A SIMPLE EXAMINATION OF THE EMPIRICAL RELATIONSHIP BETWEEN DIVIDEND YIELDS AND DEVIATIONS FROM THE CAPM

Edwin ELTON and Martin GRUBER*

New York University, New York, NY 10006, USA

Joel RENTZLER*

City University of New York, New York, NY 10010, USA

Received October 1981, final version received March 1982

Several papers have been published in recent years dealing with both the theoretical and the empirical impact of dividend yields on security returns. Dividends have been postulated as affecting stock returns because of tax effects, agency costs and the Wealth Transfer Hypothesis.

In this paper we perform a purely empirical examination of whether and to what extent deviations from the zero beta form of the CAPM are explained by dividend yields. The paper demonstrates that dividend yield has a large and statistically significant impact on return above and beyond that explained by the zero beta form of the CAPM. This is consistent with the findings of Litzenberger and Ramaswamy. In addition our results are consistent with the findings of small firm effects.

1. Introduction

Several papers have been published in recent years dealing with both the theoretical and the empirical impact of dividend yields on security returns. Dividends have been postulated as affecting stock returns because of tax effects [see Elton and Gruber (1970, 1978), Litzenberger and Ramaswamy (1979)], information effects [see Aharony and Swary (1980), Taub (1976), Pettit (1972, 1976) and Watts (1973, 1976)], agency costs [see Jensen and Meckling (1976)], and the Wealth Transfer Hypothesis [see Kalay (1980)].

In this paper we perform a purely empirical examination of whether and to what extent deviations from the zero beta form of the CAPM are explained by dividend yields. The paper demonstrates that dividend yield has a large and statistically significant impact on return above and beyond that explained by the zero beta form of the CAPM. This is consistent with the findings of Litzenberger and Ramaswamy (1979). In addition our results are consistent with the findings of small firm effects.

*The authors would like to thank Dexter Corporation, Windsor Locks, Connecticut for funding this research.

1 The Wealth Transfer Hypothesis refers to the ability of stockholders to transfer wealth from creditors to stockholders by paying out assets in the form of dividends.
consistent with the findings of small firm effects which are documented by Roll (1980), Banz (1981) and Reinganum (1981). The next section of this paper describes our procedure while the following section presents the results.

2. Methodology

The study utilizes the monthly data on dividends, prices and returns for New York Stock Exchange securities available on the University of Chicago’s CRSP tape. The period covered by this study is from January 1927 through December 1976. Since we are interested in examining the relationship between deviations from the zero beta CAPM and dividend yield, we must calculate estimates of each of these variables.

Our first step was to estimate the zero beta CAPM. In order to reduce the error in measuring beta which is used as input in estimating the CAPM, the portfolio grouping procedure employed by several previous researchers was employed [see Black, Jensen and Scholes (1972), and Fama and MacBeth (1973)]. The beta for each stock was calculated based on five years of data (e.g., 1927–1931) and using the CRSP value weighted index. All stocks were ranked and grouped into twenty portfolios on the basis of these betas. The beta for each portfolio was then calculated using data for the next five years (e.g., 1932–1936). This beta was then employed as the best estimate of the beta of each portfolio for 1937. This procedure was repeated changing each of the 5 year periods by 1 year so that we had beta estimates for 20 portfolios for 40 yearly periods (1937–1976). For each of these years the average return on each portfolio in year t was calculated and the following cross-sectional regression was run:

\[ \bar{R}_j^t = \gamma_0^t + \gamma_1^t \beta_j^t + \bar{\epsilon}_j^t, \quad j = 1, \ldots, 20, \]

where

\[ \bar{R}_j^t = \text{average return on portfolio } j \text{ in year } t, \]

\[ \beta_j^t = \text{estimated beta for portfolio } j \text{ in year } t, \]

\[ \gamma_0^t, \gamma_1^t = \text{cross-sectional coefficients estimated for year } t, \]

\[ \bar{\epsilon}_j^t = \text{random error term in year } t. \]

The analysis in this paper was also conducted using the CRSP equal weighted index. These results were indistinguishable from those reported in this paper.

The composition and size of each of these portfolios changed from year to year. The number of stocks which had the required data over each 11 year period and so were placed into portfolios varied from a low of 419 to a high of 880 with an average number of 740.
The excess return $\alpha_i^t$ for each security $i$ in year $t$ was then defined as

$$\alpha_i^t = R_i^t - (\gamma_0^t + \gamma_i^t \beta_i^t), \quad i = 1, \ldots, N,$$

where $\beta_i^t$ is the estimated beta of the portfolio to which security $i$ belongs, $\gamma_0^t$, $\gamma_i^t$ are the estimated cross-sectional coefficients for year $t$. This value for alpha was used as the best estimate of alpha for each of the securities in a particular portfolio.

Our next step was to obtain a forecast of the dividend yield for year $t$ (year 11) where the first forecast of dividend yield was made when year 11 was 1937. In order to obtain this estimate we decided to use the dividend yield on a portfolio of stocks rather than on each individual stock. There is substantial forecast error in the estimate of a security's yield. By utilizing the dividend yield for a portfolio of stocks rather than an individual security we hoped to reduce this forecast error. Stocks were grouped by their dividend yield in year 9 — two years before the forecast year.\(^4\) Twenty groups were formed. The first 19 groups were formed by ordering all stocks which paid positive dividends from highest to lowest and then dividing these stocks into 19 equal groups. The 20th group contained all stocks that paid zero dividends during year $t-2$ (year 9).

Placing all stocks with zero dividends into one group is in contrast with the procedure used in some prior studies [e.g., Black and Scholes (1974)]. Our purpose in grouping was to reduce forecast error while at the same time obtaining portfolios with a wide range of dividend yields. In some periods stocks with zero dividend yields were so numerous that if we had an equal number of stocks in each portfolio grouping, then a large number of these portfolios would have had zero dividend yields. Table 1 shows the percentage of stocks in our sample that did not pay a dividend in year $t-2$. For example putting an equal number of stocks in each group would have resulted in about half the portfolios having zero dividend yield in the early years and two or three with this zero yield in later years. Having formed the 20 portfolios we then had to forecast the dividend yield for year $t$ (year 11). We used as our forecast the actual dividend yield in year $t-1$ (year 10).\(^5\)

The next step was to relate excess return to dividend yield. First the average excess return, $\bar{\alpha}_j$, and average forecasted dividend yield, $\bar{\delta}_j$, for each

\(^4\)The arguments behind grouping on dividend yields are directly analogous to the arguments used in grouping on betas to provide better estimates. In forming groups we defined dividend yield as dividends in year 9 divided by the ending price. Our purpose was to obtain a set of portfolios with a wide range of yields in year 10.

\(^5\)We have employed the simplest possible forecast model in this paper setting next year's dividend yield equal to last year's dividend yield. If we had found no dividend yield effect it might have been due to the use of this naive forecasting model. However, the fact that we do get strong results with this naive model highlights the importance of the dividend effect. A more sophisticated forecasting model might lead to even stronger results.
portfolio $j$ was computed. The average was taken over the 40 forecast years (1937–1976). Then a cross-sectional regression of the form

$$\bar{\alpha}_j = \alpha_0 + \alpha_1 \delta_j + \bar{\epsilon}_j, \quad j = 1, \ldots, 20$$

was run to obtain estimates of $\alpha_0$ and $\alpha_1$. This averaging process was repeated using all eight non-overlapping 5-year subperiods in our 40 year sample giving eight other estimates of $\alpha_0$ and $\alpha_1$.

As will become clear later, there were reasons to suspect that the zero dividend group of stocks would behave perversely. To test this, a regression of the following form was run:

$$\bar{\alpha}_j = \alpha'_0 + \alpha'_1 \delta_j + \alpha'_2 D + \bar{\epsilon}_j, \quad j = 1, \ldots, 20.$$  

In this regression $D$ is a dummy variable having a value of 1 for the portfolio of stocks which paid zero dividends in the grouping year and a value of 0 for all other groups.

Note that this methodology is designed so that forecasts of both the betas and the dividend yields employed in the analysis use only data which is available prior to the dates over which the returns are calculated. Thus, the existence of an excess return with respect to dividend yield indicates that an investor in the absence of transaction costs could have acted in a way to earn an extra return using available information. This would not necessarily indicate that he could have earned an excess return because of the presence of transaction costs and additional non-market risk incurred. However it does indicate an additional influence on equilibrium returns. Note also that our procedure is designed so that if there is a bias it is against finding a
dividend influence. If betas are correlated with dividend yields then our two step procedure would have a bias towards finding no dividend effect even when one was present. This was a deliberate choice on our part. As has been discussed in the literature, there is considerable disagreement on the reasonableness of the procedures used for the joint estimation of dividend and beta effects. Our procedure, while biased against observing dividend effects, avoids the estimation controversy. Thus, finding a dividend effect with this procedure would provide strong evidence in support of such an effect.

3. Results

Table 2a summarizes the statistical relationship between the average yearly excess return (alpha) and the average dividend yield in each of the 20 groups where the averages are taken over the entire forecast period 1937–1976. By examining the first row of data in table 2a we see that although there is a

<table>
<thead>
<tr>
<th>Forecast period</th>
<th>Intercept</th>
<th>Dividend coefficient</th>
<th>Dummy variable coefficient</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1937–1976</td>
<td>-2.263</td>
<td>0.344</td>
<td>0.188</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.59)</td>
<td>(2.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.774</td>
<td>0.794</td>
<td>6.189</td>
<td>0.874</td>
</tr>
<tr>
<td></td>
<td>(-10.87)</td>
<td>(9.60)</td>
<td>(9.64)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2b

Cross-sectional regressions of average excess return on dividend yield, over five year periods ($t$-values in parentheses).

<table>
<thead>
<tr>
<th>Forecast period</th>
<th>Intercept</th>
<th>Dividend coefficient</th>
<th>Dummy variable coefficient</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1937–1941</td>
<td>1.078</td>
<td>-0.091</td>
<td>-0.204</td>
<td>0.002</td>
</tr>
<tr>
<td>1942–1946</td>
<td>-13.790</td>
<td>1.076</td>
<td>25.745</td>
<td>0.853</td>
</tr>
<tr>
<td>1947–1951</td>
<td>-1.298</td>
<td>0.221</td>
<td>1.211</td>
<td>0.037</td>
</tr>
<tr>
<td>1952–1956</td>
<td>-1.082</td>
<td>0.216</td>
<td>-0.419</td>
<td>0.077</td>
</tr>
<tr>
<td>1957–1961</td>
<td>1.425</td>
<td>-0.337</td>
<td>-0.299</td>
<td>0.082</td>
</tr>
<tr>
<td>1962–1966</td>
<td>-1.448</td>
<td>0.218</td>
<td>5.588</td>
<td>0.477</td>
</tr>
<tr>
<td>1967–1971</td>
<td>-4.907</td>
<td>1.457</td>
<td>2.059</td>
<td>0.570</td>
</tr>
<tr>
<td>Mean</td>
<td>-3.986</td>
<td>0.673</td>
<td>5.551</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.075)</td>
<td>(2.062)</td>
<td>(1.862)</td>
<td></td>
</tr>
</tbody>
</table>
significant positive relationship between excess return and dividend yield over the entire period, the relationship is small. The $R^2$ for the relationship is only 0.188.

An examination of the data used in obtaining these regression results (table 3) shows that the weak positive association may be due to the presence of two distinct relationships.\(^6\) Note that portfolio 20, which was formed two years before the forecast year using all stocks which did not pay a dividend, has a high positive excess return (alpha). The return on this group is considerably higher than the return on any other group though it has by far the lowest yield in the year prior to the forecast (year 10) and so the lowest forecasted yield. These zero dividend stocks constitute an outlier group. This finding is not surprising and has been noted earlier by Blume (1979). Later in this section we will examine some possible explanations for the observed high excess return in this group. Before turning to this discussion we will first eliminate the effect this group has on the dividend yield-excess return relationship. This could be accomplished either by deleting this group from the data or by adding a dummy variable which has a value of one for the non-dividend portfolio group (number 20) and zero for all other groups. These two procedures are identical in their effects on the intercept and slope term. The advantages of using a dummy variable is that it allows a direct measurement and a test of significance for the difference of behavior of the excess return in the zero dividend group. The results of including this dummy variable in the regression for the entire forecast period are shown in the second line of table 2a. Notice that now the positive relationship between excess return and dividend yield is both large (0.794) and statistically significant at the 1% level. Furthermore the portfolio constructed from stocks which had paid zero dividends has an excess return beyond what one would expect given the behavior of the other portfolios. The value of 6.189 is statistically significant at the 1% level. The inclusion of the dummy variable also raises the coefficient of determination on the regression from 0.188 to 0.874.

As a further check on the form of this relationship, the regression including the dummy variable was run for each of the eight five year non-overlapping subperiods between 1937 and 1976. These results, reported in table 2b, are supportive of the overall relationship. The relationship between excess return and dividend yield is positive in six of the eight periods and on average the regression coefficient is significantly different from zero at the 5% level. The impact of the dummy variable is positive in five of the eight periods and its average value is positive and significant at the 5% level.

The high positive excess return associated with zero dividend stocks is important and worthy of additional attention. Is this an effect caused by low

---

\(^6\)In examining table 3 recall that group 20 has many more members than other portfolios. Thus although most groups have negative alphas, the average overall alpha is close to zero.
Table 3
Average portfolio dividend yield and average excess return over 40 year forecast period 1937–1976.

<table>
<thead>
<tr>
<th>Portfolio number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average dividend yield</td>
<td>7.48</td>
<td>7.31</td>
<td>6.89</td>
<td>6.53</td>
<td>6.41</td>
<td>6.09</td>
<td>5.86</td>
<td>5.58</td>
<td>5.37</td>
<td>5.13</td>
<td>4.80</td>
<td>4.62</td>
<td>4.40</td>
<td>4.25</td>
<td>3.89</td>
<td>3.67</td>
<td>3.32</td>
<td>3.08</td>
<td>2.40</td>
<td>0.73</td>
</tr>
<tr>
<td>Average excess return</td>
<td>1.23</td>
<td>0.61</td>
<td>-0.63</td>
<td>1.38</td>
<td>0.39</td>
<td>0.51</td>
<td>0.02</td>
<td>-0.31</td>
<td>0.11</td>
<td>-1.22</td>
<td>-0.37</td>
<td>-1.02</td>
<td>-1.41</td>
<td>-1.52</td>
<td>-1.78</td>
<td>-1.59</td>
<td>-2.23</td>
<td>-2.85</td>
<td>-2.89</td>
<td>1.99</td>
</tr>
</tbody>
</table>
dividends or is there an alternative explanation? The most obvious alternative is that the result is due to a small firm effect. Small firms make up a much higher proportion of zero dividend stocks than they do of other groups and a number of authors have found that small firms on average have positive excess returns [see Roll (1980), Banz (1981) and Reinganum (1981)]. Thus the zero dividend effect we observe may be, in part or total, a small firm effect. At present there is no adequate explanation of the small firm effect. Roll has an argument that surely explains part of the phenomena. He shows that small firms have many more non-trading days then do large firms. He then demonstrates that non-trading leads to a downward bias in the beta. A downward biased beta leads to a positive excess return if a zero excess return would exist using the unbiased beta.7

An alternative explanation of the positive excess return on the zero dividend stocks involves their price. Stocks under $5 in price are treated differently by brokerage firms and financial institutions and the zero dividend group of stocks contain a much higher proportion of these low price stocks than do the other dividend groups.8 Stocks which sell for less than $5 are not considered as appropriate collateral for margin. Thus, in figuring the amount of assets in an account for purposes of margin, stocks under $5 are eliminated. Therefore, an investor who borrows or uses margin would find these stocks less attractive unless they offered a higher return. A similar argument can be seen by examining fig. 1.

The efficient frontier with borrowing restricted to margin is $ABCD$. The position of $CD$ depends on the maximum borrowing that is available. If the borrowing limit is lowered then, ceteris paribus, the efficient frontier involving borrowing is lowered. Thus, stocks that reduce the amount that can be borrowed will be held only if their return compensates for this lowering of the frontier.

![Expected Return vs Risk](image)

**Fig. 1**

7The use of monthly data rather than daily reduces the importance of this effect.

8Many financial institutions will not hold low priced stocks due to either transaction costs or perceived risk.
Low capitalization stocks in general sell for less than $5. Thus low price can serve as a proxy for low capitalization and it is difficult to be sure which factor might be causing positive excess returns. Roll's argument concerning infrequent trading of low capitalization stocks holds equally well for stocks selling under $5. However, the margin restrictions apply to stocks selling for less than $5 and not to low capitalization stocks. Thus we decided that, in order to separate the effect of low priced small capitalization stocks from the effect of zero dividend payment, it was most relevant to eliminate stocks selling for under $5 from our population.\(^9\)

Tables 4 and 5 are analogous to tables 2 and 3 except that all stocks selling for less than $5 per share were eliminated. Examining the first row in table 4 shows that eliminating such stocks had a major effect. The impact of dividend yield on excess returns as measured by the regression coefficient has grown from 0.344 to 0.549 and is statistically significant at the 1% level. The coefficient of determination between dividend yield and excess return has changed from 0.188 to 0.471. Eliminating stocks under $5 per share reduced the unexplained excess return associated with the zero dividend group. This indicates that part of the peculiar behavior of this group was due to the inclusion of these low priced firms.

However not all of the peculiar behavior of this group is eliminated by removing $5 stocks from the sample. Table 5 shows that the return on portfolio 20 which had previously paid a zero dividend is 0.86% whereas the behavior of the other 19 groups would lead us to expect a negative excess return for portfolio 20. To test for perverse behavior on the part of the zero dividend portfolio, the regression was rerun with a dummy variable. The results again improved. As seen in the second line of table 4a, the \(t\)-value of the intercept and dividend yield coefficient both increase, the coefficient of determination increases, and the dummy variable enters with a positive and statistically significant (at the 1% level) coefficient.

\[
\begin{array}{lllll}
\text{Table 4a} \\
\begin{array}{llllll}
\text{Cross-sectional regressions of average excess return on dividend yield over 40 year forecast period with under $5 stocks eliminated (t-values in parentheses).} \\
\hline
\text{Forecast period} & \text{Intercept} & \text{Dividend coefficient} & \text{Dummy variable coefficient} & R^2 \\
1937–1976 & -2.623 & 0.549 & & 0.471 \\
 & (3.69) & (4.00) & & \\
 & -4.398 & 0.865 & 4.445 & 0.841 \\
 & (8.95) & (9.36) & (6.28) & \\
\end{array}
\end{array}
\]

\(^9\)At a number of points in the study we eliminated low capitalization stocks rather than those selling for less than $5. This didn't significantly impact on the results and so these results are not reported.
Table 4b presents the regression results for all five year non-overlapping periods with and without the dummy variable. The results continue to support the earlier findings that there is a strong positive association between excess returns and forecasts of dividend yield except for the group of stocks which had previously not paid a dividend. These stocks seem to sell at a real premium above that which one would expect based on their forecasted beta and dividend yield.

4. Conclusions

In this article we have examined the relationship between residuals from the zero beta form of the capital asset pricing model and dividend yields. We have found a persistent relationship between dividend yield and excess returns. In particular, except for those stocks which had previously paid zero dividends, the higher the dividend yield the higher the excess return. One group of stocks, those which had previously paid no dividends, had excess returns which were above what we would expect from this relationship. Part of this differential was shown to be due to the effect of low priced stocks but
Table 5
Average portfolio dividend yield and average excess return over 40 year forecast period 1937–1976 with under $5 stocks eliminated.

| Portfolio number | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  | 16  | 17  | 18  | 19  | 20  |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Average dividend yield | 7.56 | 7.28 | 6.91 | 6.55 | 6.41 | 6.10 | 5.83 | 5.59 | 5.39 | 5.15 | 4.81 | 4.60 | 4.42 | 4.27 | 3.93 | 3.69 | 3.31 | 3.11 | 2.37 | 0.94 |
| Average excess return | 2.12 | 1.29 | 0.45 | 2.15 | 1.14 | 1.64 | 1.13 | 0.41 | 0.91 | -0.64 | 0.72 | -0.67 | -0.85 | -1.17 | -1.05 | -1.00 | -1.41 | -2.20 | -2.34 | 0.86 |
an influence beyond that still seems to exist. There seems to be persistent patterns in excess returns which are related to dividend yield. Some of these differences may be due to tax effects. Others have not as yet been adequately explained.

References