Journal of Financial Economics 3 (1976) 379–402. © North-Holland Publishing Company

CAPITAL MARKET SEASONALITY: THE CASE OF STOCK RETURNS

Michael S. ROZEFF and William R. KINNEY, Jr.

The University of Iowa, Iowa City, IA 52242, U.S.A.

Received January 1976, revised version received March 1976

In this paper we present evidence on the existence of seasonality in monthly rates of return on the New York Stock Exchange from 1904–1974. With the exception of the 1929–1940 period, there are statistically significant differences in mean returns among months due primarily to large January returns. Dispersion measures reveal no consistent seasonal patterns and the characteristic exponent seems invariant among months. We also explore possible implications of the observed seasonality for the capital asset pricing model and other research.

1. Introduction

Although the existence of seasonality in stock returns has important implications for capital market theory, capital market efficiency and the nature of the distribution of stock returns, the subject has only recently received much attention in the literature [see Bonin and Moses (1974) and Officer (1975)]. These studies, which find seasonality, conflict with earlier studies which do not [see Granger and Morgenstern (1963, 1970)]. Due to the relative lack of evidence and the use of various techniques to measure seasonality, there is no consensus of opinion as to even the existence of seasonality. The result has been little or no testing of hypotheses about possible causes of seasonality and no consideration of the implications of seasonality for the finance models which ignore its presence.

In this paper we present evidence on the existence of seasonality and its implications for other research. We concentrate on the question of existence. Possible explanations for the presence of seasonality are left for future research.

After the development of models in section 2 and a review of previous studies in section 3, section 4 presents evidence for the existence of seasonality in monthly stock returns on the New York Stock Exchange for various periods since 1904. The autocorrelation functions of the time series indicate *no* consistent seasonal pattern. However, using parametric and non-parametric tests, we do generally find statistically significant differences in measures of central tendency by month of the year due primarily to consistently high rates of return in January. Differences in dispersion measures among months are significant for some periods but not others, giving no consistent pattern over the various time periods. The characteristic exponents of the monthly returns seem relatively invariant among months.

Section 5 examines some implications of stock market seasonality for finance models. Of particular concern are the efficient market model and the possible effects of seasonality on the computation of disequilibrium rates of return and betas. We find, for example, that seasonality in stock returns seems to have had a major impact on Fama and Macbeth's (1973) estimates of slope and intercept in the two-parameter capital asset pricing model, since time series of these estimates (like stock returns) have pronounced distributional differences when examined by month.

2. Statistical models of stock returns

A wealth of evidence [see Cootner (1964) and Fama (1970)] indicates that stock price behavior is well-described by a multiplicative random walk model (tildes indicate random variables),

$$(\tilde{P}_{jt}+\tilde{D}_{jt}) = P_{j,t-1}\exp(\tilde{e}_{jt}+\mu_{jt}), \tag{1}$$

where $\tilde{P}_{jt}$ is the price of stock j at time t, $\tilde{D}_{jt}$ is the dividend from $t-1$ to t, $\tilde{e}_{jt}$ is a random variable with zero mean, independently and identically distributed through time, and μ_{jt} is the stock's expected rate of return, conditional on the information set available at $t-1$, (I_{t-1}). With the additional assumption that $\mu_{jt} = \mu_j$, a time-invariant constant, and rearranging, we obtain

$$\ln\left(\frac{\tilde{P}_{jt}+\tilde{D}_{jt}}{P_{j,t-1}}\right) \equiv \ln(1+\tilde{r}_{jt}) = \mu_j+\tilde{e}_{jt}. \tag{2}$$

For $\tilde{r}_{jt} \leqq 0.15$, $\ln(1+\tilde{r}_{jt}) \approx \tilde{r}_{jt}$, so that approximately,

$$\tilde{r}_{jt} \approx \mu_j+\tilde{e}_{jt}, \tag{3}$$

and $\mu_j = E(\tilde{r}_{jt}|I_{t-1})$. The assumed time-invariant mean and the assumptions of (1) on $\tilde{e}_{jt}$ imply that autocovariances of $\tilde{e}_{jt}$ and $\tilde{r}_{jt}$ are zero, a property displayed by stock returns [see Cootner (1964)].[1]

If N individual stocks are aggregated into an arithmetic index in which each stock has an equal weight, $X_{jt} = 1/N$, we obtain

$$\sum_{j=1}^{N} X_{jt}\tilde{r}_{jt} = \sum_{j=1}^{N} X_{jt}\mu_{jt}+ \sum_{j=1}^{N} X_{jt}\tilde{e}_{jt}, \tag{4}$$

or

$$\tilde{R}_t = \mu+\tilde{e}_t, \tag{5}$$

[1] If $E(\tilde{r}_{jt}|I_{t-1})$ is not time stationary, then the variable $\tilde{z}_{jt} = \tilde{r}_{jt}-E(\tilde{r}_{jt}|I_{t-1}) = \tilde{e}_{jt}$ will have an unconditional expectation of zero *so long as* the assumptions of (1) hold with respect to $\tilde{e}_{jt}$. In this case, the variable $\tilde{z}_{jt}$ will display zero autocovariances, but $\tilde{r}_{jt}$ will not necessarily have zero autocovariances. See Fama (1970) for a discussion of this point.

where each weighted sum is denoted by its market equivalent. The multiplicative random walk is therefore applicable in principle to stock indexes as well as individual stocks.[2] Officer's (1975, p. 35) analysis shows that seasonality is more likely to be detected in an index of shares than in individual shares. For this reason, we restrict the tests in this paper to a market index.

In models (3) and (5), the random variables $\tilde{e}$ and $\tilde{r}$ seem well-described by a symmetric distribution lacking finite second and higher moments and belonging to the Stable Paretian family of distributions with characteristic exponent less than two.[3] Due to controversy over the exact nature of the distribution, we employ both non-parametric tests which assume only that the distributions are continuous and parametric tests which assume normal distributions with finite moments. For the most part, we use terms applicable to Stable distributions and refer to *location* and *scale* parameters rather than the more common terms, mean and variance.

Departures from model (5) can occur in many ways. A simple seasonal alternative to (5) of some interest assumes that $\tilde{e}_t$'s are independently distributed random variables whose distributions differ only in location parameter by season or month. Letting subscript m denote the month of the year this model is

$$\tilde{R}_{tm} = \mu + \tilde{e}_{tm}. \tag{6}$$

Letting $E(\tilde{e}_{tm}) = \lambda_m$, this can be written

$$\tilde{R}_{tm} = \mu + \lambda_m + \tilde{e}_t, \tag{7}$$

where $\tilde{e}_t$ is again independent and identically distributed with mean zero. In the modified model, expected rates of return depend upon the month or season of the year and autocovariances of rates of return are non-zero.[4] Assuming that model (5) is the true model we test the null hypothesis that expected rates of return conditional on month of the year are equal, that is,

$$H_0 : E(\tilde{R}_1) = E(\tilde{R}_2) = \ldots = E(\tilde{R}_{12}) = \mu. \tag{8}$$

We also examine the autocorrelation function of stock returns to verify the implication of (5) that serial covariances are zero.

[2]Due, however, to the fact that rates of return on many securities may not be measured simultaneously at time t, it is usually observed that indexes show larger first-order autocorrelations than do the individual securities which compose them. This spurious effect causes no difficulty in the tests undertaken below.

[3]For evidence that stock return distributions are Stable Paretian, see, for example, Mandelbrot (1963, 1967), Fama (1965), Fama and Roll (1968, 1971). Some other viewpoints may be found in Press (1967), Teichmoller (1971), Officer (1972), Barnea and Downes (1973), Blattberg and Gonedes (1974), Hsu, Miller and Wichern (1974).

[4]Officer (1975, p. 31) presents model (6) and discusses it.

The error terms, $\tilde{e}_{tm}$, in (6) may differ in scale as well as location and differences in characteristic exponent by month are also possible. We therefore test the following hypothesis:

$$H_0 : d_1 = d_2 = \ldots = d_{12} = d, \tag{9}$$

that dispersion parameters, d_i, of stock return distributions do not differ by month. We also examine the characteristic exponent, α, of each return distribution by month.

3. Previous studies

Although statistical tests of location (central tendency) and scale (dispersion) parameters of stock return distributions by month are absent from the literature, previous studies do provide evidence on seasonality by applying (I) spectral analysis [Granger and Morgenstern (1963, 1970)], (II) Bureau of the Census X-11 procedures [Shiskin (1967) and Bonin and Moses (1974)], and (III) the time series methods of Box and Jenkins [Officer (1975)].[5] Other studies provide further evidence but lack statistical tests [Wachtel (1942) and Zinbarg (1964)].

Using aggregate monthly price data from 1875–1956, Granger and Morgenstern (1970, p. 130) concluded that spectral analysis 'gave no evidence of a seasonal (12-month) peak in the spectra although small peaks corresponding to seasonal harmonics (4, 3, 2.4 months) were quite frequently observed'. This conclusion is in basic agreement with the evidence from the autocorrelation function of returns presented below especially in view of (i) the fact that Granger and Morgenstern first removed a trend in price with moving averages, a procedure which makes implicit assumptions about the stochastic process and can lead to non-comparability with the autocorrelation function obtained from data on returns; and (ii) the difficulties in estimating a spectrum and establishing significance tests for individual peaks.[6]

Bonin and Moses look for seasonality in the 30 individual Dow–Jones Industrial stocks using monthly price data adjusted for capital changes over the period 1962–1971. Using the Census X-11 program, they remove price trends via moving average procedures and adjustment of extreme values. Due to these initial procedures, the analysis of variance tests performed on the residual variation must be interpreted with caution. Aware of this, the authors apply several other criteria – comparisons with other time series and tests on a

[5]See Box and Jenkins (1970). Note that spectral analysis and Box–Jenkins methods should reach the same conclusions since both employ the autocorrelation function as a basic descriptive device. The Census X-11 procedures might reach different conclusions since (unlike Box–Jenkins procedures) they impose strong a priori assumptions about the stochastic process generating the time series.

[6]Their time period also includes, as we see below, a sub-period during which the dispersion of the generating process noticeably increased, 1929–1940.

holdout period – before accepting seasonality. They conclude that 7 of the stocks display significant and persistent seasonal patterns.

Officer studies aggregate Australian stock returns over 1958–1970 and develops a mixed autoregressive and moving average linear stochastic model which includes seasonal elements. He then shows that forecast errors using the seasonal model are lower in a holdout period than forecast errors using a simple random walk model. The predictive test is a helpful complement to tests of significance for individual autocorrelations since these are based on the assumption of normal distributions, not strictly true for stock returns. The correspondence of the prediction test and the correlation tests may indicate that it is useful to examine sample statistics on the usual normal distribution assumptions.

In contrast with Granger and Morgenstern, Officer finds a 9-month, 6-month and lesser 12-month seasonal in the autocorrelation function. Below we briefly compare the autocorrelation function for the United States with Officer's Australian and Granger and Morgenstern's U.S. results; however, our main concern is the direct examination of stock return distributions by month.

In other less rigorous studies of seasonality, Zinbarg notes the tendency of advances to outnumber declines in the months of January and July from 1918–1962 and, similarly, Wachtel finds a marked tendency for January price advances in the overlapping period, 1927–1942. In addition, Wachtel formulates an explanation based on year-end tax selling.

4. Statistical tests

4.1. The data

We examine aggregate rates of return on the New York Stock Exchange for the period January, 1904 through December, 1974. Rates of return from 1904–1909 are computed from the aggregate Cowles Commission (1938) price indexes. From 1910–1925 they are derived from price relatives of Standard & Poor's aggregate index, omitting 1914 during which the Exchange was closed for four months. From 1926–1974 the data are equally-weighted arithmetic rates of return inclusive of dividends and adjusted for capital changes for all common stocks on the New York Stock Exchange, computed from Standard & Poor's 1975 version of the CRSP file.[7] For all statistical tests reported below, we use the natural logarithm of 1 plus the rate of return $[\ln(1+R_{tm})]$.

Visual inspection of the monthly rates of return since 1904 shows that returns seem to have been generated by a stochastic process that is mean stationary (see fig. 1). The mean rate of return of the (unlogged) time series is 0.009 per month. However, the dispersion increases noticeably during the Great Depression,

[7]All tests were also carried out on Fisher's (1966) Investment Performance Index substituted for the years 1926–70 with essentially the same results.

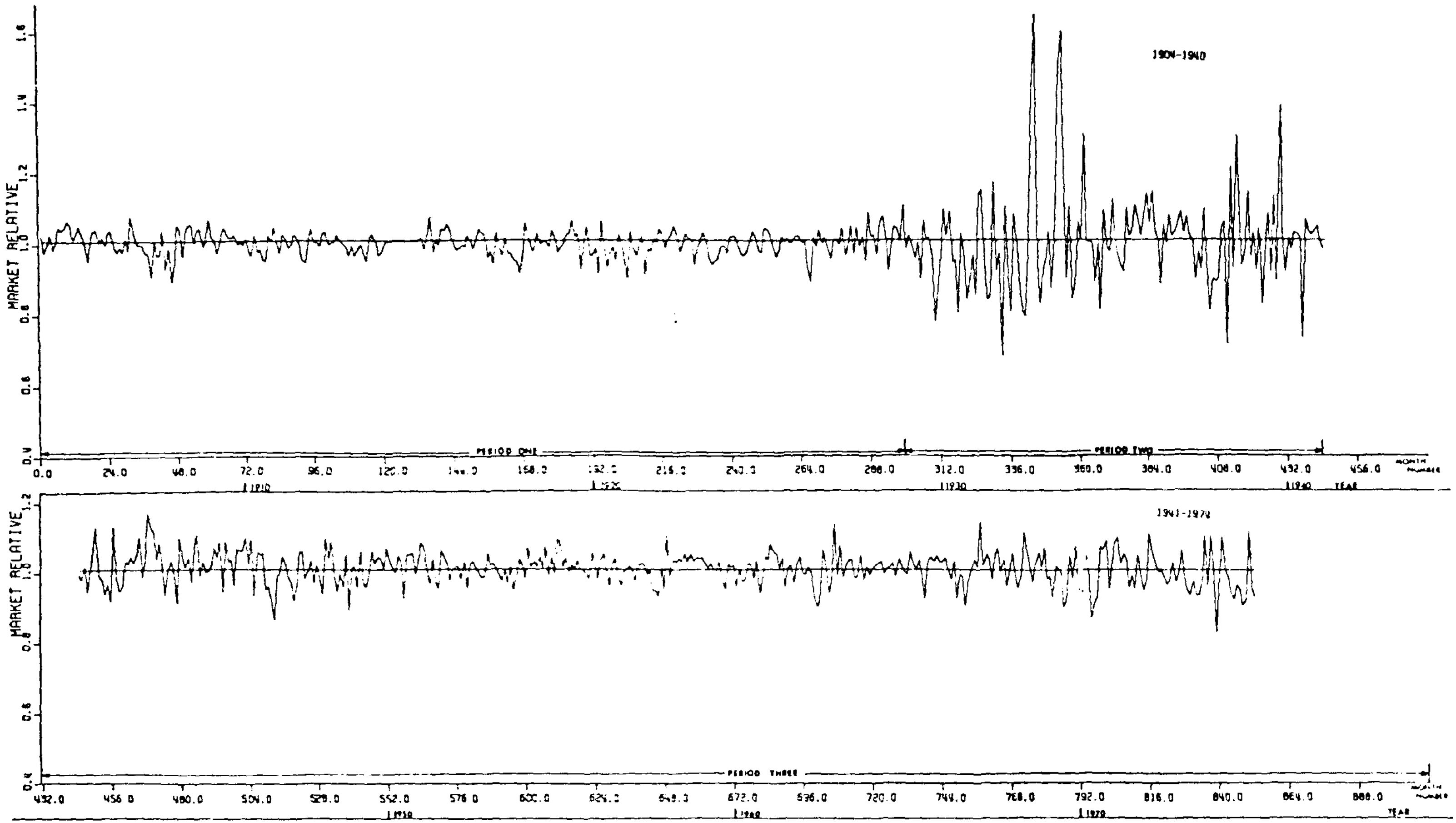

Fig. 1. Market return monthly-relatives [$\ln(1+R_t)$], 1904–1974.

implying drawings from a different distribution. The period of high variability is neatly bracketed by two months October, 1929 and May, 1940 which lie more than 3 standard deviations from the mean. Outside this time period, no observations lie more than 3 standard deviations from the mean. Since all the statistical tests could be influenced by observations during this period, we analyze the time series over 4 time periods: 1904–1928 (months 1–288); 1929–1940 (months 289–432); 1941–1974 (months 433–840); and 1904–1928 *combined* with 1941–1974 (months 1–288 and months 433–840). We denote these periods 1 through 4 respectively. For completeness, statistics are also computed for the entire time period, 1904–1974, denoted period 5.

4.2. *Autocorrelation function*

The autocorrelation function, which consists of serial correlations at various lags, provides a starting point to compare the U.S. data to that of the Australian market and to the spectral analysis of Granger and Morgenstern of U.S. data. For this comparison, use of raw rates of return or natural logarithms of 1 plus the rate of return give essentially identical results. The autocorrelation functions in the natural logarithm form for all 5 periods and Officer's time period are presented in fig. 2. Two standard errors of the estimated autocorrelation coefficients [see Box and Jenkins (1970, p. 34)] are shown by the horizontal lines on either side of each function.

The significance levels of individual lags in the autocorrelation function must be interpreted with some caution due to the lack of strict normality. Given this warning, we see that sample autocorrelations are significant at lag 1 for all periods, an artifact (as noted in footnote 2) of index measurement which, of course, has nothing to do with seasonality. Generally, the sample autocorrelations are quite small with very few greater than two standard deviations from zero. The position of significant lags depends to some extent on the time period studied. In period 1, the first homogeneous subperiod (1904–1928), lags 1 and 6 are significant. Over period 3 (1941–1974), *no* autocorrelations (other than at lag 1) are more than two standard deviations from zero, indicating a random walk model with no (multiplicative) seasonal effects. Autocorrelations at lags 3, 9 and 12 are unimportant in both periods. In combined period 4, (1904–1928 and 1941–1974), the data have significant autocorrelations at lags 4 and 6 but lags 3, 9 and 12 remain unimportant.[8] These results are in broad agreement with those of Granger and Morgenstern, namely, a flat function with a tendency to peaks at lags of 4 months and less. However we also find signs of a 6 month seasonal. Certaintly, longer seasonals of 9 or 12 months seem absent.

Comparing the U.S. and Australian sample autocorrelations over the identical

[8]The autocorrelation function of the time series over the entire period is not emphasized due to the nonhomogeneity of the series inclusive of 1929–40. For those who are curious, lags 1, 3, and 9 are significant with logged data; in addition, lag 10 is significant with unlogged data.

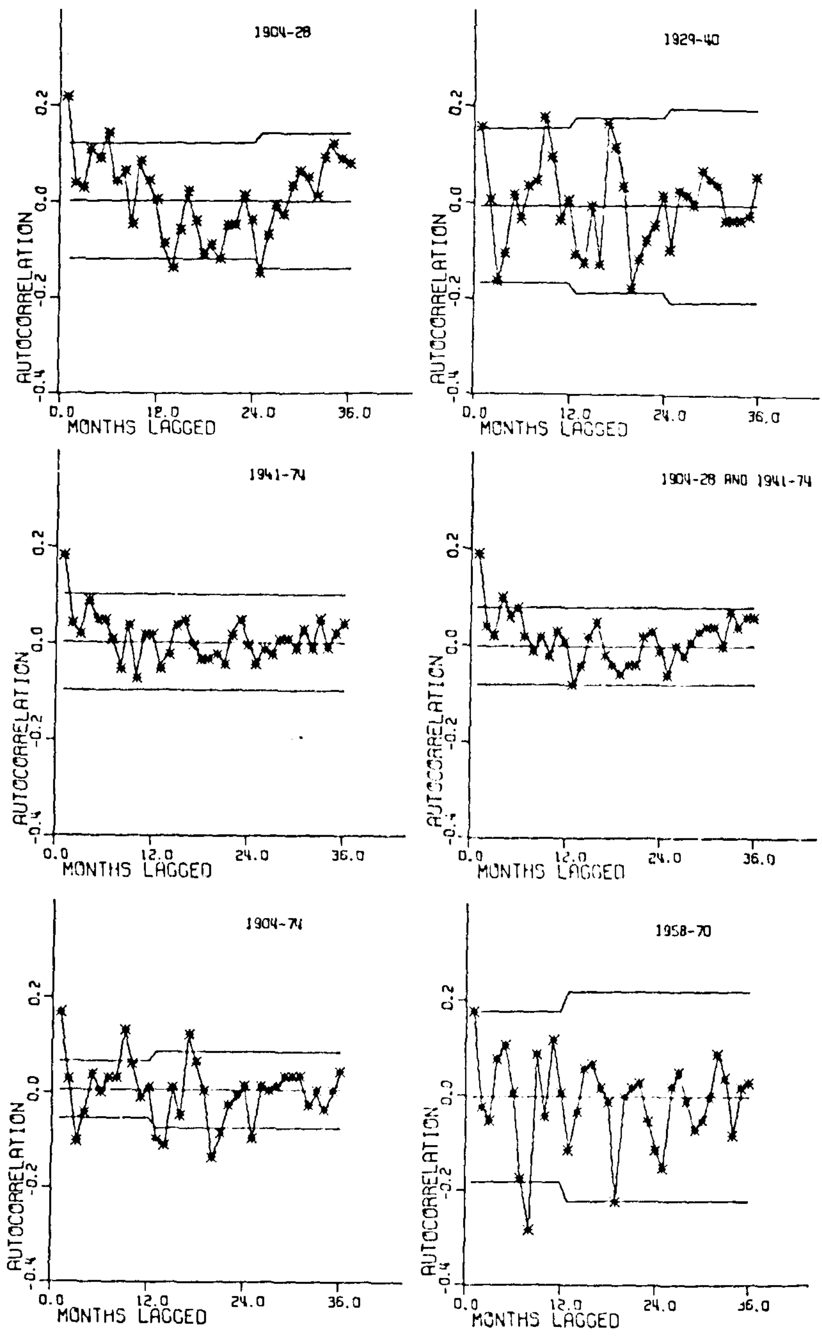

Fig. 2. Autocorrelation function for monthly market return relative for various periods. Solid horizontal lines indicate two standard errors of the estimated autocorrelation coefficients.

time period used by Officer (2/1958-6/1970), we find that lags 6 and 9 are unimportant in the U.S. data in contrast to the Australian data, while lag 8 with a sample autocorrelation of -0.28 shows up significant in the U.S. data. Such a finding may be due to sampling error.

The concept of seasonality implied by significant lagged autocorrelations differs from that expressed by our hypotheses presented in (8) and (9). To see the difference, consider, for example, a six-month seasonal which appears as a significant autocorrelation at lag 6. This seasonal means that on average each month's observation is linked or related to that of the sixth preceding month, January with July and February with August and so on. Since the sample autocorrelation for lag 6 is computed or averaged over all six pairs of months, note that for the autocorrelation between observations six months apart to appear significant requires that the effect be strong enough for those pairs for which it occurs so as not to be hidden by its possible absence in other pairs. The sample autocorrelation function is best at uncovering systematic homogeneous cyclicalities between months but might fail to reveal peculiarities of individual months (or even pairs, triples, etc. of months) if the remaining months display no relationship.

Furthermore, the sample autocorrelation function provides no direct evidence about the distributions of stock returns by month. In the next several subsections, we examine stock return distributions by month directly, The tests can be interpreted as tests for seasonality in the sense that some month(s) have different distributions than other months.

4.3. Sample statistics

Summary sample statistics of the stock return distributions calculated *by month*, over the several time periods are presented in Table 1. Also included for comparison are statistics for the pooled sample of all months combined within each time period. The 8 panels of table 1 contain 8 sample statistics of interest, arranged by months 1–12 (January–December) and 'all' months. Within each panel, lines 1–5 correspond to periods 1–5 defined in section 4.1 above. Period 3 (1941–74) has special interest since it is the longest and most recent contiguous homogeneous subperiod.

Panels a – c contain three location measures: the arithmetic mean, the median and the 75% truncated mean.[9] The most consistent feature of these statistics is the high January rate of return. This appears in *all* time periods and in *all* the statistics. For example, for period 3 the 75% truncated mean for January was

[9]The 75% truncated sample mean is the arithmetic average of the middle 75% of the observations ordered from smallest to largest. The properties of these three statistics as estimators of the location parameter of *stable distributions with characteristic* exponents between 1 and 2 are examined by Fama and Roll (1968). Given the sample values of α (the characteristic exponent) found in 4.6 below and the Fama-Roll results, the 75% truncated mean is probably the most efficient location estimator.

Table 1
Summary descriptive statistics for returns by month.

Period		Month 1	2	3	4	5	6	7	8	9	10	11	12	all
Location measures														
a. Mean	1	0.0130	−0.0097	0.0017	0.0122	0.0033	−0.0058	0.0090	0.0102	0.0101	0.0042	0.0127	0.0010	0.0052
	2	0.0663	0.0191	−0.0528	0.0080	−0.0331	0.0403	0.0497	0.0487	−0.0510	−0.0510	−0.0123	−0.0314	0.0019
	3	0.0391	0.0056	0.0142	0.0016	0.0017	−0.0063	0.0153	0.0056	0.0002	0.0085	0.0100	0.0201	0.0096
	4	0.0283	−0.0008	0.0090	0.0060	0.0024	−0.0061	0.0127	0.0075	0.0043	0.0067	0.0111	0.0122	0.0078
	5	0.0348	0.0026	−0.0016	0.0063	−0.0037	0.0018	0.0190	0.0146	−0.0052	0.0007	0.0071	0.0047	0.0068
b. Median	1	0.0136	−0.0135	0.0010	−0.0011	−0.0031	0.0021	0.0071	0.0227	0.0068	0.0058	0.0228	0.0021	0.0062
	2	0.0421	0.0250	−0.0328	0.0014	−0.0729	0.0399	0.0383	0.0220	−0.0436	−0.0212	−0.0369	−0.0175	0.0074
	3	0.0393	0.0115	0.0127	0.0106	0.0080	−0.0068	0.0115	0.0057	−0.0112	0.0014	0.0147	0.0210	0.0138
	4	0.0224	0.0038	0.0127	0.0078	0.0076	0.0013	0.0113	0.0209	0.0051	0.0058	0.0207	0.0151	0.0107
	5	0.0247	0.0085	0.0048	0.0054	0.0036	0.0035	0.0135	0.0220	0.0010	0.0039	0.0107	0.0149	0.0107
c. Trunc. mean	1	0.0133	−0.0099	0.0051	0.0090	0.0035	−0.0016	0.0101	0.0150	0.0125	0.0081	0.0131	0.0033	0.0070
	2	0.0598	0.0212	−0.0395	−0.0114	−0.0558	0.0424	0.0442	0.0194	−0.0569	−0.0235	−0.0153	−0.0227	−0.0014
	3	0.0379	0.0055	0.0133	0.0046	0.0046	−0.0054	0.0158	0.0111	0.0021	0.0069	0.0153	0.0217	0.0107
	4	0.0247	−0.0008	0.0097	0.0081	0.0045	−0.0039	0.0131	0.0126	0.0068	0.0073	0.0145	0.0124	0.0091
	5	0.0297	0.0030	0.0063	0.0053	−0.0012	−0.0011	0.0169	0.0117	0.0021	0.0058	0.0099	0.0092	0.0082
Scale measures														
d. Std. dev.	1	0.0238	0.0347	0.0423	0.0321	0.0374	0.0332	0.0272	0.0387	0.0267	0.0397	0.0486	0.0338	0.0355
	2	0.0931	0.0654	0.1191	0.1696	0.1955	0.1316	0.1398	0.1577	0.1719	0.1192	0.0930	0.0816	0.1348
	3	0.0523	0.0427	0.0379	0.0516	0.0484	0.0467	0.0498	0.0431	0.0518	0.0418	0.0602	0.0452	0.0486
	4	0.0445	0.0400	0.0399	0.0445	0.0438	0.0413	0.0418	0.0410	0.0432	0.0406	0.0552	0.0416	0.0473
	5	0.0568	0.0454	0.0643	0.0789	0.0887	0.0669	0.0690	0.0748	0.0818	0.0617	0.0631	0.0526	0.0685
e. M.A.D.	1	0.0201	0.0272	0.0326	0.0268	0.0322	0.0267	0.0214	0.0296	0.0207	0.0289	0.0391	0.0261	0.0280
	2	0.0686	0.0409	0.0884	0.1160	0.1288	0.0956	0.0968	0.0978	0.1188	0.0997	0.0778	0.0624	0.0960
	3	0.0419	0.0320	0.0304	0.0411	0.0391	0.0372	0.0414	0.0337	0.0417	0.0324	0.0421	0.0339	0.0380
	4	0.0343	0.0315	0.0314	0.0342	0.0362	0.0329	0.0332	0.0323	0.0334	0.0309	0.0409	0.0316	0.0339
	5	0.0414	0.0343	0.0413	0.0482	0.0541	0.0443	0.0459	0.0419	0.0491	0.0430	0.0485	0.0373	0.0444
f. Statistic $\hat{c}$	1	0.0201	0.0246	0.0295	0.0250	0.0357	0.0296	0.0191	0.0279	0.0193	0.0222	0.0364	0.0307	0.0246
	2	0.0635	0.0152	0.0784	0.0849	0.0768	0.0873	0.1010	0.0476	0.0917	0.0979	0.0510	0.0734	0.0731
	3	0.0437	0.0279	0.0251	0.0397	0.0411	0.0338	0.0438	0.0266	0.0403	0.0247	0.0310	0.0277	0.0308
	4	0.0273	0.0244	0.0297	0.0289	0.0346	0.0305	0.0294	0.0234	0.0293	0.0217	0.0294	0.0234	0.0277
	5	0.0290	0.0291	0.0298	0.0338	0.0419	0.0319	0.0345	0.0296	0.0328	0.0321	0.0394	0.0274	0.0305
Higher moments														
g. Skewness	1	−0.0000	−0.0000	−0.0015	0.0007	−0.0001	−0.0009	−0.0002	−0.0012	−0.0006	−0.0017	−0.0002	−0.0009	−0.0006
	2	0.0052	−0.0021	−0.0152	0.0297	0.0472	−0.0008	0.0060	0.0503	0.0089	−0.0039	0.0032	−0.0068	0.0113
	3	0.0005	0.0005	0.0004	−0.0014	−0.0010	−0.0002	−0.0001	−0.0016	−0.0011	0.0007	−0.0038	−0.0004	−0.0006
	4	0.0013	0.0004	−0.0005	−0.0010	−0.0007	−0.0004	0.0001	−0.0015	−0.0012	−0.0002	−0.0025	−0.0003	−0.0005
	5	0.0040	0.0003	−0.0098	0.0100	0.0105	0.0039	0.0058	0.0213	−0.0051	−0.0042	−0.0019	−0.0036	0.0029
h. Kurtosis	1	−1.1180	−0.4543	0.5994	−0.6433	−1.2009	−0.2389	−0.5552	−0.1418	0.2650	1.3774	−0.5364	0.5534	0.5299
	2	0.0958	0.8796	0.4122	0.4984	1.6019	−0.6185	0.3889	3.4558	0.4786	−1.2582	−1.4327	−0.3025	2.3956
	3	−0.6002	0.5485	−0.2128	0.2144	−0.6171	−0.4674	−0.8126	0.0651	0.1870	−0.0344	2.0174	−0.2251	0.6269
	4	0.2616	0.5611	0.6525	0.9156	−0.5401	−0.1462	−0.2309	0.1211	1.1822	0.6865	1.7488	0.1783	0.8904
	5	3.2340	1.3334	9.5048	9.3797	12.1698	4.1685	7.9625	23.9807	8.4130	2.8396	0.3562	3.3301	11.1597

0.0379 (panel c, line 3), which is nearly 3/4 again as large as the next highest return, December, at 0.0217. Panels d–f contain three scale measures: standard deviation, mean absolute deviation and the statistic $\hat{c}$, defined as (72nd fractile − 28th fractile)/2(0.827), the Fama–Roll (1971) dispersion measure.[10] For these scale measures, the differences among months do not seem to be as pronounced or consistent as for the location measures; nevertheless they are substantial. Note, for example, that for statistic $\hat{c}$, regarded as the most stable dispersion measure (panel f), for period 3 the statistics span 0.0247 to 0.0438. Also note that the sample standard deviations by month are not wildly erratic in comparison with the order statistic $\hat{c}$. Again referring to period 3, the *orderings* of months from high dispersion to low which are achieved by $\hat{c}$ and by standard deviation are highly correlated: the Spearman's rho is 0.75. This feature of the data is particularly important if we are to have confidence in the parametric analysis of variance tests which rely on the standard deviation as a dispersion measure.

Finally, panels g and h measure sample skewness and kurtosis.[11] No departures from symmetry are evident as indicated by the skewness measures close to zero. In panel h positive entries imply positive kurtosis (usually due to a peaked center and fat tails). Note that for the pooled samples in the last column, where sample size is at least 144, kurtosis is always positive as others have found by inspecting the density function. For individual months, a fair number of negative statistics appear, usually when the sample size is 50 or less. This behavior of sample kurtosis drawn from non-normal distributions parallels Mandelbrot's (1967) finding for samples of size fifty drawn from daily spot prices of cotton, 1900–1905.[12]

In the next two subsections the location and scale hypotheses [eqs.(8) and (9)], are tested using both non-parametric and parametric tests. We begin with the non-parametric tests since they require less restrictive distributional assumptions. Parametric tests follow for completeness and to allow testing of some relevant subhypotheses.

4.4. *Non-parametric tests*

The Kruskal–Wallis test statistic, fully described in Conover (1971, p. 157), is a test which uses 'ranks' and requires no distributional assumptions other than that the random variables are continuous and measurable on an ordinal scale. The Kruskal–Wallis test statistic is used to test the hypothesis that all 12 of the populations from which the 12 samples are drawn have identical popula-

[10] For the α values of the samples, $\hat{c}$ is probably the most efficient scale estimator.

[11] Let x_i denote the ith sample value, $\bar{x}$ the sample mean of n observations and $M_k = \Sigma_{i=1}^{n} (x_i - \bar{x})^k/n$. Then sample skewness is $M_3/(M_2\sqrt{M_2})$, and sample kurtosis is $(M_4/M_2{}^2) - 3$.

[12] Geary's kurtosis measure, the ratio of the mean absolute deviation to the standard deviation is very poorly behaved and usually indicates negative kurtosis even in large samples.

tion distributions.[13] The test statistic is approximately distributed as chi-square with 11 degrees of freedom and a one-tailed rejection region is appropriate.[14]

The results are shown in table 2 in panel a. For period 3 the null hypothesis can be rejected at the 10% level; in period 4, the combined period, the hypothesis can be rejected at the 5% level. Given the conservative nature of the test,

Table 2

Summary non-parametric test statistics for returns by month.

a. *All months combined*

Period	Kruskal–Wallis statistic	Siegel–Tukey statistic
1. 1904–28	11.46	20.69[b]
2. 1929–40	17.51[a]	16.72
3. 1941–74	18.04[a]	17.30[a]
4. 1904–28 and 1941–70	21.30[b]	13.71
5. 1904–74	23.19[b]	13.73

b. *Pairwise comparisons*

Probability level	Period	Location	Dispersion
0.99	3. 1941–74	None	None
	4. 1904–28 and 1941–74	Jan. > June	None
0.95	3. 1941–74	Jan. > June	None
	4. 1904–28 and 1941–74	Jan. > June, Feb.	None
0.90	3. 1941–74	Jan. > all except Dec. Dec. > June	Feb. < Jan., July, Sept., Nov., Dec.
	4. 1904–28 and 1941–74	Jan. > all except July, Nov., Dec. July > Feb., June Nov. > Feb., June Dec. > Feb., June	None

[a]Significant at the 0.90 level.
[b]Significant at the 0.95 level.

this is rather convincing evidence for seasonality in stock rates of return, in that at least one of the population distributions from which the samples are drawn

[13]Since the Kruskal–Wallis statistic is designed to be sensitive to differences in population means, the test may more loosely be regarded as testing the hypothesis that all the distributions have identical means.

[14]Kruskal and Wallis found that for 10% significance levels or less, the *true* significance level is *smaller* than that given with the chi-square distribution [see Conover (1971)]. In other words, the chi-square gives a conservative hypothesis test. This feature of the test is consistent with the higher significance levels generally found in the parametric tests.

differs from some of the rest in location. In contrast, the autocorrelation function gave no signs of seasonality in period 3.

In order to find the months that are responsible for the result, we conduct multiple sequential pairwise comparisons among months using the rank sums for each month and investigating the differences in rank sums. The sequential procedure is due to Hartley (1955); Wilcoxon and Wilcox (1964) provide the standard deviation of rank sum totals required in the test. Note that the sequential nature of the test protects against logical contradictions that arise in less powerful methods of making pairwise comparisons when many treatments are being considered.

Restricting attention to the longer homogeneous periods 3 and 4, we find that January has a significantly greater mean rank than June at the 1% level in period 4 and at the 5% level in period 3. At the 10% level, other differences are detectable. For example, in period 4 January has a greater mean rank than all other months except November, July and December which are the months with the highest rank totals (and 75% truncated means) after January. The sequential procedure then tests these months against the remaining months. November, July and December all have significantly greater mean ranks (at the 10% level) than February and June.

Referring back to the statistics in table 1, we see that the rank tests mirror the differences in location statistics. For example, in panel c, line 4, January has the highest truncated mean of 0.0247 while June and February have the lowest means of -0.0039 and -0.0008, respectively. By the usual interpretation of significance levels differences at the 10% level would be considered as 'weak', but in numerical terms they are quite large. December's mean monthly rate of return in period 4 of 0.0124 (about 15% on an annual basis) compares for example to February's 0% annual return. Since the 75% truncated mean eliminates extreme observations and the non-parametric tests do not give extra weight to extreme observations, it is difficult to accept the view that these differences are merely the result of sampling error and not the result of factors giving rise to seasonal differences.

For a distribution-free test of differences in scale, we use the Siegel-Tukey statistic [see Conover (1971, p. 229)], a rank test designed to detect differences in dispersion. The test requires that the populations be aligned according to location; lacking knowledge of the true population location, the usual procedure is to adjust the samples by a sample measure of location such as mean or median, adjusting to the sample in the middle. The location measure we use to make the adjustment is the 75% truncated mean. The chi-square table is again appropriate. Referring to column 3 of table 2, we see that over 1941-1974, the statistic is just significant at the 10% level; it is also significant in period 1. In combined period 4, however, it is not significant, being smaller in contrast with the larger Kruskal-Wallis statistic. This result is consistent with the relatively lower heterogeneity in scale measures compared to location measures noted above.

Taking the larger sample size as the more reliable, it appears that differences in scale among months are at best rather weak. These results are consistent with Bartlett's (parametric) Homogeneity of Variances tests reported in the next section.

As a possible explanation for seasonality in the Australian stock index, Officer suggests (1975, p. 47) that months with high returns may also have a higher market factor risk and vice versa, but finds no evidence for this hypothesis. To examine this hypothesis we make sequential pairwise comparisons of dispersion differences despite the weak significance of such differences when all months are considered together. For period 4 there are no differences that are statistically significant even at the 10% level. For period 3 which has a Siegel–Tukey statistic that is significant at the 10% level, we find that dispersion in February is significantly less than in January, July, September, November and December. This may be partial verification of Officer's hypothesis since February is a low return month and January, July, November and December are high return months. On the other hand, September (a high dispersion month) is also a *low* return month and June (a low return month) does not have low dispersion.

Consider now the non-parametric statistics for Officer's Australian data. Using his Dex126 over 1959–1970, the Kruskal–Wallis statistic is 33.67 and the Siegel–Tukey statistic is 40.10, indicating that at least one monthly population distribution likely differs from the rest in location and similarly for scale. The highly significant scale difference is in contrast to the U.S. data. The strength of seasonal effects also is much more clearly evident in the Australian index.

The pairwise comparisons of means show that July exceeds March and September (5% level) and January exceeds March. The pairwise scale comparisons show that December has greater dispersion than all other months in the year excepting October and November. It is of considerable interest that in the U.S. data, January and July are also relatively high return months. Over 1941–1974 for example, January is highest and July third highest, measured by the truncated mean. On the other hand, the extraordinary variability of December returns in the Australian data finds no parallel in the U.S. data

In summary, seasonalfty in the U.S. aggregate market data is a statistically significant phenomenon that is clearly observable when the data are examined by month. At least one distribution is different from the rest as shown by the Kruskal–Wallis test. The pairwise tests allow the inference that January has this distribution and examination of the data indicates that January's relatively high mean rate of return is likely the source of the distributional difference.

4.5. Parametric tests

Initially the distribution test results of Fama (1965) and Fama and Roll (1968, 1971) led to skepticism concerning the usefulness of parametric statistics such as mean and standard deviation in the context of distributions with characteristic exponent less than two. Evidence which has since accumulated has shown among other things that the standard deviation as a scale measure seems well-behaved and that a representation by a normal distribution with finite variance may be adequate in many instances [see Officer (1971) and Hsu, Miller and Wichern (1974)]. In addition, we have seen that Officer's inferences (1975) from the autocorrelation function which are based on the assumption of normality are correct inferences in a holdout time period. Finally we have pointed out in examining the raw data that the various scale and location measures do seem to correspond closely with each other. For all these reasons, and the generally recognized robustness of the analysis of variance model, it seems appropriate to consider hypothesis tests based on parametric statistics.

To test the hypothesis of equal means, analysis of variance techniques in conjunction with Bartlett's test for homogeneity of variances are applied. Each month is considered as a separate 'treatment' level and tests are conducted for each of the five time periods. Using Bartlett's (1937) test, we find that homogeneity of variances can be rejected for periods 1, 2 and 5 (see table 3, panel a). We therefore report the analysis of variance tests only for the remaining periods 3 and 4. The linear model assumed is [see eq. (7)]

$$\ln(1+\tilde{R}_{tm}) = \mu + \lambda_m + \tilde{e}_{tm}. \tag{10}$$

The analysis of variance results are given in panel b of table 3. The hypothesis that the mean monthly rates of return are equal over all months can be rejected at above the 0.975 level for each period, confirming the results of the nonparametric tests. We again consider which month(s) seem to be significantly different from the rest. Such a determination may allow further understanding of the phenomenon for the eventual construction and testing of theories which explain the seasonality as well as the incorporation of such knowledge in decision models.

Two comparisons or linear contrasts are considered in panel c. These are

$$C_1: 11\lambda_1 - (\lambda_2+\lambda_3+\ \dots\ +\lambda_{12}), \tag{11}$$

$$C_2: (\lambda_4+\lambda_5+\lambda_6)-(\lambda_{10}+\lambda_{11}+\lambda_{12}). \tag{12}$$

There are strong a priori reasons to consider C_1: (1) the high return for January (month 1), is a phenomenon which others have noted [see Wachtel (1942) and Zinbarg (1964)]; (2) January marks the beginning and ending of several potentially important financial and informational events. As examples of the latter, January is the start of the tax year for investors, the beginning of the tax and accounting years for most firms[15] and the period during which

[15]For example 59.8% of the COMPUSTAT firms had accounting year closings in December for 1974.

preliminary (and in many cases final) announcements of the previous calendar (fiscal) year's accounting earnings are made. It is possible that seasonality is in some way associated with these accounting events.

Contrast C_2, on the other hand, is derived from observation of the data. The contiguous months of April, May and June seem to have below average returns while six months later the contiguous months of October, November and December have above average returns. In addition, each of these groups of months is followed by a month that has above average returns, July and January respectively.

Table 3

Summary parametric test statistics for returns by month.

a. *Bartlett's test for homogeneity of variances*

	1904–28	1929–40	1941–74	1904–28 & 1941–74	1904–74
Corrected χ^2	20.67[a]	20.48[b]	11.46	10.92	53.49[b]

b. *Analysis of variance*

Source of variation	df	Sum of squares	Mean square	F
1941–74				
Month	11	0.052210	0.004746	2.06330
Residual	396	0.910956	0.002300	
Corrected total	407	0.963166	0.002366	
1904–28 & 1941–74				
Month	11	0.045733	0.004157	2.21604
Residual	684	1.283271	0.001876	
Corrected total	695	1.329004	0.001912	

c. *Contrasts*

C_1:	1904–28 & 1941–74:	$F = 14.18$[b]
	1941–74:	$F = 13.99$[b]
C_2:	1904–28 & 1941–74:	$F = 3.99$[a]
	1941–74:	$F = 4.27$[a]

[a]Significant at the 0.95 level.
[b]Significant at the 0.99 level.

To test whether a contrast is equal to zero, the appropriate test statistic [Winer (1962, p. 69)] is

$$F = \left(\sum_{i=1}^{12} c_i T_i \right)^2 \Big/ \left(n \left(\sum_{i=1}^{12} c_i \right) MS \right),$$

where c_i is the contrast coefficient for level i, n is the number of observations per month ($n = 34$ for period 3), $T_i = n\lambda_i$ and MS is the within-treatments mean square. The statistic is compared to a theoretical $F_{\alpha,1,11n}$ value. In panel c the returns for January are shown to be significantly greater than those for the rest of the year on average and the low Spring returns are significantly smaller than the high Fall returns.

These parametric analysis of variance tests are consistent with the non-parametric tests in indicating that monthly distributional differences may be attributed to location rather than scale parameters. Again it appears that January is the responsible month. In addition, contrast C_2 gives statistical evidence of a systematic difference between second-quarter and fourth-quarter months, a difference also clearly observable in the descriptive sample statistics.

4.6. *Characteristic exponents*

Symmetric stable distributions are described by three parameters and two have been examined: the scale and location parameters. The third parameter, denoted by α, is the characteristic exponent. Individual stock return distributions usually have α's less than 2. In line with interest in seasonality, we will examine, to the extent possible, the hypothesis that the stock return index distributions by month have the same α's. Lacking firm knowledge of the sampling distribution of α, our conclusions are tentative.

Estimates for α are computed using the order statistic technique described by Fama and Roll (1971, p. 333) in combination with table 2 of their earlier (1968) paper. As they did, we use the 96th fractile in estimating α. When the order statistic required by the fractile is not integral, the next lower integer is consistently used, that is we truncate rather than round or interpolate. The results for the three homogeneous periods 1, 3 and 4 are shown in table 4. Fama and Roll also present Monte Carlo sampling distributions of α for sample sizes of 24, 49, and 74. We use these for indications of significant differences.

Our period 1 with sample size 24 by coincidence is represented in their choice of sample size. Under the null hypothesis that α is 1.5, the standard error is 0.367. It is 0.309 for $\alpha = 1.9$, and 0.25 for $\alpha = 2.0$. With the population parameter unknown, the α value for all months combined within a subperiod is used as a benchmark. For period 1 this value is 1.96. The α values estimated for individual months are quite close to 1.96 with the exception of October's 1.42 which is approximately two standard deviations away. Thus for period 1 there is no strong indication that the monthly stock return distributions have different α's.

In period 3 the sample size is 34, and in period 4 it is 58. For these sample sizes the Fama–Roll results show that a standard error of 0.30 is conservative, that is, an upper bound. Under the hypothesis that the pooled population values of α are 1.62 and 1.7 respectively in periods 3 and 4, all values of α

estimated for individual monthly samples lie well within two standard deviations or 0.60. We conclude that the evidence is consistent with the hypothesis that stock return distributions (in aggregate form) by month have the same characteristic exponents.

5. Implications of seasonality

5.1. Capital market efficiency

The fact that expected returns vary by month is not necessarily inconsistent with market efficiency. We have demonstrated in this paper that the simple

Table 4
Characteristic exponents by month.

Month	Period 1 1904–28	Period 3 1941–74	Period 4 1904–28 and 1941–74
Jan.	2.0	2.0	1.75
Feb.	1.8	2.0	1.86
March	1.7	2.0	1.75
April	2.0	2.0	1.86
May	2.0	2.0	2.0
June	2.0	2.0	2.0
July	1.93	2.0	2.0
Aug.	2.0	1.75	2.0
Sept.	1.86	1.91	2.0
Oct.	1.42	1.71	1.64
Nov.	2.0	1.31	1.49
Dec.	2.0	1.76	1.46
All	1.96	1.62	1.7

random walk model does not hold, not that the market is inefficient with respect to information regarding seasonality. Trading rule tests will have to be carried out to test for market efficiency. It is our expectation that the seasonal patterns we have found will not allow the investor to earn abnormal rates of return which are incommensurate with the degree of risk that is accepted. We hold this view for several reasons: (1) the large amount of evidence consistent with market efficiency; (2) the tendency of high return months over 1941–1974 to be also high dispersion months; and (3) the anecdotal and other evidence which indicates that seasonal effects have not gone unnoticed by Wall Street's technical analysts. It is unlikely that seasonality in stock returns will raise any serious problem for the efficient market model. Since, however, market efficiency is usually stated in terms of equilibrium expected returns, models of market equilibrium may require seasonal effects in expected returns.

5.2. Nature of the distribution of stock returns

Recent tests describing the distribution of stock returns have concentrated on the stability of the characteristic exponent, alpha, under addition. Good estimates of alpha are crucial. We have found some evidence that differences in alpha by month are explainable by sampling error, but the estimates by month do vary widely. Our findings should serve as a warning for those engaged in empirical research on stock return distributions, especially if monthly data are used. Use of daily data may perhaps bypass any possible problems although here too very little is known about distributional differences by day. Fama (1965) documents a difference between dispersion in Friday and Monday returns. Phenomena of this type which may well affect investigations of the stability of alpha should not be ignored.[16]

5.3. Expected return models

Many applications which require the computation of abnormal or disequilibrium rates of return have used as models for expected rates of return the market model and the two-parameter capital asset pricing model. Both compute expected returns conditional on market returns and the security's risk, measured by the beta coefficient of the security. Since market returns seem to depend on month of the year, the calculation of abnormal returns may be affected. The next two subsections examine the effects of seasonality on the market model and two-parameter model.

5.3.1. Market model

The market model assumes the existence of a linear regression of the jth security's rate of return, $\tilde{r}_{jt}$, with the market rate of return, $\tilde{R}_t$,

$$E(\tilde{r}_{jt}|R_t) = \alpha_j + \beta_j R_t. \tag{13}$$

The minimum variance linear estimator of $E(\tilde{r}_{jt}|\cdot)$ is

$$\hat{r}_{jt} = \widehat{E(\tilde{r}_{jt}|\cdot)} = \hat{\alpha}_j + \hat{\beta}_j R_t, \tag{14}$$

where $\hat{\alpha}_j$, $\hat{\beta}_j$ are least squares estimates of α_j and β_j.

[16]Teichmoeller (1971) was careful to use only calendar-day changes and also avoided changes over mid-week holidays. Barnea and Downes (1973) do not specify their selection procedure.

Now assume the existence of h linear regressions in h strata corresponding to distinct 'seasons' (e.g., January may form one stratum and the remaining months of the year another stratum),

$$E(\tilde{r}_{jth}|R_{th}) = \alpha_{jh}+\beta_{jh}R_{th}. \tag{15}$$

For the special case in which β_j is common to all strata, sampling theory [see Cochran (1963, pp. 200–203)] suggests the pooled least squares estimate

$$\hat{\beta}_j = \sum_h \sum_{t\epsilon h} (r_{jth}-\bar{r}_{jh})(R_{th}-\bar{R}_h)\Big/\sum_h \sum_{t\epsilon h} (R_{th}-\bar{R}_h)^2. \tag{16}$$

This estimate assumes proportional sampling in strata which will be the case for months of the year. The pooled estimate should result in a more efficient estimate of β_j than the usual least squares estimate. If the regression coefficients vary by stratum, then a separate regression estimate of β_{jh} can be computed for each stratum. A weighted average of the $\hat{\beta}_{jh}$ may, in many cases be an efficient estimate [Cochran (1963, p. 203)].

5.3.2. *Two-parameter model*

Black, Jensen and Scholes (1972) and Fama and Macbeth (1973) propose a model of expected returns of the form.

$$E(\tilde{r}_j) = E(\tilde{R}_0)+[E(\tilde{R})-E(\tilde{R}_0)]\beta_j, \tag{17}$$

where $E(\tilde{R}_0)$ is the expected return on the minimum variance portfolio which has zero covariance with market returns. The second component of the right-hand side represents a risk premium for security j.

A stochastic version of (17) is

$$\tilde{r}_{jt} = \tilde{\gamma}_{0t}+\tilde{\gamma}_{1t}\beta_j+\tilde{e}_{jt}. \tag{18}$$

To overcome error measurement in β_j, (18) has been estimated using portfolio returns. Since $E(\tilde{\gamma}_{1t}) = E(\tilde{R}_t)-E(\tilde{R}_{0t})$ and since $E(\tilde{R}_t)$ has been shown above to be conditional on the month of the year, one is led to consider possible seasonality in estimates of γ_{1t}. For generality we also at the same time examine $\tilde{\gamma}_{0t}$.

We apply the same nonparametric and parametric tests to Fama and Macbeth's monthly estimates of γ_{0t} and γ_{1t} as have been applied to R_t.[17] Selected sample descriptive statistics for each month and all months combined

[17]We thank Professor Fama for making this data available.

Table 5

Statistics for γ_1 and γ_0, estimated intercept and slope of the two-parameter model (eq. (18)).

	Month 1	2	3	4	5	6	7	8	9	10	11	12	all
a. Summary descriptive statistics													
1. 75% truncated mean													
γ_0: 1935–67	0.0006	0.0068	0.0092	0.0080	0.0107	0.0055	0.0184	0.0076	−0.0033	0.0023	0.0168	0.0142	0.0081
1941–67	−0.0033	0.0017	0.0124	0.0068	0.0061	0.0012	0.0191	0.0124	0.0019	0.0050	0.0183	0.0107	0.0078
γ_1: 1935–67	−0.0430	0.0054	0.0017	−0.0003	−0.0013	−0.0007	0.0168	−0.0080	−0.0008	0.0046	−0.0036	0.0068	0.0049
1941–67	0.0450	0.0045	0.0065	−0.0011	0.0098	−0.0014	0.0054	−0.0087	−0.0065	−0.0001	0.0020	0.0140	0.0056
2. Standard deviation													
γ_0: 1935–67	0.0283	0.0306	0.0452	0.0356	0.0520	0.0311	0.0312	0.0278	0.0479	0.0353	0.0414	0.0335	0.0376
1941–67	0.0267	0.0302	0.0296	0.0273	0.0303	0.0294	0.0296	0.0256	0.0290	0.0281	0.0282	0.0301	0.0289
γ_1: 1935–67	0.0682	0.0567	0.0554	0.0558	0.0517	0.0654	0.0590	0.0398	0.1208	0.0596	0.0607	0.0565	0.0660
1941–67	0.0595	0.0580	0.0364	0.0424	0.0432	0.0439	0.0496	0.0348	0.0457	0.0410	0.0489	0.0527	0.0486
3. Statistic c													
γ_0: 1935–67	0.0206	0.0178	0.0112	0.0227	0.0151	0.0218	0.0134	0.0146	0.0251	0.0164	0.0175	0.0226	0.0194
1941–67	0.0148	0.0194	0.0119	0.0258	0.0160	0.0189	0.0122	0.0174	0.0222	0.0118	0.0175	0.0297	0.0178
γ_1: 1935–67	0.0293	0.0133	0.0232	0.0286	0.0371	0.0272	0.0318	0.0200	0.0349	0.0232	0.0239	0.0354	0.0300
1941–67	0.0267	0.0142	0.0245	0.0286	0.0317	0.0318	0.0318	0.0177	0.0351	0.0187	0.0199	0.0286	0.0252

b. Summary test statistics

	Non-parametric tests: Kruskal–Wallis	Siegel–Tukey	Parametric tests: F	d.f.	Bartlett's test
1. γ_0: 1935–67	15.5	34.10[c]	1.25	11,384	32.58[c]
1941–67	18.51[a]	31.20[c]	1.47	11,312	1.61
2. γ_1: 1935–67	21.10[b]	12.16	1.84[b]	11,384	59.58[c]
1941–67	27.81[c]	8.71	3.16[c]	11,312	15.81[a]

[a]Significant at the 0.90 level.
[b]Significant at the 0.95 level.
[c]Significant at the 0.99 level.

are shown in table 5, panel a, for γ_{0t} and γ_{1t} for two time periods covering (1) 1935–1967, Fama and MacBeth's entire time period, and (2) 1941–1967, which we have seen is a period of more homogeneous stock returns. Panel b of table 5 contains the statistical tests.

Examining the statistical tests of location in table 5, the nonparametric and parametric tests show a close correspondence. Both tests reject the hypothesis of equal means for γ_1 at the 1% level for 1941–67. The χ^2 of 15.8 which is significant at a 10% level and indicates a lack of homogeneity of variances, disagrees slightly with the low Siegel–Tukey statistic. However, given the large difference in means which is evident in the sample values of the 75% truncated mean, we are quite confident of highly significant differences in mean γ_1 by month. In particular the January mean of 0.0450 compared to the average of 0.0056 for all months is truly extraordinary. $\hat{\gamma}_1$, it will be recalled, estimates the risk premium inherent in the security market line and has expected value $[E(\tilde{R})-E(\tilde{R}_0)]$. It appears then that possible seasonal effects in $E(\tilde{R}_0)$ do not negate those we have found earlier for $E(\tilde{R})$. To the extent that realizations from our somewhat small sample size of 27 monthly observations adequately measure expectations, it seems that the tradeoff of return for risk demanded and received by the market in January is much greater than in other months of the year. This conclusion should perhaps be somewhat tempered by observing that fully five months of the year – April, June, August, September, and October – show realized *negative* risk premiums; however, the largest of these, -0.0087 for August, is dwarfed in absolute terms by January's $+0.0450$.

Seasonal effects in $\hat{\gamma}_0$ which estimates $E(\tilde{R}_0)$ seem also to be present, approximately to the extent shown by stock returns themselves. The Kruskal–Wallis statistic is of the same order of magnitude while the analysis of variance indicates a lower level of significance (10.8%) as compared with stock returns (1% level). Seasonality in $E(\tilde{R}_0)$ therefore seems to be weaker, if anything, than that in $E(\tilde{R})$.

Concerning possible differences in scale among distribution of γ_0 and γ_1 by month, we have noted that such differences for γ_1 seem to be weak, if not absent. For γ_0 the χ^2 and Siegel–Tukey statistics shown in table 5 give conflicting results. The very low χ^2 statistic of 1.614 reflects the homogeneity of standard deviations observable in table 5, panel a, with standard deviations ranging only from 0.0256 to 0.0303. On the other hand, $\hat{c}$ ranges from 0.0118 to 0.0297 and the rank statistic seems to be reflecting the variability in this order statistic. We therefore are unable to reach a definite conclusion concerning variability in γ_0.

The two-parameter model seems to impound seasonality in the coefficient estimates. One would expect, therefore, that the abnormal returns computed from this model will be free from seasonal effects, or at least much more free than the market model residuals. While this is a plus, seasonality in the risk premium has no obvious explanation while seasonality in expected market returns may be somewhat easier to rationalize. Seasonality in the risk premium may simply be induced by that in market returns. Also possible is a capital asset pricing model which incorporates seasonal effects and gives rise to the fluctuating market risk premium observed here. Chen, Kim and Kon (1975) present a model in which the market price of risk depends upon cash demands and liquidity risk. If the latter, in turn display seasonal characteristics, perhaps such a model can explain the seasonal in the risk premium.

6. Conclusions

Seasonality on the New York Stock Exchange, undetectable with any clarity in the autocorrelation function of returns, becomes clearly evident once rates of return are tested by month. The most outstanding feature of this seasonality is the higher mean of return of the January distribution of returns compared with most other months. Other seasonal peculiarities include relatively high mean returns in July, November and December and low mean returns in February and June. Differences among months in the dispersion parameters of the probability distributions are present to a much lesser extent and can probably be disregarded for most purposes. Seasonal effects on the New York Stock Exchange seem to be weaker than those Officer found in Australian stocks. Interestingly, July and January (in that order) are the highest return months in the Australian stock market.

Seasonality is also a prominent feature of 'risk premiums' estimated from the two-parameter capital asset pricing model. Again it is January with a relatively large risk premium which differs noticeably from the other months.

Hypotheses which seek to explain seasonality are not tested in this paper, our main purpose having been to demonstrate its existence and point out a few of its possible ramifications. Promising avenues of exploration that we have noted in passing are as follows: (1) tax-selling hypothesis, (2) accounting information hypothesis, (3) stochastic cash demand hypothesis. All three we believe deserve elaboration and testing.

References

Barnea, A. and D. H. Downes, 1973, A reexamination of the empirical distribution of stock price changes, Journal of the American Statistical Association 68, 348-350.

Bartlett, M.S., 1937, Some examples of statistical methods of research in agriculture and applied biology, Journal of the Royal Statistical Society Supplement 4, 137.

Black, F., M.C. Jensen and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, in: M.C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).

Blattberg, R.C. and N.J. Gonedes, 1974, A comparison of the stable and student distributions as statistical models for stock prices, Journal of Business 47, 244-280.

Bonin, J.M. and E.A. Moses, 1974, Seasonal variations in prices of individual Dow-Jones industrial stocks, Journal of Financial and Quantitative Analysis 9, 963-991.

Box, G.E.P. and G. Jenkins, 1970, Time series analysis, forecasting and control (Holden-Day, San Francisco, CA).

Chen, A.H., E.H. Kim and S.J. Kon, 1975, Cash demand, liquidation costs and capital market equilibrium under uncertainty, Journal of Financial Economics 2, 293–308.
Cochran, W.G., 1963, Sampling techniques (Wiley, New York).
Conover, W.J., 1971, Practical nonparametric statistics (Wiley, New York).
Cootner, P.H., 1964, The random character of stock market prices (M.I.T. Press, Cambridge, MA).
Cowles, A., 1938, Common-stock indexes 1871–1937 (Principia Press, Bloomington, IL).
Fama, E.F., 1965, The behavior of stock market prices, Journal of Business 37, 34–105.
Fama, E.F., 1970, Efficient capital markets: A review of theory and empirical work, Journal of Finance 25, 383–417.
Fama, E.F. and J.D. Macbeth, 1973, Risk, return, and equilibrium: Empirical tests, Journal of Political Economy 81, 607–636.
Fama, E.F. and R. Roll, 1968, Some properties of symmetric stable distributions, Journal of the American Statistical Association 63, 817–836.
Fama, E.F. and R. Roll, 1971, Parameter estimates for symmetric stable distributions, Journal of the American Statistical Association 66, 331–338.
Fisher, L., 1966, Some new stock market indexes, Journal of Business, Security Prices: A Supplement 39, 191–225.
Granger, C.W.J. and O. Morgenstern, 1963, Spectral analysis of New York stock market prices, Kyklos 16, 1–25.
Granger, C.W.J. and O. Morgenstern, 1970, Predictability of stock market prices (D.C. Heath and Co., Lexington, MA).
Hartley, H.O., 1955, Communications on Pure and Applied Mathematics 8, 47–55.
Hsu, D., R.B. Miller and D.W. Wichern, 1974, On the stable Paretian behavior of stock-market prices, Journal of the American Statistical Association 69, 108–113.
Mandelbrot, B., 1963, The variation of certain speculative prices, Journal of Business 36, 394–419.
Mandelbrot, B., 1967, The variation of some other speculative prices, Journal of Business 40, 393–413.
Officer, R.R., 1972, The distribution of stock returns, Journal of the American Statistical Association 67, 807–812.
Officer, R.R., 1975, Seasonality in Australian capital markets: Market efficiency and empirical issues, Journal of Financial Economics 2, 29–51.
Press, S.J., 1967, A compound events model for security prices, Journal of Business 40, 317–335.
Shiskin, J., 1972, Systematic aspects of stock price fluctuations, in: J. Lorie and R. Brealey, eds., Modern developments in investment management (Praeger, New York).
Teichmoller, J., 1971, A note on the distribution of stock price changes, Journal of the American Statistical Association 66, 282–284.
Wachtel, S.B., 1942, Certain observations on seasonal movements in stock prices, Journal of Business 15, 184–193.
Wilcoxon, F. and R.A. Wilcox, 1964, Some rapid approximate statistical procedures (Lederle Laboratories, Pearl River).
Winer, B.J., 1962, Statistical principles in experimental design (McGraw-Hill, New York).
Zinbarg, E.D., 1964, The stock market's seasonal pattern, Financial Analysts Journal 20, 53.