STOCK RETURNS AND DIVIDEND YIELDS: SOME MORE EVIDENCE

by

Marshall E. Blume

Working Paper No. 1-79a

Revised

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

University of Pennsylvania

The Wharton School

Philadelphia, Pa. 19104

I. Introduction

A recent survey of 1041 individual investors[1] revealed a strong preference for dividend payout, even if retained earnings were reduced. If the proportion of corporate earnings paid out as dividends were to increase substantially, 41.8 percent of the respondents would plan to increase their stockholdings, while only 10.5 percent would plan a reduction. Moreover, the greater the preference for dividends, the greater the investor's income.

This pattern of responses is inconsistent with the usually more favorable tax treatment of capital gains to individuals. One possible explanation is that, battered by the poor market of recent years, many investors may have reached the conclusion that retained earnings bear only a tenuous relationship to subsequent capital gains and may now believe that high dividend-yielding stocks offer, on a risk-adjusted basis, greater expected returns than low dividend-yielding stocks -- sufficiently greater to offset any tax disadvantages. This explanation would gain some support, though by no means be proved, if in the recent past, risk-adjusted returns of individual stocks were positively related to their dividend yields and if investors assessed the future, at least in part, by extrapolating the recent past.

Using data from the period before the seventies, studies such as Friend and Puckett (1964) or Black and Scholes (1974) were unable to find any significant relationship between dividend policy and total returns on a risk-adjusted basis. The purpose of this study will be to reexamine these findings and extend them into the seventies. Following a brief theoretical discussion of the relationship between dividend yields and expected returns, the main empirical results will be presented.

II. Theoretical Considerations

In the perfect world of Miller and Modigliani in which, among other things, there are no transaction costs broadly construed, no taxes, and costless information, it is now generally acknowledged that dividend policy should have no effect upon the value of equity providing the investment strategy of the firm is known. If however there were taxes such that dividends were taxed at a higher rate than capital gains and at least in the extreme case in which all investors were subject to the same tax rates, a firm's cost of equity would be expected to increase with increases in dividends. Thus, one would expect that all firms would stop paying dividends,[2] so that there would be no variability in observed dividend yield and thus no observable relationship between risk-adjusted rates of return and dividend yields, just as in the no-tax case. The principal assumption underlying this argument is that firms can reduce their dividends at no cost: subject to certain IRS restraints, a firm could, for instance, always repurchase its own stock at a minimal cost.

In the more realistic case, investors will be taxed at different rates with some paying no taxes at all. Black and Scholes (1974) have argued in this case that there could be differential dividend yields and yet still no observable relationship between these yields and risk-adjusted returns. The argument relies on "clientele" effects. Those paying no taxes would be the predominant holders of high-yielding stocks and those paying high taxes would be the predominant holders of low-yielding stocks. Obviously diversification considerations would enter; but if in some sense, the array of different dividend-paying stocks were exactly equal to that demanded by the different clienteles, one would be unable to observe any relationship between risk-adjusted returns and dividend yields. If the array were not equal to that demanded, corporations would have the incentive to alter their dividend policy until the array of different dividend-paying stocks was in line with that which was demanded.

In sum, unless there are substantial costs associated with changes in dividend policy, one would expect in an efficiently operating capital market in which firms were optimizing their stockholders' interests that there would be no observable relationship between risk-adjusted returns measured before stockholders' taxes and those dividend yields which actually prevail in the market--even if stockholders faced different tax rates. A clear finding to the contrary would suggest one or more of the following: there are substantial costs associated with changes in dividend policy; firms are not optimizing stockholders' interests; the model used in adjusting returns for risk is incorrect; or for some reason the actual performance of individual companies did not correspond to investors' prior assessments over the period of study.[3]

III. The Preliminaries

Previous empirical work on the capital asset pricing model[4] is generally consistent with a linear relationship between realized returns on common stocks in any specific cross section and the corresponding beta coefficients. However, the slopes and intercepts of the linear relationship appear to vary randomly from one cross section to another and do not always correspond to those which would be predicted by the Sharpe-Lintner model.

A simple extension of this type of relationship to include anticipated dividend yields would be

$$r_{it} = a_t + b_t \beta_{it} + c_t \delta_{it} + \varepsilon_{it}, \tag{1}$$

where r_{it} is the total realized return on stock i over period t, β_{it} is the relevant beta coefficient, δ_{it} is the dividend yield anticipated over period t, and ε_{it} is a mean-zero disturbance uncorrelated with β_{it} and δ_{it}. The coefficients a_t, b_t, and c_t are allowed to vary randomly from one period to the next and are assumed to be drawn from normal stationary distributions with respective expected values a, b, and c. The null hypothesis that expected returns are unrelated to anticipated dividend yields is that c be zero.[5]

The model, as formulated in (1), involves two critical assumptions: The first is that ε_{it} is uncorrelated with the independent variables β_{it} and δ_{it}. The economic meaning of this assumption is that there are no variables omitted from (1) except possibly variables which are orthogonal to β_{it} and δ_{it}.

The second critical assumption is that the values of a, b, and c implicit in the investors' subjective distributions of r_{it}, β_{it}, and δ_{it} did not vary in any systematic way from their true underlying values over the estimation period. If this were not the case, it might be possible to observe a relationship empirically even though investors subjectively anticipated no relationship. For example, investors might have misassessed the returns from reinvested earnings either as to level or risk, and thus have mispriced securities as a function of their dividends. As investors perceived their errors, the prices of the securities would adjust, leading to perhaps a statistically significant relationship between average realized returns and anticipated dividend yield.

Significance Tests: In the spirit of Blume and Friend (1973) and Fama and MacBeth (1973), (1) will be estimated on a specific cross section and then again on another cross section and so on. Thus will be obtained a vector of estimates for a_t, a vector for b_t, and a vector for c_t. Except in the special case in which the variance-covariance matrix of the ε_{it}'s in each cross section were proportional to the identity matrix, these estimates would generally be inefficient and the associated t-values biased. The primary tests in this paper will not use these t-values and will properly adjust for any inefficiencies.

To describe these tests, consider a vector of T estimates of the c_t's obtained from estimating (1) on each of T cross sections. The mean of these estimates, $\bar{c}$, provides an estimate of c, and the sample standard deviation of these estimates provides an estimate of the standard deviation of the distribution of the c_t's which, upon dividing by $\sqrt{T}$, gives an estimate of the standard error of $\bar{c}$, $\sigma(\bar{c})$.

It should be noted that $\sigma(\bar{c})$ takes into account not only the random fluctuations in the true c_t's from one period to the next but also any errors associated with the estimates of the c_t's, so that it properly adjusts for any inefficiencies in the estimates of the c_t's. Under the assumed distributional properties of the c_t's, the null hypothesis that c equals some constant, like k, can be tested with the t-statistic defined as the ratio of $(\bar{c}-k)$ to $\sigma(\bar{c})$. The estimates of a_t and b_t can be treated in an exactly analogous fashion.

Expected Dividend Yield: In their tests, Black and Scholes (1974) measured the anticipated future dividend yield by the ratio of the dividends paid over the previous 12 months to the price at the end of these 12 months. If dividends were extremely sticky, this ratio would be expected to yield highly accurate forecasts of the future yield. However, consider a company with an historically low dividend yield which had such a bad fourth quarter that it became clear to everybody that future dividends would have to be curtailed and thus the price dropped. The Black-Scholes measure of dividend yield would in this instance substantially overstate the expected yield.

An alternative measure of dividend yield would be the ratio of the dividends paid over the previous twelve months to the price at the beginning of these twelve months. This alternative might be more accurate than the Black-Scholes measure if companies tended to adjust dividend levels quickly to maintain a fixed payout ratio and the price-earnings ratio was relatively sticky.

This alternative measure can be further adjusted for market-wide changes in the level of dividend yields over the prior 12 months by assuming that the beginning-of-period price had increased or decreased the same percentage as the market average over these 12 months and using this adjusted price in the calculation of the dividend yield. For example, if a stock were priced at the beginning of the twelve months at \$100 and paid a dividend of \$6.60 over these twelve months, the yield before adjustment for market movements would be calculated as 6.6 percent. If the market price index had increased 10 percent over these 12 months, the adjusted yield would be calculated as 6.0 percent.

Depending upon the stickiness of dividends, this alternative measure of dividend yield may or may not yield more accurate forecasts of future dividend yields than the Black-Scholes measure. But which actually is the more accurate measure overall is ultimately an empirical question.

To answer this question, the Black-Scholes measure and the measure proposed in this paper adjusted for market movements[6] were calculated for 1927 and evaluated as forecasts against the actual dividend yield in 1928. There were 509 NYSE stocks with the necessary data in both 1927 and 1928, and all were included. The mean squared error associated with the measure proposed in this paper was 0.00144, while that associated with the Black-Scholes measure was 0.00814.

Similar mean squared errors were calculated for each pair of years beginning in 1928 and ending with the 1975-1976 pair. Including the comparison for the 1927-1928 pair, 49 mean squared errors were calculated for each of the two measures. The average of these mean squared errors for the measure proposed in this paper was 0.00257, while that for the Black-Scholes measure was 0.00540.[7]

Because of its greater forecasting accuracy, the tests of this paper will utilize as the measure of anticipated dividend yield, the ratio of dividends paid over a 12 month period to the beginning-of-period price adjusted for general market movements rather than the Black-Scholes measure. Some preliminary work for this paper suggested that the Black-Scholes measure would attenuate the relationships reported subsequently, although it would not elimate them -- a result quite consistent with measured error.[8]

The Cross-Sectional Regressions: Although monthly returns were available, the cross-sectional regressions reported in this study were estimated with quarterly returns. If there were a measurable tax effect associated with dividend yield, the effect might differ as between months or periods in which a stock went ex-dividend and those in which it did not. Since most dividend-paying stocks pay dividends quarterly, the use of quarterly returns should make the estimated regression less sensitive to any such possible differential tax effects.[9]

The first cross-sectional regression was estimated using the quarterly total returns for the first calendar quarter of 1936 and data prior to 1936 to estimate β_{it} and δ_{it}. The second cross-sectional regression was estimated using returns for the second calendar quarter of 1936 and data prior to this quarter to estimate β_{it} and δ_{it}. This process was repeated for each and every calendar quarter through the end of 1976 for a total of 164 quarterly cross-sectional regressions.

To reduce the magnitude of measurement errors and the impact of non-stationarities upon the estimates of β_t and δ_{it}, the securities were partitioned into groups; and the simple averages of the r_{it}'s, β_{it}'s, and δ_{it}'s within each group were used as the basic data in estimating the cross-sectional regressions.[10] The grouping was done as follows: for the regression run in the first quarter of 1936, all NYSE securities with continuous data from 1926 through March of 1936 were ranked from low to high by their beta coefficients as estimated over the period 1926 through 1930 using the Standard & Poor's Composite Index.[11] These ranked securities were then partitioned into five groups of an equal number of securities, give or take a security.[12] Thus, the first group contained those securities with the smallest estimates of beta and the last group those with the largest estimates.

The next step was to divide each of the five groups of securities in turn into five subgroups of roughly equal numbers of securities according to the adjusted 1935 dividend yield. This process yielded a total of 25 groups or portfolios. The beta for each portfolio as estimated over the 1931-35 period and the dividend yield as estimated in 1935, reexpressed on a quarterly basis,[13] were used as forecasts of the betas and dividend yields in the regression using the returns for the first quarter of 1936 as the independent variable.

The same process was repeated for the second quarter of 1936 using the prior ten years of data beginning in April 1927 and ending in March 1936 to form the groups and the required estimates of β_{it} and δ_{it}. This process was repeated again for each subsequent quarter through the fourth quarter of 1976.

By grouping first on betas and then on dividend yields, one would in general obtain less variability in dividend yields than if one first grouped on dividend yields and then on beta, leading perhaps to less precise estimates of the coefficients on dividend yield than possible and thus biasing the tests against finding a dividend effect. To ascertain the importance of this potential bias, all the cross-sectional regressions were reestimated in exactly the same way as above except that the grouping was done first on dividend yield and then on beta. The results turn out to be robust to the grouping technique.

IV. The Cross-Sectional Regressions

Over the 1936-76 period, these cross-sectional regressions reveal a positive and significant relationship on the average between the quarterly realized rates of return and both the beta coefficients and the anticipated quarterly dividend yields. Thus, for the portfolios grouped first on beta and then on dividend yield, the average coefficient on dividend yield for the 164 cross-sectional regressions estimated over these 41 years was 0.5232 with a t-value of 2.07, and the average coefficient on beta was 0.0215 with a t-value of 2.34 (Table 1).[14] Taken at face value, these numbers imply that for a given beta, realized quarterly returns increased an average of 0.52 percent for each 1 percent increase in the anticipated quarterly dividend yield.

The results are virtually the same for the cross-sectional regressions using portfolios grouped first on dividend yield and then on beta. The average coefficient on dividend yield was 0.5938 with a t-value of 2.45, and the average coefficient on beta was 0.0233 with a t-value of 2.55. Thus, the results appear fairly robust to the order of grouping.

The significance of the dividend yield variable does, however, vary over time. In the two decades from 1937 through 1946 and from 1957 through 1966, the average coefficients on dividend yield were positive but not significant at any usual level of significance. But in the decade from 1947 through 1956, the average coefficient on dividend yield was 0.8743 with a highly significant t-value of 4.27. In the last decade ending in 1976, the average coefficient on dividend yield was 1.1272 with a t-value of 1.93, indicating a level of significance of between 5 and 10 percent.[15]

To permit comparison to previous studies, cross-sectional regressions were estimated in which beta was the only independent variable and also in which dividend yield was the only one. In the overall period, the average coefficient on beta is insignificant in the regressions in which it is the only independent variable, but significant in the regressions in which dividend yield is included. This finding suggests some form of interaction between betas and dividend yields. While not always significant, the average coefficent on beta in any of the four decades has a larger t-value when dividend yield is included.

This interaction as well as a detailed visual examination of the data suggested that there might be a non-linear difference between dividend paying and non-dividend paying stocks. To test this possibility, a dummy variable, denoted by ξ, was added to each cross-section and the cross-sectional quarterly regressions were reestimated. The value assigned to this dummy variable for a particular cross-section was the proportion of stocks in the portfolio which were expected to pay dividends and would correspond to that which would have been obtained if, before grouping and averaging, each stock were assigned a value of 1.0 if it were dividend-paying or 0.0 if not.

When these dummy variables are included in the cross-sectional regression, the dividend effect generally becomes more significant, which is consistent with the existence of a non-linear relationship of return to dividend yield. This increase is particularly pronounced for the decade ending in 1946 when the t-value of the average of the cross-sectional coefficients on dividend yield increases from 0.04 to 3.49.[16] Yet, even for the overall period, the t-value increases from 2.07 to 4.26.[17]

Since the average coefficient on this dummy variable is negative in the overall period, it would be expected that non-dividend paying stocks would have averaged for a given beta a greater quarterly return than dividend paying stocks up to some breakeven level of dividend yield and then a lesser return. For the overall period, this level would be 2.06 percent per quarter, calculated as the negative of the ratio of the average coefficient on the dummy variable (-0.0226) to the average coefficient on the dividend yield (1.0950). A quarterly indicated dividend yield of 2.06 percent per quarter is a high dividend yield in comparison to the average indicated quarterly dividend yield for all the portfolios over these 41 years of 1.16 percent. Thus, over the entire period studied, the average quarterly returns on non-dividend-paying stocks for a given beta exceeded the quarterly returns on all dividend-paying stocks as a group. Only those stocks with extremely high dividend yields might have averaged greater returns.

The analysis by decade shows that most of the superiority of non-dividend-paying stocks to dividend-paying stocks as a group stemmed from the decade from 1937 through 1946. In this period, the breakeven level of dividend yield to overcome the negative sign on the dummy variable was 4.07 percent per quarter which compares to the average indicated dividend yield of 1.18 percent. In each of the next three decades, the breakeven level of dividend yield was approximately the same or less than the average indicated levels. For the decade ending in 1956, the breakeven level was 0.90 percent and the average indicated dividend yield was 1.50 percent; for the decade ending in 1966, the corresponding figures were 0.93 percent and 0.97 percent; and for the last decade ending in 1976, the corresponding figures were 1.02 percent and 0.98 percent.

In these three decades, there was little difference in the average quarterly returns for a given beta as between non-dividend-paying stocks and all dividend-paying stocks as a group. However, the returns on those stocks with anticipated dividend yields in excess of the mean anticipated yield tended on average to exceed for a given beta those on non-dividend-paying stocks. As a mirror image, the returns on those dividend-paying stocks with low yields tended on average to be less than those on non-dividend-paying stocks.

V. Possible Explanations

Tax effects, at least by themselves, would seem unable to explain the finding that over the thirty years beginning in 1947 and ending in 1976, the quarterly returns on non-dividend-paying stocks exceeded on average the returns on low-yielding stocks. A tax system in which dividends are taxed in aggregate at a greater rate than capital gains would imply that expected before-tax returns should increase smoothly with increases in dividend yields, not first decrease and then increase.

Specifically, if dividend yields have no effect upon expected returns except through tax effects, the standard capital asset pricing model or its zero-beta equivalent would imply the after-tax relationship

$$E[(1-\tau_c)r_{it}^c + (1-\tau_d)r_{it}^d] = \gamma_o + \gamma_1\beta_{it}, \qquad (2)$$

where r_{it}^c and r_{it}^d are the before-tax returns due to capital gains and dividends respectively, τ_c and τ_d are appropriate aggregate tax rates for capital gains and dividends respectively, and γ_o and γ_1 are constants. The beta should be on after-tax basis, but if the uncertainty attached to dividend yield is small relative to the uncertainty attached to the capital gain return, the after-tax beta will be approximately the same as the before-tax beta.[18] Thus, the subsequent empirical tests will use the before-tax beta as if it were the after-tax beta.

With some rearragement of terms, (2) can be rewritten as

$$E(r_{it}) = \frac{\gamma_o}{1-\tau_c} + \frac{\gamma_1}{1-\tau_c}\,\beta_{it} + \frac{\tau_d-\tau_c}{1-\tau_c}\,E(r^d_{it}) \qquad (3)$$

If τ_d is greater than τ_c, there would be on average a positive linear relationship between quarterly returns and expected dividend yields, holding beta constant -- a relationship which is inconsistent with the observed non-linearity between returns and dividend yield.

Another possible explanation is that investors' expectations about the levels or risks of future dividends did not properly anticipate the actual levels or risks. For example, if the future growth rates of dividends for high-yielding stocks exceeded on average those which investors expected, there could be a positive relationship between average quarterly returns and anticipated dividend yield with beta held constant, even though investors expected no such relationship. Even if true, such a finding would not imply irrational expectations on the part of investors unless it could be demonstrated that investors could have made more accurate assessments of the future on the basis of information available to them at the time than they did.

If price-dividend ratios are expected to remain unchanged, the expected quarterly return on a stock before taxes will be the sum of the expected dividend yield and the expected rate of growth of dividends, say $E(g_{it})$. Assuming that the anticipated yield δ_{it} is a good measure of $E(r^d_{it})$ and after substituting $E(g_{it})$ for $E(r^c_{it})$, (2) can be rewritten as

$$E(g_{it}) = \frac{\gamma_o}{1-\tau_c} + \frac{\gamma_1}{1-\tau_c}\beta_{it} - \frac{1-\tau_d}{1-\tau_c}\delta_{it} \qquad (4)$$

$$= \gamma_o' + \gamma_1'\beta_{it} + \gamma_2'\delta_{it},$$

where γ_o', γ_1', and γ_2' are constants implicitly defined by (4).

If there were no taxes, γ_2' would be -1.0. If there were appropriate clientele effects, as Black and Scholes (1974) have proposed, such that the aggregate effective tax rates on capital gains and dividends were the same, γ_2' would still be -1.0. If, however, τ_d exceeded τ_c, γ_2' would be somewhere between 0.0 and -1.0. As an example, if τ_d were 0.30 and τ_c were 0.15, γ_2' would be -0.82.

If investors misassessed the growth rate of future dividends as a function of dividend yield, the coefficient γ_2' would differ from -1.0 or, if τ_d were not equal to τ_c, from $-(1-\tau_d)/(1-\tau_c)$. To test whether γ_2' equals these theoretical values, various measures of subsequent growth rates of dividends were calculated and regressed upon β_{it} and δ_{it}.[19] Because of both the difficulty in measuring growth rates of dividends for stocks with anticipated yields of zero and the observed non-linearity between dividend and non-dividend paying stocks, this analysis was confined to those stocks with positive anticipated yields. Again to reduce measurement errors and to control for non-stationarities, these stocks were grouped quarterly, first on beta and then on yield, into 25 portfolios using the exact same procedure as used for the entire sample.

To visualize how the various measures of subsequent dividend growth were calculated, consider an investor who had made an initial investment of one dollar at the end of 1935 in, for instance, the highest yield portfolio of those five portfolios with the lowest betas as determined at the end of 1935 and then at the end of the first quarter of 1936 had rolled this investment, including any dividends, over into the comparable portfolio in terms of beta and yield as determined at that point in time and so on, quarter by quarter. The ratio of the dividends which he would have received in the first quarter of 1936 to those which he would have received in the first quarter of 1935, expressed as a quarterly compounded rate, provides an estimate of $E(g_{it})$ for the end of 1935. By moving time ahead one quarter to the second quarter of 1936, similar estimates of $E(g_{it})$ can be derived for an initial investment of one dollar at the end of the first quarter of 1936, and so on.[20]

Five cross-section regressions on β_{it} and δ_{it}, one for each of the various measures of growth, were estimated for each and every quarter from the end of 1935 to as close to 1976 as data would permit. As an illustration of the results, Table 2 presents estimates of γ_2' at five year intervals beginning from 1935 for the five year estimates of subsequent growth of dividends. The t-values for γ_2' are measured from -1.0; and unlike those in Table 1, these t-values are based upon standard regression theory. Similar estimates covering three, five and seven, and ten subsequent years were also calculated.

The theoretical discussion indicated that γ_2' should be -1.0 except for the possible effect of differential taxes in which case it should be somewhat closer to 0.0. The estimates of γ_2' shown in Table 2 are, except for 1935, positive and significantly different from -1.0. The same pattern holds for the other estimates of $E(g_{it})$ with a few of the earlier estimates negative and the later ones positive. These generally positive values of γ_2' are clearly inconsistent with a tax effect, but are consistent with the hypothesis that market participants often underestimated over this period the subsequent growth of dividends for high-yielding stocks relative to low-yielding stocks.

Before concluding this section, two possible technical explanations will be evaluated. First, it could be that anticipated dividend yield is acting as a surrogate for changes in risk as perceived by the market but not fully captured by the historically estimated beta coefficient. To examine this possibility, the betas used in the cross-sectional regressions in Table 1 were compared to betas estimated subsequently to the original estimation period. These subsequent betas were estimated using the portfolio returns which an investor would have actually experienced if he had invested in a particular dividend-yield quintile within a particular beta quintile immediately after the historical estimation period and every quarter thereafter had rolled his resulting investment, including dividends, into the corresponding beta-dividend portfolio of the next grouping period. Repeating this process for twenty quarters would yield sixty monthly portfolio returns which were then regressed on the market index.

The subsequent betas were estimated for every group of portfolios from those grouped on data ending in December 1935 through those grouped on data ending in December 1971, the last date with sufficient subsequent data. Because of the possible non-linearity between dividend and non-dividend paying stocks, similar types of betas were also estimated for those portfolios used in the analysis of subsequent dividend growth which contain only stocks with non-zero anticipated yields. The qualitative results were the same.

For space reasons, only one set of these pairs of historically estimated and subsequently estimated portfolio betas will be presented and this set was arbitrarily picked as the one based upon the most recent data available for the entire sample. (Table 3). The remaining sets will be summarized through regression analysis.

The most striking characteristic of these data is the negative relationship within any of the five beta groups between the dividend yields and the beta coefficients which were used in the cross-sectional regressions. This negative relationship is particularly surprising since in the grouping period all five portfolio within each beta group were picked so as to have similar portfolio betas. Indeed, the portfolio betas as calculated from the grouping period, 1962-1966, indicate that there was nothing peculiar in the grouping period which might explain this negative relationship. Moreover, the changes in beta from 1962-1966 to 1967-1971 do not appear consistent with the usual order bias as discussed in Blume (1975). To provide a formal test of this visual relationship, the following regression was run on the total sample using the data from Table 3:

$$\beta_{i2} - \beta_{i1} = \underset{(3.23)}{0.112} - \underset{(4.18)}{30.31}\,(\delta_i - \bar{\delta}_k),\ \bar{R}^2 = 0.41, \qquad (5)$$

where β_{i2} is the beta used in the cross-sectional regression, β_{i1} is the beta implied by the grouping period, δ_i is the anticipated dividend yield for portfolio i, $\bar{\delta}_k$ is the average anticipated yield within a beta group, and the numbers in parentheses are t-values. Similar regressions were estimated for every quarter going back to the portfolios formed on the 1926-1930 betas on both the dividend-paying and total samples. The average coefficient of $(\delta_i - \bar{\delta}_k)$ over all the quarterly regressions was -15.60 for the total sample and -10.94 for the dividend-paying sample.[21]

Of greater relevance for this paper is the behavior of the beta coefficients as between those used in the cross-sectional regressions and the subsequent betas. To examine this behavior, regressions of the same form as (5) were run quarterly except that the independent variable was replace by $\beta_{i3} - \beta_{i2}$, where β_{i3} is the subsequent beta for portfolio i. The average coefficient over all the 145 quarterly regressions was -4.10 for the total sample and -4.24 for the dividend-paying sample.[22] This relationship is less pronounced than that observed between the grouping and estimation betas, and could well be insignificant. However, if taken at face value, these generally negative coefficients imply that higher yielding stocks tended to have lesser subsequent betas and lower yielding stocks, greater subsequent betas. This relationship is, of course, the reverse of that required to explain the results of this paper.

A second possible technical explanation of the results is associated with Roll's observation (1977) that tests involving the capital asset pricing model may produce biased results if the market return of all assets is approximated by the stock market return. In the special case in which dividend yields are uncorrelated with the betas calculated with the returns on the stock market and also uncorrelated with the true betas calculated with the returns on the true market, no bias would be introduced.[23] However, the empirical results would suggest that the betas and dividend yields used in the regressions are negatively correlated.[24]

If the stock betas and dividend yields were negatively correlated, one could obtain virtually any bias with appropriate assumptions as to the relationship of the returns of individual stocks to the returns on the market of all assets other than stock. Nonetheless, under some plausible assumptions as to this relationship, the estimated coefficients on dividend yield would be unbiased. For example, if the beta coefficients calculated with respect to the stock market were a linear function of the beta coefficients calculated with respect to the entire market, there would be no bias in the estimates of the coefficients on dividend yields. The estimates of the intercept and the coefficient on beta however would typically be biased. Such a linear relationship between the stock and true betas is consistent with the two-factor models which appear to describe stock returns quite well.[25] Thus, the use of the stock market as a proxy for the entire market has probably not introduced any substantial bias in the estimates of the effect of dividend yield.

VI. Summary

The analyses in this paper suggest a considerably more complicated relationship between returns realized on common stocks and dividend yield than has been revealed in prior work. In the three decades ending in 1976, those stocks with anticipated yields in excess of the mean market-wide yield outreturned non-dividend paying stocks at each level of beta. This finding helps to rationalize the survey results reported in the Introduction that many individuals currently prefer increased dividends with a corresponding reduction of retained earnings when tax considerations might suggest the opposite.

Throughout the forty-one years ending in 1976, the risk-adjusted returns on dividend-paying stocks increased monotonically with anticipated dividend yield. Over the thirty years ending in 1976, the average returns on all dividend-paying issues were about the same as on non-dividend paying issues. Over the 1937-46 period, the total returns on non-dividend paying stocks tended to exceed on average the returns on most dividend paying stocks.

Various possibilities were considered in an attempt to explain the monotonic relationship of the risk-adjusted returns on dividend paying stocks to anticipated dividend yield. Of those considered, the most plausible one is that the market failed to anticipate the greater relative growth of dividends for high yielding stocks as compared to lower yielding stocks. As pointed out in the text, this result, even if true, is not sufficient to conclude that market expectations were irrational.

The empirical results of this paper have been interpreted within the context of the Sharpe-Lintner capital asset pricing model in which expected returns of individual assets are solely a function of their beta coefficients. Another interpretation of the results is that the Sharpe-Lintner model is too restrictive and that dividend yield is acting a surrogate for some unspecified variables omitted from this model. Indeed, there is a growing body of evidence consistent with this proposition.[26] Whether this more general interpretation of the results is warranted must await further theorectical and empirical work.

Footnotes

[1]Blume and Friend (1978).

[2]Merton Miller (1977) has recently argued that both the retention of earnings as well as the payment of dividends involve tax liabilities. If these liabilities are equal, the optimal dividend policy would be indeterminate, so that essentially the no-tax case holds.

[3]It has sometimes been argued that dividend policy per se may provide information of value to stockholders about management's assessment of the company's future prospects. If so, a corporation might decide to pay dividends despite the tax disadvantages. In this last case, there might be some observable relationship, but without further specification as to the type of information conveyed by dividends, the exact functional form of the relationship would be unknown.

[4]Cf. Black, Jensen and Scholes (1972), Blume and Friend (1973), and Fama and MacBeth (1973).

[5]Even if c were zero, it might be noted that realized returns could be related to anticipated dividend yields in any cross section, but in an unpredictable way. The null hypothesis that realized returns were unrelated to anticipated dividend yield would be that c_t were zero for all t.

[6]The Standard and Poor's Composite Index was used to make this adjustment.

[7]A closer examination of the mean squared errors showed that the difference in the relative accuracy of the two measures was greater in the first half than in the second half of the period. The average mean squared errors for the measure of this paper and that of Black and Scholes were respectively for the 25 pairs of years for 1927 through 1952, 0.00407 and 0.00941; and for the 24 pairs of years from 1952 through 1976, 0.00100 and 0.00123.

[8]Some of these preliminary results are discussed in footnote 17.

[9]Preliminary work for this paper was done with monthly returns and the results are qualitatively the same as those reported below. These results using monthly data are available in a working paper from the author.

[10]Grouping techniques are, of course, not the only way to handle measurement errors and non-stationarities. If one were willing to make explicit assumptions about the distribution of the measurement errors and the nature of any non-stationarities, the estimated regression coefficients could be adjusted directly for any bias. The difficulty in applying this approach is that it is often difficult to verify the explicit assumptions and it may not be known how sensitive the results are to small violations of the assumptions. Although the grouping procedure may yield less efficient estimates than possible alternative procedures, it does provide a fairly robust way of handling measurement errors and non-stationarities.

[11]The data tapes contained no returns for January, 1926, so that these betas were based upon 59 months of data. All other betas were based upon 60 months of data.

[12]In order to use all securities, even those in excess of the maximum multiple of 5, the portfolios were not always exactly of equal number. If there were one more security than a multiple of 5, the number in the lowest beta group was increased by one; if two more, the number in the lowest and next lowest beta group were increased by one; and so on. The same type of adjustment was also used in the grouping by dividend yield.

[13]If δ^a were the annual dividend yield adjusted for overall price movements, the equivalent quarterly dividend yield was defined as $[(1+\delta^a)^{\frac{1}{4}}-1]$.

[14]An examination of the autocorrelation function of each of the time series of a_t, b_t, and c_t for the overall period and subperiods did not find any obvious violations of the independence assumption necessary in applying the standard t-test.

[15]Estimated on the portfolios which were first grouped on dividend yield, the average coefficients on dividend yield were in chronological order for the four decades: -0.0155 with a t-value of -0.03, 0.8459 with a t-value of 4.03, 0.6825 with a t-value of 1.64, and 1.2195 with a t-value of 2.09.

[16]The significance level of the average coefficient on beta decreases slightly with the addition of these dummy variables, suggesting that beta may be in part a surrogate for the dummy variable.

[17]Some analyses not reported in this paper attempted to replicate as closely as possible the empirical tests of Black and Scholes. These analyses, which are available in a preliminary working paper, provide evidence on the differential effect of measuring dividend yield with beginning- rather than end-of-period price. In this replication in conformity with the Black-Scholes analysis, the cross-sectional regressions were based upon monthly data and dividend yield was measured by the ratio $(\delta_{it} - \delta_{mt})/\delta_{mt}$, where δ_{it} is the dividend paid over the prior twelve months for portfolio i divided by some measure of price and δ_{mt} is the average of the δ_{it}'s. Black and Scholes measured dividend yield with end-of-period price and obtained an estimate of the coefficent on their dividend yield measure over the 1936-66 period of 0.0009 (0.94), where the t-value is indicated in parentheses. Using the data base available for this study, the comparable coefficient would have been estimated as -0.0006 (-0.78) using end-of-period price and 0.0002 (0.21) using beginning-of-price -- virtually the same numbers.

In the preliminary work for this paper, an allowance was made for the potential non-linearity between those stocks which paid a dividend and those which did not by analyzing separately those stocks with positive anticipated dividends. This sample, which is described in Section V, was used in estimating (4). Again using monthly cross-sections, the average of the estimated coefficients on the Black-Scholes measure of dividend yield was 0.0009 (1.00) using end-of-period price and 0.0024 (2.87) using beginning-of-period price. These results are quite consistent with a non-linear relationship and a simple attenuation bias due to an independent measurement error. The dummy variable approach used in the text is a more satisfactory way of examining non-linearity in that it allows a significance test.

[18]Specifically, the beta coefficient on an after-tax basis will be given by the ratio

$$\frac{\mathrm{Cov}[(1-\tau^c)r_{it}^c + (1-\tau^d)r_{it}^d, (1-\tau^c)r_{mt}^c + (1-\tau^d)r_{mt}^d]}{\mathrm{Var}[(1-\tau^c)r_{mt}^c + (1-\tau^d)r_{mt}^d]},$$

where the m subscript denotes the market return. On the assumption that $\mathrm{Var}(r_{it}^d)$ is small relative to $\mathrm{Var}(r_{it}^c)$, this ratio is approximately equal to $\mathrm{Cov}(r_{it}^c, r_{mt}^c)/\mathrm{Var}(r_{mt}^c)$, which is approximately the beta coefficient on a before-tax basis.

[19]The various estimates of $E(g_{it})$ would, of course, contain measurement error. If this measurement error is uncorrelated with β_{it} and δ_{it} and if β_{it} and δ_{it} are measured without error, the estimated coefficients in the regressions of estimates of $E(g_{it})$ on β_{it} and δ_{it} would be unbiased.

[20]In addition to these geometric estimates, estimates of the continuously compounded rate of return were obtained by regressing the natural logarithms of the quarterly dividend yields on a time index. The results based upon these estimates are similar to those based upon the geometric estimates which are reported in the text.

[21]The t-value of these averages are -6.39 and -10.08 respectively on the assumption that the succesive cross-sectional regression coefficients are independent over time. Since there is substantial positive correlation between successive values of these estimates, these t-values are undoubtedly overstated. Nonetheless, it is interesting to note in the case of the total sample that 81 out of the 145 regression coefficients were negative and significant at the 5 percent level, while none were positive and significant at this level.

[22]The t-value of these averages are -4.64 and -5.04 assuming the independence of successive estimates, an assumption which does not appear to hold. Nonetheless, 59 out of the 145 regression coefficients for the total sample were negative and significant at the 5 percent level, while only 14 were positive and significant at the same level.

[23]In this case, the estimate of the coefficient on the dividend yield, in the probability limit, would be the ratio of $\mathrm{Cov}(r_t, \delta_{it})$ to $\mathrm{Var}(\delta_{it})$ regardless of which beta were used.

[24]In the regression of r_{it} on dividend yield, the probability limit of the estimate of the slope c would be

$$c = \frac{\mathrm{Cov}(r_{it}, \delta_{it})}{\mathrm{Var}(\delta_{it})} .$$

The probability limit of the estimate of the slope coefficient holding β_{it} constant, say c', would be

$$c' = \frac{c - \dfrac{\mathrm{Cov}(\beta_{it}, \delta_{it})\mathrm{Cov}(r_{it}, \beta_{it})}{\mathrm{Var}(\beta_{it})\mathrm{Var}(\delta_{it})}}{1 - \dfrac{[\mathrm{Cov}(\beta_{it}, \delta_{it})]^2}{\mathrm{Var}(\beta_{it})\mathrm{Var}(\delta_{it})}}$$

To be consistent with the estimates in the total sample, $\mathrm{Cov}(\beta_{it}, \delta_{it})$ must be negative if $\mathrm{Cov}(r_{it}, \beta_{it})$ is positive.

[25] The empirical work contained in Black, Jensen, and Scholes (1972) and Blume and Friend (1973) is consistent with the hypothesis that stock returns appear to be generated by a function of the form

$$r_{it} = E(r_{it}) + \pi_{1t} + b_{it}\pi_{2t} + \varepsilon_{it},$$

where π_{1t} and π_{2t} are mean-zero common factors, not necessarily independent; b_{it} is a response coefficient appropriate to asset i at time t; and ε_{it} is a mean-zero independent disturbance. If the market return on common stocks r^s_{mt} is approximated by $E(r_{mt}) + \pi_{1t} + \pi_{2t}$, it can be shown that the beta of a stock with respect to r^s_{mt} is a linear function of b_{it}. The return on the total market r^T_{mt} will be given as $E(r^T_{mt}) + k_1\pi_{1t} + k_2\pi_{2t} + \pi_{3t}$, where π_{3t} is any common market factor not affecting stock returns. The constants k_1 and k_2 are required since the definition of r^s_{mt} implicitly implied a scaling of π_{1t} and π_{2t} and there is no guarantee that the coefficients on π_{1t} and π_{2t} would be 1.0 for the overall market. The beta of a stock with respect to r^T_{mt} can be shown to be a linear function of b_{it}.

Thus, the beta coefficient with respect to r_{mt}^{S} must in turn be a linear function of the beta coefficient with respect to r_{mt}^{T}.

[26]Basu (1977) and more recently Reingarum (1979) have presented evidence that the price-savings ratio may convey some additional information not contained in the beta coefficient. Friend, Westerfield, and Granito (1978) have concluded that the residual variance is needed to explain expected return, but Roll and Ross (1979) have reached the opposite conclusion.

References

1. Basu, S., "The Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficent Market Hypothesis," Journal of Finance 32 (June 1977).

2. Black, Fischer, Michael Jensen, and Myron Scholes, "The Capital Asset Pricing Model: Some Empirical Tests," in Michael C. Jensen; (ed.), Studies in the Theory of Capital Markets (New York: Praeger Publishing, 1972).

3. Black, Fischer, and Myron Scholes, "Effects of Dividend Yield on Stock Prices," Journal of Financial Economics 1 (May 1974).

4. Blume, Marshall E., "Betas and Their Regression Tendencies," Journal of Finance 30 (June 1975).

5. Blume, Marshall E., and Irwin Friend, "A New Look at the Capital Asset Pricing Model," Journal of Finance 28 (March 1973).

6. Blume, Marshall E., and Irwin Friend, The Changing Role of the Individual Investor (New York: John Wiley, 1978).

7. Fama, Eugene F., and James D. MacBeth, "Risk, Return and Equilibrium: Empirical Tests," Journal of Political Economy 81 (May 1973).

8. Friend, Irwin, and Marshall Puckett, "Dividends and Stock Prices," American Economic Review 54 (September 1964).

9. Friend, Irwin, Randolph Westerfield, and Michael Granito, "New Evidence in the Capital Asset Pricing Model," Journal of Finance 33 (June 1978).

10. Miller, Merton, "Debt and Taxes," Journal of Finance 32 (May 1977).

11. Reingarum, Marc, "Earnings-Price Ratio, Unexpected Earnings Forecasts, and Models of Capital Market Equilibrium: An Empirical Investigation," (Unpublished Ph.D. dissertation, University of Chicago, 1979)

12. Roll, Richard, "A Critique of the Asset Pricing Theory's Tests; Part I On Past and Potential Testability of the Theory," Journal of Financial Economics 4 (March 1977).

13. Roll, Richard, and Stephen A. Ross, "An Empirical Investigation of the Arbitrage Pricing Theory," unpublished manuscript (1979).

Table 1

Summary of Cross-Sectional Regressions
of Quarterly Returns on Betas, Dividend Yields, and Dummy Variables
Various Dates[1]

Date	$R_{it} = a_t + b_t\beta_{it} + c_t\delta_{it}$			$R_{it} = a_t + b_t\beta_{it}$		$R_{it} = a_t + b_t\delta_{it}$		$R_{it} = a_t + a'_t\xi_{it} + b_t\beta_{it} + c_t\delta_{it}$			
	$\bar{a}$	$\bar{b}$	$\bar{c}$	$\bar{a}$	$\bar{b}$	$\bar{a}$	$\bar{b}$	$\bar{a}$	$\bar{a}'$	$\bar{b}$	$\bar{c}$
1/36 - 12/76	0.0074 (1.01)	0.0215 (2.34)	0.5232 (2.07)	0.0187 (2.98)	0.0158 (1.54)	0.0383 (2.97)	0.1746 (0.47)	.0278 (2.63)	-.0226 (-2.46)	.0139 (1.72)	1.0950 (4.26)
1/37 - 12/46	0.0006 (0.04)	0.0359 (1.53)	0.0247 (0.04)	-0.0008 (-0.04)	0.0350 (1.19)	0.0548 (1.38)	-0.3962 (-0.37)	.0539 (1.68)	-.0565 (-1.88)	.0136 (.77)	1.388 (3.49)
1/47 - 12/56	0.0112 (1.69)	0.0153 (1.38)	0.8743 (4.27)	0.0271 (5.37)	0.0108 (0.97)	0.0284 (2.22)	0.7533 (3.34)	.0199 (2.21)	-.0104 (-1.10)	.0125 (1.32)	1.1605 (3.95)
1/57 - 12/66	0.0138 (0.88)	0.0092 (0.53)	0.5062 (1.20)	0.0237 (2.44)	0.0031 (0.20)	0.0315 (1.96)	0.0692 (0.15)	.0213 (1.36)	-.0070 (-0.55)	.0063 (0.36)	.7532 (1.14)
1/67 - 12/76	-0.0031 (-0.15)	0.0230 (1.08)	1.1272 (1.93)	0.0257 (1.68)	0.0070 (0.30)	0.0274 (0.97)	0.7784 (0.88)	.0062 (0.28)	-.0140 (-0.87)	.0227 (1.11)	1.3761 (2.28)

[1] The t-value of the average is shown in parentheses below the average.

Table 2

Summary Statistics for the Regression

$$E(g_{it}) = \gamma_o' + \gamma_1' \beta_{it} + \gamma_2' \delta_{it}$$

Date of Portfolio Formation	γ_2'	$t(\gamma_2')$[1]	$\bar{R}^2$
December 1935	-0.685	1.03	0.13
December 1940	0.043	4.58	0.00
December 1945	0.215	4.95	0.51
December 1950	0.368	9.15	0.15
December 1955	0.369	3.09	0.23
December 1960	0.482	3.99	0.19
December 1965	1.379	3.92	0.41
December 1970	1.208	3.33	0.07

[1]The t-values are measured from -1.0 as discussed in the text.

Table 3

Average Beta Coefficients
For Portfolios Formed on the Basis of Data through December, 1971

Sample	Type of Beta	Beta Group	Dividend Yield Group				
			Low	2	3	4	High
Total	Grouping Period (1962-1966)	Low	0.452	0.626	0.621	0.612	0.584
		2	0.861	0.872	0.863	0.866	0.860
		3	1.033	1.042	1.045	1.039	1.039
		4	1.263	1.294	1.256	1.273	1.254
		High	1.798	1.729	1.646	1.629	1.608
	Estimation Period (1967-1971)	Low	1.127	1.054	0.773	0.734	0.651
		2	1.399	1.129	1.045	0.921	0.855
		3	1.324	1.257	1.186	1.080	1.039
		4	1.659	1.320	1.292	1.267	1.159
		High	1.840	1.569	1.411	1.494	1.392
	Subsequent Period (1972-1976)	Low	1.080	0.929	0.866	0.729	0.774
		2	1.263	1.140	0.943	0.849	0.828
		3	1.400	1.215	1.168	1.064	0.989
		4	1.513	1.266	1.251	1.153	1.066
		High	1.611	1.493	1.293	1.264	1.072