WISCONSIN WORKING PAPER

4-75-20

ON THE EVIDENCE SUPPORTING THE EXISTENCE OF RISK PREMIUMS IN THE CAPITAL MARKET

BY

ROBERT A. HAUGEN AND A. JAMES HEINS

DECEMBER 1972

GRADUATE SCHOOL OF BUSINESS

UNIVERSITY OF WISCONSIN-MADISON

Business

Auy1282

On the Evidence Supporting the Existence of Risk Premiums in the Capital Market

Robert A. Haugen and A. James Heins

Strides have been made recently in the discovery and refinement of theoretical models which purport to describe the relationship between asset prices and their risk attributes. (See especially Lintner [26,27,28], Sharpe [37], Mosin [32,33] and Fama [8 ,9 ,10].) Although the theoretical arguments are based on assumptions which are admittedly unrealistic, the models have gained wide spread acceptance because of their intuitive appeal and because most reported empirical evidence [1,4,5,22,38,39] alledgedly supports their predictive value. It is our purpose to critically analyze the nature of this evidence, reveal its inherent weakness, and to design an alternative test to overcome many of the problems of prior studies.[1] In doing so we conclude that the supportive nature of previous results can be accounted for by sampling and measurement procedures which bias the evidence in favor of the "risk premium hypothesis". After observing the performance of an extremely large number of issues over long periods of time, we find little support for the notion that risk premiums have, in fact, manifested themselves in realized rates of return.

---

[1] We shall restrict our critique to the many cross section tests of the risk return relationship. Black, Jensen, and Scholes [3], and Fama and MacBeth [11] have employed time series analysis to examine the relationship. While many of the problems we shall subsequently discuss bear importantly on tests of this type, a complete analysis of their effects is best reserved for a separate paper.

I. The Nature of Empirical Tests of the Risk-Return Relationship

In conventional tests of the risk-return relationship, expected return is generally taken to be the mean of the possible outcomes, and risk is typically measured by the variance of a well diversified portfolio, or by the covariance of an individual issue with a surrogate for the market portfolio.[2] A major problem here (and a problem that really cannot be overcome) is that we cannot know what the expectations of investors are. Nor can we know what rate of return is appropriate. We can only know present and past values.

Customarily, what is done is to wave the hands and say that over sufficiently long periods, investors' expectations are borne out on the average, and the sample distribution of results for any issue reflects the probability distribution existing in investors' minds. That is, ex post figures are used to derive estimates of ex ante expectations.

For example, suppose that the ex ante probability distributions of rates of return to efficient portfolios of financial assets for two risk classes are structured as in Figure 1a. If these distributions are stationary in their parameters and are subjected to repeated sampling over four periods, as shown for the portfolio in risk class A in Figure 1b, then we could expect the sample means, $\bar{r}$, and variances σ^2_r, over a cross section of portfolios to plot as shown in the scatter diagram in Figure 1c. That is, if the real probability distributions are shown in Figure 1a, and for each portfolio we employ the procedure of Figure 1b, then the plotting of sample results can be expected to look like Figure 1c.

---

[2] In the case of individual issues the relevant relationship is not between the mean and variance as only the systematic, or undiversifiable, component of an issue's variance (as measured by its covariance with the market) is relevant to pricing under the capital asset pricing model. For well diversified portfolios of issues, the diversifiable or residual variance, has been all but eliminated, and any relationship which exists between their means and variances is economically interesting.

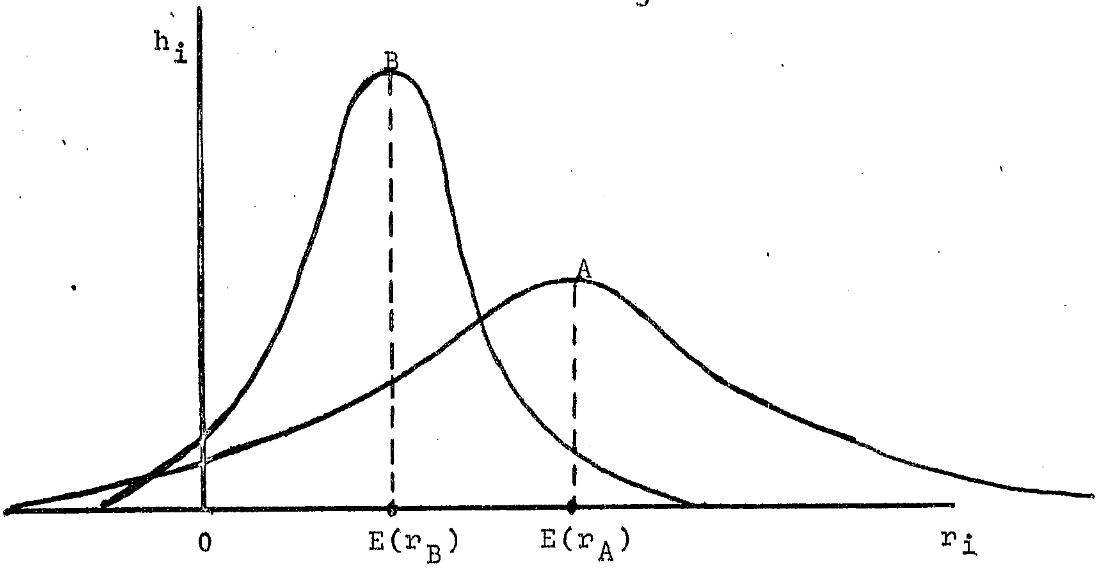


Figure 1a

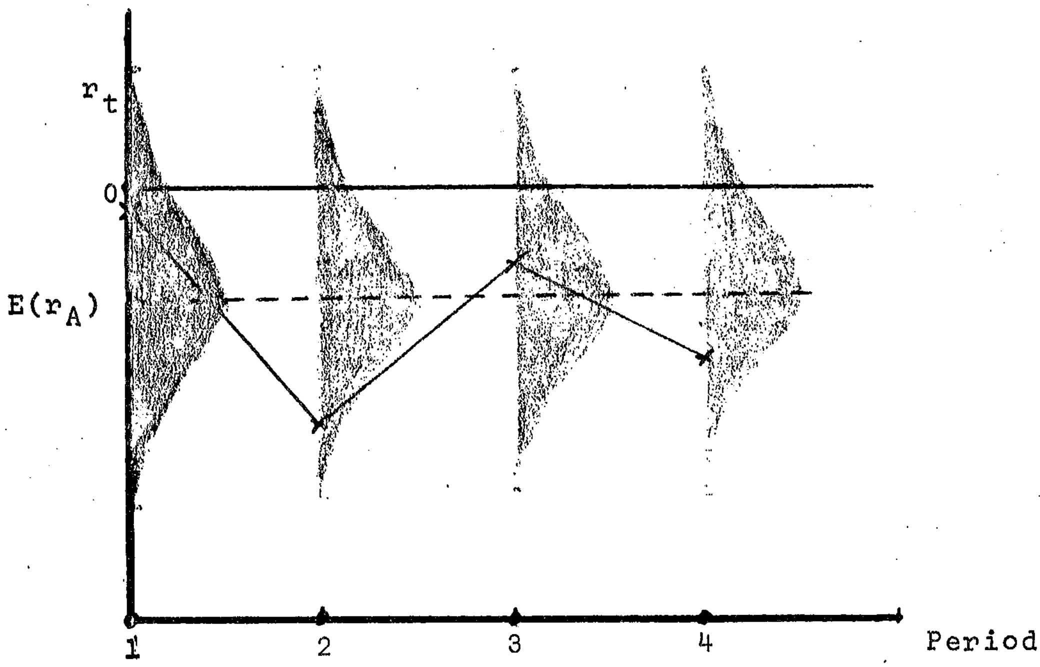


Figure 1b

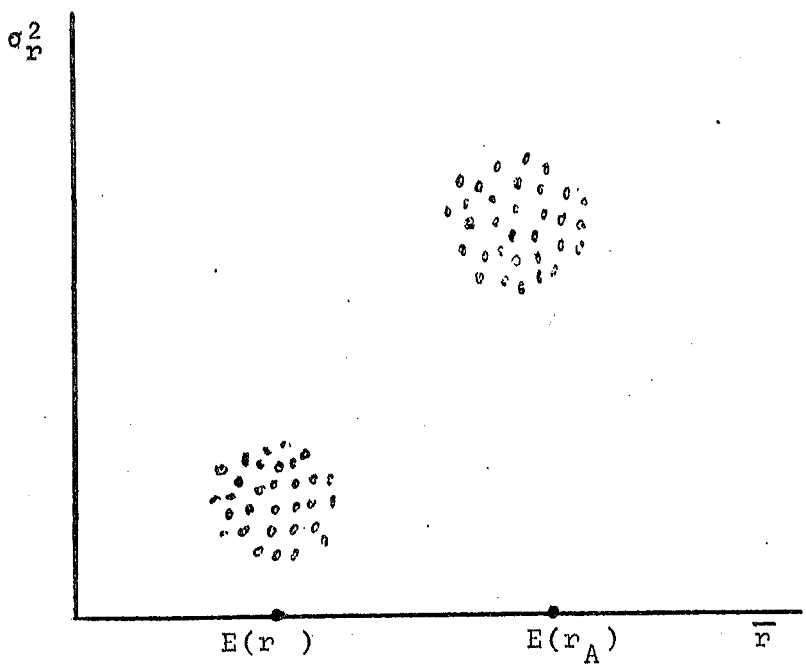


Figure 1c

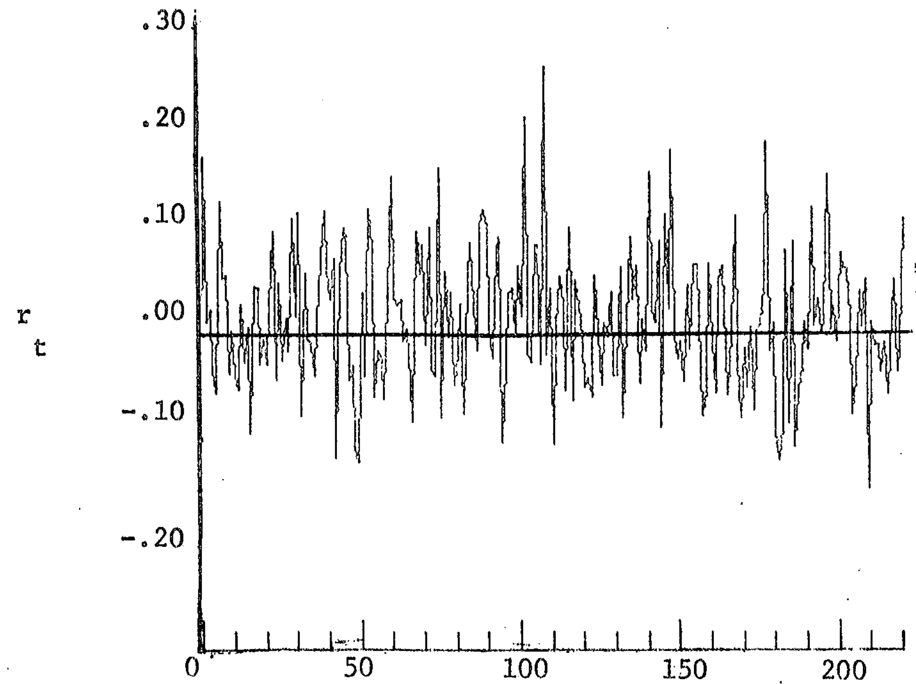


Figure 2A

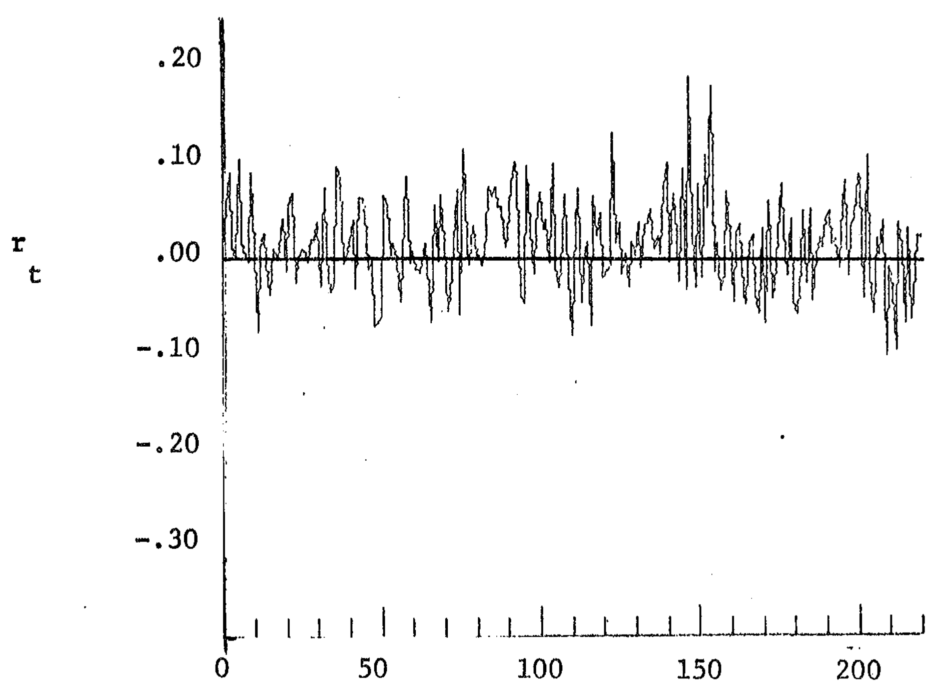


Figure 2B

To illustrate the procedure we have constructed Figures 2a and 2b from actual data. These figures are very similar in nature to Figure 1b. They chart the time series of the continuously compounded monthly rates of return (1946 through 1966) for two portfolios of common stock, A and B respectively.[3] It would appear from casual observation that, at least over this interval, we are indeed sampling from an underlying probability distribution of some form which is stationary in its mean and variance.[4] Obviously, portfolio A exhibits the greater variability, but it is difficult to determine visually which has the greater average rate of return. To test for a risk-return relationship over a large number of portfolios, one need only repeat this procedure over a broad cross section, and observe the nature of any systematic relationship which exists between average rates of return and other parameters of the sample distributions typically associated with risk.

While this technique has been widely used and has a great deal of intuitive appeal, it suffers from a number of potential problems, at least one of which is unavoidable. These problems are explored in detail in the next section where we review and critically examine the available evidence supporting the existance of risk premiums in security returns.

II. Pitfalls in Sampling and Measurment

A. Selection of a Time Period

Timing is critical in any test of the relationship between risk and expected rates of return. If the assumption that investors' expectations are

---

[3]Each of the portfolios is equally invested in 25 stocks selected from the New York Stock Exchange.

[4]While these distributions appear to remain stationary over this particular interval, we shall later show that the stationarity assumption does not hold up over all intervals.

borne out on average is violated, a systematic error crops up which tends to bias the results in one direction or the other.

To prove this we need only assume, in the convention of Sharpe, Lintner and Fama, that rates of return to financial assets are bivariate normal, and thus, are linearly related to the rate of return to the market portfolio, p,

$$r_{j,i} = a_j + \beta_j p_i + e_{j,i}$$

where $a_j = E(r_f)(1-\beta_j)$ and $E(r_f)$ is the expected rate of return accruing to an issue devoid of systematic risk.[5]

Under these assumptions if we sample from the stationary distributions over a number of periods, t = 1, 2,..., n, each of which is equal to the investors' horizon[6], our sampling mean can be expressed by,

$$\bar{r}_j = 1/n \sum_{t=1}^{n} [a_j+\beta_j p_t+e_{jt}] = a + \beta_j\bar{p} + \bar{e}_j. \tag{1}$$

If the sample covariance between e_{jt} and p_t, is small, the sample estimate of systematic risk is given by,[7]

$$\hat{\beta}_j = \frac{1/(n-1) \sum_{t=1}^{n} [r_{jt}- \bar{r}_j][p_t- \bar{p}]}{\sigma_p^2} \simeq \beta_j \tag{2}$$

---

[5]This may also be equal to the rate of return accruing to an asset with zero variance. However, given the arguments presented in Section II C, one may want to consider the possibility of unequal risk free rates in the bond and stock markets.

[6]for a discussion of the problem of misspecification of the horizon, see Jensen [22, p. 186].

[7]If the risk free rate, R_f, is not constant over the sampling period, one is likely to obtain a biased estimate of β_j with (2) unless the risk free rate is subtracted in each year from both the independent and the dependent variables. Our results show, however, that the bias is likely to be small.

where σ_p^2 is the sample variance of the rates of return to the market portfolio. By substituting (1) and the approximation of (2) into the definitional form of the cross-section covariance between $\hat{\beta}_j$ and $\bar{r}_j$ we get,

$$\mathrm{Cov}_{\hat{\beta}_j, \bar{r}_j} = 1/m-1 \sum_{j=1}^{m} (\hat{\beta}_j - \bar{\hat{\beta}})(\bar{r}_j - \bar{\bar{r}}) \simeq \bar{p}\, \sigma^2_{\beta_j} + \mathrm{Cov}_{a_j, \beta_j} + \mathrm{Cov}_{\bar{e}_j, \beta_j}$$

and since $a_j = E(r_f)\ (1-\beta_j)$,

$$\mathrm{Cov}_{a_j, \beta_j} = 1/m-1 \sum_{j=1}^{m} (a_j - \bar{a})(\beta_j - \bar{\beta}) = -\,E(r_f)\ \sigma^2_{\beta_j}.$$

Thus, with $\hat{\beta}_j \simeq \beta_j$, for all j, the slope coefficient of the cross-section regression of risk on return is found to be approximated by,

$$b_\beta = \frac{\mathrm{Cov}\hat{\beta}_j, \bar{r}_j}{\sigma^2_{\hat{\beta}_j}} \quad [E(p) - E(r_f)] + [\bar{p} - E(p)] + \frac{\mathrm{Cov}_{\bar{e}_j, \beta_j}}{\sigma^2_{\beta_j}} \tag{3}$$

where E(p) is the <u>expected</u> rate of return to the market portfolio.

Since, for a representative cross section of well diversified portfolios all highly correlated with the market portfolio, we can assume that the final term of (3) is small, in a test across these portfolios the following approximation should hold up well.

$$b_\beta \simeq [E(p) - E(r_f)] + [\bar{p} - E(p)].$$

The first term in brackets can be taken to be the slope of the true risk-return function; the second term is the estimation error. In a bear market, where, on average, results exceed expectations, we <u>underestimate</u> the slope of the true function; in a bull market we overestimate. Thus, we are faced with the unfortunate situation that the nature of our empirical result is determined by the nature of the market we sample in.

The single means of attacking the problem is to assume stationarity in the underlying probability distributions over long periods of time and to sample over these intervals, hoping to obtain an accurate picture of the distributions by increasing the sample size. In light of this we would suggest that the timing of the empirical tests conducted thus far has been unfortunate since they have sampled a relatively small number of observations within the bullish market, 1953 through 1968.[8] It may prove interesting to sample over other long run periods when the possibility that stock returns have generally exceeded expectations is less distinct. We shall direct our attention to the design of such a test in part III.

B. The Posterior Selection Problem

The technique generally employed to sample mutual funds or portfolios may serve to bias the empirical result in favor of the risk premium hypothesis. The samples employed for the empirical analyses of Sharpe [38], Jensen [22], and Soldofsky-Miller [39] are composed exclusively of portfolios which were in existence throughout the entire period of their study. Those in existence at the inception but which terminate operations thereafter are systematically eliminated from their studies through the use of this sampling method. We suspect that a great many funds which cease operating do so because of poor performance and if mapped, might occupy positions to the northwest on the mean variance mapping of Figure 1c.

---

[8]During this period the geometric average rate of return to all stocks listed on the New York Stock Exchange was 15.09 percent.

To support our position we select at random 150 stock issues from those listed on the New York Stock Exchange in 1926. The performance of each issue is observed until it is delisted. In the month in which it leaves the exchange, it is assumed sold and the proceeds reinvested in another randomly selected NYSE stock. The same procedure is followed for the replacement. Following this process, we generate 150 series of annual rates of return through 1969 some of which are broken and some of which are continuous in the sense that they represent the rates of return to a single firm throughout the entire period. Of the 150 firms initially selected, only 60 managed to survive through 1969. The average geometric mean (the compounding interval is continuous) and the average standard deviation of the 60 "successful" firms is presented in Table 1 along with the same statistics for the 90 broken series. As a class the successful firms are characterized by greater means and smaller standard deviations than those series which were broken at least once. If "conventional" sampling procedures are employed, the firms represented by the broken series would of course, be eliminated from consideration in a cross section test of the relationship between the two variables.

TABLE 1

	$\bar{G}_R$	$\bar{\sigma}_r$
90 Broken Series	.040	.557
60 Successful Firms	.073	.264

The argument is also germane to tests of the performance of the portfolios of mutual funds. One might expect that those funds which exhibit greater year to year variability and lower average rates of return should suffer the exodus of their shareholders and eventually face termination of their operations. These funds would again be excluded from consideration under conventional sampling procedures. These contentions are supported by Nantell [34] who examines the relative performance of mutual funds which were in existence in 1954 but which failed to survive the remainder of the period examined by Sharpe. Matching each fund with another randomly selected from Sharpe's sample of "successful" funds, he observes their relative performance during the remaining period of the "unsuccessful" fund's existence. Nantell concludes that, as a group, the funds which Sharpe systematically ignores are characterized by higher standard deviations and lower average rates of return than those which he observes.[9]

The posterior selection problem is avoided if one constructs his sample from funds or securities in existence at the inception of his period of study and works forward rather than sampling at the termination and working backward. Unfortunately, given the nature of available data (see, for example, Weisenberger's Manual of Investment Companies or Standard and Poor's Compustat

---

[9]The average mean and standard deviation for the failure sample was .10 and .21 respectively. For the Sharpe sample, over identical periods, the respective figures were .12 and .16.

Tapes), it is much easier to adopt the latter procedure than the former.[10]

C. Testing Across Markets Which May Be Segmented

Others have searched for risk premiums in realized rates of return by including in their samples the instruments of both the bond and stock markets. Sharpe and Jensen measure the performance of mutual funds of the type that invest large proportions of their funds in bonds as well as those which invest exclusively in equity issues. Similarly, the risk premium curves of Soldofsky and Miller are constructed by relating the risk attributes of bonds, preferred and common stock to their respective realized rates of return.

At the same time preliminary evidence has been introduced by Farrell [12], Roll [36] and Regan [35] supporting the notion that the markets for prior and residual claims may be separate rather than fully integrated and the differentials between their respective rates of return may be explained by factors other than risk. Farrell concludes that the remarkably large difference in his measured return between equities and fixed-interest securities may be partially explained by the existence of large groups of institutions for which equities represent very poor substitutes for prior claims, and which are therefore unlikely to switch very heavily to equities, except in the most extreme circumstances. Similarly, Roll, using rather ingenious methods of

[10] In [31] Miller and Scholes replicate a cross section test of the risk-return relationship conducted earlier by Lintner over the period 1959 through 1963. Unfortunately, they restrict their sample to stocks listed continuously over the period 1946 to 1966. Incidentally, Lintner's original sample consisted of 301 firms selected from the Compustat File. The Compustate File is an extremely poor population from which to sample in an empirical test of this type. The file is designed for use by professional investment analysts and not surprisingly it is "populated" by large firms and small companies which have exhibited distinctive market performance in the past. Firms which have failed to test of survival, of course, are not represented. Hawk [17], has shown that average rates of return to the equities on the tape are far in excess of yearly rates of return to some of the more widely used market indexes. In spite of the fact that the stocks on the tape are hardly representative, the tape has been widely used in the search for risk premiums. See for example Arditti [3] and Douglas [5].

analysis to measure correlations between measure riskless and market rates and their true but unobservable counterparts, finds evidence which suggests that there may be some non-portfolio motives for holding Treasury Bills that are not accounted for by the model of SLM. Regan, in a dissertation under Haugen, using sampling and measurement procedures similar to those described in Part III of this paper, examines the ex post relationships between average rates of return and various measures of risk for portfolios of corporate bonds and stocks traded on the New York Stock Exchange between 1927 and 1968. His results reveal a dramatic difference in the ex post risk-return relationships for prior and residual claims.

Since a vast amount of capital is managed by intermediaries which operate within regulatory constraints which restrict investment in common stock, the possibility that these constraints manifest themselves in relative prices and rates of return cannot be ruled out. Let us again assume, for the moment, that expected rates of return to securities are unaffected by risk, but that institutional rigidities and regulatory constraints serve to inflate bond prices relative to stocks. In the presence of these assumptions, all securities classified as stocks might sell at prices to yield an identical expected rate of return which is higher than that associated with all bonds.

If, under these assumed conditions, we empirically measure the performance of portfolios which are exclusively invested in stocks (△ portfolios), and those invested exclusively in bonds (+ portfolios) we might report results similar to those which are reported by Regan in Figure 3.[11] In the figure we note that we appear to be sampling from two distinct and separate

---

[11]Each portfolio contains ten securities, randomly selected from those listed on the New York Stock Exchange. If for any reason a security is delisted from the Exchange it is replaced by another issue similarly selected. In the figure risk is measured as the standard deviation of the yearly rates of return is the continuously compounds geometric mean. The portfolios are annually adjusted so that an equal dollar amount is invested in each stock.

populations with respect to risk and return. As Regan reports, the two clusters have statistically signifcant differences in intercept and slope. It would clearly be inappropriate to regress risk on return across both clusters. It would hardly be more appropriate to fill the void by constructing hypothetical portfolios which balance their investments in bonds and stocks, run the regression, and then presume that the derived relationship adequately describes the relationship which exists between risk and return within the population of stocks and within the population of bonds. In effect, however, this is precisely what has been done in most studies of risk-return relationships in the mutual fund industry.

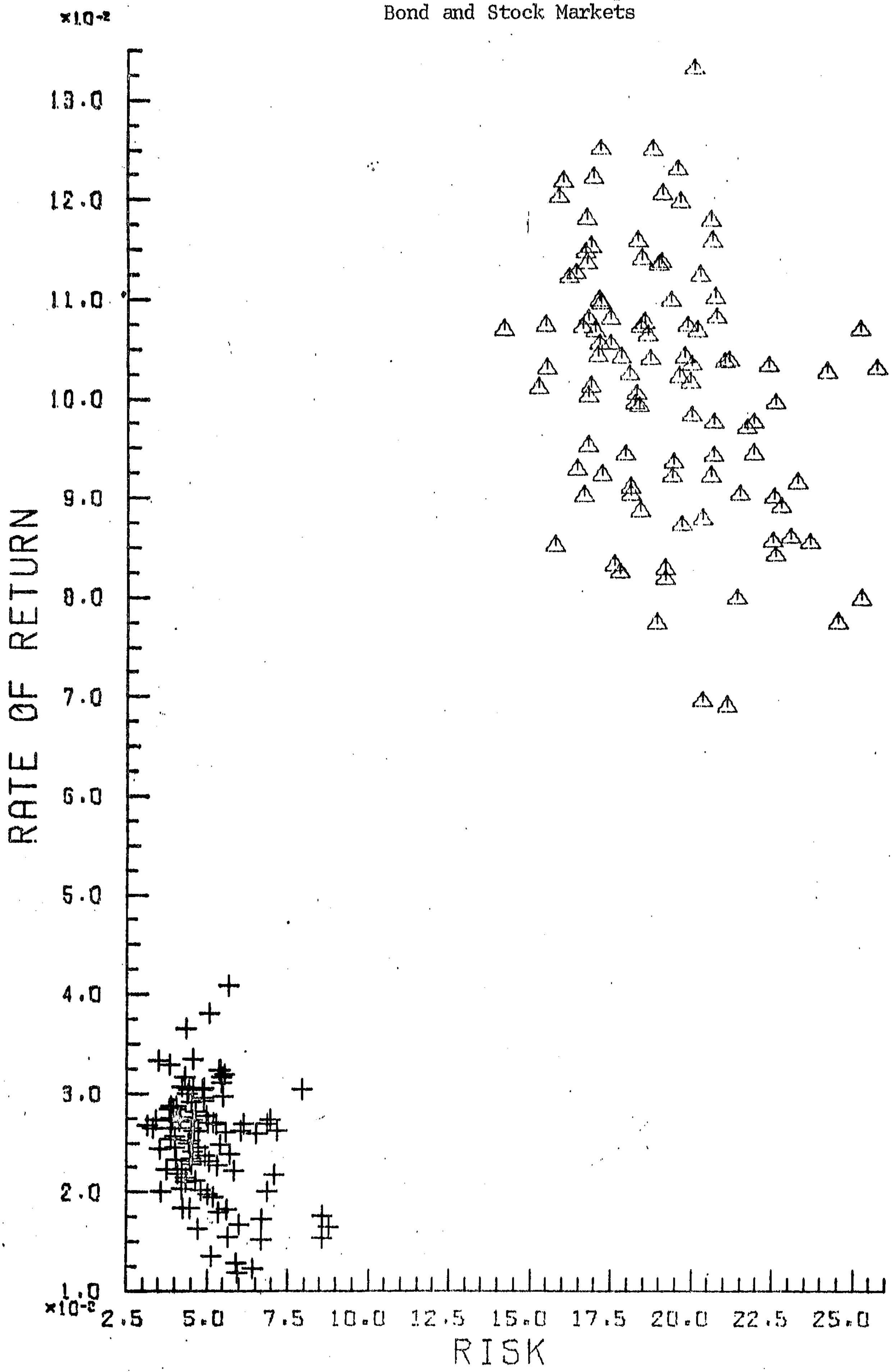
One may question the extent to which the empirical results of these and other studies rest on observations taken from the bond market. After deleting the income and balanced mutual funds from the studies of Sharpe[12] and Jensen and the bonds and preferred stock from the work of Soldofsky and Miller, we are left with the results of Table 3. Deleting the observations based on bond performance not only removes the significant positive relationships between risk and return but, in the case of Jensen and Soldofsky and Miller, leaves significant negative relationships in their place. Taken in this light, it is difficult to find support for the risk premium hypothesis in these results, at least as it relates to the stock market. Given the results, with what confidence can we say that an increase in the risk of an <u>equity</u> is likely to be accompanied by an <u>increase</u> in the expected rate of return?

---

[12] Sharpe originally regressed the arithmetic mean, $\bar{r}$, on σ_r, however as has been pointed out first by Merton Miller and then by Jensen [22, p. 220] the covariance between samples estimates of the mean and variance of any distribution can be directly related to the third moment of that distribution. Since the distributions of stock returns can be shown to be positively skewed, large <u>sample</u> means can be expected to be associated with large <u>sample</u> variances even though there may be no such relationship in the underlying probability distributions. To correct for this spurious correlation problem we have logged the yearly price relatives of the mutual funds examined by Sharpe. The funds examined include all these in the original sample which constrained investment in fixed income securities to less than one-third of the value of the portfolio.

Figure 3

Risk Return Relationships in the Bond and Stock Markets



D. The Problem of Non-Stationarity

In statistical inference the characteristics of an unknown population are inferred from the characteristics of an observed sample. If by necessity the sample must be taken over long periods of time, one must make the working assumption, as we have done, that the population parameters are stationary over the sampling interval. If this assumption is violated the sample is likely to produce an inaccurate picture of the underlying probability distribution in existence at any point within the interval.

If, for example, we unknowingly sample from a normal distribution which is non-stationary in its variance, we are likely to infer that we are sampling from a non-gausian distribution which is heavy tailed. Moreover, if the distribution is non-stationary in its mean as well as its variance, we may infer the existence of distributions of any number of peculiar forms even though the actual populations may be purely gausian in character.

It is unfortunately the case that we cannot count on our assumption of parameter stationarity to hold up over the long run. This is strikingly revealed in Figure 4 where we chart the time series of the continuously compounded monthly rates of return (February 1926 - December 1971) to all common stock issues on the New York Stock Exchange using Fisher's Link Relative Index. The distribution appears to be non-stationary in its variance at least over the period February 1926 through 1945. Casual observation, however, would suggest that the stationarity assumption holds as a reasonable approximation over the interval, January 1946 through December 1971. We should note here that while the distribution of monthly returns exhibits some degree of stationarity over this period, the distribution of daily returns does not. This fact is revealed in Figure 5 which charts the time series of continuously compounded daily rates of return (September, 1959 - September,

TABLE 3

Regressions, Risk and Return
(T Ratios in Parentheses)

Sharpe

$\bar{r}^* = a + b\sigma_r$

Variant	b	R^2
All 34 Funds 1954-1963	.474 (8.61)	.70
Common Stock Funds Only 1954-1963	.146 (.81)	.04

Soldofsky and Miller

$G_r = a + b\sigma_r$

Variant	b	R^2
All Securities 1950-1966	.878 (11.01)	.91
Common Only 1950-1966	-.869 (-4.31)	.82

Jensen

$\bar{r}^* = a + b\beta$

Variant	b	R^2
All 115 Funds 1955-1964	.243 (3.28)	.09
Common Stock Funds Only 1955-1964	-.259 (-2.24)	.06

where:

$$\bar{r}^* = 1/n \sum_{i=1}^{n} r_t^*,$$

$$r^*_t = \text{loge } (1+r_t)$$

$$G_p = \left(\prod_{t=1}^{n} (1+r_t)\right)^{1/n} - 1$$

$$\beta = \frac{1/n-1 \sum_{t=1}^{n} (r_t^* - \bar{r}^*)(p^*_t - \bar{p}^*)}{\sigma_{p^*}^2}$$

p^*_t = the logged rate of return to the Standard and Poor's Composite Index.

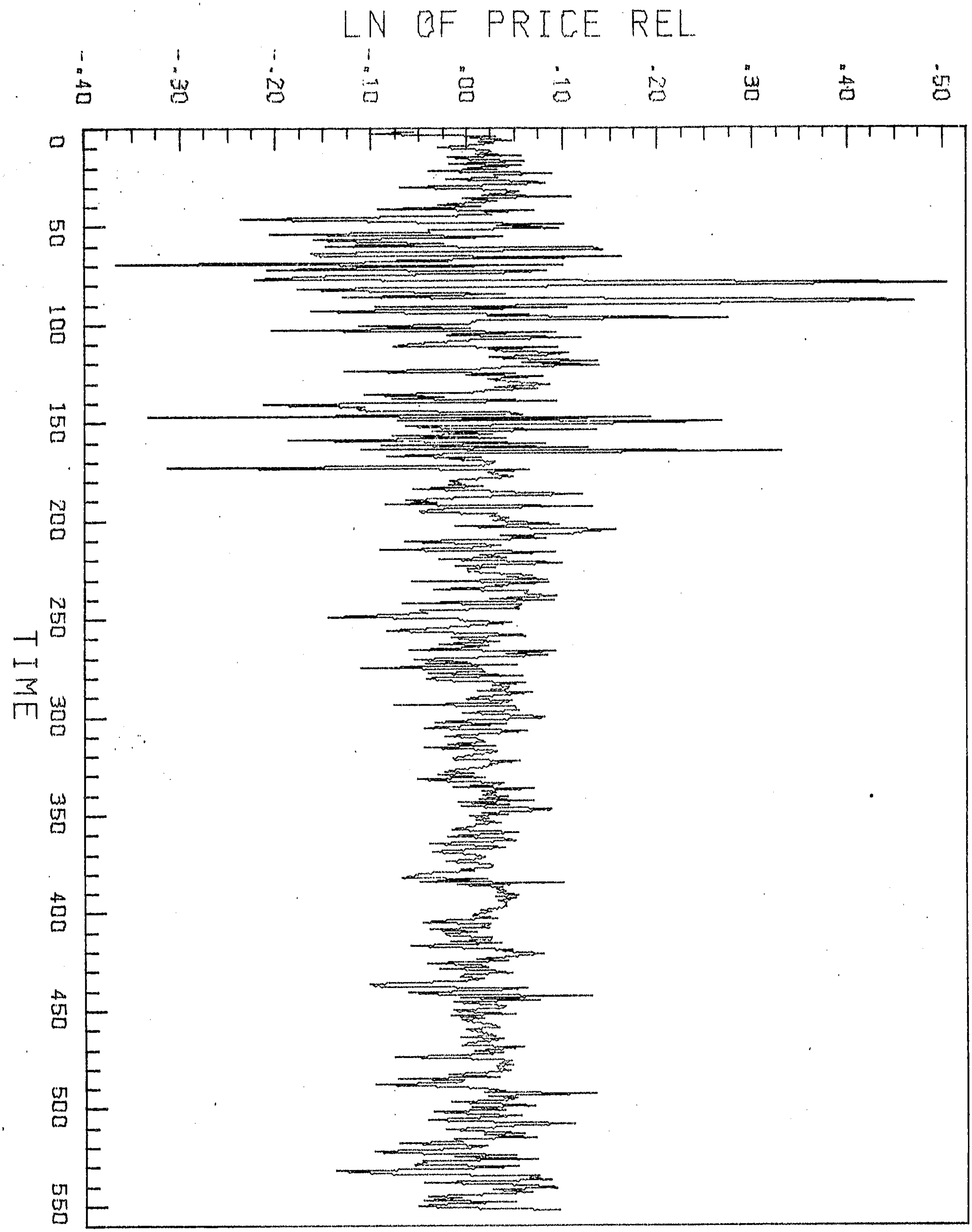

LN OF PRICE REL
-.40
-.30
-.20
-.10
.00
.10
.20
.30
.40
.50
0
50
100
150
200
250
300
350
400
450
500
550
TIME

Figure 4

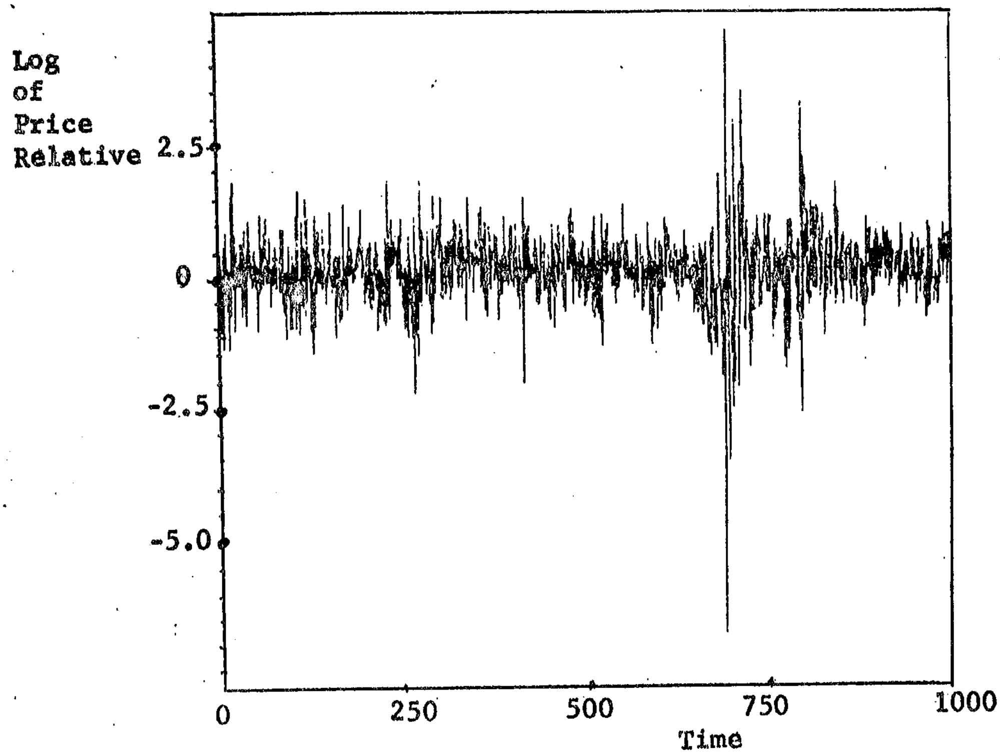


Figure 5

1963) to the Standard and Poor's Composite Index of 500 stocks. Examination of this same series over the total interval 1926-1971 reveals that the variance of the series changes repeatedly. Moreover, dramatic increases in the variance appear to accompany dramatic downward movements in the time series of stock prices.

It is difficult to specify the effect which this type of nonstationarity may have on a cross section test of the risk-return relationship. Examination of the data points to the conclusion that the change in variance is a phenomenon which simultaneously sweeps the distributions of all stocks. While there is no apriori reason to believe that such a violation of the basic assumption of the empirical test will serve to bias the results in one direction or another, one must at least allow for the possibility of distortion in the estimate of the underlying function.

III. Continuing the Search for Risk Premiums

A. Design of the Test

In order to resolve the posterior selection problem, we construct sample portfolios from stocks selected from those listed on the New York Stock Exchange in 1926.[13] No attempt is made to pre-screen the stocks to assure their survival over the period observed.[14] Each portfolio consists of 25 stocks, and 114 such portfolios are constructed.[15] Monthly performance relatives are calculated for each portfolio from February 1926 to December 1971 by taking the arithmetic mean of the performance relatives for the 25 stocks in each portfolio. This is tantamount to assuming that a given number of dollars is divided equally among each of the 25 stocks in a portfolio at the beginning of a month and held until the end of the month, at which time the value of the portfolio (with distributions and adjustments) determines the performance relative for the portfolio. If a stock in any portfolio is de-listed from the Exchange, for any reason, a new stock is selected to take its place in the portfolio at the time of de-listing.

From the monthly performance relatives for the 114 portfolios, we

---

[13] The tapes which were originally developed by the Center for Research on Security Prices (CRISP) at the University of Chicago, were revised and updated by the Standard and Poor Company. The source of all our data is the CRISP tapes.

[14] A comparison of the breadth of the risk spectrum exhibited by our portfolios and that exhibited by portfolios held by mutual funds, reveals little, if any, discernable difference.

[15] The number of portfolios is determined by constraints on computer time. In measuring the rates of return to all the portfolios, the compounding interval is continuous.

calculate the geometric mean of the monthly rates of return and standard deviation of the monthly returns over the entire 46-year period and nine shorter periods of five years. (The last period is six years, to be precise.) The period 1946-1971 is of special interest because it is a period over which the variance of the monthly performance relatives is relatively constant and the problem of non-stationarity noted earlier is apparently at a minimum.

We also make separate calculations of the β coefficent for each portfolio (when regressed on the average return for all stocks on the CRISP tapes), in addition to the coefficients of skewness and kurtosis. Since the portfolios are well diversified, residual variance can be taken to be small relative to the portfolios own variance.[16]

B. <u>What to Expect</u>

Before proceeding to the results, we should take a moment to reconsider what one might expect in light of the alternative theories available. If, risk premiums are non-existent, the long periods should show little or no relationship between portfolio returns and portfolio standard deviations (or between β coefficients and portfolio returns). That is, if:

$$E(r_i) = E(r_j) = E(p) = E(r_f), \text{ for all } i \text{ and } j$$

and:

$$\bar{p} = E(p)$$

---

[16]This is confirmed by the fact that the product moment correlation coefficient between the rates of return to the portfolios and the average return for all stocks falls within the range .90 to 1.00 over all the 114 portfolios. In every case we correlate the rates of return as opposed to the risk premiums ($r_{jt}-r_f$ and p_t-r_f; where r_f is the yield to maturity on a one period risk-free bond at the beginning of period t). While an improved estimate of the β coefficient can be obtained by relating the risk premiums as opposed to the total rates of return, Miller and Scholes [31] have shown that the difference between the estimates is small. Our analysis confirms this finding.

then:

$$\text{Cov } \bar{r}_j, \sigma^2 r_j = 0,$$

$$\text{Cov } \bar{r}_j, \hat{\beta}_j = 0$$

This merely states that if the sampling period is one over which the sample mean of market returns is close to expectations, then there should be no correlation between observed portfolio rates of return and observed variances or β coefficients.

For shorter periods, if $\bar{p} \neq E(p)$, we expect to observe that:

$$\text{when } \bar{p} > E(p), \quad \text{Cov } \bar{r}_j, \sigma^2 r_j > 0$$

$$\text{when } \bar{p} < E(p), \quad \text{Cov } \bar{r}_j, \sigma^2 r_j < 0$$

That is, during bull markets (results better than expectations), we expect to observe that portfolio rates of return are positively related to portfolio variances, and vice-versa during bear markets.

C. The Results

Our principal results are shown in Tables 4 and 5 and Figures 6, 7 and 8. Table 4 and Figures 6 and 7 show the regression relationships between sample means and standard deviations and β coefficients for two long periods 1926-71, and 1946-71. The significance of the latter period becomes manifest by referring to Figure 8 in which we show a composite scatter diagram of returns and standard deviations for all five-year periods. The cluster in the bottom center represents the results for all five-year periods between 1946 and 1971. It is self evident that the variance of monthly returns is relatively stable over this

TABLE 4

Regression Statistics for Long Periods

Statistics	1926-71	1946-71
	Geometric means regressed on standard deviations	
b_σ coefficient	-.0353	-.0945
T-ratio	-4.87	-5.82
	Geometric means regressed on $\hat{\beta}$ coefficients	
b_β coefficient	-.0031	-.0043
T-ratio	-4.79	-5.42
Average of portfolio returns	.0094	.0089
Average of Standard Deviations	.0893	.0460

TABLE 5

Comparison of Simple Regression of Portfolio Rates of Return on Portfolio Variances for Five-Year Periods with Changes in Market Performance

Period	Regressions		Average* Market Rate	Change from Last Ten Years
	b_σ-coef	T-ratio		
1926-30	-.035	-0.73	-.0048	-.0003**
1931-35	-.045	-1.70	.0173	-.0063**
1936-40	-.098	-5.58	.0012	-.0049
1941-45	+.212	+7.12	.0258	+.0166
1946-50	-.162	-4.08	.0063	-.0071
1951-55	-.234	-4.46	.0142	-.0018
1956-60	-.101	-2.55	.0085	-.0023
1961-65	+.292	+6.35	.0119	+.0006
1965-71***	-.041	-1.33	.0072	-.0040

*Geometric mean of monthly performance with all stocks on the CRISP tapes.
**Change from previous ten years using the continuously compounded monthly average rate of growth in the Standard and Poor's Index of 425 industrial stocks for both the earlier and later periods.
***Six-year period

Figure 6

1926-71

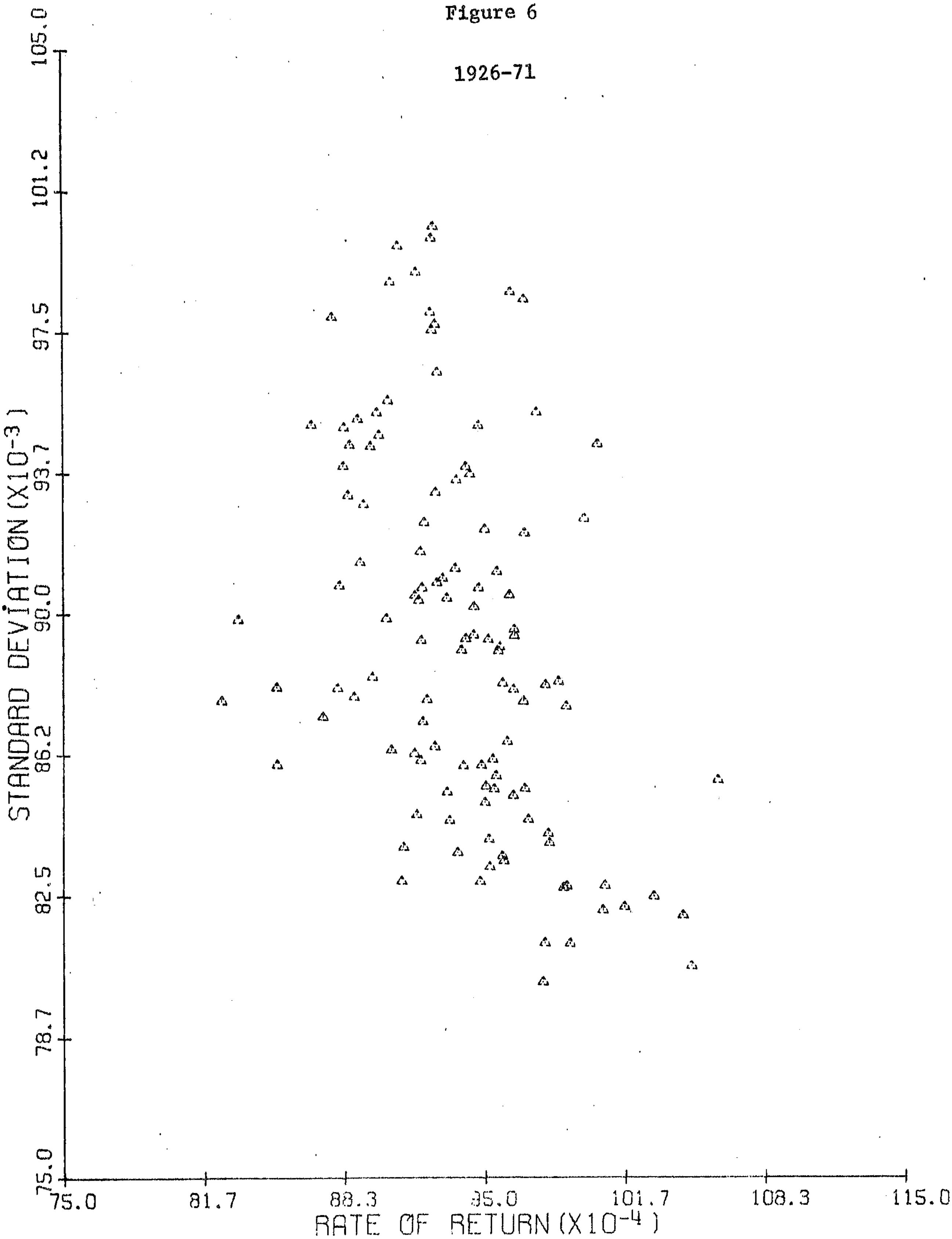


Figure 7

1946-71

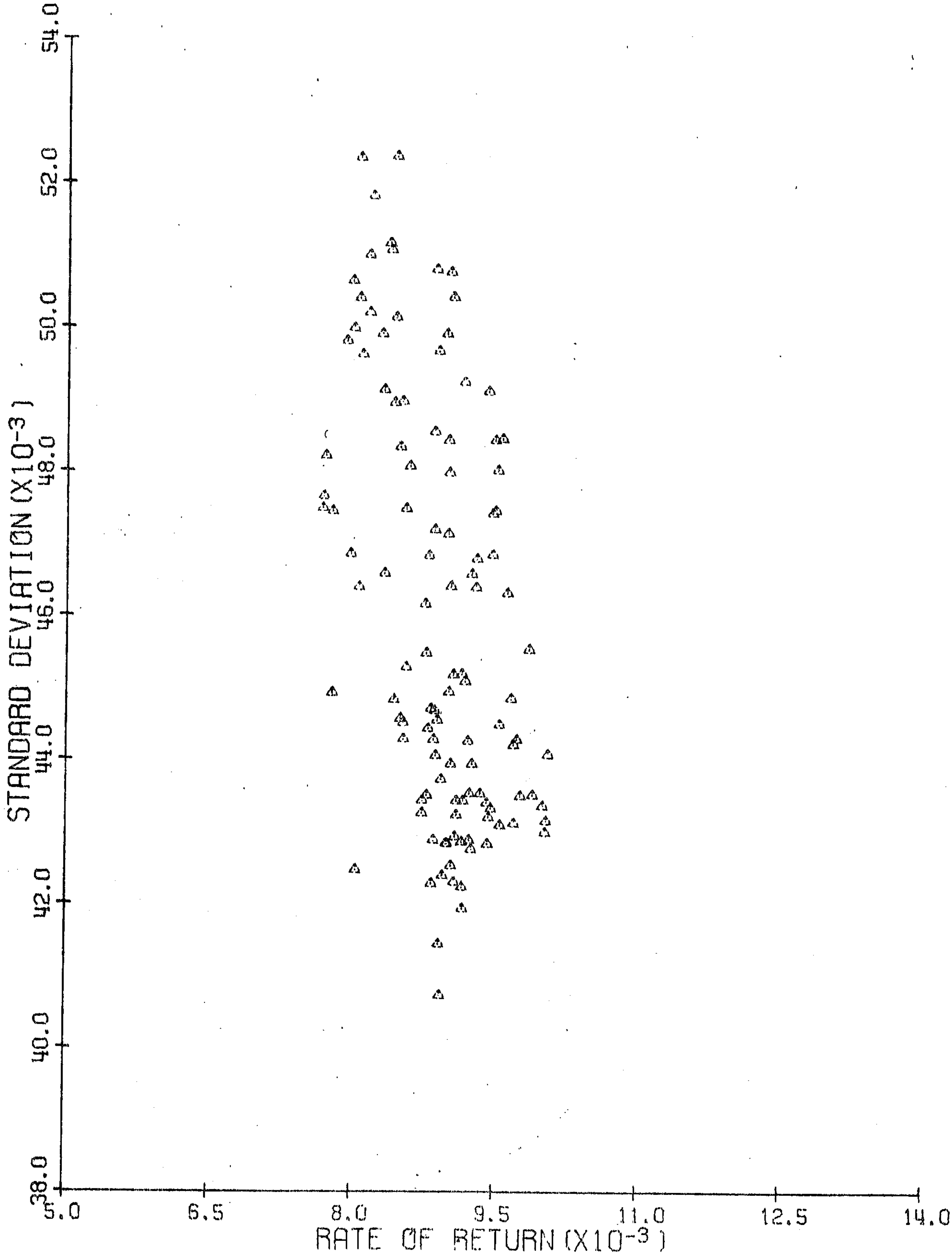
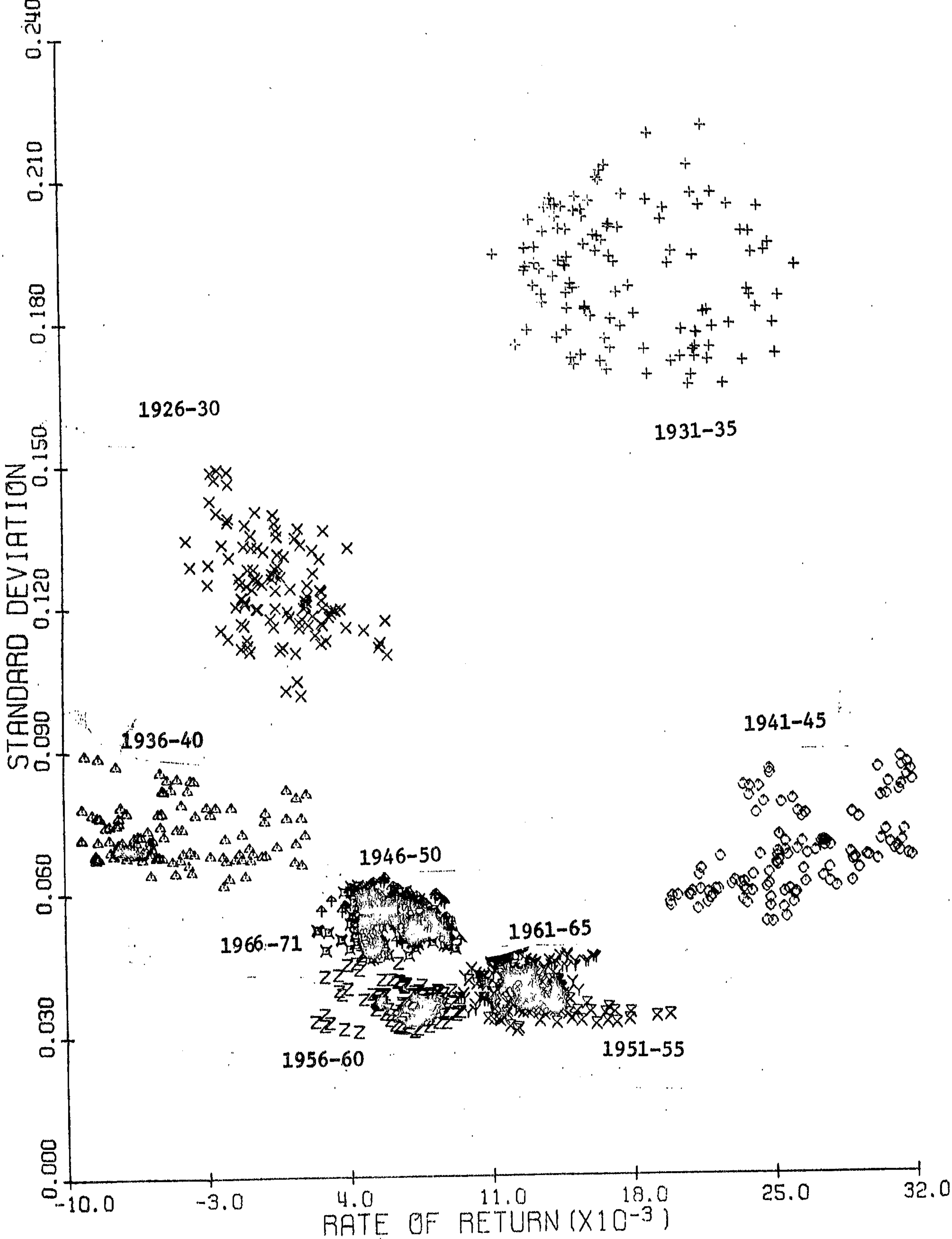


Figure 8


0.240
0.210
0.180
0.150
0.120
0.090
0.060
0.030
0.000
STANDARD DEVIATION
-10.0
-3.0
4.0
11.0
18.0
25.0
32.0
RATE OF RETURN (X10^-3)
1926-30
1931-35
1936-40
1941-45
1946-50
1951-55
1956-60
1961-65
1966-71

period when compared with the other periods represented.[17]

The effects of the time-period problem are apparent when looking at the results for the five year periods, Table 5 shows the results of regressions of returns on standard deviations for each of the periods. In addition, in each case we compare the average performance in the market with market performance over the ten years preceeding the period in question. The latter comparison is significant because it may provide some notion of how market performance in any period compared with expectations based on past performance. Note that in all cases, when the market performance during a period exceeded the previous ten-year performance (indicated by a + sign in the last column), the relationship between returns and standard deviations is positive. And, when market performance was less than the previous period, the relationship between returns and standard deviations is negative. These results coincide strongly with the bull-bear market hypothesis.[18]

We note one other result. Earlier we pointed out that sampling from a distribution that has non-stationary variance may lead

---

[17]Because of the attractiveness of the stable period 1946-71, we might note a few more statistics for this period. The arithmetic mean of portfolio returns is also negatively related to the standard deviation (T-ratio = -2.97). And, a multiple regression of the geometric mean on the standard deviation, skewness and kurtosis yields the following (T-ratios in parentheses):

$$\bar{r} = .015 - .098\sigma_r - .039SKEW - .004KUR$$
$$(11.33)\ (-6.01)\ (-0.48)\ (-1.43)$$

[18]Earlier we referred to a independent experiment in which we construct sample portfolios from a much smaller sample of 150 stocks selected at random from those listed on the New York Stock Exchange in 1926, and compute annual rates of return for the period 1926-69. All of the results of that effort are consistent with results presented here. The relationship between annual rate of return and standard deviations is negative over the long periods, and alternatively positive and negative during shorter periods of either bull or bear markets.

to the inference that the distribution is a non-gaussian distribution that is heavy tailed or peaked. This shows up clearly in our calculations of the kurtosis of the distribution of returns for each of the portfolios. Over the period 1946-71, a period of stable variance, kurtosis values for the 114 portfolios ranged from 2.7 to 3.9. For the period 1926-71, a period during which the variance of stock returns varied widely as shown in Figure 8, kurtosis values for the portfolios ranged from 12 to 25 indicating abnormally peaked - heavy tailed - distributions. These results tend to indicate that the distribution of monthly rates of return may not be heavy tailed, but rather characterized by shifting variance over time.[19]

---

[19] This possibility was originally suggested to us by Der-Ann Hsu (Princeton) and Dean W. Wichern (University of Wisconsin).

IV. Summary and Conclusions

Our purpose here has been twofold. First, we point out the conceptual shortcomings of previous empirical efforts that generally support the concept of a risk premium and measure the trade-off between risk and return. Uncontained, each of the four problems outlined in Part II of this paper is capable of generating the inference that premiums are awarded for risk-taking in the stock market when in fact no risk premiums exist.

Second, we design a method of empirically measuring the risk-return relationship that circumvents the four problems. The results of our empirical efforts do not support the conventional hypothesis that risk - systematic or otherwise - generates a special reward. Indeed, our results indicate that, over the long run stock portfolios with lesser variance in monthly returns have experienced greater average returns than their "riskier" counterparts.

The implications of all this are not abundantly clear. But, it would seem that the search for new theories of relative asset pricing in the face of the differential attributes characterized as "risk" is not over.

BIBLIOGRAPHY

1. Arditti, F.D., "Risk and the Required Return on Equity," The Journal of Finance, XXII (March 1967), pp. 19-36.

2. Black, F. "Capital Market Equilibrium with Restricted Borrowing." Journal of Business, Vol. XXXXV, (July 1972), pp. 444-54.

3. Black, F., Jensen, M.C., and Scholes, M. "The Capital Asset Pricing Model: Some Empirical Tests," in M.C. Jensen, ed., Studies in the Theory of Capital Markets [23].

4. Briscoe, G., J.M. Samuels and D.J. Smyth, "The Treatment of Risk in the Stock Market," The Journal of Finance, XXIV (September 1969), pp. 707-713.

5. Douglas, G.W. "Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency." Yale Economic Essays, Vol. 9 (Spring 1969), pp. 3-45.

6. Duesenberry, James S., Income, Saving and the Theory of Consumer Behavior, (Harvard University Press, Cambridge, Massachusetts, 1967).

7. Evans, John L. and Stephen H. Archer, "Diversification and the Reduction of Dispersion: An Empirical Analysis," The Journal of Finance, XXIII (December 1968), pp. 761-767.

8. Fama, Eugene F., "Risk, Return and Equilibrium: Some Clarifying Comments," The Journal of Finance, XXIII (March 1968), pp. 29-40.

9. _____. "Multiperiod Consumption-Investment Decisions." The American Economic Review, Vol. 60, No. 1 (March 1970), pp. 163-74.

10. _____. "Risk, Return and Equilibrium." Journal of Political Economy, Vol. 79, No. 1 (January/February 1971), pp. 30-55.

11. _____ and MacBeth, J. "Risk, Return and Equilibrium: Empirical Tests," Unpublished manuscript, University of Chicago, July 1972.

12. Farrell, M.J., "On the Structure of the Capital Market," The Economic Journal, (December 1962), pp. 830-844.

13. Fisher, L., "Outcomes for Random Investments in Common Stock Listed on the New York Stock Exchange," The Journal of Business, XXXVIII (September 1965), pp. 148-161.

14. _____ and James H. Lorie, "Rates of Return on Investments in Common Stocks: The Journal of Business, XXXVII (January 1964), pp. 1-21.

15. _____ and James H. Lorie, "Rates of Return on Investments in Common Stocks: The Year by Year Record, 1926-65," The Journal of Business, XVI (July 1968), pp. 291-316.

16. Fraine, Harold G. and Robert H. Mills, "Effect of Defaults and Credit Deterioration on Yields of Corporate Bonds," The Journal of Finance, XVI (September 1961), pp. 423-439.

17. Hawk, Stephen, "Sampling Bias on the Computstat Tape", unpublished paper, University of Wisconsin, 1973.

18. Hirschleifer, J., "Investment Decision Under Uncertainty: Choice-Theoretic Approaches," Quarterly Journal of Economics, LXXIX, (November, 1965), pp. 509-536.

19. Hersog, John P., "Investor Experience in Corporate Securities: A New Technique for Measurement," The Journal of Finance, XIX (March 1964), pp. 46-62.

20. Hickman, W. Braddock, Corporate Bond Quality and Investor Experience, (Princeton University Press, 1959).

21. Jacob, N. "The Measurement of Systematic Risk for Securities and Portfolios: Some Empirical Results." Journal of Financial and Quantitative Analysis, Vol. 6. No. 1 (March 1971), pp. 815-34.

22. Jensen, Michael C., "Risk, the Pricing of Capital Assets and the Evaluation of Investment and Portfolios," The Journal of Business, XVII (April 1969), pp. 167-247.

23. _____. "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios. Journal of Business, Vol. 42, No. 2 (April 1969), pp. 167-247.

24. _____, ed. Studies in the Theory of Capital Markets. New York: Praeger Publishers, forthcoming.

25. Levy, Haim, "A Utility Function Depending on the First Three Moments," The Journal of Finance, XXIV (September 1969), pp. 715-719.

26. Lintner, John, "Security Price, Risk and Maximal Gains from Diversification," The Journal of Finance, XIX (December 1965), pp. 587-615.

27. _____. "The Valuation of Risk Assets and the Selection of Risky Investments on Stock Portfolios and Capital Budgets," The Review of Economics and Statistics, XLVII (February 1965), pp. 13-37.

28. _____. "The Aggregation of Investors' Diverse Judgment and Preferences in Purely Competitive Securities Markets." Journal of Financial and Quantitative Analysis, Vol. 4, No. 4 (December 1969), pp. 347-400.

29. Litzenberger, Robert H., "Equilibrium in the Equity Market Under Uncertainty," The Journal of Finance, XXIV (September 1969), pp. 663-671.

30. Markowitz, Harry, "Portfolio Selection,: The Journal of Finance, VII (March 1952), pp. 77-91.

31. Miller, M. H. and Scholes, M. "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings," in M.C. Jensen, ed., Studies in the Theory of Capital Markets [23].

32. Mossin, J., "Equilibrium in a Capital Asset Market," Econometrica, (October 1966), pp. 768-783.

33. _____. "Security Pricing and Investment Criteria in Competitive Markets." The American Economic Review, Vol. 59, No. 5 (December 1969), pp. 749-56.

34. Nantell, Tim. "Sampling Bias in the Empirical Tests of the Sharpe-Lintner Mossin Capital Asset Pricing Model", unpublished manuscript, University of Wisconsin, 1971.

35. Regan, William. "The Segmented Nature of the Bond and Stock Markets", unpublished Ph.D. dissertation, University of Wisconsin, 1972.

36. Roll, Richard, "Bias in Fitting The Sharpe Model to Times Series Data," The Journal of Financial and Quantitative Analysis, (September 1969).

37. Sharpe, William F., "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," The Journal of Finance, XVIII (September 1964), pp. 425-442.

38. Sharpe, William F., "Risk Aversion in the Stock Market, Some Empirical Evidence," The Journal of Finance, XIX (September 1965), pp. 416-422.

39. Soldofsky, Robert M. and Roger L. Miller, "Risk Premium Curves for Different Classes of Long-Term Securities, 1950-1966," The Journal of Finance, XXIV (June 1969), pp. 429-445.

40. Tobin, J., "Liquidity Preference as Behavior Towards Risk," The Review of Economic Studies, XXVI (February 1958), pp. 65-86.

41. Arthur Wiesenberger Services, Investment Companies: Mutual Funds and Other Types, (New York: Arthur Wiesenberger Services).