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 forecastability may be attributed to long range dependence, or long memory, in the returns time series.


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 \( \rho_j \) at lag \( j \). McLeod and Hippel (1978) describe \( r_t \) as having long memory if

\[
\lim_{n \to \infty} \sum_{j=-n}^{n} |\rho_j| = \infty
\]

A popular method of capturing the type of behaviour defined in Equation 1 is the fractionally differenced time
series model of Granger (1980), Granger and Joyeaux (1980), and Hosking (1981). In this case, \( r_t \) satisfies

\[
(1 - B)^d r_t = \varepsilon_t, \quad \varepsilon_t \sim IID(O, \sigma^2_\varepsilon) \tag{2}
\]

where \( B \) is the backshift operator, \( B r_t = r_{t-1} \). 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 \( \mu \), and a trend \( \tau \). The test is based upon

\[
\eta_t = T^{-2} \sum S_t^2 / \sigma_T(q)
\]

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

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

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

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

GPH propose estimation of \( d \) by regressing the periodogram \( I_\theta(\theta_j) \) at frequencies \( \theta_j = 2\pi j / T \), where \( 0 < k_1 < j \leq K \ll T \), against a constant and \( \ln \left[ 4 \sin^2 \left( \frac{\theta}{2} \right) \right] \). 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 \( \varepsilon_t \) 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

\[
F(\theta) = \int f(\lambda)d\lambda \tag{7}
\]

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

\[
H_q = \{1 - 2\log (q)\}^{-1} \log \{F(qw_m)/F(w_m)\}
\]

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 \( \varepsilon_t \). 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, \( \delta \) is a constant and \( \nabla^d \) 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. The original share price indices were rebased by the RBA such that January 1985 = 100. The data were transformed into continuously compounded returns \( r_{1,t} = \log (p_{i,t}/p_{i,t-1}) \), 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
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 $\eta_\alpha$ and $\eta_r$. 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 $\eta_\alpha$ test at the 5% level. For the $\eta_r$ 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 $E(r_{t+1}) = 0$, 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.
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 \( d \) 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 \( \varepsilon \) 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 \((1 - B)^{\hat{d}}\hat{r}_t\). The sampling distribution of this approach is unknown, however given the degree of nonnormality 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 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


