An Econometric Analysis of Long Horizon Mean Reversion in UK Stock Prices

Stewart D Hodges

and

Sanjay Yadav

Financial Options Research Centre,
Warwick Business School, The University of Warwick
Coventry, CV4 7AL

August 1995

Funding for this work was provided by Corporate Members of the Financial Options Research Centre: HSBC Markets Ltd., Morgan Grenfell & Co., Swiss Bank Corporation, Mitsubishi Finance International Ltd. and UBS Ltd.

We are grateful to Les Clewlow, Michael Selby, Chris Strickland and Gary Xu for helpful comments.

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Telephone: (01203) 524118
Fax: (01203) 524167

FORC Preprint: 95/58

Abstract

This paper is concerned with examining the mean reversion hypothesis for UK stock prices for the 1925-1991 period. Specifically, this paper incorporates recent innovations in time series techniques which are not subject to the flaws which have plagued other studies.

To the best of our knowledge, the phenomenon of mean reversion in UK stock prices has not been studied using a data set which covers the period 1925-1991. Most UK studies use data that extends from the mid 1950s onwards.

In this paper, we eschew reliance on asymptotic distributions to conduct inference. Instead, our preferred strategy is to use simulations to generate the empirical distributions of test statistics in order to conduct inference. We use Monte Carlo as well as Randomisation methods to obtain significant levels.

Using somewhat related testing methodologies, we do not find any evidence that is consistent with mean reverting behaviour in UK stock prices. This conclusion is robust with respect to heteroscedasticity in stock returns and different choice of sample periods. By and large, the random walk model for stock prices is accepted. There are a few instances when we come close to rejecting the random walk model, but in every case, rejection is in favour of mean aversion or positive serial correlation in stock returns. Unlike for the US, there is no evidence of statistically significant mean reversion in UK stock prices.

1. Introduction

Fama (1970) investigated the autocorrelation of short horizon returns i.e., daily and weekly stock returns and found that the autocorrelations were close to zero. Recently, however, many researchers, in particular Poterba and Summers (1988) and Fama and French (1988) have cast doubt upon the earlier finding that stock returns are serially independent. By concentrating on the autocorrelations of long horizon returns, they discovered that stock returns possess large predictable components. Poterba and Summers employed Cochrane's (1988) variance ratio methodology to test for predictable or mean reverting components in stock prices; Fama and French used a closely related regression specification to test the same phenomenon. Their results suggest that stock returns exhibit positive autocorrelation for return horizons under a year and negative autocorrelations for return horizons in excess of one year. Statistically significant positive as well as negative autocorrelations are consistent with the rejection of the random walk model for stock prices. Negative autocorrelations has been interpreted as evidence supportive of mean reversion in stock prices. Mean reversion implies that a given change in price tends to be reversed by a predictable change in the opposite direction in due course.

Kim, Nelson and Startz (1991) and Richardson and Stock (1989), however, are critical of the empirical results reported by Fama and French and Poterba and Summers. Based on simulation experiments that is appropriate if the underlying return generation process is nonnormal and conditionally heteroscedastic, they argue that the statistical evidence in favour of mean reversion in the two aforementioned papers has been overstated. Additionally, Kim, Nelson and Startz found that the strongest evidence in favour of mean reversion in US stock prices was entirely concentrated in the pre WWII period.

This paper is concerned with examining the mean reversion hypothesis for UK stock prices for the 1925-1991 period. This paper incorporates recent innovations in time series techniques which are not subject to the flaws which have plagued other studies. Specifically, this paper improves earlier research on this topic and reexamines mean reversion in UK stock prices with at least four extensions.

To the best of our knowledge, the phenomenon of mean reversion in UK stock prices has not been studied using a data set which covers the period 1925-1991. Most UK studies use data that extends from the mid 1950s onwards. The addition of pre WWII data to the sample and the consequent enlargement of the sample size is crucial in view of Poterba and Summers finding that tests utilising overlapping return observations have low power against alternatives to the random walk model. Also, use of pre WWII data in the study is important in order to detect whether the well documented evidence in favour of mean reversion in US stock prices for the 1926-46 subperiod carries over to the UK and is perhaps an international phenomenon. Additionally, isolating periods of mean reverting behaviour in stock prices will assist researchers in identifying the factors responsible for mean reversion.

Researchers investigating mean reversion in UK stock prices have not conducted joint tests over multiple return horizons. The sole exception is Jegadeesh (1991). However, the sample period used in his study extends from 1955 to 1988 and so his sample size is considerably smaller than the one used in this study. More importantly, his data set does not include the pre WWII period. Richardson (1989) is critical of studies which focus on individual return horizon statistics, typically the return horizon which provides strongest evidence consistent with deviations from a random walk model, on the grounds that individual test statistics are not independent and so reliance on them may yield misleading results. In this paper we implement joint tests over multiple measurement intervals.

We eschew reliance on asymptotic distributions to conduct inference. Instead, our preferred strategy is to use simulations to generate the empirical distributions of test statistics in order to conduct inference. We use Monte Carlo as well as Randomization methods to obtain significance levels. Our intention is to find out whether the difference between the prob values obtained using Monte Carlo simulations and Randomization methods are substantial, as reported by Kim, Nelson and Startz.

The superior power properties of the variance ratio test compared with regression based test against non-factional, i.e., ARIMA alternatives is well documented in Poterba and Summers (1988). It is worthwhile investigating whether the same conclusion carries over to the case of fractional, that is, ARFIMA alternatives to the random walk. This is particularly important in view of recent findings that certain macro and financial time series, such as, GNP (Diebold and Rudebusch (1989), exchange rates (Cheung (1993)) and interest rates (Shea (1991)) can be represented as fractional processes. Consequently, to aid the ability of researchers to detect a mean reverting or mean averting process against a random walk, this paper uses extensive Monte Carlo simulations to analyse the size and the power of random walk tests against fractional alternatives.

The remainder of this paper is organised as follows. Section 2 contains a brief outline of the different testing methodologies adopted in this study. Section 3 contains the Monte Carlo study regarding the size and the power of the random walk tests considered in this paper. Section 4 describes the data and presents the empirical results. We offer our conclusions and suggestions for further research in Section 5.

2. Methodology

The tests considered in this study are the variance ratio (VR) test of Cochrane (COCH), the extension of this test by Lo and MacKinlay (1988) (both the homoscedastic (LOMAC) and heteroscedastic (LOMACH) versions), the long horizon autocorrelation test of Fama and French (FF) and finally a modified version of the FF test, JEG which is due to Jegadesh. The tests are summarised in Table 1.

The variance ratio test exploits the fact that the variance of returns should be proportional to the sampling interval if share prices (P_t) follow a random walk. In other words, the variance of P_t - P_{t-k} is k times the variance of P_t - P_{t-1} if P_t follows a random walk.

Cochrane shows that VR(k) can be approximated by:

$$VR(k) = 1+2 \sum_{j=1}^{k-1} \frac{k-j}{k} \hat{p}(j), \ where \ \hat{p}(j) = j^{th} \text{ autocorrelation coefficient.}$$

It is clear that a variance ratio of less than one implies negative serial correlation in stock returns.

Lo and Mackinlay (1988) have derived a modified version of the variance ratio statistic which utilises unbiased estimators of Var ($R_{t,k}$) and Var (R_t), where R_t is the one period return, and thus yields a more useful test.

Since it is generally recognised that returns are conditionally heteroscedastic, Lo and Mackinlay derive a version of the variance ratio statistic that is robust to heteroscedasticity and is asymptotically standard normally distributed. On the basis of extensive Monte Carlo simulation evidence, Lo and Mackinlay suggest that their test statistic (Z(k) and Z*(k)) is standard normally distributed even in small samples.

Another methodology for testing the hypothesis of long term mean reversion in stock prices has been proposed by Fama and French and entails regressing $R_{t,t+k}$ on $R_{t,t-j}$. Under the random walk null hypothesis, the theoretical value of the slope coefficient is equal to zero. Fama and French set the return measurement interval for the dependent as well as the independent variable to be equal i.e., j=k. Jegadeesh has called into question the widespread practice of setting j=k on statistical grounds. He argues, on the basis of evidence gleaned from simulation experiments, that setting j=1 results, asymptotically, in the highest power against Summer's (1986) mean reverting alternative model to the random walk. Put succinctly, it is desirable, on statistical grounds, to use one month returns as the dependent variable and aggregated returns as the independent variable in the regression model.

The regression model suffers from three shortcomings that renders the task of drawing statistical inference pertaining to mean reversion rather difficult. It is well documented in the literature that OLS regression models that use overlapping returns data induce serial correlation in the residuals which leads to biased and inconsistent estimates of the standard errors. Second, stock returns are heteroscedastic and so under the null hypothesis of a random walk, the residuals will be heteroscedastic. Finally, it is generally acknowledged that returns are not normally distributed and so, under the null hypothesis of a random walk, the disturbance term will also not be normally distributed.

Fama and French employ the Hansen and Hodrick (1983) procedure to adjust the standard errors of the slope coefficients for residual autocorrelation. Jegadeesh uses White's (1980) heteroscedasticity consistent estimator to obtain heteroscedasticity robust standard errors of the slope coefficients. However, neither adopts a testing procedure that is robust with respect to the non normality of stock returns. For instance, both conduct Monte Carlo simulations assuming that returns are drawn from a normal distribution. We adopt a radically different approach to tackle the problem of non normal returns. It is well known that stock returns are non normal and conditionally heteroscedastic. The actual distribution of stock returns is, of course, unknown. Randomization (or shuffling) methods are valid when the population distribution is unknown.

Randomization methods rely on resampling the data to estimate the distribution of sample statistics. Randomization is based on the premise that under the null hypothesis one variable is distributed independently of another. The null hypothesis we are interested in is that returns are distributed independently of their ordering in time. Randomization shuffles the data to destroy any time dependence and then recalculates the test statistic for each shuffle to estimate its distribution under the null. To obtain significance levels we count the number of times the calculated statistic after randomization falls below the value of the actual historical statistic. Significance levels below (above) 0.05 (0.95) indicate mean reversion (aversion) in stock prices. For the sake of comparison, we also report significance levels using Monte Carlo methods under normality.

UK stock return variances were higher in the stock market boom in 1975, the crash around September 1981, and the crash in October 1987. It is also apparent that volatility reverts to pre crash/boom levels after these periods. Thus, one would expect that the return series is heteroscedastic and that the regression errors exhibit variation through time. Radomisation, however, destroys the temporal pattern of heteroscedasticity present in the data. Consequently, under randomisation we are forced to assume that returns are homoscedastic undeer the null hypothesis. It would be interesting to investigate how the presence of heteroscedasticity under the null alters inference. To preserve the pattern of heteroscedasticity present in the actual return series we carry out a stratified randomisation of the total return data. We partition the sample into seven regimes or strata, where each regime displays a different level of return variance. To generate an artificial return series, the return data are sampled without replacement from each regime.

A range of tests have recently been developed which test the null hypothesis of a random walk over multiple return horizons, k. These tests are extensions of the single horizon test set out in Table 1. Chow and Denning (1993) developed the LOMAC tests, denoted JLOMAC and JLOMACH, Eckbo and Liu (1993) developed the Cochrane test, denoted JCOCH, and Jegadeesh (1991) developed the Jegadeesh test, JJEG. These test are presented in Table 5.

3. Monte Carlo Study

3.1 Fractionally Integrated Models

We can distinguish between:

(i) Stationary ARMA(p1,p2) series, represented as

$$A(L)y_t = B(L)\,\varepsilon_t, t = 1, \ldots, T \qquad (1)$$

where, $A(L) = 1 - \phi_1 L - \phi_2 L^2 - .. - \phi_{p1} L^{p1}$ and $B(L) = 1+\theta_1 L+\theta_2 L^2 + .. + \theta_{p2} L^{p2}$.

This y_t series is said to be integrated of order zero, that is, $y_t \sim I(0)$, and is characterised as having a finite memory, an autocorrelation function which decays to zero, the roots of A(L) lying outside the unit circle, and a finite spectral density at the origin and

(ii) Nonstationary ARIMA(p_1, d, p_2) series, represented as

$$A(L)y_t \equiv (1-L)^d A^*(L)y_t = B(L)\,\varepsilon_t,\ t = 1, \ \ldots, T \qquad (2)$$

where, d is an integer number, often taken to be unity. This y_t series is said to be integrated of order, d, that is, $y_t \sim I(d)$, and is characterised by having infinite memory, an autocorrelation function which does not decay, at least one of the roots of A(L) on the unit circle, and an infinite spike in the spectral density at the origin. A non-stationary series, y_t, as represented in equation (2), is transformed into a stationary series by appropriately differencing y_t d times.

Following the paper by Granger and Joyeux (1980) the range of models available were expanded by introducing fractionally integrated series. Fractionally integrated series broke the dichotomy between I(0) and I(I) series, by permitting the integration parameter, d, in equation (2), to take on non-integer values. For non-integer values of d equation (2) is an ARFIRMA(p_1,d,p_2) model.

For $-1.5 < d < 0.5$, the series, y_t, is stationary with an invertible ARMA representation. This ARFIMA model exhibits more persistence with the autocorrelation function decaying much slower than for the corresponding I(0) series with similar ARMA parameters. For $d > 0.0$ the spectral density has an infinite spike at the origin, while for $d < 0.0$ the value of the spectral density at the origin is zero.

For $0.5 < d < 1.0$, the series, y_t, is nonstationary with a non-invertible ARMA representation. However, even though the series is non-stationary, the autocorrelation function will still decay to zero, implying that the memory of the process is finite and that given a shock the process will tend to revert to its mean, that is mean-reverting. For $d > 1.0$ the series is mean-averting, and a shock to the process will cause the series to deviate away from its original starting point.

Fractionally integrated processes have autocorelation functions that decays at a much slower rate than for the corresponding I(0) series. The persistence in these series implies that unit root tests, which attempt to estimate the extent of persistence, lose power quickly as the fractional parameter, d, approaches unity. This point is demonstrated in Diebold and Rudebusch (1991).

For a simple ARFIMA(0,d,0) process, equation (2) is written as

$$(1-L)^d \; y_t \; \equiv [1-dL+\frac{d(d-1)}{2!} \; L^2 \; - \frac{d(d-1)(d-2)}{3!} \; L^3 \; + \ldots]y_t \; = \varepsilon_t, t = 1,\ldots,T \qquad (3)$$

where $\varepsilon_t \sim \in(0,\sigma^2)$, the corresponding autocorrelation coefficients are calculated as

$$\rho_\tau \; = \frac{\Gamma(1-d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1-d)}, \quad \tau = \pm 1, \pm 2, .., \qquad (4)$$

and autocovariances as,

$$\gamma_\tau \; = \frac{\sigma^2_\varepsilon \; \Gamma(1-2d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(1-d)\Gamma(\tau+1-d)}, \; \tau = \pm 1, \pm 2, ... \quad \gamma_o = \frac{\sigma^2_\varepsilon \; \Gamma(1-2d)}{\Gamma(1-d)\Gamma(1-d)} \qquad (5)$$

3.2 Monte Carlo Simulations

We consider fractionally integrated DGP's for the series y_t, for a range of integration parameters, d = 0.7, 0.8, 0.9, 0.95, 1.00, 1.05, 1.1, 1.2, and 1.3. In order to generate the series y_t for these values of d, we actually generate $(1-L)y_t = \Delta y_t$ using the fractionally integrated parameters, $\tilde{d} = d - 1 = -0.3$, -0.2, -0.1, -0.05, 0.0, 0.05, 0.1, 0.2, 0.3, and then use a staring value of $y_o = 0.0$ to construct the series $(y_1, y_2, \ldots, y_T)$. The series $(1-L)y_t$ is constructed along the lines used by Diebold (1989) and Sowell (1990). A vector ε_t of T N(0,1) deviates is constructed using the Nag routine G05DDF. Using equation (5) we calculate the desired TxT autocovariance matrix, Σ. Using NAG routine F07FDF a Choleski decomposition of Σ, $\Sigma = P P'$ is performed, where P is lower triangular. Finally, $(1-L)y_t$ is formed as $(1-L)y=P\varepsilon$. As Diebold (1989) and Granger and Joyeux (1980) note, the attraction of this method for generating the series is that the series y is formed without dependence on startup values. The number of replications undertaken for the power results is N = 1000 and for the size results is N = 10,000.

The test statistics are calculated for a range of different sample sizes and a range of k. Following Lo and MacKinlay (1989) our sample sizes are taken as T = 32, 64, 128, 256, 512, and 1024. For the two tests of Lo and MacKinlay, k is allowed to take on values up to 50% of the sample size, k =2,4,8, ..., (1/2)T. For the tests of Cochrane and Jegadeesh, k is allowed to be as large as 25% of the sample size, that is, k = 2, 4, ..., (1/4)T, the smaller choice of k is because Kim, Nelson and Startz (1991) use a smaller number for k compared with Lo and MacKinlay (1989), when calculating the Cochrane statistic. For the sake of brevity, only results for d = 1.1 and 0.9 are presented, but results for other values of d are qualitatively similar to what is presented here.

3.3 Power Results

Tables 2 and 3 report the power of the tests when the fractional integration parameter, d = 1.1 and 0.9 respectively. For each Table, power results are reported at the 5% significance level against the two-sided alternative of a fractional process.

Tables 2 and 3 show that regardless of the sample size or the parameter d, the power of the LOMAC test initally increases with k up to around k = 4 and then declines quite rapidly for larger values of k. In fact, across all values of d and all sample sizes, a value of k = 4 or 8 seems to yield highest power. A small optimal value of k in these fractionally integrated models is in contrast with the much larger optimal values found by Lo and MacKinlay. This test exhibits high power and correctly rejects the null hypothesis of a random walk up to 95.4% of the time at k = 4 and T = 1024. This test performs slightly asymmetrically, for d < 1 and d > 1. There is evidence that power does not fall quite as quickly for d> 1.0 with increases in k, compared with the results for d < 1.0. Overall, it seems this test is capable of detecting mean aversion better than mean reversion, although this disparity disappears as the sample size increases. The LOMACH test has power properties that are similar to the LOMAC test. However, the LOMACH test is marginally less powerful, which is not surprising given that the errors are homoscedastic.

For the COCH test, the optimal choice of k appears to be either k = 4 or 8, with a slight tendency for larger values of k compared with the Lo and MacKinlay tests. The power results for this test are slightly better than those for the Lo and MacKinlay tests, although the difference between the two is negligible. Once again, the asymmetry between d < 1.0 and d > 1.0 exists, with this test better able to detect mean aversion than mean reversion.

The results from the JEG test appears to show that the optimal lag length is $k = 2$, with power falling after $k = 4$. This test exhibits less power than either the LOMAC, LOMACH or the COCH tests for $d < 1.0$, although this test has comparable power to the other tests for $d > 1.0$.

3.4 Size Results

As the LOMACH and JEG tests are designed to allow for heteroscedastic errors, we now report the finding when we assume a near integrated ARCH model of the form $\varepsilon_t \sim N(0,h_t)$, $h_t = 0.95h_{t-1} + \eta_t$. The size results are presented in Table 4. The empirical size results for the LOMAC and the COCH tests are poor with a tendency to over-reject the correct null hypothesis for each sample size T for small k. The size results of the LOMACH test are close to their theoretical values. With, if anything, a tendency to under-reject the null hypothesis, especially for small k. The JEG test performs well, with the empirical size results close to their theoretical values.

3.5 Multiple Horizon Random Walk Tests

Table 6 contains the power probabilities for the joint tests for $d = 1.1$ over all sample sizes and all values of k. While the power of the best multiple horizon test, JJEG, is not a large as the maximum power of the LOMAC test at $k = 4$, the power can be markedly larger than that associated with the single horizon tests at non-optimal values of k. The relative ordering of the alternative multiple horizon random walk tests would be as follows: JJEG at the top, followed by JLOMAC, JLOMACH and JCOCH.

4. Data and Empirical Results

The data used in this study was obtained from R. Watsons and Sons, a firm of Actuaries. We are grateful to David Wilkie for supplying us with the data. The return series is the logged change in a total return index (or accumulation index) that is constructed on the basis of two main indices: the Actuaries Indices (1930-1962) and the Financial Times Actuaries All Share Index (1962-1991). The total return index incorporates dividend income gross of tax and reinvested free of expenses.

The variance ratios, M(k)+1, for increasing holding periods are presented in Tables 7 and 8 along with their p-values, computed using Monte Carlo, Randomization and Stratified Randomization methods. For the full sample period the variance ratios, M(k)+1, are greater than one. The variance ratio rises up to 16 months and then declines. None of the p-values obtained using Randomization methods are greater than 0.95 or less than 0.05, which implies that the random walk null hypothesis is accepted for the sample period 1925-1991. Compared with Randomization, p-values obtained using Stratified Randomization are somewhat lower, although they always lie between 0.05 and 0.95. Monte Carlo p-values are very similar to those obtained using Randomization. Likewise, p-values obtained using Monte Carlo simulations are remarkably similar, regardless of whether a random walk model with homoscedastic or heteroscedastic disturbances is assumed under the null.

For the 1925-1946 subperiod, there is no evidence of statistically significant mean reversion or mean aversion. We therefore accept the random walk null hypothesis. Unlike for the US, there is no evidence of statistically significant mean reversion in stock prices. For the 1947-1991 subperiod the random walk model is accepted for all return horizons tested. The results from conducting joint tests over multiple return horizons suggests that there are no departures from a random walk model for the entire sample period or for subperiods.

5. Conclusions

We used Monte Carlo simulations to investigate the relative performance of random walk tests against fractional alternatives. The two tests of Lo and MacKinlay and the Cochrane test perform similarly, with high power against fractional alternatives. Compared to these tests, Jegadeesh's test performs slightly worse. The superior power properties of the variance ratio test compared with regression based test against non-fractional, that is, ARIMA, alternatives carries over to the case of fractional, that is, ARFIMA, alternatives.

All tests detect mean aversion with slightly higher power than mean reversion. For all tests, the optimal lag length was small (k = 2 or 4) and comparable to those obtained by Lo and MacKinlay (1989) for the ARIMA (1,1,0) model. This lag length seems shorter than those traditionally used in the empirical finance literature.

In the presence of heteroscedastic (ARCH) errors, the LOMAC and COCH tests exhibit substantial size bias and are therefore unreliable. The LOMACH and JEG tests show no sign of size bias.

The multiple horizon random walk tests do not perform as well as the best simple horizon random walk test for optimal k. However, power can be significantly greater than that associated with the single horizon randon walk test for non-optimal values of k.

Using somewhat related testing methodologies, we do not find any evidence that is consistent with mean reverting behaviour in UK stock prices. This conclusion is robust with respect to heteroscedasticity in stock returns and different choice of sample periods. By and large, the random walk model for stock prices is accepted. Unlike for the US, there is no evidence of statistically significant mean reversion in UK stock prices.

Randomization methods and Monte Carlo simulations assuming a random walk model under the null yield very similar p-values, contrary to what is reported in Kim, Nelson and Startz.

There is scope for further work in this area that would aid our understanding of the phenomenon of mean reversion in stock prices. Firstly, we need to investiage whether evidence in favour of mean reversion is concentrated entirely in the month of January. Also, the Jegadeesh regression model is suitable for out of sample forecasting of stock returns. Our finding of no mean reversion or predictable components in stock prices would be reinforced if out of sample forecasts of stock returns based on the Jegadeesh model proved to be very poor.

References

Cheung, Y.W. (1993) "Long Memory in Foreign Exchange Rates", *Journal of Business and Economics Statistics*, 11, 93-101.

Chow, K.V. and Denning, K.C. (1993) "A Simple Multiple Variance Ratio Test", *Journal of Econometrics*, 58, 385-401.

Cochrane, J.H.(1988), "How Big is the Random Walk in GNP?", *Journal of Political Economy*, 96, 893-920.

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

Diebold, F..X. and Rudebusch, G.D. (1991) "On the Power of Dickey-Fuller Tests Against Fractional Alternatives", *Economics Letters*, 35, 155-160.

Diebold, F.X. (1989) "Random Walks Versus Fractional Integration: Power Comparisons of Scalar and Joint Tests of the Variance Time Function", in B. Raj (ed) *Advances in Econometrics and Modelling*, Kluwer Academic Publishers.

Eckbo, B.E. and Liu, J. (1993) "Temporary Components of Stock Prices: New Univariate Results", *Journal of Financial and Quantitative Analysis*, 28, 161-176.

Fama, E.F. and French, K.R.(1988), "Permanent and Temporary Components of Stock prices", *Journal of Political Economy*, 96, 246-273.

Fama, E.F.(1970), "Efficient Capital Markets:A Review of Theory and Empirical Work", *Journal of Finance*, 25, 383-416.

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

Hansen, L.P. and Hodrick, R.J.(1983), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates:An Econometric Analysis", *Journal of Political Economy*, 88, 829-853.

Jegadeesh, N.(1991) "Seasonality in Stock Price Mean Reversion:Evidence from the US and the UK", 46, 1427-1444, *Journal of Finance*.

Kim, M.J., Nelson, C.R. and Startz, R.(1991), "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence", *Review of Economic Studies*, 58, 515-528.

Lo, A. and Mackinlay, C.(1988), "Stock Prices Do Not Follow Random Walks:Evidence from a Simple Specification Test", *Review of Financial Studies*, 1, 41-66.

Lo, A.W. and MacKinlay, A.C. (1989) "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation", *Journal of Econometrics*, 40, 203-238.

Marriot, F.H.C. and Pope, J.A.(1954), "Bias in Estimation of Autocorrelations", *Biometrica*, 41, 390-402.

Poterba, J. and Summers, L.(1988), "Temporary Components of Stock Prices:Evidence and Implications", *Journal of Financial Economics*, 22, 27-59.

Richardson, M. and Stock, J.H.(1989) "Drawing Inferences from Statistics Based on Multiyear Asset Returns", *Journal of Financial Economics*, 25, 323-348.

Richardson, M.(1989), "Temporary Components of Stock Prices:A Skeptic's View", Unpublished Working Paper, Stanford University.

Shea, G.S. (1991) "Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest Rate Term Structure", *Empirical Economics*, 16, 287-312.

Sowell, F. (1990) "The Fractional Unit Root Distribution", *Econometrica*, 58, 495-505.

Summers, L.(1986) "Does the Stock Market Rationally Reflect Fundamental Values", *Journal of Finance*, 41, 591-601.

White, H.(1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity", *Econometrica*, 48, 817-838.

Table 1: Random Walk Tests

Name	Test	Null Hypothesis[1]
LOMAC	$Z(k) = \frac{\sqrt{T}M(k)}{\sqrt{2(2k-1)(k-1)/3k}}$, where $M(k) \equiv \frac{\Omega^{-1}\sum_{q=k}^{T}(P_q - P_{q-k} - k\hat{\mu})^2}{(T-1)^{-1}\sum_{q=1}^{T}(P_q - P_{q-1} - \hat{\mu})^2}$, $\Omega = k(T-k+1)(1-\frac{k}{T})$	$Z(k) = 0$
LOMACH	$Z^*(k) = \frac{\sqrt{T}M(k)}{\sqrt{\theta}}$, where $\theta(k) = \sum_{j=1}^{k-1}\left[\frac{2(k-j)}{k}\right]^2 \left[\frac{T\sum_{q=j+1}^{T}\left(P_q - P_{q-1} - \hat{\mu}\right)^2\left(P_{q-j} - P_{q-j-1} - \hat{\mu}\right)^2}{\left(\sum_{q=1}^{T}\left(P_q - P_{q-1} - \hat{\mu}\right)^2\right)^2}\right]$	$Z^*(k) = 0$
JEG[2]	$t_k = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$, where $\Delta P_t = \alpha_k + \beta_k \sum_{i=1}^{k} \Delta P_{t-i} + u_t$	$\beta_k=0$
COCH	$VR(k) = 1 + 2\sum_{j=1}^{k-1}\left[\frac{k-j}{k}\right]\hat{\rho}(j)$, where $\hat{\rho}(j) = j^{th}$ autocorrelation coefficient	$VR(k)=1$

1. Simulated critical values are based on 10,000 replications of a standard normal distribution.
2. Standard errors obtained using White's (1980) procedure.

Table 2: Power Results for Random Walk Tests; d = 1.1

T	k	LOMAC	LOMACH	JEG	COCH
32	2	7.1	7.6	7.3	7.5
	4	7.2	8.1	7.6	8.4
	8	6.3	6.6	7.6	7.7
	16	6.2	6.2	7.1	7.2
64	2	9.8	8.4	10.3	9.7
	4	11.5	11.4	10.2	11.3
	8	9.5	10.2	10.4	10.9
	16	10.9	11.2	9.0	12.2
	32	11.1	11.8	8.4	11.8
128	2	18.5	17.4	18.4	19.1
	4	22.2	23.3	20.0	23.2
	8	21.0	21.7	19.3	21.1
	16	19.6	19.8	15.8	23.0
	32	15.8	16.8	11.7	19.0
	64	14.7	14.7	9.4	16.5
256	2	35.9	37.1	40.3	37.2
	4	44.1	45.9	38.3	46.0
	8	43.4	44.4	36.0	45.5
	16	38.2	39.1	28.8	40.6
	32	29.1	30.6	22.0	33.3
	64	22.1	22.3	15.4	26.6
	128	19.3	19.1	12.1	22.1
512	2	63.5	63.0	68.8	62.8
	4	73.7	73.1	66.0	73.4
	8	74.2	74.1	59.2	75.0
	16	65.0	64.4	49.9	66.0
	32	53.1	53.7	33.8	55.7
	64	38.7	39.0	23.5	42.6
	128	26.3	26.3	16.0	29.6
	256	19.3	19.2	12.2	22.1
1024	2	89.8	90.2	92.3	90.9
	4	95.4	95.3	90.5	95.7
	8	94.6	94.5	84.0	94.7
	16	89.5	89.9	74.5	90.0
	32	79.0	79.5	57.1	80.5
	64	64.9	64.5	43.3	66.6
	128	48.5	47.6	31.4	51.3
	256	33.4	33.5	22.0	39.4
	512	28.4	28.5	19.1	27.4

Table 3: Power Results for Random Walk Tests; d = 0.9

T	k	LOMAC	LOMACH	JEG	COCH
32	2			5.3	
	4	6.5	6.2		6.5
	8				
64	2			8.4	
	4	9.1			9.4
	8		8.6		
128	2			17.0	
	4	21.6	19.9		20.8
	8				
256	2			33.5	
	4	38.5	40.9		41.9
	8				
512	2			57.7	
	4				
	8	68.5	66.4		66.1
1024	2			87.7	
	4				
	8	95.0	94.5		94.6

Reported power statistics are for optimal values of lag length k.

Table 4: Size Results for Random Walk Tests

T	K	LOMAC	LOMACH	JEG	COCH
32	2	11.0	4.0	5.3	11.4
	4	9.0	4.3	5.2	9.7
	8	6.8	4.1	4.9	7.5
	16	5.7	4.3	4.7	6.9
64	2	13.6	4.5	5.7	13.6
	4	10.0	4.4	5.3	10.0
	8	7.3	4.2	5.1	7.4
	16	5.8	4.4	4.9	6.4
	32	5.9	4.8	4.6	6.0
128	2	14.3	4.6	5.5	14.5
	4	11.3	4.8	5.4	11.6
	8	7.8	4.4	5.2	8.0
	16	6.3	4.3	4.9	6.6
	32	5.4	4.6	4.7	5.7
	64	5.0	5.0	5.1	5.4
256	2	16.1	5.0	5.7	16.4
	4	11.6	5.0	5.5	12.6
	8	8.4	5.0	5.1	8.6
	16	6.6	4.7	5.1	7.1
	32	5.7	4.8	5.1	5.9
	64	5.5	5.0	4.8	5.7
	128	5.3	4.9	5.1	5.4
512	2	15.9	4.9	5.1	15.4
	4	11.8	4.8	5.0	11.7
	8	8.6	4.6	4.9	8.2
	16	7.1	4.7	4.9	6.7
	32	5.8	4.7	4.8	5.6
	64	5.1	4.8	5.0	5.2
	128	5.1	4.9	4.7	5.2
	256	5.1	4.9	4.6	5.1
1024	2	16.5	5.0	5.0	16.8
	4	12.4	4.7	4.7	12.1
	8	9.4	4.8	5.0	8.9
	16	6.7	4.9	4.8	6.7
	32	6.0	4.6	4.9	5.7
	64	5.5	4.8	5.1	5.3
	128	5.1	4.7	5.3	5.4
	256	5.2	5.2	4.9	5.6
	512	4.8	4.8	4.6	5.2

Table 5: Multiple Horizon Random Walk Tests

Name	Test	Null Hypothesis
JLOMAC	$Z(k) = \max_{1 \le i \le m} \lvert Z(k_i) \rvert$	$Z(k) = 0$
JLOMACH	$Z^*(k) = \max_{1 \le i \le m} \lvert Z^*(k_i) \rvert$	$Z^*(k) = 0$
JJEG	$\overline{\beta}_k = \sum_{i=1}^{m} \beta_{ki} / m$	$\overline{\beta}_k=0$
JCOCH	$N\left(\hat{V}R_m - 1_m\right)' AA'\left(\hat{V}R_m - 1_m\right)$, where $A = 2\begin{pmatrix} (1-1/t_1) & \dots & \left(1-\frac{t_1-1}{t_1}\right) & 0 & \dots & 0 \\ \vdots & & & & & \\ (1-1/t_m) & \dots & \dots & \dots & \dots & \left(1-\frac{t_m-1}{t_m}\right) \end{pmatrix}$	$VR(k)=1$

Table 6: Power Results for Multiple Horizon Random Walk Tests; d = 1.1

T	JLOMAC	JLOMACH	JJEG	JCOCH
32	7.8	8.7	6.9	7.1
64	10.4	12.6	10.8	8.9
128	21.5	21.8	22.3	15.7
256	35.4	33.9	42.8	30.3
512	60.1	57.3	73.0	59.4
1024	85.6	84.3	94.2	87.5

Table 7

Sample variance ratios $M(k)+1$, Heteroscedasticity robust $Z^*(k)$ statistics and the marginal significance levels (p-values) associated with the $Z^*(k)$ statistics for UK stock returns data for the period 1925-1991.

	Measurement Horizon (Months)							
k	2	4	8	16	32	64	128	Joint
$M(k)+1$	1.102	1.107	1.141	1.231	1.201	1.101	1.105	
$Z^*(k)$	1.643	0.922	0.769	0.846	0.508	0.179	0.130	1.643
Marginal significance levels (p-values) based on Randomization								
	0.944	0.799	0.776	0.902	0.715	0.622	0.632	0.651
Marginal significance levels (p-values) based on Stratified Randomization								
	0.939	0.794	0.740	0.762	0.635	0.466	0.412	0.657
Marginal significance levels (p-values) based on Monte Carlo, homoscedastic increments								
	0.943	0.795	0.778	0.910	0.725	0.619	0.636	0.646
Marginal significance levels (p-values) based on Monte Carlo, heteroscedastic increments								
	0.942	0.791	0.776	0.912	0.724	0.620	0.638	0.636

Randomization and Stratified Randomization p-values are based on 1000 shuffles of the data. To preserve the pattern of heteroscedasticity present in the actual return series, we carry out a Stratified Randomisation of the total return data. We partition the sample into seven regimes or strate, where each regime displays a different level of return variance. To generate an aritificial return series, the return data are sampled without replacement from each regime. Monte Carlo p-values are based on 25,000 replications. Data for Monte Carlo experiments is drawn from a standard normal distribution and the sample size for each replication is 803. To allow for heteroscedasticity in the return generation process, the following model for stock returns is used: $R_t = \sigma_t \, \varepsilon_t$ where R_t is the one period return, ε_t is $iid\, N(0,1)$ and $\log(\sigma_t^2) = 1 + 0.1\left|\frac{R_{t-1}}{\sigma_{t-1}}\right| + 0.8\log\sigma_{t-1}^2 + 0.1\left(\frac{R_{t-1}}{\sigma_{t-1}}\right)$.

The joint test results presented in this table are the results obtained from applying the JLOMACH test (extension of the heteroscedasticity robust Z*(k) statistic for multiple horizons) to the data.

Marginal significance levels (p-values) below (above) 0.05 (0.95) indicate mean reversion (aversion) in stock prices.

Table 8

Sample variance ratios, $M(k)+1$, heteroscedasticity robust $Z^*(k)$ statistics and the marginal significance levels (p-values) associated with the $Z^*(k)$ statistics for UK stock returns data for the subperiods 1925-1946 and 1947-1991.

Subperiod 1925-1946

Measurement Horizon (Months)

k	2	4	8	16	32	64	128	Joint
$M(k)+1$	1.059	1.007	1.037	1.506	2.012	1.640	1.131	
$Z^*(k)$	0.475	0.030	0.100	0.924	1.276	0.564	0.081	1.276

Marginal significance levels (p-values) Based on Randomization

	0.655	0.523	0.537	0.801	0.860	0.759	0.648	0.383

Subperiod 1947-1991

Measurement Horizon (Months)

k	2	4	8	16	32	64	128	Joint
$M(k)+1$	1.113	1.133	1.165	1.155	0.974	0.816	0.870	
$Z^*(k)$	1.603	1.004	0.791	0.497	-0.057	0.286	-0.142	1.603

Marginal significance levels (p-values) based on Randomization

	0.950	0.832	0.766	0.657	0.493	0.430	0.549	0.613

Randomization p-values are based on 1000 shuffles of the data. The joint test results presented in this table are the results obtained from applying the JLOMACH test (extension of the heteroscedasticity robust Z*(k) statistic for multiple horizons) to the data. Marginal significance levels (p-values) below (above) 0.05 (0.95) indicate mean reversion (aversion) in stock prices.