Applied Financial Economics, 2002, **12**, 725–729

Long memory in stock returns: some international evidence

ÓLAN T. HENRY

Department of Economics, The University of Melbourne, Victoria 3010, Australia
E-mail: oth@unimelb.edu.au

Recent empirical studies suggest that long horizon stock returns are forecastable. While this phenomenon is usually attributed to time varying expected returns, or speculative fads, it may also be due to long memory in the returns series. Long range dependence is investigated using parametric and semiparametric estimators in a sample of nine international stock index returns. The results provide evidence of long memory in the German, Japanese, South Korean and Taiwanese markets.

I. INTRODUCTION

Recent evidence on the presence of long memory in stock returns is mixed. Whereas many studies find evidence for long horizon predictability in stock returns (see Fama and French, 1988; Porterba and Summers, 1988; Mills, 1993 *inter alia*), other authors (Goetzman and Jorion, 1993; Nelson and Kim, 1993; and Malliaropoulos, 1996 *inter alia*) argue that the converse is true. Where evidence of long horizon predictability is found, it is usually attributed to time variation in expected returns, or speculative bubbles, periods of marked divergence between the price of an asset and its fundamental, or true value. Campbell (1991) and Campbell and Ammer (1993) argue that persistence in equity returns underlies this long horizon forecast power. That is, this observed forcastabilty may be attributed to long range dependence, or long memory, in the returns time series.

The existing work on long memory in asset returns derives largely from the pioneering work of Hurst (1951). Greene and Fielitz (1977) and Aydogan and Booth (1988) both test for long memory using the rescaled range statistic of Hurst (1951). Lo (1991), using a modified rescaled range (R/S) statistic, finds no evidence of long memory in a sample of US stock returns. Mills (1993), using the modified R/S statistic and the semiparametric approach of Geweke and Porter-Hudak (1983), hereafter GPH, finds weak evidence of long memory in a sample of monthly UK stock returns. Lobato and Savin (1997) find no evidence of long memory in daily Standard and Poor 500 returns over the period July 1962–December 1994. Interestingly, Lobato and Savin (1997) find some evidence of long memory in the squared return data, which supports the conclusions of Ding *et al.* (1993).

This paper makes two main contributions. First, the paper aims to address some of the concerns regarding the robustness of the various estimation strategies by applying a range of parametric and semiparametric techniques to monthly index return data on nine countries, namely the USA, Japan, Germany, Great Britain, Hong Kong, Taiwan, South Korea, Singapore and Australia. Secondly, the disparities in capitalization, sophistication and market microstructure across the nine countries reduces concerns about sample or market specific results.

This paper is divided into four sections. Section II outlines the methods used to detect and model long memory in time series. The third section describes the data and presents the empirical results. The final section provides a brief summary and conclusion.

II. DETECTING AND MODELLING LONG MEMORY IN TIME SERIES

Consider a time series of stock returns, r_t, with an autocorrelation function ρ_j at lag j. McLeod and Hippel (1978) describe r_t as having long memory if

$$\lim_{n\to\infty}\sum_{j=-n}^{n}|\rho_j| = \infty \quad (1)$$

A popular method of capturing the type of behaviour defined in Equation 1 is the fractionally differenced time

Applied Financial Economics ISSN 0960–3107 print/ISSN 1466–4305 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/09603100010025733

series model of Granger (1980), Granger and Joyeaux (1980), and Hosking (1981). In this case r_t satisfies

$$(1-B)^d r_t = \varepsilon_t \qquad \varepsilon_t \sim IID(\mathrm{O}, \sigma_\varepsilon^2) \tag{2}$$

where B is the backshift operator, $Br_t = r_{t-1}$. Granger and Joyeaux (1980) show that

$$(1-B)^d r_t = \sum_{k=0}^{\infty} A_k r_{t-k} = \varepsilon_t \tag{3}$$

where the AR coefficients A_k are in terms of the gamma function

$$A_k = (-1)^k \binom{d}{k} = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} \tag{4}$$

For $|d| > 1/2$, the variance of r_t is infinite and hence r_t is nonstationary. For $-1/2 < d < 1/2$ the process is stationary and invertible. The autocorrelations of such a process decline at a hyperbolic rate to zero, a much slower rate of decay than the exponential decay of the ARMA process. Indeed the autocorrelation of such fractionally integrated processes remain significant at long lags, giving rise to the 'long memory' label. Baillie (1996) provides a survey and review of the theoretical and applied econometric literature on long memory processes.

Lee and Schmidt (1996) propose the test of Kwiatkowski *et al.* (1992), hereafter KPSS, as a test for the null of stationarity against the alternative hypothesis of fraction integration. The KPSS test involves regressing r_t against a constant μ, and a trend τ. The test is based upon

$$\eta_\tau = T^{-2} \sum S_t^2 / \sigma_T(q)$$

where S_t represents the partial sum of the residuals and $\sigma_T(q)$ represents an estimate of the long run variance of the residuals. Lee and Schmidt (1996) present Monte Carlo evidence that the KPSS test has power properties similar to the modified rescaled range statistic of Lo (1991).

The GPH test of Geweke and Porter-Hudak (1983), involves a semiparametric estimate of d, the degree of fractional integration. The spectral density of r_t is given by

$$f_r(\theta) = |1 - \exp(-i\theta)|^{-2d} f_\varepsilon(\theta) = \left[4\sin^2(\theta/2)\right]^{-d} f_\varepsilon(\theta) \tag{5}$$

where $f_\varepsilon(\theta)$ is the spectral density of ε_t. It follows that

$$\ln[f_r(\theta)] = \ln[f_\varepsilon(\theta)] - d\ln\left[4\sin^2(\theta/2)\right] \tag{6}$$

GPH propose estimation of d by regressing the periodogram $I_T(\theta_j)$ at frequencies $\theta_j = 2\pi j/T$, where $0 < k_1 \le j \le K \ll T$, against a constant and $\ln\left[4\sin^2(\theta/2)\right]$. Despite the fact that the GPH test is simple to apply and potentially robust to non-normality, Agiakloglou *et al.* (1992) argue that it is biased and inefficient when ε_t is AR(1) or MA(1), and the AR or MA parameter is relatively large.

An alternative semiparametric approach is that of Robinson (1992), who considers the discretely averaged periodogram

$$F(\theta) = \int f(\lambda)\mathrm{d}\lambda \tag{7}$$

Robinson estimates the Hurst coefficient, or equivalently the degree of fractional integration, d, as

$$H_q = \{1 - 2\log(q)\}^{-1} \log\{F(qw_m)/F(w_m)\} \tag{8}$$

As with the rescaled range and GPH approaches, there is a substantial amount of evidence documenting the poor performance of the Robinson's semiparametric estimator in terms of bias (see Baillie, 1996).

Sowell (1992) derives the exact Maximum Likelihood Estimator of the ARFIMA(p,d,q) process with unconditionally normally distributed disturbances ε_t. However the Sowell estimator is computationally burdensome and the approach taken here is the frequency domain based likelihood technique of Fox and Taqqu (1986). This approach simultaneously estimates all of the parameters in

$$\alpha(B)\nabla^d r_t = \delta + \beta(B)\varepsilon_t \tag{9}$$

where $\alpha(B)$ and $\beta(B)$ are polynomials in the lag operator, δ is a constant and ∇^d is the fractional differencing filter.

III. EMPIRICAL RESULTS

The data used in this study consist of monthly stock index observations over the period January 1982 to September 1998. The data under consideration are: USA – Dow Jones Industrial Index, denoted as *US*; Japan – Nikkei 225 Index, *JP*; Germany – Commerz-Bank Index, *GR*; United Kingdom – FT Industrial Index, *UK*; Hong Kong – Hang Seng Index, *HK*; Taiwan – Weighted Index, *TW*; South Korea – Composite Index, *SK*; Singapore – Straits Times Index, *SN* and the Australian All Ordinaries Index, *AU*. The data are published in Table F.06 of the *Reserve Bank of Australia Bulletin*, and are originally sourced from the *Nomura Research Institute Quarterly Economic Review*, the *Asian Wall Street Journal* and *Electronic News Services*.[1] The original share price indices were rebased by the RBA such that January 1985 = 100. The data were transformed into continuously compounded returns $r_{i,t} = \log(p_{i,t}/p_{i,t-1})$, where $p_{i,t}$ represents the value of index i at time t. This yields a total of 200 return observations. Table 1 displays summary statistics for the data.

[1] The data were downloaded from the Australian Bureau of statistics WWW page, gopher://trent.abs.gov.au:70/11/PUBS/reser/RBABF

Table 1. *Summary statistics for the data*

	Mean	Variance	Skewness	Excess kurtosis	Normality $\sim\chi^2(2)$	ADF	KPSS η_μ	KPSS η_τ
US	0.012	0.002	−1.374	8.195	585.23 [0.000]	−13.733	0.067	0.081
JP	0.004	0.004	−0.423	1.238	18.231 [0.000]	−13.716	0.410*	0.084
GR	0.009	0.003	−0.921	2.978	96.059 [0.000]	−12.839	0.109	0.095
UK	0.009	0.003	−1.430	7.783	545.123 [0.000]	−13.048	0.071	0.027
HK	0.013	0.007	−1.552	9.020	712.759 [0.000]	−13.351	0.108	0.066
TW	0.015	0.015	−0.440	3.066	79.710 [0.000]	−12.657	0.124	0.086
SK	0.009	0.004	0.316	0.310	3.888 [0.143]	−14.143	0.465*	0.148*
SN	0.005	0.005	−2.860	20.663	3569.41 [0.000]	−12.188	0.047	0.063
AU	0.009	0.004	−3.951	35.088	10133.25 [0.000]	−14.041	0.069	0.094

Notes: Marginal significance levels displayed as [.]. Normality is the Bera–Jarque (1980) test for normality distributed as $\chi^2(2)$. ADF is the Augmented Dickey–Fuller (1981) unit root test, with a 5% critical value of − 3.435. KPSS is the Kwiatkowski, Phillips, Schmidt and Shin test with 5% critical value of 0.463 and 0.146 respectively. * indicates significance at the 10% level, ** indicates significance at the 5% level.

The data appear extremely non-normal. All of the return distributions are negatively skewed, possibly due to the large negative return associated with the Market Crash of October 1987. The data also display a high degree of excess kurtosis. Such skewness and kurtosis are common features in asset return distributions, which are repeatedly found to be leptokurtic. With the exception of South Korea, the data fail to satisfy the null hypothesis of normality of the Bera–Jarque test at the 5% level. The results of the ADF unit root test indicate that all of the returns series are stationary. However, stationarity does not preclude the possibility of long-memory in the returns data.

Table 1 also presents the results of the KPSS tests η_μ and η_τ. Lee and Schmidt (1996) argue that KPSS test is consistent against stationary long memory alternatives and may therefore be used to distinguish between short and long memory processes. The KPSS test is similar to Lo's modified rescaled range statistic in both power and construction (see Lee and Schmidt, 1996 and Baillie, 1996 for further details). Again there is some evidence of long memory in the Japanese and South Korean stock indices. The remaining series satisfy the null hypothesis of stationarity about a (possibly non-zero) mean of the η_μ test at the 5% level. For the η_τ test, the Japanese, Hong Kong, South Korean and Australian returns fail the null hypothesis of stationarity around a trend. Given that efficient markets theory predicts $E_t(r_{t+1}) = 0$, the presence of a trend in the returns is unlikely. In short, the evidence from the KPSS tests is against the presence of long memory in the majority of the returns series.

Baillie (1996) comments that the GPH estimator is potentially robust to non-normality. Table 2 displays the results of the GPH estimates of d, the degree of fractional integration. A concern in the application of the GPH estimator is the choice of m, the number of spectral ordinates

Table 2. *Semiparametric estimates of d*

	GPH estimates $k = 0.475$	GPH estimates $k = 0.500$	GPH estimates $k = 0.525$	Robinson's estimator
US	−0.232 (0.512)	0.326 (0.424)	−0.400 (0.363)	0.10
JP	0.481 (0.214)	0.405 (0.202)	0.157 (0.0.227)	0.04
GR	0.318 (0.159)	0.258 (0.137)	0.215 (0.120)	0.08
UK	0.055 (0.180)	−0.058 (0.163)	−0.194 (0.161)	−0.03
HK	0.180 (0.217)	0.330 (0.217)	0.236 (0.211)	−0.05
TW	0.646 (0.175)	0.462 (0.182)	0.293 (0.185)	0.090
SK	0.701 (0.199)	0.696 (0.163)	0.589 (0.160)	0.16
SN	−205 (0.180)	0.031 (0.234)	−0.094 (0.220)	0.01
AU	0.068 (0.283)	−0.073 (0.247)	−0.106 (0.213)	−0.01
	{0.256}	{0.230}	{0.211}	{0.060}

Notes: OLS standard errors displayed as (.). Asymptotic standard errors displayed as {.}.

from the periodogram of r_t, to include in the estimation of d. Here results are presented for $m = T^k$, where $k = 0.475$, 0.5, 0.25, where T represents the sample size. Asymptotic and OLS standard errors for the GPH estimates of d are also presented. The results reveal little evidence of long memory in the US, UK, Hong Kong, Singapore and Australian returns. On the other hand, the GPH estimator provides some evidence of long memory in the German, Japanese and Taiwanese markets. The evidence for the South Korean market is far stronger, with $\hat{d}$ suggesting that returns are borderline nonstationary.

Table 2 also displays the estimates of d calculated using the Gaussian semiparametric estimator of Robinson (1992). There is a weak suggestion of long memory in the German and Taiwanese returns, and strong evidence for South Korea. The estimate of d for South Korea is in the stationary region. The Robinson approach provides no evidence for long memory in the Japanese returns, which contrasts with the GPH evidence. The remaining series display little evidence of long memory. Interestingly a number of the estimates of d are negative which is suggestive of antipersistence. However none of the negative estimates are statistically significant.

Agiaklogou *et al.* (1992) raise concerns about the GPH estimator when ε_t is not i.i.d. Scholes and Williams (1977) argue that stale price can induce an MA(1) error in stock index returns. It therefore appears prudent to obtain joint estimates of any short run ARMA parameters as well as d for our sample of stock index returns. The two-step procedure of Diebold and Rudebusch (1989), involves obtaining $\hat{d}$ using the GPH estimator, and in the second step, estimating the ARMA parameters from the filtered series $(1 - B)^{d} r_t$. The sampling distribution of this approach is unknown, however given the degree of non-normality in the data inappropriate inference is likely. Subject to a caveat about the non-normality of the data, the parameters of the ARFIMA process (9) are estimated using the frequency domain estimator of Fox and Taqqu (1986).

Table 3 reports $\hat{d}$ for various models from ARFIMA $(0, d, 0)$ to ARFIMA $(2, d, 2)$. The short run ARMA parameters are not reported to conserve space but are available from the author upon request. Schmidt and Tschernig (1995) discuss the identification of ARFIMA models using information criteria, highlighting the usefulness of the Schwarz (1978) Bayesian information criterion (BIC) in this task. Table 3 also highlights the ARFIMA parameterizations selected using the BIC. In contrast with the results of the semiparametric estimators, the Japanese, Taiwanese and South Korean markets appear to be consistent with short memory processes. The BIC selects an ARFIMA$(0, d, 0)$ or fractional noise model for all markets except the UK where an ARFIMA$(0, d, 1)$ is chosen. For the UK the BIC value was 156.62 and 156.08 for the ARFIMA(0,d,0) and ARFIMA$(0, d, 1)$ respectively. Thus

Table 3. *Frequency domain maximum likelihood estimates of d*

	(0,d,0)	(0,d,1)	(1,d,1)	(1,d,1)
US	−0.019*	−0.068	−0.107	−0.507
	(0.066)	(0.104)	(0.158)	(0.569)
JP	0.019*	0.059	0.051	0.820
	(0.058)	(0.110)	(0.084)	(0.255)
GR	0.044*	−0.007	0.002	0.005
	(0.063)	(0.113)	(0.177)	(0.062)
UK	−0.064	−0.243*	−0.295	−0.258
	(0.070)	(0.083)	(0.128)	(0.116)
HK	−0.028*	−0.127	−0.155	−0.143
	(0.068)	(0.092)	(0.121)	(0.150)
TW	0.061*	0.026	0.024	0.024
	(0.062)	(0.093)	(0.104)	(0.097)
SK	0.050*	0.279	0.173	0.760
	(0.051)	(0.216)	(0.078)	(0.785)
SN	0.063*	−0.054	−0.265	−0.386
	(0.070)	(0.097)	(0.264)	(0.282)
AU	−0.048	−0.102	−0.108	−0.106
	(0.063)	(0.099)	(0.118)	(0.115)

Notes: ARFIMA Parameterisation selected using BIC denoted as *

the choice between the two parameterizations is highly marginal. The ARFIMA$(0, d, 1)$ estimate of d for the UK is significant and negative, which indicates antipersistence. The estimate of d from the fractional noise model is insignificantly different from zero. In both cases the returns to the UK market are unforecastable. There is no evidence of fractional integration in the other series, which contrasts with the results of the semiparametric estimators reported in Table 2.

IV. SUMMARY AND CONCLUSION

There is a body of existing empirical evidence to suggest that long horizon stock returns are forecastable. Such forecastability is often attributed to time variation in expected returns or speculative fads, however it may also arise from the possibility of long memory in the data. This paper tests a set of monthly stock index returns for long memory. A wide range of parametric and semiparametric estimators were employed in an effort to obtain inference that are robust to the non-normality in the returns data. As a further measure to ensure robustness the study considered markets which differ widely in terms of capitalization, sophistication and microstructure. The semiparametric approaches provide strong evidence of long memory in the South Korean returns and some evidence of long range dependence in the German, Japanese and Taiwanese returns. The other returns series are broadly consistent with short memory processes. Frequency domain maximum likelihood estimation of the ARFIMA suggests that the returns are short memory processes.

The results suggest that long horizon predictability in stock returns for the UK, USA, Hong Kong, Singapore and Australia is more likely to arise from time variation in expected returns, or speculative bubbles, than from long memory. Furthermore the weak evidence of antipersistence in these markets suggests that such long horizon forecastability is unlikely. However for Japan, Germany, South Korea and Taiwan long range dependence may be a source of long horizon predictability. Relative transaction costs are greater for trading strategies based on short horizon predictability than for those strategies based on long horizon predictability. Thus the long horizon strategy may represent an unexploited profit opportunity. The tradeoff, of course, is the increased exposure to market risk as the investment horizon increases into the future.

ACKNOWLEDGEMENTS

I am grateful to Nigel Wilkins who provided the GAUSS routine to estimate the ARFIMA(p,d,q) models. Discussions with Peter Summers and Chris Brooks also contributed to the development of this paper. The usual disclaimer applies to all remaining errors and omissions

REFERENCES

Agiakloglou, C., Newbold, P. and Wohar, M. (1992) Bias in an estimator of the fractional difference parameter, *Journal of Time Series Analysis*, **14**, 235–46.

Akaike, H. (1974) A new look at statistical model identification, *IEEE Transactions on Automatic Control*, **AC-19**, 716–23.

Baillie, R. T. (1996) Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, **73**, 5–59.

Aydogan, K. and Booth, G. G. (1988) Are there long cycles in common stock returns? *Southern Economic Journal*, **55**, 141–9.

Bera, A. and Jarque, C. (1980) Efficient tests for normality, heteroscedasticity and serial independence of regression residuals, *Economics Letters*, **6**, 255–9.

Campbell, J. Y. (1991) A variance decomposition for stock returns, *The Economic Journal*, **101**, 157–79.

Campbell, J. Y. and Ammer, J. (1993) What moves stock and bond markets? A variance decomposition for long term asset returns, *The Journal of Finance*, **43**, 3–37.

Dickey, D. and Fuller, W. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, **49**, 1057–72.

Diebold, F. X. and Rudebusch, G. D. (1989) Long memory and persistence in aggregate output, *Journal of Monetary Economics*, **24**, 189–209.

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

Fama, E. F. and French, K. R. (1988) Permanent and temporary components in stock prices, *Journal of Political Economy*, **96**, 246–73.

Fox, R. and Taqqu, M. S. (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Annals of Statistics*, **14**, 517–32.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long memory time series models, *Journal of Time Series Analysis*, **4**, 221–38.

Goetzman, W. N. and Jorion, P. (1993) Testing the power of dividend yields, *Journal of Finance*, **48**, 663–79.

Granger, C. W. J. (1980) Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics*, **14**, 227–38.

Granger, C. W. J. and Joyeaux, R. (1980) An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis*, **1**, 15–39.

Greene, M. and Fielitz, B. (1977) Long term dependence in common stock returns, *Journal of Financial Economics*, **4**, 339–49.

Hosking, J. R. M. (1981) Fractional differencing, *Biometrica*, **68**, 165–76.

Hurst, H. E. (1951) Long term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, **116**, 770–99.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?, *Journal of Econometrics*, **54**, 159–78.

Lee, D. and Schmidt, P. (1996) On the power of the KPSS test of stationarity against fractionally-integrated alternatives, *Journal of Econometrics*, **73**, 285–302.

Lo, A. W. (1991) Long term memory in stock market prices, *Econometrica*, **59**, 1279–313.

Lobato, I. N. and Savin, N. E. (1997) Real and spurious long memory properties of stock market data, Discussion Paper, Department of Economics, University of Iowa.

Malliaropoulos, D. (1996) Are long horizon stock returns predictable? A bootstrap analysis, *Journal of Business Finance and Accounting*, **23**, 96–106.

Mills, T. C. (1993) Is there long-term memory in UK stock returns?, *Applied Financial Economics*, **3**, 303–6.

McLeod, A. I. and Hippel, A. W. (1978) Preservation of the rescaled adjusted range, 1: a reassessment of the Hurst phenomenon, *Water Resources Research*, **14**, 491–508.

Nelson, C. R. and Kim, M. J. (1993) Predictable stock returns: the role of small sample bias, *The Journal of Finance*, **48**, 641–61.

Porterba, J. M. and Summers, L. H. (1988) Mean reversion in stock prices: evidence and implications, *Journal of Financial Economics*, **22**, 27–59.

Robinson, P. M. (1992) Semiparametric analysis of long-memory time series, *Annals of Statistics*, **22**, 515–39.

Scholes, M. and Williams, J. (1977) Estimating betas from non-synchronous data, *Journal of Financial Economics*, **5**, 309–27.

Sowell, F. B. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, **53**, 165–88.

Schmidt, C. M. and Tschering, R. (1995) The identification of fractional ARIMA models, discussion paper, Humboldt University.

Schwarz, G. (1978) Estimating the dimensions of a model, *Annals of Statistics*, **6**, 461–4.