By J. P. Botha

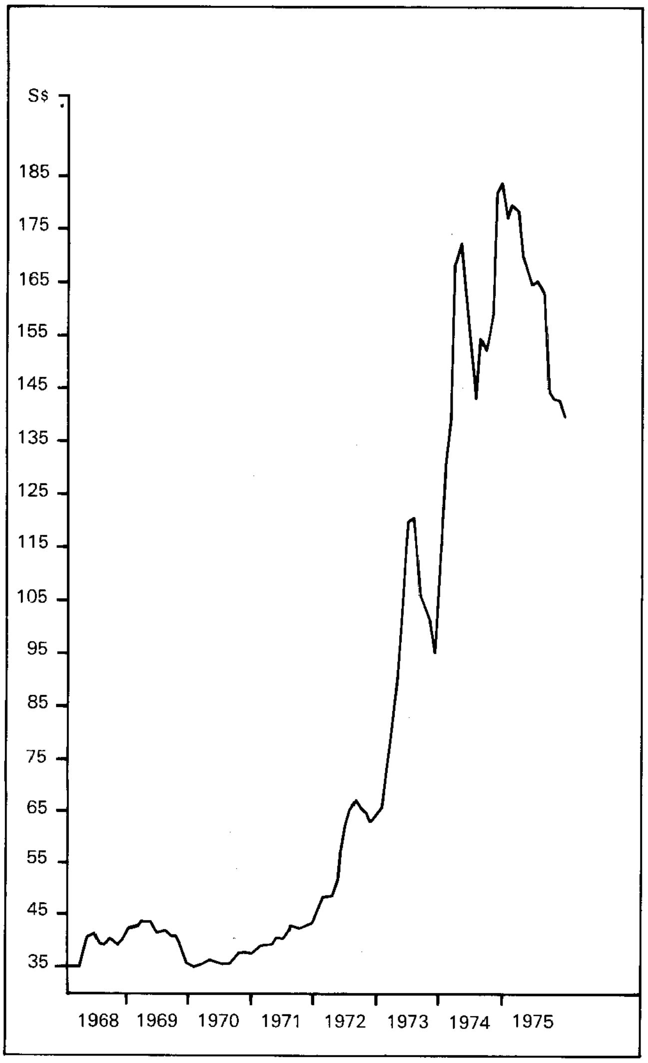
The random walk model and the behaviour of gold prices: a note

I INTRODUCTION

During recent years a considerable amount of evidence has been produced to demonstrate that in a highly competitive and organised market, price changes will be close to random.[1] Prices in such markets are said to be "weakly efficient", a phrase used to describe a certain kind of statistical behaviour in a price's time series. A price is weakly inefficient if the historical price series suggests information about this series in the future; weakly efficient if the price follows a random walk.

Thus, the underlying economic rationale for the random walk hypothesis is that in an efficient market, characterized by many well-informed, profit maximizing participants, competing actively with one another to predict future market values, a series of price changes has no memory – that one cannot predict future price changes on the basis of the history of price behaviour.

Figure 1: Gold price 1968–1975



Most of the tests of the random walk model have been performed on stock markets, but a model of this nature is of interest in the analysis of gold markets. In an efficient gold market, the price of gold can be expected to approximate its "intrinsic" value because of events that have occurred in the past and as a result of events which are expected to take place in the future. Essentially then, the random walk model implies that the gold price will fluctuate randomly around its intrinsic value. This intrinsic value may change across time. If the market is efficient then the price will rapidly adjust to wander about its new intrinsic value.

Since April 1968 (with the exception of a few months in 1970) the price of gold has been free to vary (Fig. 1) and up to September 1975, the average annual rate of appreciation was about 17%. We wish, then, to examine the nature of daily gold price changes in London over this sample period and to ask if the gold market is efficient in the technical sense of that phrase.

II THE MODEL

The random walk is a stochastic process in which each successive change in the variable is drawn independently from a probability distribution with mean zero. The variable Z_t then evolves according to:

$$Z_t - Z_{t-l} = U_t$$

or

$$Z_t = Z_{t-l} + U_t$$

where, for our purposes, Z_t is the price of gold at time t and where U_t is a random variable with mean zero and is drawn independently every period, thus making each successive step taken by Z random. A drift is built into the random walk model by adding a constant, a, each period

$$Z_t = Z_{t-l} + U_t + a$$

which means that on the average the process will tend to move in the direction given by the sign of the constant.

In order to examine the random walk model in the context of the gold market we have to consider the hypothesis that successive price changes are independent and that the price changes conform to some probability distribution.

* The author, who is a graduate student at the University of Pennsylvania, Philadelphia, gratefully acknowledges financial assistance from the H.S.R.C. The views expressed in this paper do not necessarily reflect those of the H.S.R.C. He also wishes to thank Mr J. F. Affleck-Graves and Dr D. Vining for valuable suggestions in the preparation of a revised version of this paper.

[1] P. H. Cootner (Ed.), *The Random Character of Stock Market Prices* (Cambridge, Mass.: MIT Press, 1964); E. F. Fama, "The Behaviour of Stock Market Prices", *Journal of Business,* vol. 38, 1965 pp. 34 – 105; J. F. Affleck-Graves and A. H. Money, "A Note on the Random Walk Model and South African Share Prices", *The South African Journal of Economics,* 43, Sept. 1975) pp. 382 – 388; E. F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance,* vol. 25 (May 1970) pp. 383 – 417.

III THE DATA

In March 1968 unprecedented demand by speculators and investors led to the immediate liquidation of the official Gold Pool. The Washington agreement by Central Bank members of the Gold Pool resulted in the segregation of monetary gold stocks and the re-introduction of a free market. From the end of 1969 to the end of 1973 the IMF put a floor under the price of gold by agreeing to make purchases of newly mined South African gold when the free market quotation was $35 or below for amounts necessary to meet the Republic's need for foreign exchange. But except for that agreement, which proved necessary only for a few months in the beginning of 1970, the gold markets have been free for the past eight years. Gold prices in London are "fixed" twice daily at 10h00 and 15h00, as from April 1, 1968, to reflect the balance of supply and demand on the books of five leading bullion houses at those hours.

The data[2] which will be used consist of daily prices at 15h00 and the time period runs from April 1, 1968 to September 26, 1975. The tests are performed on the data for each individual year from 1968 through 1975 as well as on the data covering the time period April 1968 to September 1975.

The actual tests in this paper are not performed on the daily prices themselves, but on the variable U_t defined as:

$U_t = \log_e Z_t - \log_e Z_{t-1} - V \quad t = 1,2,\ldots N$

where Z_t is the price of gold per ounce at day t over N sample observations, U_t is a random variable and V is a constant. It is assumed that $E(U_t) = o$ and $E(U_t; U_{t+k}) = o$ for lag = o if the random walk hypothesis is true.

The transformation to logarithms has been argued in the literature to be more appropriate because:[3]

(i) the change in the logarithm of price is the yield, with continuous compounding, from holding the share for that day;

(ii) the absolute magnitude of price changes is an increasing function of the price level of the share and by taking logarithms we neutralize most of the price level effect;

(iii) for relatively small price changes the change in the logarithm of price is close to the percentage price change.

IV THE EMPIRICAL RESULTS

Our approach relies primarily on statistical tools such as autocorrelation coefficients and analysis of runs of consecutive price changes of the same sign. In addition, we examine the distributional evidence concerning the probability distribution governing the price changes.

Autocorrelation

The autocorrelation coefficient is a measure of the relationship between the value of a random variable at time t and its value k periods earlier. An estimate of the k^{th} lag autocorrelation is [4]

$$r_k = C_k / C_o$$

$$\text{where } C_k = l/N \sum_{t=l}^{N-k} (U_t - \bar{U})(U_{t+k} - \bar{U}) \quad k = o, l, \ldots k$$

is the estimate of the autocovariance and $\bar{U}$ is the mean of the time series.

The autocorrelation coefficients were computed for lags of k = 1,2...20 days to determine if any correlation exists and are shown in Table 1 for lags of up to ten days. For the random walk hypothesis to hold, each correlation coefficient should not be significantly different from zero and there should be no pattern to the sign or the size of these coefficients.

Examination of these results show that the sample autocorrelation coefficients are relatively small in absolute values and that only 16 of the 180 computed autocorrelations were more than twice their computed standard error. Since a significance level of approximately 5% (2 standard errors) was used, it is reasonable to expect that 5% of the computed autocorrelations would be more than 2 standard errors from zero, under the null hypothesis. Thus one would expect about 9 of the computed autocorrelations to be "significant". However, 16 were observed which might indicate a very slight dependence.[5]

As far as the signs of the coefficients are concerned we find that of the ninety coefficients reported in Table 1, forty-six have negative signs and forty-four have positive signs. There is some evidence that for the yearly coefficients (but not for the whole sample period coefficients) the first four lags resulted in more negative

[2] The data used in this study were collected from Metal Bulletin, Richmond, Surrey: Simpson and Co. Ltd. 1968-1975.

[3] See e.g. Fama (1965) op. cit.

[4] G..E. P. Box and G. J. Jenkins, *Time Series Analysis: Forecasting and Control* (San Francisco: Holden - Day Inc., 1970).

[5] ($1/N^{\frac{1}{2}}$) supplies an approximate upperbound for the standard errors (Box and Jenkins, op. cit.) which implies that for a sample of almost 1 900 observations, that is the whole sample period, a coefficient as small as ,05 is more than twice its standard error so that statistically "significant" deviations from zero may not necessarily be a basis for rejecting the independence assumption.

Table 1
DAILY AUTOCORRELATION COEFFICIENTS FOR LAGS ONE TO TEN DAYS

	Lag (Days)									
Year	1	2	3	4	5	6	7	8	9	10
1968	–,062	0,029	–,054	–,015	–,002	0,005	0,100	0,039	–,096	0,031
1969	–,153*	–,053	0,134*	–,130*	0,068	0,118	–,080	0,108	0,104	0,026
1970	–,095	0,042	–,054	0,080	–,004	–0,060	–,136*	–,038	0,023	0,044
1971	0,010	0,021	–,021	–,074	0,066	0,030	–0,048	–,070	0,029	0,067
1972	–,029	–,157*	–,081	0,022	0,037	0,084	0,012	–,069	–,044	–,008
1973	0,133*	–,091	–,087	–,008	0,013	0,141*	0,049	–,074	–,035	–,028
1974	–,040	–,077	–,041	–,001	0,056	–,033	0,070	0,025	–,028	0,040
1975	–,063	–,145*	0,075	–,052	0,002	–,034	0,001	0,071	–,023	0,021
1968-75	0,032	–,072*	–,030	0,0	0,026	0,056*	0,054*	–,012	–,013	0,012

The star (*) indicates sample estimates more than twice their computed standard error.

terms than positive terms. However, since the coefficients for the 1968-1975 sample period do not show a marked preponderance of positive or negative signs we conclude that any "pattern" is only due to random fluctuations. Furthermore, since neighbouring estimates of the autocorrelations generally covary, the sample autocorrelations display a rather smooth cyclical tendency that is not present in the true autocorrelation function.

It is worth noting that the time span 1968-1975 exhibits a lack of autocorrelation greater than that for each individual year.

The coefficients are very small and close to zero for all the lags.

In sum, the evidence produced by the autocorrelation coefficients seem to indicate that dependence in successive price changes is very slight with no marked preponderance of negative or positive signs. This suggests that the assumption of zero autocorrelation is valid for the behaviour of gold prices over the lags and the time period considered.

Runs test

Very little is known about the distribution of the serial correlation coefficients when price changes do not follow a normal distribution and to test the randomness of the sequence of price changes we use a nonparametric method – the runs test. A run is defined as a period of time over which the change in prices is in the same direction, either upward or downward. The purpose of this test is to determine if the number of consecutive days of price changes in any direction conforms to that merely due to random fluctuations. This test is thus testing the same aspect of the random walk model (viz the independence of successive price changes) as the previous analysis of the autocorrelation coefficients.

Under the hypothesis of independence and for a large number of observations, N, the expected number of runs, R, of all signs can be computed as[6]

$$E(R) = (2n_1n_2/n_1 + n_2) + 1$$

where n_1 is the number of price changes in one direction and n_2 is the number of price changes in the other direction. The variance of R is

$$\text{var}(R) = \frac{2n_1n_2\,(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

and for large N the sampling distribution of R is approximately normal.

Table 2 shows the total expected and actual number of runs for each year and for the time span 1968-1975 with a one day differencing interval. The amount of dependence implied by the runs test is given by the difference between the expected and actual number of runs. This difference is expressed by the standardized normal variable Z, with mean zero and variance one. The column labelled (R–E(R))/E(R) gives the difference between the actual and expected number of runs as proportions of the expected numbers. Except perhaps for 1972 these percentage differences are quite small.

The hypothesis that the number of runs was realized by a random process is rejected only for 1972 where the actual number of runs is more than two standard errors less than the expected number of runs. In sum, the results of the runs analysis tend to substantiate the hypothesis of independence of successive price changes.

Table 2
RUNS ANALYSIS

Year	Expected R	Actual R	Z values	(R–E(R))/E(R)
1968	97	92	–0,73	–0,5
1969	124	136	1,14	,09
1970	127	114	–1,69	–,10
1971	124	118	–0,87	–,04
1972	123	104	–2,55	–,15
1973	126	118	–1,01	–,06
1974	127	125	–0,27	–,01
1975	89	101	1,76	,13
1968-1975	943	912	1,43	–,03

Distributional evidence

A last issue of the random walk model has centered on the nature of the distribution of price changes. If transactions are fairly uniformly spread over time and if the number of

[6] A. M. Mood and F. A. Graybill, *Introduction to the Theory of Statistics* (McGraw-Hill Co. Inc., New York, 1963). The runs test used allows for only two types of movement (i.e. up or down) and all zero changes are regarded as "ties" and are ignored.

Table 3: Unit normal and empirical relative frequencies

	0–0,5S	0,5–1,0S	1,0–1,5S	1,5–2,0S	
Unit normal	0,3830	0,2996	0,1838	0,0881	
Annual average	0,5702	0,2421	0,0853	0,0449	
1968-1975	0,6624	0,1629	0,0757	0,0429	
Fama	0,4667	0,2802	0,1378	0,0631	

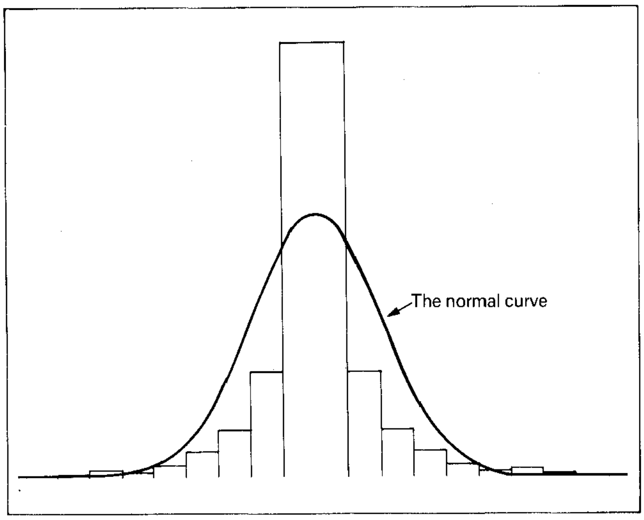
	2,0–2,5S	2,5–3,0S	3,0–4,0S	4,0–5,0S	>5,0S
Unit normal	0,0331	0,0097	0,002638	0,0000614	0,0000006
Annual average	0,0276	0,0096	0,010768	0,0039668	0,0055396
1968-1975	0,0211	0,0096	0,014288	0,0073416	0,0037037
Fama	0,0278	0,0130	0,008359	0,0018778	0,0011632

transactions each day is large then, appealing to the Central Limit Theorem, we would expect that the daily price changes will have a Gaussian distribution.

We performed the usual tests for normality such as goodness of fit and analysis of extreme areas. In the goodness of fit test our hypothesis of normality was rejected for every period considered.

Table 3 shows frequency distributions for changes in the price of gold. The empirical proportions of price changes between certain given standard deviations from the mean are reported in row three for the period 1968-1975, while row two gives the empirical average annual proportions of price changes over the sample period.

Figure 2: The normal curve and the empirical distribution of price changes 1968–75



To compare these findings with the expected proportions of price changes if the distribution was exactly normal, row one gives the proportions for the unit normal distribution. Row four gives Fama's[7] findings of the distribution of stock price changes on the New York Stock Exchange. If the number in row one is smaller than those in the other rows, it should be interpreted as an excess of relative frequency in the empirical distribution over what would be expected for the given interval if the distribution was normal. Similarly, if the number in row one is larger than those in the other rows, it should be interpreted as a deficiency of relative frequency in the empirical distribution within that given interval.

A consistent departure from normality is the excess of observations near the mean so that the empirical distributions are more peaked in the centre and have longer tails than the normal distribution. These findings are in the direction of the results obtained by Fama but with an even more marked degree of leptokurtosis suggesting that some non-normal distribution will provide a better description of daily price changes than the normal (Fig. 2). This aspect of the random walk model, while of interest, is not of great importance. The importance of the random walk model stems from the fact that future price changes cannot be predicted from the history of price changes alone, no matter what their distribution.

V CONCLUSION

The purpose of this paper has been to test empirically the random walk model of gold price behaviour. Our basic conclusion is that the random walk model offers a satisfactory explanation of the movement of daily price changes. The data show almost no tendency for dependence as is evidenced by examining sample autocorrelation functions – a finding which is in general substantiated by an examination of runs. The distributional evidence suggests that the distribution of daily price changes is not normal and that there is some (unspecified and perhaps unknown) non-normal distribution which is more suitable.

If successive price changes are dependent then the sequence of price changes prior to any given day is important in predicting the price change for that day. Thus the way to predict prices, under the assumption that history tends to repeat itself, is to develop a familiarity with past patterns of price behaviour in order to recognize situations of likely recurrence.

The empirical results indicate considerable support for the random walk model. The price reflects all the information available to participants in the market and all changes in the price are independent of any past history of the price. Except for a trend related to the desired rate of return, future prices could just as well be estimated by the flip of a coin (unless private information is available) as by any elaborate and detailed analysis of past prices.

[7] Fama (1965) op. cit. Table 1.