RANDOM WALKS, MARTINGALES AND THE OTC

ROBERT L. HAGERMAN AND RICHARD D. RICHMOND*

I. INTRODUCTION

SECURITY price movements have been characterized by the "random walk" hypothesis. Basically, this hypothesis states that changes in a security's price are independent over time. If this hypothesis is correct, it implies that no trading rule based on past prices will earn an economic profit. Because many investors do use trading rules, tests of the random walk hypothesis are of considerable practical, as well as academic, interest.

The independence implication of the "random walk" hypothesis has been tested in several studies of securities traded on organized markets.[1] The evidence overwhelmingly shows that security returns are independent over time. The purpose of this paper is to extend the investigation to securities traded over-the-counter.

The motivation for doing this is that the OTC market is considerably different from the organized markets, such as the N.Y.S.E., so the results of previous studies may not hold for this market. The OTC differs from organized exchanges since it is composed of geographically dispersed firms which act as principals in security transactions with investors by buying and selling from inventory. Since many investors use past prices to make investment decisions in this market, it is of interest to determine if there is empirical evidence that such strategies will be profitable.

To test this possibility, we examined the monthly returns of 253 securities for serial independence. Serial correlation coefficients were computed for each security and tested for significance assuming both normal and non-normal symmetric stable underlying distributions. In addition, a distribution-free runs test was used to test the returns for serial independence. The results indicate that price changes are serially independent which is consistent with the "random walk" hypothesis.

II. RANDOM WALKS AND MARTINGALES

The primary reason for the interest in the "random walk" hypothesis is its relation to the concept of an efficient market. In an efficient market, the current price of a security is an unbiased estimator of its intrinsic value which means that information, relevant to the value of the security, is reflected in the current price. A necessary condition for efficiency then is that the information contained in past security price movements be reflected in the current price of the security.

* Duke University and the University of Rochester, respectively. We are grateful to J. Long, D. Mayers and M. Rao and especially Michael C. Jensen, Eugene F. Fama and Richard Roll for guidance and to George Benston for the data.

1. Cootner [1] has compiled the major earlier studies. Fama [4] and Smidt [12] provide excellent summaries of the issues and the evidence.

The simplest formation of an empirically testable efficient market model is the "fair game" model which for a weak form of an efficient market can be expressed as:

$$E(P_{t+\tau}|P_t, P_{t-1}, P_{t-2}, \ldots, P_{t-n}) = E(P_{t+\tau}|P_t). \tag{1}$$

This is a weak form of the "fair game" model in the sense that the information set under consideration is limited to the set of past prices. The submartingale and random walk models frequently discussed in the literature are successive subsets of the fair game model.

The random walk model which assumes that successive returns are independent and identically distributed imposes the strongest conditions for tests of a weakly efficient market. In other words, the random walk model is a sufficient but not necessary condition for an efficient market. The tests which follow are "random walk" tests of the form found in the earlier literature but are not complete tests of the random walk hypothesis since they only examine returns for serial independence.[2]

III. The Data

The data are monthly bid prices of 253 securities from January 31, 1963 to December 31, 1967. These prices were obtained from the National Quotation Service's *National Monthly Stock Summary*. Capital changes and dividends for these securities were taken from Moody's *Dividend Record*. The securities in the sample were selected at random from those firms traded over-the-counter that had more than 500 shareholders and assets in excess of one million dollars. Initially 387 securities were selected but incomplete data reduced this number to 253. Potential errors in the data were identified as month-to-month price changes of greater than 15% and as bid prices less than 95% of the related asked price. The prices, dividends and capital changes associated with the potential errors were traced to independent sources such as *Barron's* and the data corrected if necessary. From the 14,927 observations, approximately 2300 potential errors were identified of which only 85 needed corrections.

IV. Tests of Serial Correlation

The serial correlation coefficient is a popular statistic for determining the amount of serial dependence. This statistic is defined as:

$$\rho = \frac{\text{Covariance } (U_t, U_{t-1})}{\sigma(U_t) \cdot \sigma(U_{t-1})} \tag{2}$$

which can be approximated by:

$$\rho \approx \frac{\text{Covariance } (U_t, U_{t-1})}{\text{Variance } (U_t)} \tag{2a}$$

where U_t is a log price relative.[3] A log price relative is:

2. See Fama [4] for a comprehensive discussion of the issues in this section.

3. The log price relative is used in this study because it approximates the one period return and eliminates the skewness in price distributions which result because prices are bounded at zero.

$$U_t = \ln\left(\frac{P_t + D_t}{P^*_{t-1}}\right) \tag{3}$$

where:

P_t is the price of the security at time t.
D_t is the dividend paid between $t - 1$ and t.
P^*_{t-1} is the price of the security at $t - 1$ adjusted for capital changes between $t - 1$ and t.

For large samples the estimate of the serial correlation coefficient is equal to the estimate of the slope coefficient in the regression model:

$$U_{j,t+1} = \alpha_j + \beta_j U_{j,t} + e_{j,t+1} \tag{4}$$

where $U_{j,t}$ is the log price relative of the j'th security at time t. In the model, β_j is the effect of the return from $t - 1$ to t on the return from t to $t + 1$. α_j is the average continuously compounded monthly return on security j if β_j is zero which is implied by serial independence.

Regressions in the form of equation (4) were computed for all 253 companies for the entire 60 months, the first 30 months and the second 30 months. A summary of the results is given in Table 1. A frequency distribution of the correlation coefficients is presented in Table 2.

TABLE 1
SUMMARY OF RESULTS FROM ORDINARY LEAST SQUARE REGRESSIONS

	Full Period	First Half Period	Second Half Period
Mean Absolute Value of Sample Serial Correlation Coefficient	0.1308	0.1735	0.1590
Mean Sample Serial Correlation Coefficient	−0.0762	−.0994	−.0737
Standard Error of Mean	.1483	.1893	.1885
Number of Sample Serial Correlation Coefficients Significant at 5%			
Positive	2	2	3
Negative	29	21	19
Total	31	23	22
Number of Sample Serial Correlation Coefficients Significant at 1%			
Positive	0	0	1
Negative	15	6	6
Total	15	6	7

For the entire period, 31 serial correlation coefficients, 12.3% of the total, were significant at the 5% level under normality assumptions. For the first 30 month subperiod and for the second 30 month subperiod, 23 and 22 coefficients were significant at the 5% level. At the 1% level of significance 15

TABLE 2
FREQUENCY DISTRIBUTION OF SERIAL CORRELATION COEFFICIENTS

Class	Full Period	First Half Period	Second Half Period
.40 to .50	0	2	2
.30 to .40	2	2	6
.20 to .30	6	13	10
.10 to .20	18	22	24
.00 to .10	52	38	44
−.10 to .00	69	51	57
−.20 to −.10	54	49	45
−.30 to −.20	31	40	38
−.40 to −.30	18	19	14
−.50 to −.40	3	13	10
−.60 to −.50	0	4	2
−.70 to −.60	0	0	1
	253	253	253

coefficients were significant in the full period sample, 6 in the first subperiod and 7 in the second subperiod. Each of these numbers of significant coefficients is different at the 5% level from the number expected under the null hypothesis, i.e., using the binomial distribution, the probability of occurrence of each of the above outcomes is less than .05 if the null hypothesis is true.

These results indicate the existence of serial dependence of price changes for securities traded OTC. However, even though the number of significant coefficients is statistically different from the expected number, this dependence is not stable. No security had significant dependence in both subperiods and less than half of the securities with significant dependence during the full period had significantly dependence in either subperiod.[4] This lack of stable dependence greatly increases the likelihood that the results are unreliable, so possible causes of spurious correlation and bias must be examined.

One possible cause of misleading results is bias in the coefficients due to errors in the variables. We may examine the errors in variables problem in the following manner. Take as the measured value of the log price relative,

$$\begin{aligned} U_t &= \ln(P^{*\prime}_t \cdot k'_t) - \ln(P^*_{t-1} \cdot k_{t-1}) \\ &= U^*_t + \ln k'_t - \ln k_{t-1} \\ &= U^*_t + C'_t - C_{t-1}, \end{aligned} \tag{5}$$

where $C_t = \ln k_t$, k_t is the error in measured price expressed as a multiplicative factor of the true price P^*_t, primes denote adjustments for dividends and P_{t-1} is adjusted for capital changes. U^*_t is the true value of the log price relative at time t. The serial correlation model is,

$$\begin{aligned} U^*_{t+1} &= \alpha + \beta U^*_t + \xi_{t+1} \\ \xi_{t+1} &\sim N(0, \sigma^2), \end{aligned} \tag{6}$$

4. Of the 31 securities with coefficients significant for the full period, only 8 and 15 had significant coefficients in the first and second subperiods, respectively.

which can be expressed in deviation form as,

$$u^*_{t+1} = \beta u^*_t + \xi_{t+1}. \tag{7}$$

The estimate of the serial correlation coefficient is

$$\hat{\beta} = \frac{\Sigma u_{t+1} u_t}{\Sigma u_t^2}, \tag{8}$$

but

$$u_{t+1} = u^*_{t+1} + c'_{t+1} - c_t$$

and

$$u_t = u^*_t + c'_t - c_{t-1}$$

so that

$$\hat{\beta} = \frac{\Sigma u_{t+1} u_t}{\Sigma u_t^2} = \frac{\Sigma (u^*_{t+1} + c'_{t+1} - c_t)(u^*_t + c'_t - c_{t-1})}{\Sigma (u^*_t + c'_t - c_{t-1})^2}. \tag{9}$$

Then, assuming that the errors are independent of the log price relatives and assuming

$$E(c_t) = E(c'_t) = 0 \qquad \text{for all t}$$

$$E(c_{t+j} c_{t+k}) = E(c_{t+j} c'_{t+k}) = E(c'_{t+j} c'_{t+k}) = 0 \qquad \text{for } j \neq k$$

and $$E(c_t^2) = \sigma_c^2 \approx E(c_t \cdot c'_t) \approx E(c'_t \cdot c'_t)$$

we have

$$\operatorname*{plim}_{n\to\infty} \hat{\beta} = \operatorname*{plim}_{n\to\infty} \frac{\Sigma (u^*_{t+1} \cdot u^*_t - c_t \cdot c'_t)}{\Sigma (u^*_t + c'^2_t + c_{t+1}{}^2)} \tag{10}$$

$$\approx \frac{\operatorname{cov}(u^*_{t+1} u^*_t) - \sigma_c^2}{\sigma_{u*}{}^2 + 2\sigma_c^2}.$$

This means that to the extent errors in the data exist, the serial correlation coefficient will be biased.

We may obtain an upper limit on the size of the induced coefficient due to errors in variables by assuming that $\operatorname{cov}(u^*_{t+1} u^*_t)$ and σ^2_{u*} are equal to zero. Then the expression in equation (10) assumes the value $-\frac{1}{2}$. Of course, the value of σ^2_{u*} will not be zero so that the value of the expression will be much closer to zero.[5] Unfortunately, little additional information about possible bias

5. An alternative formulation of $\hat{\beta}$ in terms of β, under the same assumptions, can be shown to yield

$$\operatorname*{plim}_{n\to\infty} \hat{\beta} \approx \beta \left(1 - \frac{2\sigma_c^2}{\sigma_u^2}\right) - \frac{\sigma_c^2}{\sigma_u^2}$$

which shows that there are two elements of bias due to errors. The first is a bias which pushes $\hat{\beta}$ toward zero as $2\sigma_c^2$ approaches σ_u^2 and then beyond zero, reversing the sign of $\hat{\beta}$, as $2\sigma_c^2$ exceeds σ_u^2. The second is a negative bias due to the term $-\frac{\sigma_c^2}{\sigma_u^2}$.

can be directly deduced with some measure of the size of σ^2_u, σ^2_{u*} and σ^2_e. Of course, these are not known.

The magnitude of the bias due to errors, if it exists, is likely to be small. The existence of a sizable bias is dependent upon having σ^2_e a large fraction of σ^2_{u*}. This requires that there be at least a few large errors or many small errors in the price series. Neither possibility is very likely. Large errors were eliminated by the initial data screening. Many small errors might be possible due to inaccurate reporting or printing in the National Quotation Service reports. Errors of this type are unlikely, however, because we used the median bid prices and most securities in the sample had multiple dealer listings.

A second apparent source of difficulty is the non-fixed regressor problem. The independent variable U_{jt} in equation (4) is a random variable rather than a fixed variable as in the usual linear regression model. This introduces a finite sample bias in the estimate of β so that, when the true value of β is zero, the expected value of $\hat{\beta}$ is $-\dfrac{1}{n-1}$ to the first order of approximation.[6] This effect, then, can be expected to introduce a small negative bias into the tabulated results.

A third statistical problem is the effect of the market factor. It is well documented that security prices move together so there will be a tendency for individual security returns to move with the "market" return.[7] The predominance of serial correlation coefficients with negative signs may be caused partly or wholly by this market factor effect.

The existence of a common element in a sample of security returns, of course, invalidates the assumption that the estimated serial correlation coefficients are independent. Therefore, the result using the binomial distribution to test whether the actual number of serial correlation coefficients is significantly different from the expected number loses the impact of the interpretation that there is significant statistical dependence in individual security returns. However, comparing the sample mean serial correlation coefficient in each period with its standard errors shows the mean always less than its standard error so the mean is not significantly different from zero. This uncertainty plus the interaction of the other statistical problems make interpretation of sample serial correlation coefficients extremely difficult.

A fourth, possibly significant, source of error is the assumption of normality in tests of significance of the regression coefficients. There is considerable evidence supporting the theory that the distribution underlying price changes is a symmetric stable distribution with characteristic exponent less than two rather than a normal distribution.[8] Although most of this evidence is derived from studies of organized exchanges, it is likely to be applicable to price changes for OTC securities.

6. For a discussion of the autoregressive non-fixed variable problem see Johnston [7, pp. 211-215]. The approximation of the bias follows from the derivation in a note by Kendall [8].

7. King [9] presents evidence on this point and Fama [3, pp. 73-74] provides a full discussion of this issue.

8. See for example Fama [3], and Roll [11].

The family of symmetric stable distributions is defined by the log characteristic function

$$\ln \phi_x(k) = \ln\left[\int_{\infty}^{\infty} e^{ikx} dF(x)\right] = i\delta k - \gamma|k|^{\alpha}, \tag{11}$$

where k is a real number, and by the three parameters: α, the characteristic exponent, δ, the location parameter, and γ, the dispersion parameter. Symmetric stable distributions have two important properties: (1) they exhibit stability or invariance under addition and (2) they are the limiting distributions for sums of independent, identically distributed random variables so that the class of symmetric stable distributions generalize the central limit theorem. If α, the characteristic exponent, equals two, the distribution is normal; if α is less than two the second and higher moments do not exist.[9]

The behavior of the F statistic, therefore, is not defined when the distribution has a characteristic exponent less than two, so significance tests based on this statistic would be inappropriate. Thus a different statistic is required to provide evidence on possible serial dependence.

The approach taken was to apply a significance test for the ordinary least squares regression results assuming the underlying distribution is symmetric stable with characteristic exponent, α, less than two.[10] Since the form of the density function for symmetric stable variables is not known except in special cases (Cauchy distribution, $\alpha = 1$, and normal distribution, $\alpha = 2$), there is no established body of statistical procedures for hypothesis testing when the distribution is of this form. However, Fama and Roll have constructed a table of cumulative distribution functions for standardized symmetric stable distributions based on approximate densities. Since the ordinary least squares estimate of the correlations coefficient is a linear function of log price relatives, a standardized variable for the estimate of the coefficient can be constructed and compared with the Fama-Roll table in order to make probability statements or tests of significance about the estimates.

The value of the characteristic exponent, α, was not known in this case and the data were insufficient to construct estimates with any degree of validity. Therefore, the test was conducted with assumed values of α from 1.1 through 2.0 in increments of 0.1.

The test was constructed as follows: From the ordinary least squares regression we have the estimate of the serial correlation coefficient,

$$\begin{aligned} \hat{\beta} &= \frac{\Sigma u_t \cdot u_{t+1}}{\Sigma u_t^2} \qquad (12) \\ &= \frac{\Sigma u_t(\beta u_t + \epsilon_{t+1})}{\Sigma u_t^2} \\ &= \beta + \frac{\Sigma u_t \epsilon_{t+1}}{\Sigma u_t^2}. \end{aligned}$$

9. A detailed treatment may be found in Feller [6], Fama [2] or Mandelbrot [10].

10. The ideas and tests developed here are either taken directly from or are extensions of the work of Fama and Roll [5].

If we assume that the ε_t's are independent, identically distributed with log characteristic function,

$$\ln \phi_\epsilon(k) = \ln\left[\int_{-\infty}^{\infty} e^{ik\epsilon} dF(\epsilon)\right]$$
$$= i \cdot \delta(\epsilon) \cdot k - \gamma(\epsilon) \cdot |k|^\alpha, \tag{13}$$

where $i = \sqrt{-1}$, k is a real number, $F(\varepsilon)$ is the cumulative distribution function, α is the characteristic exponent, δ is the location parameter and c is the scale parameter ($\gamma = c^\alpha$), and with $\delta(\varepsilon) = o$, then $\hat{\beta}$ is symmetric stable with log characteristic function

$$\ln \phi_{\hat{\beta}}(k) = i \cdot \delta(\hat{\beta}) \cdot k - \gamma(\hat{\beta}) \cdot |k|^\alpha \tag{14}$$

where

$$\delta(\hat{\beta}) = \delta\left[\beta + \frac{\Sigma u_t \cdot \epsilon_{t+1}}{\Sigma u_t^2}\right] \tag{15}$$
$$= \beta + \frac{\Sigma u_t \cdot \delta(\epsilon_{t+1})}{\Sigma u_t^2}$$
$$= \beta^{11}$$

and

$$\gamma(\hat{\beta}) = \gamma\left[\beta + \frac{\Sigma u_t \cdot \epsilon_{t+1}}{\Sigma u_t^2}\right] \tag{16}$$
$$= \gamma(\epsilon) \cdot \Sigma\left|\frac{u_t}{\Sigma u_t^2}\right|^\alpha$$
$$= [c(\epsilon)]^\alpha \cdot \Sigma\left|\frac{u_t}{\Sigma u_t^2}\right|^\alpha .^{12}$$

For our test, the null hypothesis is $\beta = 0$ and the alternate hypothesis is $\beta \neq 0$. The estimate of the serial correlation coefficient, $\hat{\beta}$, is provided by the ordinary least squares regression analysis. The estimate of the dispersion parameter is

$$\hat{c}(\hat{\beta}) = [\hat{\gamma}(\hat{\beta})]^{1/\alpha}$$
$$= \left[\hat{\gamma}(\epsilon) \cdot \Sigma\left|\frac{u_t}{\Sigma u_t^2}\right|^\alpha\right]^{1/\alpha}$$
$$= \hat{c}(\epsilon) \cdot \frac{[\Sigma|u_t|^\alpha]^{1/\alpha}}{\Sigma u_t^2}.$$

$c(\varepsilon)$ is estimated from the measured residuals as

$$\hat{c}(\epsilon) = \frac{1}{1.654}[\hat{\epsilon}(.72) - \hat{\epsilon}(.28)] \tag{18}$$

11. See Footnote 12.

12. This is correct only if u_t is a fixed variable. The nonfixed regressor problem, to be discussed later, makes the terms rather complex.

where $\hat{\varepsilon}(.72)$ and $\hat{\varepsilon}(.28)$ represent the .72 and .28 fractile measured residuals. This method for estimating $c(\hat{\varepsilon})$ was suggested by Fama and Roll [5] since a range estimate avoids the second moment and the specified fractiles minimize the potential bias in the estimates caused by the unknown value of α.

From the estimates of β and $c(\beta)$ we have a standardized variable,

$$\mathcal{Z} = \frac{\hat{\beta} - \beta}{\hat{c}(\hat{\beta})} \left\{ = \frac{\hat{\beta}}{\hat{c}(\hat{\beta})} \text{ under the null hypothesis, } \beta = 0 \right\} \tag{19}$$

for each assumed value of α. The test applied to the standardized variable, $\mathcal{Z} = \dfrac{\hat{\beta}}{\hat{c}(\hat{\beta})}$ is entirely analogous to the test of significance when a normal distribution is assumed. The computed values of $\mathcal{Z}$ were compared to the values of $\mathcal{Z}$ for the appropriate α value at .975 of the cumulative distribution functions as tabulated by Fama and Roll for a two-tailed test of significance at the 5% level. Since we have an estimate of the dispersion parameter, we used the Fama-Roll tabulated distribution on the basis of a large sample $(n = 59)$ as the cumulative Normal distribution would be used for the t-distribution with a sample of this size.

The number of coefficients, total, positive and negative, significant at the 5% level for each assumed value of α, is given in Table 3. The table shows, as expected, that the number of significant coefficients declined with α. Approximately 5% of the 253 coefficients are significant at the 5% level when $\alpha = 1.6$. The results indicate that, if α is in the range of 1.7 to 1.9, the range indicated by the work of Fama [3] for listed securities, then the data provide weak support for the existence of dependence. If α is less than 1.7, then the existence of dependence is ruled out.

The results when α is assumed to be 2.0 are quite interesting. The number of coefficients significant at the 5% level is 61, while under the assumption of normality, 31 coefficients are significant at 5%. When the true value of α is two, the number significant under the two approaches should be the same since $\gamma = c^2 = \dfrac{\sigma^2}{2}$. When α is actually less than two, then $c^2 < \dfrac{\sigma^2}{2}$ since the sample variance gives more weight to the tails of the distribution which reflect greater dispersion. Therefore, when $\alpha < 2.0$ more coefficients will be found significant by the symmetric stable test for α assumed equal to 2.0 than under the F test which assumes normality. This result, then, is consistent with the hypothesis that the underlying distribution is non-normal.

As mentioned above, however, these results are based on the incorrect assumption that the lagged independent variable U_t is fixed. When the assumption of fixed regressors is relaxed, the esimates are no longer simple linear combinations of U_t. Since we have no estimators of the location and scale parameters in this case other than approximations derived from the fixed regressor case, it is not possible to determine precise estimates nor to evaluate those obtained assuming a fixed independent variable.

The results based on serial correlation coefficients to determine independence are mixed. There is some evidence that the security returns in the OTC are

TABLE 3
NUMBER AND PERCENTAGE OF COEFFICIENTS SIGNIFICANT AT 5%
UNDER THE ASSUMPTION OF A SYMMETRIC STABLE DISTRIBUTION

Characteristic Exponent	Number Significant at 5%		
	Total	Positive	Negative
1.1	0	0	0
1.2	0	0	0
1.3	1	0	1
1.4	1	0	1
1.5	2	0	2
1.6	12	2	10
1.7	18	3	15
1.8	27*	5	22*
1.9	42*	8	34*
2.0	61*	11	50*

Characteristic Exponent	Percent Significant at 5%		
	Total	Positive	Negative
1.1	0.0	0.0	0.0
1.2	0.0	0.0	0.0
1.3	0.4	0.0	0.4
1.4	0.4	0.0	0.4
1.5	0.8	0.0	0.8
1.6	4.7	0.8	3.9
1.7	7.1	1.2	5.9
1.8	10.7*	2.0	8.7*
1.9	16.6*	3.2	13.4*
2.0	24.1*	4.3	19.8*

* Statistically significantly larger than expected at 5%.

serially dependent but this result depends on the assumption that the returns are contemporaneously independent and that α, the characteristic exponent, is greater than 1.7. These results are also subject to bias due to errors in the variables and to the lack of a fixed regressor. The existence and interaction of these factors make the tests based on the serial correlation coefficient suspect.

V. DISTRIBUTION FREE TESTS OF SERIAL INDEPENDENCE

In order to provide more evidence about the movement of security prices in the over-the-counter market non-parametric or distribution free statistics were used to eliminate any assumptions concerning the underlying distribution of the sample. To examine the independence of successive price changes, we employed runs tests. Each price change was classified as a member of one of three groups: +, if the price change is positive; —, if negative; and 0, if no change occurs. The runs tests take the observed proportion of positive, negative, and zero changes as given and then determine whether there is any significance to the observed ordering of these changes, i.e., whether there is a pattern deviating from a random ordering.

A run in a time series over a trinary population is an unbroken series of elements of one type which begins and ends with an element of a different type or with the beginning or ending of the series. For example, in the series $++---+$, there are three runs, a run of pluses of length two, a run of minuses of length three and a plus run of length one.

Runs by company were examined in the following manner. Let n_{1j} be the number of positive price changes, n_{2j} be the number of zero price changes, n_{3j} be the number of negative price changes. $n_j = \sum_{i=1}^{3} n_{ij}$, the total number of price changes for company j. Then $p_{ij} = \frac{n_{ij}}{n_j}$, $i = 1, 2, 3$, can be taken as the sample proportions for company j. Using these proportions, and assuming independence of price changes, the number of total runs expected were calculated for each security for comparison with the actual number.

Given the assumption of independence, the distribution of the total number of runs is approximately normal[13] with mean,

$$\mu_j = \frac{n_j(n_j + 1) - \sum_i n_{ij}^2}{n_j}, \tag{20}$$

and variance,

$$\sigma_j^2 = \frac{\sum_i n_{ij}^2[\sum_i n_{ij}^2 + n(n + 1)] - 2n_j \sum_i n_{ij}^3 - n_j^3}{n_j^2(n - 1)}. \tag{21}$$

The unit normal deviate, $Z = \frac{R_j + \frac{1}{2} - \mu_j}{\sigma_j}$, where R_j is the actual total number of runs for company j, was calculated for each company for the full period and for each half period. For the full period and first and second half periods, respectively, the number of Z values significant at 5% were 19, 17 and 16 and the number significant at 1% were 2, 2, and 2. None of these values is statistically different from expected at the 5% level.

This evidence supports the random walk hypothesis. It is possible, however, that the securities exhibit no statistically significant dependence when examined individually, but the sample as a whole does, i.e., the distribution is skewed or has a non-zero mean. To test this possibility the distribution of the deviates $Z = \frac{R_j - \mu_j}{\sigma_j}$, without the continuity correction, was compared with the unit normal distribution by means of the Kolmogorov-Smirnov Test. Poorness of fit implies the sample has statistically significant dependence. The test showed that for the full period and the two subperiods, the null hypothesis could not be rejected at the 5% level of significance. The distribution of the Z_j's is shown in Table 4.

The runs tests on individual securities and the Kolmogorov-Smirnov Test unequivocally show that the "random walk" hypothesis cannot be rejected. This evidence is particularly strong since it is free of distributional assumptions.

13. Wallis and Roberts [13, p. 571].

TABLE 4
FREQUENCY DISTRIBUTION OF Z VALUES FOR TOTAL NUMBER OF RUNS

Interval	Full Period	First Half Period	Second Half Period
3.0 to 3.5	0	0	0
2.5 to 3.0	1	2	1
2.0 to 2.5	5	4	3
1.5 to 2.0	12	8	14
1.0 to 1.5	27	30	24
0.5 to 1.0	49	36	51
0.0 to 0.5	38	50	43
−0.5 to 0.0	33	46	45
−1.0 to −0.5	50	36	29
−1.5 to −1.0	22	20	23
−2.0 to −1.5	7	16	8
−2.5 to −2.0	7	3	10
−3.0 to −2.5	1	3	2
−3.5 to −3.0	1	0	0
Total	253	253	253

VI. Summary and Conclusion

Based on our random sample, the evidence indicates that monthly returns of stocks traded over-the-counter are serially independent. This conclusion is supported by the results of our distribution free tests while the results of the serial correlation tests must be discounted because of the possible bias involved in the estimation procedure.

The serial independence of over-the-counter security returns implies that mechanical trading rules based on linear dependencies will not be able to earn extra-normal profits. This result then is consistent with the hypothesis that the over-the-counter market is a "weakly" efficient market.

REFERENCES

1. Paul Cootner, (ed.), *The Random Character of Stock Market Prices,* (Cambridge: M.I.T., 1964).
2. Eugene F. Fama. "Mandelbrot and the Stable Paretian Hypothesis," *Journal of Business,* 36 (October, 1963), 420-429. Also reprinted in [1].
3. ————. "The Behavior of Stock Market Prices," *Journal of Business,* 38 (January, 1965), 34-105.
4. ————. "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance,* 25 (May, 1970), 383-417.
5. Eugene F. Fama and Richard Roll, "Some Properties of Symmetric Stable Distributions," *Journal of the American Statistical Association,* 63 (September, 1968), 817-36.
6. William Feller. *An Introduction to Probability Theory and Its Applications,* Vol. II (New York: John Wiley and Sons, Inc., 1966).
7. J. Johnston. *Econometric Methods,* (New York: McGraw Hill Book Co., 1963).
8. M. G. Kendall. "Note on Bias in the Estimation of Auto-correlation," *Biometrika,* 41 (December, 1954). 403-404.
9. Benjamin F. King. "Market and Industry Factors in Stock Price Behavior," *Journal of Business,* 39 (Special Supplement, January, 1966), 139-90.
10. Benoit Mandelbrot. "The Variation of Certain Speculative Prices," *Journal of Business,* 36 (October, 1963), 394-419.

11. Richard Roll. "The Efficient Market Model Applied to U. S. Treasury Bill Rates," Unpublished Ph.D. thesis, Graduate School of Business, University of Chicago, 1968.
12. Seymour Smidt. "A New Look at the Random Walk Hypothesis," *Journal of Financial and Quantitative Analysis,* 3 (September, 1968), 235-262.
13. W. Allen Wallis and Harry V. Roberts. *Statistics: A New Approach* (Glencoe, Illinois: Free Press, 1956).
14. John Wise. "Linear Estimators for Linear Regression Systems Having Infinite Variances," (Unpublished paper presented at the Berkeley-Stanford Mathematical Economics Seminar, October, 1963.)