian, see Fox and Taqqu (1986), Dahlhaus (1989). However, if the parametric model is misspecified, such estimates of d will be inconsistent. Sowell (1992b), in his application, considered a sequence of parameterizations to guard against misspecification, choosing a model based on procedures such as AIC, although the theoretical behaviour of the determination procedures in fractional models for $d > 0$ has yet to be studied. We exploit the large value of n available to employ semiparametric procedures which are not $\sqrt{n}$-consistent but can be justified as consistent in the absence of parametric model assumptions. Our procedures also have some computational advantage over Gaussian maximum likelihood estimates in that they are given in closed form.

Our estimates are functions of the periodogram

$$I(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} y_t e^{it\lambda} \right|^2.$$

We estimate $I(\lambda)$ only at frequencies for $\lambda = \lambda_k = 2\pi k/n$, for $k = 1, \ldots, m$, where $m < n$. For such λ, $I(\lambda)$ is invariant to location, so that mean-correction of y_t is unnecessary. The integer m is a user-chosen 'bandwidth number'. In asymptotic theory for our estimates, m is regarded as increasing as n tends to infinity, but at a slower rate.

To define our first estimate of d, introduce a user-chosen 'trimming number' τ, which is a small non-negative integer in practice but in the theory is regarded as tending to infinity slower than m, and a user-chosen 'pooling number' J, which is also a small integer but it says fixed in the asymptotic theory. Now introduce

$$W_k^{(J)} = \log\left\{\sum_{j=1}^{J} I(\lambda_{k,j})\right\}, \quad k = \tau + J, \tau + 2J, \ldots, m,$$

where we have implicitly assumed that $(m - \tau)/J$ is an integer, though end effects when this is untrue are asymptotically negligible. Now write

$$W_k^{(J)} = c^{(J)} - d(2 \log \lambda_k) + U_k^{(J)}, \quad k = \tau + J, \tau + 2J, \ldots, m, \tag{3.1}$$

where $c^{(J)} = \log C + \psi(J)$ and ψ is the digamma function, C being the scale factor in (1.4). Define $\hat{d}_1$ to be the least squares estimate of d based on the 'regression model' (3.1), where the $U_k^{(J)}$ are regarded as unobservable disturbances.

On choosing $\tau = 0$ and $J = 1$, and replacing $2 \log \lambda_k$ by $\log(4 \sin^2 \lambda_k/2)$, $\hat{d}_1$ becomes the estimate of Geweke and Porter-Hudak (1983). These authors attempted a proof of asymptotic properties of this estimate only under the assumption $-\frac{1}{2} < d < 0$ (so that there is no long memory) but their proof was incomplete, as shown by Robinson (1995). Robinson (1995) provided a hopefully correct proof which allows $-\frac{1}{2} < d < \frac{1}{2}$ but depends on the trimming number τ (possibly relating to the large bias in Geweke and Porter-Hudak's estimate found by Agiaklogou et al. (1993)), and also on Gaussianity of y_t. The pooling innovation was also introduced by Robinson (1995): asymptotic efficiency of $\hat{d}_1$ increases monotonically in J. The central limit theorem established by Robinson (1995) implies that

$$\frac{\hat{d}_1 - d}{SE(\hat{d}_1)} \to_d N(0, 1), \tag{3.2}$$

where $SE(\hat{d}_1)$ is the usual least squares standard error from the regression (3.1). Notice that there is no asymptotic difference in using $\log(4 \sin^2 \lambda_k/2)$ in place of $2 \log \lambda_k$ in (3.1).

To define a second estimate, introduce the 'averaged periodogram',

$$\hat{F}(\lambda) = \frac{2\pi}{n} \sum_{k=1}^{[n\lambda/2\pi]} I(\lambda_k),$$

where [.] means that the integer part is taken. Robinson (1994a) proposed estimating d by

$$\hat{d}_2 = \tfrac{1}{2}[1 - \log\{\hat{F}(q\lambda_m)/\hat{F}(\lambda_m)\}/\log q],$$

where q is a fixed constant chosen by the user to be within the interval (0, 1). Robinson (1994a) showed that $\hat{d}_2$ is consistent for d under conditions that are far weaker than Gaussianity, requiring little more than (1.4) and the requirement that $m \to \infty$ slower than n. Indeed consistency obtains also if (1.4) replaces C by a slowly varying function of λ that is possibly of unknown from, so that a more general class than the $I(d)$ class is covered. Asymptotic distribution theory for $\hat{d}_2$ is discussed by Lobato and Robinson (1996).

4. Empirical results

We employ daily data from 2 January 1975 (750102) to 31 July 1991 (910731) obtained from the Financial Database, National University of Singapore. Two stock indices are used for analysis together with twenty-six blue-chip stocks. The two indices are the Singapore Stock Exchange of Singapore (SES) All-Share Price Index and the Straits Times Industrial Index (STI). The twenty-six stocks are the component stocks of the STI index and details are listed in the Appendix.

The SES All-Share Price Index is based on 100% coverage of the prices of all the listed companies. A conventional value-weighted method is used to calculate the SES index, that is, the index number is expressed as a percentage of the current aggregate market value to the base aggregate market value. The weight reflects the importance of the share as reflected by their paid-up capital. The index is adjusted for rights issue, new listing and delisting.

The Straits Times Industrial Index has a much longer history than the SES Index. The STI was first introduced in 1948 while the SES Index was officially launched in 1984. The STI is unweighted and is adjusted for new listing and delisting.

Denote by p_t the stock price or index level at time t. We shall apply the methods of the previous section to the returns $y_t = \log(p_t/p_{t-1})$.

Fig. 1(a) and (b) are plots of the periodogram and a weighted autocovariance estimate of the spectral density using a Bartlett window for the returns for the y_t computed from the SES Index from 750102 to 870930. In Fig. 1(b), we can see that as frequency approaches zero, the spectral density estimate rapidly increases. The shape displayed in Fig. 1(b) is the 'typical spectral shape' of many observed economic time series (Granger, 1966). Taking the first difference of the same return series, we observe in Fig. 1(c) that the spectral density estimate is zero at zero frequency and then tends to increase with λ. These results are consistent with a model for the returns of form (1.2) and (1.3) with $0 < d < \frac{1}{2}$. Corresponding plots for the STI Index are given in Fig. 2(a)–(c).

There are empirical results which suggest that the detection of long memory in stock prices can be strongly influenced by a period of uncertainty. Kim et al. (1991), using variance-ratio tests, have found no evidence of long memory in stock prices after World War II. Using data from 1871 to 1987, McQueen (1992) has also found that there is no evidence of long memory in stock indices, citing the large variance of stock prices during periods of uncertainties as the reason; two particular events. World War II and the Great Depression, are the source of the heteroscedasticity. Using generalized least squares randomization tests, the random walk hypothesis is not rejected. Thus in addition to estimating d from the full sample, we use also a subsample from 2 January 1975 through 31 September 1987 in order to avoid the October 1987 Crash.

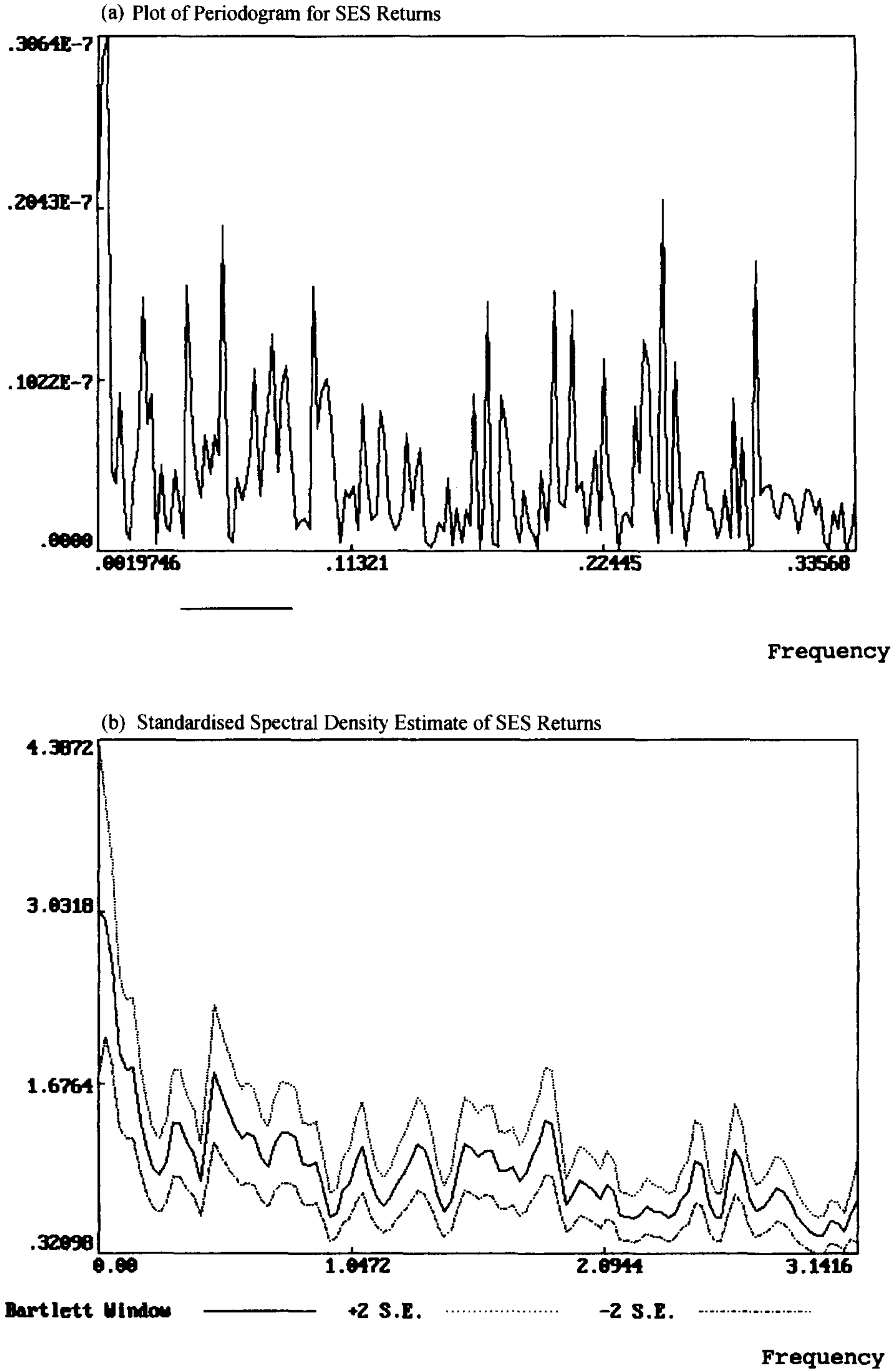
(a) Plot of Periodogram for SES Returns
.3064E-7
.2043E-7
.1022E-7
.0000
.0019746
.11321
.22445
.33568
Frequency
(b) Standardised Spectral Density Estimate of SES Returns
4.3872
3.0318
1.6764
.32098
0.00
1.0472
2.0944
3.1416
Bartlett Window
+2 S.E.
-2 S.E.
Frequency

Fig. 1.

(c) Standardised Spectral Density Estimate of Change of SES Returns

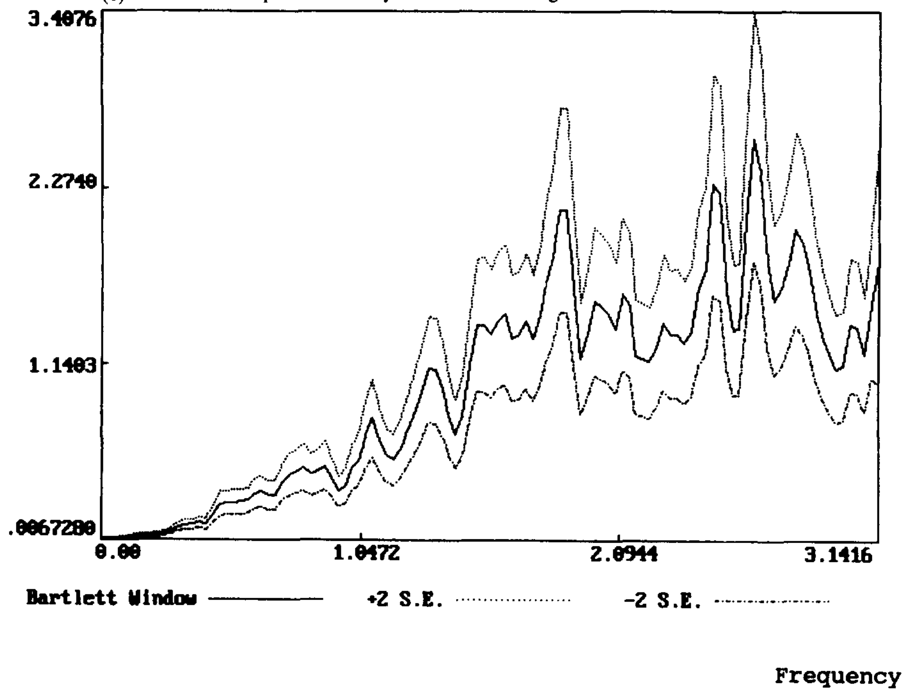

Fig. 1. Continued.

Results for the individual stocks are presented in Tables 1–3. In Table 1 we report $\hat{d}_1$ as well as the least squares estimate of $c^{(J)}$ in (3.1). Throughout we take $J = 4$ and $\tau = 4$. In choosing m we make use of the results of Robinson (1994b). He showed that for $0 < d < \frac{1}{4}$, a bandwidth that asymptotically minimizes a mean squared error criterion is

$$m = \left\{ \frac{(1 - 2d + \alpha)^2}{E_\alpha^2 (2\pi)^{2\alpha} (1 - 4d)} \right\}^{1/2\alpha + 1} n^{2\alpha/2\alpha + 1}, \tag{4.1}$$

where α and E_α are given by the relation

$$\frac{f(\lambda)}{C\lambda^{-2d}} = 1 + E_\alpha \lambda^\alpha + o(\lambda^\alpha), \quad \text{as } \lambda \to 0^+,$$

for the largest possible α in (0, 2]. Thus α is a sort of smoothness parameter and we fix it at its maximal value of 2. We choose E_α arbitrarily to be 2. We also employ the 'rule of thumb' bandwidth $m = n^{1/2}$, which typically turns out to be three or four times smaller than (4.1).

For most of the stocks, $\hat{d}_1$ falls in the interval (0.01, 0.3), as is consistent with most of the point estimates given in previous studies where long memory has been detected.

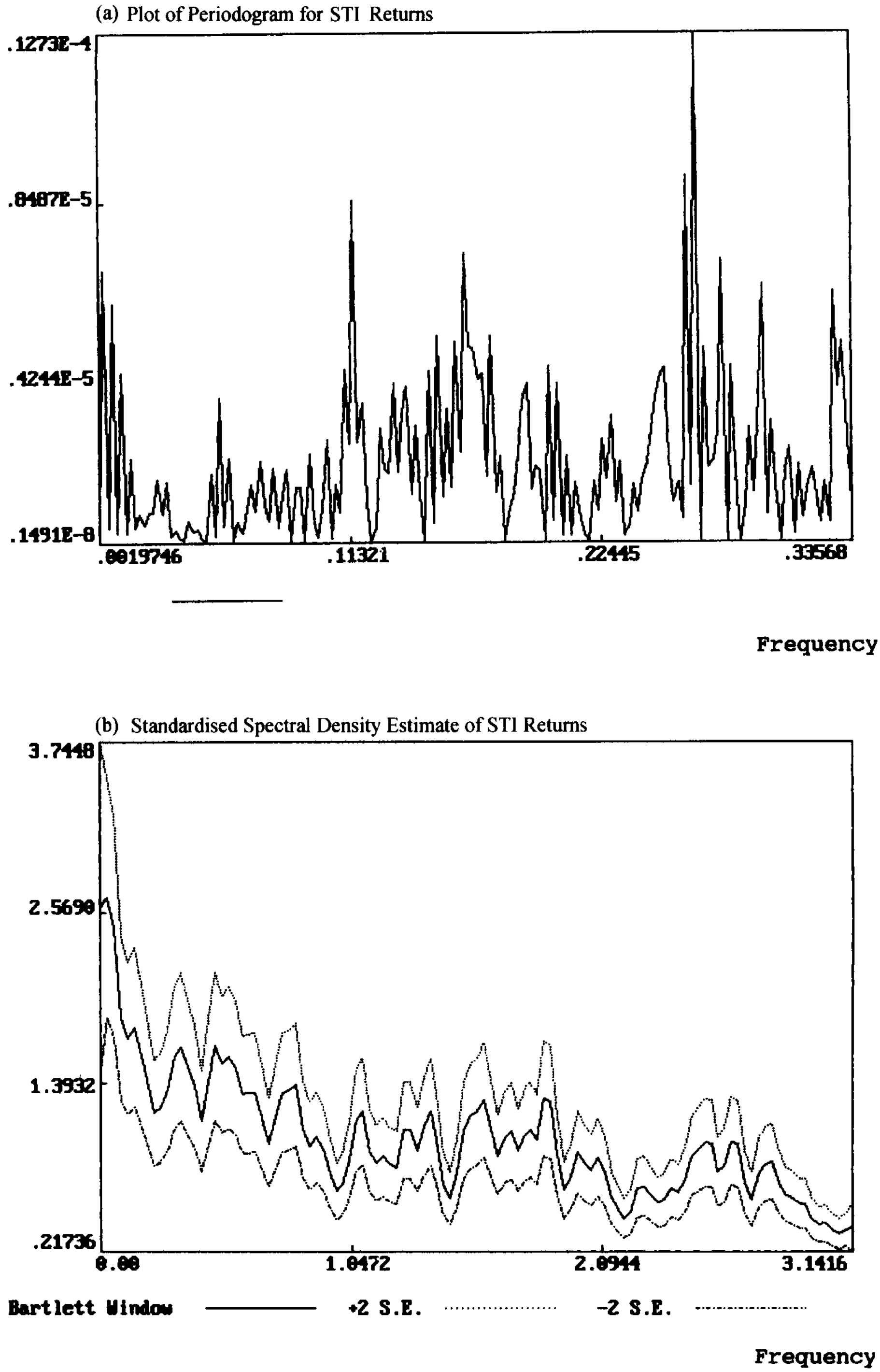
(a) Plot of Periodogram for STI Returns
.1273E-4
.8487E-5
.4244E-5
.1491E-8
.0019746
.11321
.22445
.33568
Frequency
(b) Standardised Spectral Density Estimate of STI Returns
3.7448
2.5690
1.3932
.21736
0.00
1.0472
2.0944
3.1416
Bartlett Window
+2 S.E.
-2 S.E.
Frequency

Fig. 2.

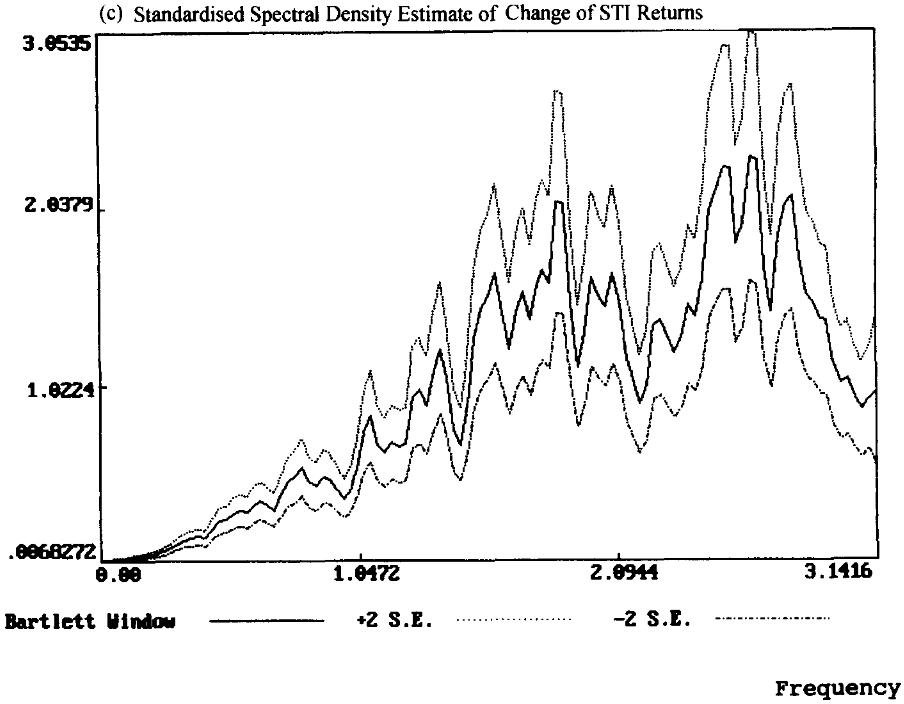

Fig. 2. Continued.

Table 1
$\hat{d}_1$ Estimates

No.	Full sample				Sub sample			
	m	$(m-\tau)/J$	$\hat{d}_1$	Test $(\hat{d}_1)$	m	$(m-\tau)/J$	$\hat{d}_1$	Test $(\hat{d}_1)$
1	32	7	0.148	1.37				
	68	16	0.159	4.48				
2	64	15	0.215	4.22	56	13	0.241	5.95
	216	53	0.080	3.26	176	43	0.056	1.26
3	44	10	0.102	1.49	32	7	0.057	0.19
	120	29	− 0.015	0.11	72	17	− 0.062	0.79
4	24	5	0.262	2.97				
	52	12	− 0.091	1.11				
5	64	15	0.095	0.40	56	13	0.055	0.12
	216	53	0.000	0.00	176	43	− 0.032	0.38
6	64	15	− 0.024	0.04	56	13	0.007	0.00
	216	53	0.010	0.05	176	43	0.013	0.10
7	64	15	0.116	0.98	56	13	0.156	1.22
	216	53	0.085	3.41	176	43	0.224	13.99

Table 1 Continued.

No.	Full sample				Sub sample			
	m	$(m-\tau)/J$	$\hat{d}_1$	Test $(\hat{d}_1)$	m	$(m-\tau)/J$	$\hat{d}_1$	Test $(\hat{d}_1)$
8	64	15	−0.009	0.00	56	13	−0.069	0.65
	216	53	0.124	6.46	176	43	0.066	1.60
9	52	12	0.431	21.05	40	9	0.319	7.69
	152	37	0.107	3.85	108	26	0.171	4.70
10	40	9	0.309	14.50	24	5	0.379	5.09
	100	24	0.183	7.44	52	12	−0.051	0.23
11	64	15	0.129	2.25	56	13	0.250	11.80
	216	53	0.067	2.16	176	43	0.096	4.11
12	48	11	0.426	13.54	40	9	0.069	0.09
	144	35	0.137	4.93	100	24	0.017	0.04
13	64	15	0.029	0.09	56	13	0.062	0.21
	216	53	0.022	0.29	176	43	0.001	0.00
14	32	7	0.078	0.13				
	76	18	0.045	0.30				
15	64	15	0.184	5.43	56	13	0.216	2.19
	216	53	0.071	3.21	184	45	0.071	2.11
16	36	8	0.187	2.20				
	88	21	0.048	0.57				
17	52	12	0.104	3.08	44	10	0.016	0.27
	160	39	0.030	0.39	116	28	0.076	2.11
18	40	9	0.232	3.06	24	5	0.112	0.23
	100	24	0.043	0.43	52	12	−0.084	0.60
19	56	13	0.270	8.71	48	11	0.141	1.72
	180	44	0.047	0.82	136	33	−0.041	0.37
20	64	15	0.043	0.15	56	13	−0.109	0.45
	216	53	−0.138	12.37	176	43	−0.210	17.71
21	24	5	−0.075	0.41				
	44	10	−0.060	0.40				
22	64	15	0.012	0.01	56	13	0.111	0.83
	216	53	0.116	4.92	176	43	0.203	11.00
23	64	15	−0.070	1.00	56	13	−0.101	1.71
	216	53	−0.039	0.94	176	43	−0.113	5.70
24	64	15	−0.108	1.02	56	13	−0.271	5.26
	216	53	−0.505	168.5	176	43	−0.619	218.9
25	24	5	0.264	1.35				
	44	10	0.121	1.12				
26	64	15	−0.040	0.28	56	13	−0.022	0.11
	216	53	0.020	0.25	176	43	0.056	1.41

Note: Refer to the appendix for the corresponding stock and period. The starting period is 750 102 or the day the stock is listed whichever is later. The end period for the full sample for each stock is 910 731, and for the subsample is 870 930.

Table 2
d_2 Estimates

No.	Full sample m	Full sample $\hat{d}_2$	Sub sample m	Sub sample $\hat{d}_2$	No.	Full sample m	Full sample $\hat{d}_2$	Sub sample m	Sub sample $\hat{d}_2$
1	32	0.001			14	34	0.006		
	70	0.105				76	0.050		
2	65	0.082	57	0.119	15	65	0.124	59	0.149
	216	0.029	176	0.022		216	0.002	186	0.029
3	45	0.060	33	0.031	16	37	0.001		
	120	− 0.049	72	− 0.127		90	− 0.059		
4	27	− 0.048			17	54	0.000	44	− 0.034
	52	− 0.088				162	− 0.039	118	0.051
5	65	0.071	57	0.081	18	41	0.072	26	0.156
	216	− 0.046	176	− 0.012		102	− 0.039	52	− 0.105
6	65	− 0.050	57	0.005	19	58	0.114	49	0.137
	216	− 0.038	176	0.012		180	− 0.041	136	− 0.093
7	65	0.051	57	0.068	20	65	0.055	57	− 0.010
	216	0.051	176	0.210		216	− 0.196	176	− 0.203
8	65	0.071	57	− 0.126	21	24	0.171		
	216	0.054	176	0.056		44	0.026		
9	52	0.243	42	0.186	22	65	0.004	57	− 0.003
	152	0.032	108	0.094		216	0.061	176	0.193
10	41	0.080	26	0.212	23	65	− 0.129	57	− 0.133
	102	0.038	52	− 0.126		216	− 0.075	176	− 0.138
11	65	0.106	57	0.098	24	65	− 0.087	57	− 0.367
	216	− 0.013	176	0.032		216	− 0.692	176	− 0.784
12	51	0.176	40	0.061	25	24	0.134		
	146	0.068	100	0.059		44	0.115		
13	65	− 0.068	57	0.061	26	65	− 0.165	57	− 0.105
	216	− 0.019	176	0.004		216	− 0.023	176	0.025

Note: Refer to the appendix for the corresponding stock and period. The starting period is 750 102 or the day the stock is listed, whichever is later. The end period for the full sample for each stock is 910 731, and for the sub-sample is 870 930.

There is often considerable sensitivity to the choice of m. We also report the test statistics $\chi^2 = \{\hat{d}_1/SE(\hat{d}_1)\}^2$; in view of (3.2), the null hypothesis of $d = 0$, or no longer memory in the first differences of the logged stocks, is rejected when χ^2 is significantly large relative to the χ_1^2 distribution. In more than half the stocks, we cannot reject the hypothesis that the returns are $I(0)$, that is that the logged stocks are $I(1)$. However, some returns exhibit stationary long memory, while (undifferenced) stock 24 may have the just non-stationary $d = \frac{1}{2}$, thus exhibiting the so-called $1/f$ noise long memory behaviour. Our results also suggest that the 1987 October crash did not make much difference to our conclusion regarding long memory behaviour. The $\hat{d}_2$ estimates are reported in Table 2(a) and (b), and do not differ qualitatively from the $\hat{d}_1$ estimates. The estimate for stock 24 is again negative and close to $-\frac{1}{2}$.

Table 3
Lo's rescaled range test

	No.	Fullsample	Subsample	No.	Fullsample	Subsample	No.	Fullsample	Subsample
Q(90)	1	0.98		10	5.23	6.34	19	1.19	1.43
Q(180)		1.42			7.36	8.35		1.08	1.27
Q(270)		1.72			8.58	10.07		1.05	1.19
Q(360)		1.66			9.10	10.87		1.06	1.17
Q		1.01			0.75	0.75		1.11	1.25
Q(90)	2	1.17	1.32	11	1.44	1.83	20	3.89	4.07
Q(180)		1.19	1.26		1.41	1.69		4.66	4.80
Q(270)		1.25	1.28		1.45	1.69		5.20	5.27
Q(360)		1.39	1.40		1.52	1.74		5.10	5.06
Q		1.54	1.79		2.02	2.75		0.78	0.76
Q(90)	3	0.95	1.75	12	0.99	1.46	21	0.95	
Q(180)		1.05	1.67		1.05	1.34		1.31	
Q(270)		1.13	1.54		1.09	1.30		1.86	
Q(360)		1.17	1.47		1.16	1.28		1.93	
Q		1.01	1.73		1.34	1.93		1.21	
Q(90)	4	1.20		13	1.01	1.44	22	1.35	1.49
Q(180)		1.50			1.07	1.34		1.55	1.61
Q(270		1.96			1.09	1.30		1.62	1.61
Q(360)		2.39			1.07	1.26		1.56	1.53
Q		1.07			0.99	1.35		1.56	1.79
Q(90)	5	0.72	1.16	14	1.14		23	1.30	1.52
Q(180)		0.85	1.28		1.15			1.35	1.56
Q(270)		0.94	1.38		1.19			1.37	1.54
Q(360)		0.91	1.31		1.22			1.40	1.56
Q		0.79	1.24		1.48			0.73	0.74
Q(90)	6	1.11	1.21	15	1.47	1.47	24	2.96	3.24
Q(180)		1.17	1.21		1.42	1.39		3.83	4.32
Q(270)		1.23	1.20		1.37	1.32		4.44	4.96
Q(360)		1.23	1.17		1.36	1.31		4.96	5.44
Q		1.05	1.09		1.99	2.04		0.72	0.73
Q(90)	7	1.47		16	1.33		25	1.18	
Q(180)		1.52			1.36			1.65	
Q(270)		1.52			1.36			1.71	
Q(360)		1.52			1.39			1.80	
Q		2.07			1.42			1.59	
Q(90)	8	1.13	1.74	17	1.12	1.43	26	1.09	1.09
Q(180)		1.24	1.75		1.14	1.31		1.12	1.12
Q(270)		1.35	1.72		1.21	1.29		1.14	1.14
Q(360)		1.35	1.64		1.32	1.35		1.17	1.17
Q		1.39	2.01		1.10	1.68		1.01	1.01
Q(90)	9	1.54	1.70	18	1.12	1.20			
Q(180)		1.49	1.56		1.41	1.42			
Q(270)		1.49	1.52		1.55	1.63			
Q(360)		1.51	1.50		1.75	1.72			
Q		2.15	2.50		1.04	1.17			

Note: We can reject the null hypothesis with 95% level of confidence if the test statistic falls outside the range [0.809, 1.862].

Table 4
Analysis for SES and ST indices

Stock Exchange of Singapore (SES) index

Full sample				Subsample			
m	$(m-\tau)/J$	$\hat{d}_1$	Test($\hat{d}_1$)	m	$(m-\tau)/J$	$\hat{d}_1$	Test($\hat{d}_1$)
64	15	0.122	1.99	56	13	0.195	2.25
208	51	− 0.121	5.94	168	41	0.230	19.05
m		$\hat{d}_2$		m		$\hat{d}_2$	
64		0.022		56		0.081	
210		− 0.268		170		0.202	

Lo's test	Full sample	Subsample
Q(90)	1.21	1.62
Q(180)	1.31	1.60
Q(270)	1.35	1.54
Q(360)	1.39	1.53
Q	0.90	2.78

Straits Times Industrial (STI) index

Full sample				Subsample			
m	$(m-\tau)/J$	$\hat{d}_1$	Test($\hat{d}_1$)	m	$(m-\tau)/J$	$\hat{d}_1$	Test($\hat{d}_1$)
64	15	0.004	0.00	56	13	− 0.338	1.73
208	51	− 0.118	1.57	168	41	− 0.129	1.42
m		$\hat{d}_2$		m		$\hat{d}_2$	
64		− 0.232		56		− 0.183	
210		0.008		170		0.018	

Lo's test	Full sample	Subsample
Q(90)	0.66	0.95
Q(180)	0.90	1.26
Q(270)	1.31	1.58
Q(360)	1.21	1.48
Q	0.39	0.57

In Table 3 we report Lo (1991)'s augmented rescaled range test for comparison. We present the test statistics for 3 months (Q(90)), 6 months (Q(120)), 9 months (Q(180)), 1 year (Q(360)) and the original rescaled range statistics (Q). Although there is evidence that the test is not very powerful (see the Monte Carlo results of Lo (1991)), the results still suggest that there are a few stocks that rejected the null hypothesis. Large test statistics are reported for stocks 10, 20 and 24 suggesting that these stocks are not generated by a random walk process.

The results for the stock indices are reported in Table 4. Those for SES are interesting. For the full samples period, $\hat{d}_1$ is significantly negative at the 5% level.

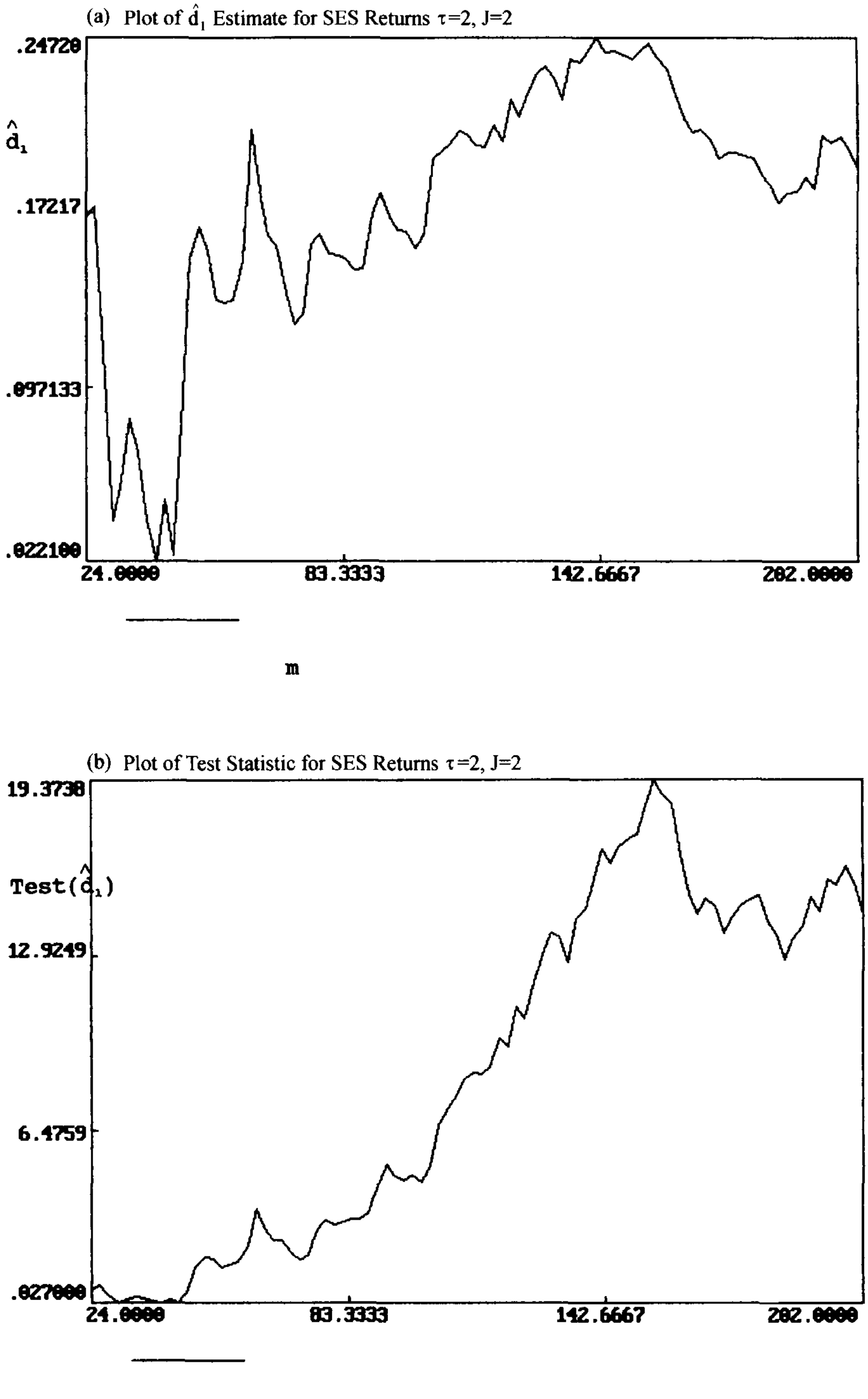
(a) Plot of $\hat{d}_1$ Estimate for SES Returns τ=2, J=2
.24720
$\hat{d}_1$
.17217
.097133
.022100
24.0000
83.3333
142.6667
202.0000
m
(b) Plot of Test Statistic for SES Returns τ=2, J=2
19.3738
Test($\hat{d}_1$)
12.9249
6.4759
.027000
24.0000
83.3333
142.6667
202.0000
m

Fig. 3.

However, for the reduced sample period, the estimate is significantly positive. This suggests that taking the 1987 October Crash into account has made the price series less non-stationary. In other words, the results suggest that prices revert towards the expected mean but 'not all the way'. The $\hat{d}_2$ estimates exhibit the same pattern while Lo's test is inconclusive. As for the STI index, there is no evidence that the series is not $I(0)$ from the χ^2 tests. This is further confirmed by the insignificant test statistics for Lo's test.

To obtain further evidence of the sensitivity to bandwidth, the plots of the estimates $\hat{d}_1$ ($\tau = 2, J = 2$), $\hat{d}_1$ ($\tau = 4, J = 4$) and $\hat{d}_2$ against m for the SES index are given in Figs. 3–5, respectively. The χ^2 test statistics are plotted in Figs. 3(b) and 4(b) to gauge the sensitivity of our conclusions to bandwidth choice.

5. Conclusion

Deseasonalised time series can be viewed as consisting of two components, a long memory and a short memory component, as opposed to an earlier perception that the series has stationary and non-stationary components. We have produced some evidence of long memory of a stationary nature in stock returns in this study. We have presented estimates of d and tests that $d = 0$, of a semiparametric character. It is perhaps not surprising that some returns exhibit long memory behaviour while others do not,

(a) Plot of $\hat{d}_1$ Estimate for SES Returns τ=4, J=4

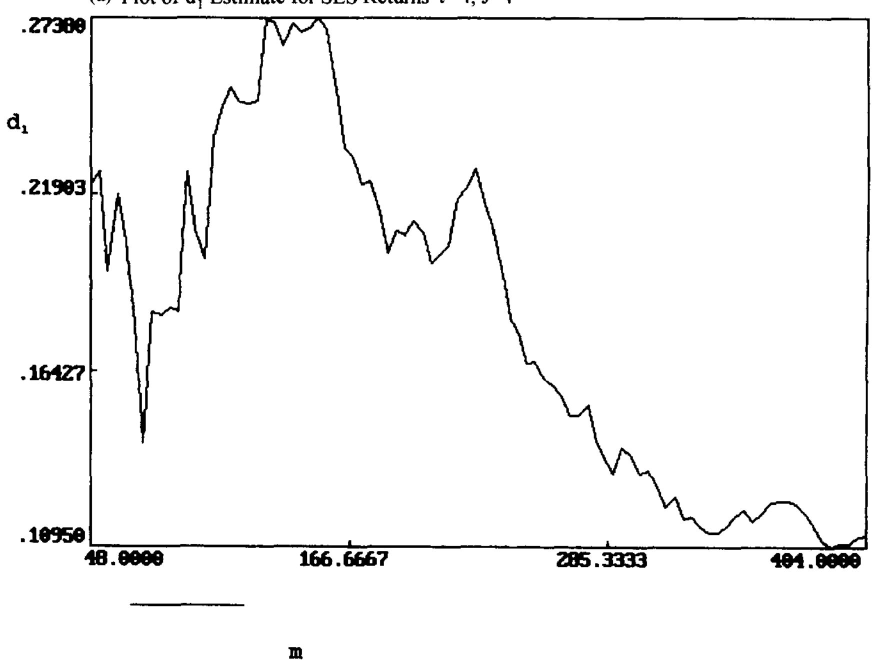

Fig. 4.

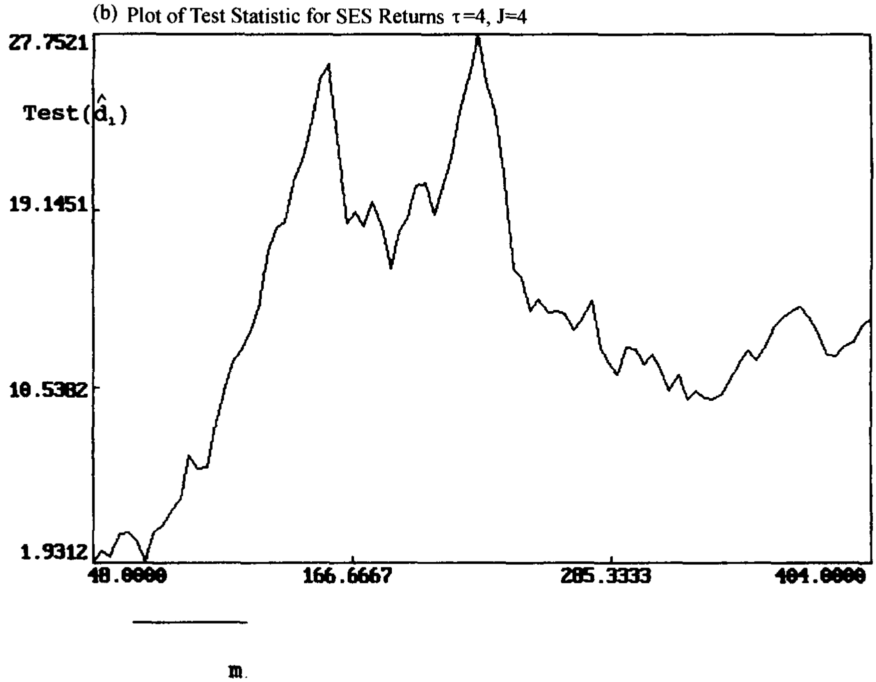

Fig. 4. Continued.

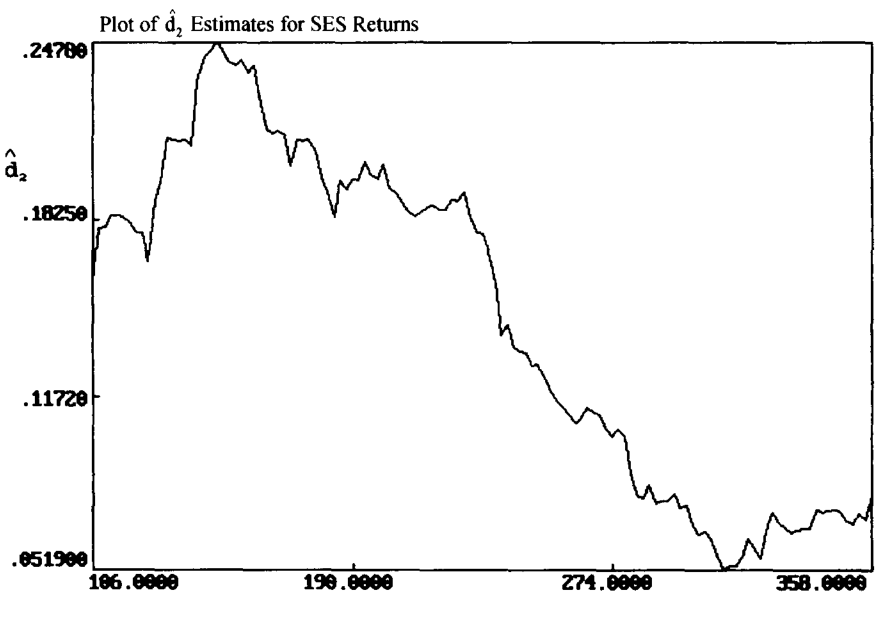

Fig. 5.

but it is interesting to note that the SES index is mean reverting. Note that the usual unit root rests, including ones which correct for short-memory autocorrelation, are not robust to long memory.

A semiparametric estimate of d can be the first step in building a parametric time series model, such as a fractional ARIMA, for use in forecasting, though there is as yet little evidence of the performance of these models in forecasting. An eventual parametric model is especially suitable in case of macro series, and policy makers in developing countries will also be interested in findings regarding the possible presence of long memory in their macro series, to add to the evidence for industrialised countries. Research in this direction will be fruitful and will certainly contribute to the public and economic policy debate and planning in developing and newly industrialised countries.

Appendix

Names of the companies and sample period are listed in Table 5.

Table 5

No.	Name of company	No. of obs.	First obs.
1	Avimo Singapore Limited	1073	870507
2	Cycle & Carriage Limited	4293	750102
3	Cerebos Pacific Limited	2073	830707
4	Cold Storage Holdings Limited	751	880801
5	Fraser & Neave Limited	4293	750102
6	Haw Par Brothers International Limited	4293	750102
7	Inchcape Bhd	4293	750102
8	Intraco Limited	4292	750103
9	Keppel Corporation	2777	801024
10	Lum Chang Holdings Limited	1687	841228
11	Metro Holdings Limited	4292	750103
12	Neptune Orient Lines Limited	2630	810519
13	Natsteel Limited	4293	750102
14	Resources Development Corporation Limited	1157	870109
15	Sembawang Shipyard Limited	4293	750102
16	Singapore Airlines Limited	1434	851218
17	Sime Singapore Limited	2992	791218
18	Singapore Press Holdings Limited	1693	841220
19	Singapore Bus Service Limited	3386	780626
20	Straits Trading Company Limited	4293	750102
21	Times Publishing Limited	593	890317
22	United Engineers Limited	4292	750103
23	United Industrial Corporation Limited	4293	750102
24	Wearne Brothers Limited	4293	750102
25	Wing Tai Holdings Limited	611	890221
26	Yeo Hiap Seng Limited	4289	750108

Note: The number of observations refers to the full sample period. We have 30 Blue-Chip stocks in the ST1 index. Four of these are newly listed stocks, which have been excluded from our study. For the sub-sample, we have 968 observations fewer. For those stocks having fewer than 1500 observations for the full sample period, we decided not to run the results for the subsample period.

References

Agiakloglou, C., P. Newbold and M. Wohar (1993). Bias in an estimator of the fractional difference parameter. *J. Time Series Analysis* **14**, 235–246.

Booth, G., F. Kaen and P. Koveos (1982a). Presistent dependence in gold prices. *J. Financial Res.* **4**, 85–93.

Booth, G., F. Kaen and P. Koveos (1982b). R/S analysis of foreign exchange rates under two international monetary regimes. *J. Monetary Econom.* **10**, 407–415.

Box, G. and G. Jenkins (1970). *Time Series Analysis: Forecasting and Control.* Holden Day, San Francisco.

Campbell, J. and G. Mankiw (1987). Are output fluctuations transitory? *Quart. J. Econom.* **102**, 857–880.

Cochrane, J.H. (1988). How big is the random walk in GNP. *J. Political Econom.* **96**, 893–920.

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar process. *Ann. Statist.* **17**, 1749–1766.

Diebold, F., S. Husted and M. Rush (1991). Real exchange rates under the gold standard. *J. Political Econom.* **99**, 1252–1271.

Diebold, F. and G. Rudebusch (1989). Long memory and persistence in aggregate output. *J. Monetary Econom.* **24**, 189–209.

Fong, W.M. (1992). Long term memory in UK stock returns: R/S analysis and fractional integration. Preprint.

Fox, R. and M. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14**, 517–532.

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

Granger, C. (1966). The typical spectral shape of an economic variable. *Econometrica* **34**, 150–161.

Greene, M. and B. Fielitz (1977). Long-term dependence in common stock returns. *J. Financial Econom.* **4**, 339–349.

Haubrich, J. and A. Lo (1989). The sources and nature of long-term dependence in the business cycle. NBER Working Paper No. 2951.

Helms, B., F. Kaen and R. Rosenman (1984). Memory in commodity futures contracts. *J. Futures Markets* **4**, 559–567.

Kim, M., Nelson, C. and R. Startz (1991). Mean reversion in stock prices? A reappraisal of the empirical evidence. **58**, 515–528.

Lo. A. (1991). Long Term Memory in Stock Prices. *Econometrica* **59**, 1279–1313.

Lo, A. and MacKinlay (1988). Stock prices do not follow random walk: evidence from a simple specification test. *Rev. Financial Stud.* **1**, 41–66.

Lobato, I. and P.M. Robinson (1996). Averaged periodogram estimation of long memory. Forthcoming, *J. Econometrics.*

McQueen, G. (1992). Long horizon mean-reverting stock prices revisited. *J. Financial Quantitative Anal.* **27**, 1–18.

Nelson, C. and C. Plosser (1982). Trends and random walks in macroeconomic time series. *J. Monetary Econom.* **10**, 139–162.

Porter-Hudak, S. (1990). An application of the seasonal fractional difference model to the monetary aggregate. *J. Am. Statist. Assoc.* **85**, 338–344.

Poterba, J. and L. Summers (1988). Mean reversion in stock returns: evidence and implications. *J. Financial Econom.* **22**, 27–60.

Robinson, P.M. (1994a). Time series with strong dependence. Invited Paper, 6th World Congress of the Econometric Society, Barcelona. *Adv. Econometrics: Sixth World Congress*, Cambridge University Press, Cambridge, 47–96.

Robinson, P.M. (1994b). Semiparametric analysis of long-memory time series. *Ann. Statist.* **22**, 515–539.

Robinson, P.M. (1994c). Rates of convergence and optimal spectral bandwidth for long range dependence. *Probab. Theory Related Fields* **99**, 443–473.

Robinson, P.M. (1995). Log-periodogram regression of time series with long range dependence. *Ann. Statist.* **6**, 1048–1072.

Shea, G. (1991). Uncertainty and implied variance bounds in long-memory models of the interest rate term structure. *Empirical Econom.* **16**, 287–312.

Sowell, F. (1992a). Maximum likelihood estimation of stationary univariate fractional integrated time series models. *J. Econometrics* **53**, 165–188.

Sowell, F. (1992b). Modelling long-run behaviour with the fractional ARIMA model. *J. Monetary Econom.* **9**, 277–302.