

The Effects of Dividends on Common Stock Prices Tax Effects or Information Effects?

ROBERT H. LITZENBERGER and KRISHNA RAMASWAMY*

I. Introduction

THERE HAS BEEN considerable controversy concerning the effect of dividend yields on common stock returns. The controversy centers on whether or not the positive association between common stock returns and dividend yields reported in a number of empirical studies can be attributed entirely to information effects. The purpose of this paper is to provide a brief critique of the theory and of the available empirical evidence (Section II), and to present some new empirical results (Section III). It is shown that there is a positive and non-linear relationship between common stock returns and expected dividend yield. The prediction rule for expected dividends is based solely on information that would have been available to the investor *ex-ante*. These results cannot, therefore, be attributed to the favorable or unfavorable information that would be present in a proxy for expected dividend yield that anticipates the occurrence (or lack thereof) of a dividend.

II.1 Review of Theory

Brennan (1970) was the first to develop an After-tax Capital Asset Pricing Model. The model was derived under the assumptions of unlimited borrowing and lending at the risk free rate of interest and unrestricted short sales. The dollar dividends paid by corporations were assumed to be certain and known to investors. The equilibrium relationship derived is given by

$$E(\tilde{R}_i - r_f) = b_0\beta_i + c_0(d_i - r_f) \quad (1)$$

where \tilde{R}_i is the before tax total rate of return on asset i , β_i and d_i are the systematic risk and the dividend yield on asset i respectively, and r_f is the risk free rate of interest. Note that the structural parameters b_0 and c_0 in this pricing relationship are not dependent on the level of the dividend yield. The parameter c_0 is a weighted average of the marginal tax rates of investors, with the weights being proportional to the individuals' global risk tolerances at the optimum. Thus $c_0 > 0$, and since by assumption individuals are risk averse, $b_0 > 0$.

Litzenberger and Ramaswamy (1979) extended the Brennan (1970) model to allow for margin constraints and for income related constraints on borrowing. The latter constraint serves to limit the interest deductions individuals can utilize to the amount of dividend income their portfolios generate. Those individuals for

* Stanford University and Columbia University.

whom this constraint is binding would find increased dividends desirable in that such increases serve to effectively relax the constraint. The equilibrium relationship so derived is given by

$$E(\tilde{R}_i - r_f) = a_1 + b_1\beta_i + c_1(d_i - r_f) \quad (2)$$

where $a_1 > 0$ is the risk premium on a zero beta portfolio that has a dividend yield equal to the riskless rate, and reflects the presence of the margin constraint. The model implies that c_1 is positive or negative, depending on whether the income related borrowing constraint is non binding or binding¹ for all individuals (see Litzenberger and Ramaswamy (1979, pp. 171-172). Note that in this model, as in Brennan's model, the parameter c_1 is independent of the level of the dividend yield d_i .

Miller and Scholes (1978) argue that the tax code permits strategies that enable one to escape the income tax on dividends altogether. Sufficient leverage of an equity portfolio can create interest expenses that can be used to offset the dividend income entirely. They argue that any unwanted risk in this levered position can be removed by the purchase of whole life insurance which contains a tax deferred investment component. In this model, a distinction is made between *accumulators* who are assumed to hold all the risky assets and employ the above strategy, and *non accumulators* who do not hold risky assets at all. The implication is that the effective marginal tax rate applicable to dividend income is zero and therefore the coefficient of dividend yield is zero. For a non accumulator not to hold equities at all, it must be the case that for each equity the after tax expected rate of return on the equity is less than the after tax rate of interest. This follows from the first order conditions for the standard portfolio problem of an investor evaluated at the point where all the wealth is invested in the riskless asset:

$$E[\tilde{r}_i - \tau d_i - r_f(1 - \tau)]u'(w_0 r_f(1 - \tau)) \leq 0 \quad \forall i \quad (3)$$

where \tilde{r}_i is the before tax rate of return on security i , τ is the marginal tax rate applicable to the nonaccumulator's income, w_0 his initial wealth, and $u(\cdot)$ his monotone increasing and concave von Neumann-Morgenstern utility function. If this condition does not hold for each asset, it would be optimal for the investor to hold some equities. Thus even if accumulators were able to costlessly defer the tax on the interest on their money market investments, the marginal tax bracket of non accumulators would enter any equilibrium relationship.

Ross (1977) and Bhattacharya (1979) have argued that dividend policy could be employed as a signalling mechanism, whereby firms with profitable projects are able and willing to pay higher dividends in order to segregate themselves from firms with less profitable projects. They provide a rationale for value maximizing firms paying positive dividends when the risk premiums per unit of dividend yield is positive in equilibrium. Stern (1979) has argued that such information signalling via dividends is excessively costly.

A model of asset prices in the presence of short selling restrictions, together

¹ The assertion of value maximizing behavior by firms in this context does not have a strong theoretical basis: see Litzenberger and Ramaswamy (1979), fn 2.

with a much simplified taxation scheme with individuals in diverse but constant marginal tax brackets, was derived in Litzenberger and Ramaswamy (1980). The implication of the model is that the differences in tax brackets in the presence of short selling restrictions would induce dividend clienteles, with the tendency of low (high) tax bracket individuals to hold high (low) dividend yield stocks: covariances among individual securities as well as the levels of yields determine the clientele that holds a given security. For the proper subset of all stocks that are held by a given clientele, the equilibrium relationship indicates that the before-tax risk premium on a stock is linearly related to its beta (measured relative to the clientele's optimal risk asset portfolio) and to the dividend yield. However, across groups the coefficient on dividend yield is a decreasing function of yield. Thus the existence of short selling restrictions tends to mitigate the tax effects of dividend changes since a corporation that attempts a sizable dividend cut would affect the clientele that holds the stock, and the associated coefficient on dividend yield would increase.

II.2 Review of Empirical Evidence and Relationship to the Present Study

In a pioneering empirical test of the effects of dividend yields on common stock returns Black and Scholes (1974) concluded that

it is not possible to demonstrate that the expected returns of high yield stocks differ from the expected returns on low yield stocks either before *or after taxes*. (emphasis added)

In spite of the ambiguous implication for the after-tax CAPM, the Black and Scholes study has frequently been cited as providing evidence against the existence of tax effects (see, for example, Miller and Scholes (1978)). Rosenberg and Marathé (1978) attribute the ambiguity of the conclusion in Black and Scholes (1974) to (a) the loss of efficiency which arises from grouping stocks into portfolios and (b) the inefficiency of their estimation procedures, which are identical to ordinary least-squares. Using a two stage generalized least-squares procedure that accounts for the problem of errors in variables, and using a more complete specification of the covariance matrix of the disturbance terms, Rosenberg and Marathé find a positive and significant relationship between dividend yields and common stock returns. The difference between these results and those of Black and Scholes cannot be attributed to the use of different dividend yield variables. Both studies use an average dividend yield over the prior twelve month period as a surrogate for the expected dividend yield.

Neither the Black-Scholes study nor the Rosenberg-Marathé study distinguishes between ex-dividend and non ex-dividend months in developing their proxies for the expected dividend yield. Presumably the rationale for ignoring the distinction is that in a world of transactions costs the effect of dividend yields on required return may occur in more than a single month. The recent work by Green (1980) provides some theoretical support for the position that dividend yield effects would be spread over time. Litzenberger and Ramaswamy (1979)

used yield variables that distinguished between ex-dividend months and non ex-dividend months and found significant positive coefficients in both ex-months and non ex-months. However, the coefficient in ex-months was more than twice as large as the coefficient in non ex-months. Litzenberger and Ramaswamy (1979, fn 8) note that

It might be argued that the persistent dividend effect is due to the fact that the dividend variable used incorporates knowledge of the ex-dividend month, which the investor may not have.

They then introduce a dividend yield variable that does not incorporate knowledge of the ex-dividend month except when the announcement occurred in a prior month. The test using this variable implicitly assumes that the effect of dividend yields on common stock returns is distributed uniformly throughout the quarter, and is therefore similar in spirit to the Rosenberg-Marathé and the Black-Scholes tests. The coefficient of the dividend yield variable in this test was positive and statistically significant.

Recently, Miller and Scholes (1981) have argued that the observed relationship between common stock returns and dividend yields is attributable to the favorable information contained in the knowledge that a firm will actually declare any dividend. They note that there is a group of

zero dividend paying firms consisting of those unfortunates who would have paid a dividend in month t on their regular quarterly schedule, but whose directors voted to omit the dividend. As the old story goes, there may be an important clue in the fact that a dog does not bark! Although these firms have declared and paid a dividend within the same month, they have declared a dividend of zero and hence are not recorded on the CRSP tapes as having declared a dividend. They are placed for test purposes in the complementary zero-dividend group where their adverse information effect serves to pull down the mean excess return of the zero-dividend firms. An upward twist is thereby imparted to the slope coefficient relating realized returns to dividend yields. (p. 13)

There are nine post Black and Scholes studies cited in Miller and Scholes which examine the relationship between returns on NYSE common stocks and dividend yields. The eight we were able to obtain all reported a significant and average relationship between returns and dividend yields. These results² are summarized in Table 1. It should be noted that only three studies, namely Litzenberger and Ramaswamy (1979, 1980) and Hess (1979), use a dividend yield variable that depends on prior knowledge of ex-dividend months: the same three studies also report that the effect of yields is non-linear. Thus, the Miller-Scholes explanation cannot be invoked for the remaining studies. While some of the authors cited in Table 1 do not attribute the significant yield effects to taxes, with the possible exception of Bradford and Gordon (1980), these results cannot be attributed to an information effect. Both Bradford and Gordon and Morgan (1981) employ

² Note that there was no test of an average linear relationship in Litzenberger and Ramaswamy (1980), where a test of the Tax induced CAPM was presented: therefore, there is no entry in Table 1. There is also no test reported by Hess (1979) which conforms to a test of the type reported in Table 1. We thank Professor Stone for providing us with the updated numbers reported here.

Table 1
**Summary of Results of Tests of Average Relationship between
 Common Stock Returns and Dividend Yields**

Author(s)	Test Period and Interval	Estimated Coefficient on Yield
Blume (1980), p. 571	1936-76, quarterly	0.52 (2.07)
Bradford and Gordon (1980), p. 127	1926-78, monthly	0.1762 (8.51)
Litzenberger and Ramaswamy (1979)	1936-77, monthly	0.236 (8.62)
Morgan (1980)	1936-77, monthly	0.209 (11.0)
Rosenberg and Marathé (1979), pp. 203-206	1931-66, monthly	0.395 (1.88)
Stone and Bartter (1979)	1947-70, monthly	0.56 (2.00)

Note. *t*-values are in parentheses under each coefficient.

sophisticated prediction rules to develop an expected yield variable. The Bradford and Gordon prediction rule is based on a pooled time series and cross-section regression of dividend yields on recent past dividends, market returns and yield, the recent capital gain on the stock, the riskless rate of interest and a constant. However, in estimating the parameters of this rule they use data from the entire sample period: thus the estimated relationship between returns and yields is based on data that would not have been available to the investor *ex-ante*. The Morgan prediction rule is based on a Box-Jenkins time series model fitted to grouped data and only uses data that would have been available to the investor *ex-ante*. Neither Bradford and Gordon nor Morgan examine the linearity of the relationship between returns and predicted yields.

Recently Miller and Scholes (1981) have examined the relative coefficients on predicted dividend yield (conditional on it being an ex-month) for stocks that announced their dividends prior to the ex-month versus stocks that announced their dividend during the ex-month. They found that the coefficient for stocks that announced prior to the ex-month was substantially smaller than the coefficient for stocks that announced during the ex-month. They interpret this evidence as supporting their thesis that the dividend yield effects are attributable entirely to the information effect. This fails to recognize that the average number of days from the beginning of an ex-month to the ex-date is greater for stocks that announce within the ex-month than for stocks that announce prior to the ex-month. From the work of Green (1980) it follows that the effect of yield on common stock returns would not just occur on the ex-day. If it is hypothesized that the tax effect occurs uniformly over a two week period, the effect would *ceteris paribus*, be less for a stock that goes ex-dividend on the second day of the month than for a stock that goes ex-dividend in the third or fourth week of the month. Thus, when examining the relative impact of dividend yield between those stocks that announced prior to the month and those that announced and

went ex-dividend within the same month, it would be important to account for the number of days in the month until the ex-date, expressed as a fraction of the two week period.

III. Empirical Tests

In this section the econometric procedures are described briefly: a more complete description is available in Litzenberger and Ramaswamy (1979, 1980). Then the development³ of the dividend yield variables follows, and the results of tests with these variables are presented. Consistent with the Litzenberger and Ramaswamy (1980) study,⁴ the empirical tests presented here assume that individuals fall into five tax clienteles and that each clientele holds one-fifth of the market value of all New York Stock Exchange (NYSE) stocks. The portfolios of the clienteles are assumed to correspond to the optimum portfolios in market equilibrium under certainty: that is, having ranked available stocks in a given year by their (past) annual dividend yield, five portfolios are formed by proceeding down this ranking until a fifth of the market value of stocks is reached, and then until two-fifths is reached, and so on. The first (group) portfolio is then a value weighted portfolio of the lowest dividend yield stocks, comprising a fifth of the market value of all stocks. The next portfolio (a fifth of the market value) contains the next lowest dividend yield stocks, and so on. This procedure ignores the influence of covariances on the tax related clientele and should only be viewed as an approximation to the true optimal portfolios. The underlying tax related clientele model is

$$E(\tilde{R}_i) - r_f = b_g \beta_{ig} + T_g(d_i - r_f), \quad \forall i \in g, \quad g = 1, 2, \dots, 5 \quad (4)$$

where β_{ig} is the beta of the i^{th} security with respect to the optimal portfolio of group g and T_g the marginal tax bracket of group g . Following the earlier work in Litzenberger and Ramaswamy (1980), and justified under a condition described there, the beta used in the tests is the standard beta with respect to the return on the market portfolio. Thus the structural model estimated is

$$\tilde{R}_{it} - r_{ft} = \gamma_{0g} + \gamma_{1g}\beta_{ig} + \gamma_{2g}(d_{it} - r_{ft}) + \tilde{\epsilon}_{it}, \quad \forall i \in g, \quad g = 1, 2, \dots, 5 \quad (5)$$

This is the basic model estimated and presented below. The econometric techniques are described in Litzenberger and Ramaswamy (1979). The Maximum Likelihood Estimator (MLE) developed there is used in each cross-section to arrive at estimates $\{\hat{\gamma}_{0gt}, \hat{\gamma}_{1gt}, \hat{\gamma}_{2gt}\}$ in month t .

The computational procedure employed took the standard steps. First,

³ Common stock return data were obtained from the monthly returns tape file provided by the Center for Research in Security Prices (CRSP) at the University of Chicago. The data on dividend distributions, the announcement dates and ex-dates, together with other relevant data are also provided on the master file by CRSP. The same service also provides the return series on a value-weighted index of all NYSE stocks. This series was used as a proxy for the returns on the 'market' portfolio (\tilde{R}_m). The riskless return series (r_{ft}) was constructed from the returns on prime commercial paper and the returns on U.S. Treasury bills.

⁴ In Litzenberger and Ramaswamy (1980), stocks were ranked by yield at the beginning of the calendar year, so that the composition of the groups did not change for a year. In this study we have ranked and formed the groups every month. Aside from this, there are no differences in the estimating procedures.

estimating the betas (and their standard errors) of all available securities using 60 months of data prior to month t . Second, running a cross sectional regression in month t using the MLE procedure. Third, finding the time series average of the estimated coefficients $\{\hat{\gamma}_{0gt}, \hat{\gamma}_{1gt}, \hat{\gamma}_{2gt}, t = 1, 2, \dots T\}$ from the T cross sections.

The measurement errors in the betas are correlated over time because 58 months of overlapping data are used to estimate security betas employed in successive cross-section regressions. This induces autocorrelation in the time series of estimated coefficients. The t -values reported in the tables are computed under the assumption that these coefficients follow a first order auto regressive process. The magnitude of the first order autocorrelation coefficient is generally small, so higher order schemes are ignored.⁵

III.2 Results for the Tax-Clientele CAPM: Dividend Variable d_1

The first procedure employed to estimate the expected dividend yield d_{1it} (subscript 1 refers to the *first* procedure) is identical to the yield variable used in the earlier studies (Litzenberger and Ramaswamy (1979, 1980)). This provides a basis for comparison with subsequent tests.

$$d_{1it} = \begin{cases} 0 & \text{if month } t \text{ was not an ex-dividend month for security } i; \text{ or if it was,} \\ & \text{it was not a regular dividend announced prior to } t. \\ D_{it}/P_{it-1} & \text{if month } t \text{ was an ex-dividend month with } D_{it} \text{ the dollar} \\ & \text{dividend per share announced prior to the month.} \\ \hat{D}_{it}/P_{it-1} & \text{if the security went ex in the month and this was a recurring} \\ & \text{dividend.} \end{cases}$$

Here \hat{D}_{it} is the previous (going back at most 12 months) recurring, taxable dividend per share adjusted for any changes in the number of shares outstanding in the interim, and P_{it-1} is the price at the end of month $t - 1$. The use of this variable assumes that the investor had prior knowledge of the ex-dividend months, though the surrogate for the dividend is based on information that would have been available to the investor *ex-ante*. The results using this variable are presented in Table 2. These results are consistent with the predictions of the Tax-clientele CAPM and indicate a pronounced non-linear effect of yields on common stock returns.

Because the dividend yield variable d_1 employed in these tests incorporates knowledge of the ex-dividend months, the results may suffer from the biases discussed at length in Miller and Scholes (1981). Thus the observed positive but non-linear association between common stock returns and yields could arise from this "information" effect. There are two simple procedures for purging the coefficient of potential information effects. The first is to construct an expected dividend yield variable based on information the investor has prior to the test

⁵ In each table below, the mean and the associated t -value are

$$\bar{y}_j = \sum_{t=1}^T \hat{y}_{jt} / T, \quad \text{and} \quad t(\hat{y}_j) = \hat{y}_j / (\hat{\sigma}_u / N(1 - \rho)) [N - 2\rho / (1 - \rho^2)]^{1/2},$$

where ρ is the estimated first order autocorrelation coefficient and $\hat{\sigma}_u$ is the standard error of the regression of \hat{y}_{jt} on \hat{y}_{jt-1} .

Table 2
Pooled Time Series and Cross Section Test of Tax
Clientele CAPM, 1940-80

5 groups: dividend variable d_1 used			
$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1\beta_i + \gamma_2(d_{1i} - r_f), \forall i \in g, g = 1, 2, \dots, 5$			
Group	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
I (Low yield)	0.00478 (2.26)	0.00518 (2.12)	0.665 (3.91)
II	0.00217 (1.07)	0.00459 (1.94)	0.516 (4.83)
III	0.00338 (1.70)	0.00422 (1.77)	0.415 (7.05)
IV	0.00159 (0.88)	0.00663 (2.68)	0.274 (7.04)
V (High Yield)	0.00327 (1.74)	0.00631 (2.66)	0.125 (3.98)

month, and the second is to use a sample of stocks known not to incorporate unavailable information for the cross-sectional regressions.

Expected Dividend Yield Variable Based on Prediction Rule (d_2)

In assessing the cash return in the future from purchasing a common stock, an investor would incorporate information regarding the periodicity of past payments as well as their (possibly changing) magnitude over time. Past studies (see for example Lintner (1956) and Fama and Babiak (1968)) have examined the payout behavior of U.S firms but these studies have ignored the within year timing of payouts and so are not immediately applicable here. While the majority of NYSE firms pays dividends on a quarterly schedule, there are several that pay dividends semi-annually or annually, and at least one that has paid regular monthly dividends. It is clear then that a prediction rule for expected dividend yield for a given firm based solely on past payments data for that firm would be free of these differences in payment schedules. For reasons of computational ease, however, this study uses a prediction rule that is based on the payment data of all firms for five years prior to the cross-sectional test period. Thus the expected yield so constructed reflects the payment behavior of the average firm, and is not the most efficient construct, though it is expedient computationally. The forecast dividend yield, labeled d_{2it}^e for stock i in period t , is constructed as follows. Using data from 60 months prior to month t , a pooled time series-cross section regression is used to estimate the parameters of the following model:

$$\tilde{D}_{i\tau}/P_{i\tau-1} = \alpha_{0t} + \sum \alpha_{jt}\delta_{ij}(D_{i\tau-j}/P_{i\tau-1}) + \tilde{u}_{i\tau} \quad (j = 3, 4, 6, 7, 12, 13)$$

$$\tau = t - 1, t - 2, \dots, t - 60 \quad i = 1, 2, \dots, N \quad (8)$$

where

$D_{i\tau}$ = regular dividend to security i in period τ , if any.

$D_{i\tau}$ = the most recent regular dividend to security i prior to period τ , if any in last 12 months adjusted for changes in number of shares outstanding in the interim.

δ_{ij} = 1 if period $\tau - j$ was a regular ex-dividend month, and 0 if otherwise.

$P_{i\tau-1}$ = the closing price in month $\tau - 1$.

Note that the dividend $D_{i\tau}$ on the right hand side of the equation is the most recent regular dividend: thus the RHS variable $D_{i\tau}/P_{i\tau-1}$ corresponds to the naive yield explanatory variable based on the most recent dividend, going back at most 12 months. The lags $\{j = 3, 4, 6, 7, 12, 13\}$ were chosen because although firms may slip forward or backward in their payment schedule, it is firms that are late in the announcement (these had a regular dividend 4 or 7 or 13 months ago) that are likely to announce and go ex-dividend within the same month, and it is precisely these firms that can cause unanticipated surprises and disappointments.⁶ The forecast dividend yield \hat{d}_{2it}^e was then found as

$$\hat{d}_{2it}^e = \hat{\alpha}_{0t} + \sum \hat{\alpha}_{jt} \delta_{ij} (D_{i\tau}/P_{i\tau-1}), \quad (j = 3, 4, 6, 7, 12, 13) \quad (9)$$

and the variable that was used in the cross-sectional regression in month t is d_{2it} , defined as either D_{it}/P_{it-1} if taxable dividend D_{it} was announced prior to month t , or \hat{d}_{2it}^e otherwise. Thus the expected dividend yield variable d_{2it} incorporates only information the investor would possess at the end of month $t - 1$.

Table 3 presents results from a test of the Tax-Clientele CAPM based on the expected yield variable d_{2it} . The coefficients on the dividend yield variables are positive and significant for all the groups except the last (highest yield) group. Furthermore these coefficients decline with the level of yield as predicted by the model.⁷ In comparison with the coefficients obtained in Table 2, which used the yield variable d_{1it} , the coefficients on yield in Table 3 are approximately 8 basis points lower. One explanation for this is that there is information contained in the prior knowledge of the ex-dividend month which biases the slopes ($\hat{\gamma}_2$) in Table 2 upward. Another explanation is that the prediction rule employed is not the most efficient in which case the coefficients in Table 3 would be biased downward.

In Panel A of Table 4, the Tax-clientele CAPM is estimated with the coefficient on beta constrained to be the same across the five groups: this corresponds to the

⁶ Clearly alternative rules are possible, with lags at $j = 2, 5$ and 11 : we have tried only this structure.

⁷ The MLE estimating procedure produces linear estimators: for example, the estimate of $\hat{\gamma}_2$ in a monthly cross-sectional regression for a given group is a weighted combination of the monthly rates of return of the stocks in that group. Since the weights sum to zero, this estimator is the rate of return on a self-financing portfolio. The MLE procedure is designed to produce an estimator of $\hat{\gamma}_2$ that, asymptotically, has a zero beta—a requirement for the estimator to be asymptotically unbiased (see Litzenger and Ramaswamy (1979), pp. 173-81). To test for possible misspecification, the betas of the monthly estimates $\hat{\gamma}_2$ for each group were computed, and these were negative for 4 of the 5 groups. Using a procedure suggested by Sharpe (1981) the coefficients in Table 3 were then adjusted. The resulting adjusted estimators for $\hat{\gamma}_2$ changed only slightly: these are 0.599, 0.481, 0.365, 0.236, and 0.045.

Table 3
As in Table 2 above: dividend variable based on prediction rule d_2 used

Group	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
I (Low Yield)	0.00477 (2.22)	0.00502 (2.06)	0.555 (2.83)
II	0.00206 (1.01)	0.00468 (1.97)	0.486 (4.18)
III	0.00339 (1.69)	0.00427 (1.78)	0.339 (5.32)
IV	0.00176 (0.98)	0.00665 (2.70)	0.212 (4.74)
V (High Yield)	0.00365 (1.94)	0.00622 (2.62)	0.222 (0.65)

Note: t values are in parentheses under each coefficient.

Table 4
PANEL A Pooled Time Series Cross-Section Test of Tax Clientele CAPM, Five Groups, 1940-80. Slope on Beta Constrained to be same across groups. Dividend Yield Variable used: d_{2i}

$$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1 \beta_i + \sum_{g=1}^5 \gamma_{2g} \delta_{ig} (d_{2i} - r_f), \forall i$$

$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_{21}$	$\hat{\gamma}_{22}$	$\hat{\gamma}_{23}$	$\hat{\gamma}_{24}$	$\hat{\gamma}_{25}$
0.00346 (2.05)	0.00491 (2.22)	0.155 (1.18)	0.488 (3.55)	0.366 (5.29)	0.267 (5.63)	0.058 (1.45)

PANEL B Same as Panel A, but with dummy variable δ_{i0} added. $\delta_{i0} = 1$ if stock i paid zero dividend in the previous year

$$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1 \beta_i + \sum_{g=1}^5 \gamma_{2g} \delta_{ig} (d_{2i} - r_f) + \gamma_3 \delta_{i0}, \forall i$$

$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_{21}$	$\hat{\gamma}_{22}$	$\hat{\gamma}_{23}$	$\hat{\gamma}_{24}$	$\hat{\gamma}_{25}$	$\hat{\gamma}_3$
0.00367 (1.60)	0.00466 (2.34)	0.231 (1.54)	0.482 (2.95)	0.362 (5.58)	0.268 (6.42)	0.059 (2.18)	0.00181 (0.88)

Note: t -values are in parentheses under each coefficient.

cross-sectional regression

$$\tilde{R}_{it} - r_{ft} = \gamma_0 + \gamma_1 \beta_{it} + \sum_{g=1}^5 \gamma_{2g} \delta_{ig} (d_{2it} - r_{ft}) + \tilde{\epsilon}_{it} \quad i = 1, 2, \dots, N_t \quad (10)$$

where $\delta_{ig} = 1$ if security i is in group g and zero otherwise. The results indicate that, except for the coefficient γ_1 of the lowest yield group,⁸ the coefficients are positive and declining in the predicted manner. It should be noted that the non dividend paying stocks are all in Group I, and for these stocks the predicted

⁸ Since Group I contains a large number of nondividend paying stocks, there is reason to suspect the distributional properties of $\hat{\gamma}_{21}$ relative to the others ($\hat{\gamma}_{22}$ to $\hat{\gamma}_{25}$). We are currently exploring this issue.

dividend yield d_{2it} would always be equal to $\hat{\alpha}_0$: this is clearly an inefficient and biased predictor.⁹ To test for a separate influence first noticed by Blume (1980) of non dividend paying stocks, relation (10) is estimated with a dummy variable:

$$\tilde{R}_{it} - r_{ft} = \gamma_0 + \gamma_1\beta_{it} + \sum_{g=1}^5 \gamma_{2g}\delta_{ig}(d_{2it} - r_{ft}) + \gamma_3\delta_{i0} + \tilde{\varepsilon}_{it}, \quad \forall i \quad (11)$$

where $\delta_{i0} = 1$ if the stock has no dividend in the past 12 months and 0 otherwise. The results of this test are in Panel B: they indicate that the pattern of the coefficients is as in Panel A, and that the coefficient γ_{21} increases. The coefficient on the dummy variable δ_{i0} indicates that the before-tax risk premium to non-dividend paying stocks is approximately 2.16%. This is lower than estimates reported by Blume (1980) and Litzenberger and Ramaswamy (1980), and unlike these studies, it is insignificant.

Estimation of Tax-Clientele CAPM with a Subsample of Stocks (d_3)

An alternative procedure which purges the coefficient on yield of potential information effects is to restrict the sample of stocks used to estimate the parameters of the Tax-clientele CAPM. There are several ways to restrict the sample: one is to use only those stocks that have announced a dividend prior to the ex-month. This suffers from the immediate criticism that the cross-sectional variation in dividend yields is greatly reduced and hence the test is inefficient. Another possibility is to use the whole sample but set the dividend yield for those that have no announcement prior to the test month to zero. This biases the coefficient downward, because it eliminates the correlation between expected yield and realized return for those stocks that announced and went ex-dividend in the same month. The subsample chosen here is based on the conjecture that if a firm paid a regular dividend in the previous month, it is not likely to pay a regular dividend in the current month. Thus the restricted sample consists of those stocks that have announced prior to the test month, in which case the anticipated yield D_{it}/P_{it-1} was used, and those stocks that went ex-dividend in the previous month, in which case the dividend yield was set to zero. This yield variable is written d_{3it} . This subsample is free of any potential information effects. In Table 5, the results of the Tax-clientele CAPM with this subsample are presented. Because there were some groups (notably the low yield groups) which had very few stocks, or even no stocks that paid a dividend in the early months, not all the cross-sectional regressions could be conducted. A cross-sectional regression (across firms in a given group) was conducted in a given month only if there were at least 5 stocks that had announced a dividend prior to the month, and if there were at least 20 stocks that were candidates for the subsample. The results indicate that there is a positive association between returns and yields within each group subsample. As before, the coefficient of yield in the highest yield group is not significant.

The problem with every subsample is that it throws away information, and thus reduces the efficiency of the estimator. The results in tables 3 and 4 indicate, however, that these alternative approaches to avoiding undesired 'information'

⁹ The average value of $\hat{\alpha}_0$, computed from 1936-1980, was 0.000996. In no case was $\hat{\alpha}_0$ negative.

Table 5

Pooled Time Series and Cross-Section Test of Tax Clientele CAPM, Five Groups, 1940-80
(Subsample of Stocks that have announced prior to ex-month, and those that have just gone ex-dividend in the previous month)
Dividend Yield Variable Used: d_{3i}

$$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1 \beta_i + \gamma_2 (d_{3i} - r_f), \quad \forall i \in g, \\ g = 1, 2, \dots, 5$$

Group	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	
I (Low Yield)	0.00715 (2.07)	0.00216 (0.66)	0.629 (2.56)	Obs* = 450
II	0.00008 (0.23)	0.00610 (1.60)	0.380 (2.72)	Obs = 463
III	0.00359 (1.57)	0.00306 (1.15)	0.331 (4.25)	Obs = 477
IV	-0.00207 (-1.01)	0.00866 (3.32)	0.135 (2.65)	Obs = 487
V (High Yield)	-0.00029 (-0.16)	0.00755 (3.14)	0.049 (1.38)	Obs = 489

*Obs is the number of cross-sectional months over which these estimates have been computed. There are 492 months in total possible.

Note: *t*-values are in parentheses under each coefficient.

effects provide reasonably close estimates of the effect of yield on common stock returns. These results also indicate that in light of the observed non-linear association between yields and common stock returns, the emphasis on an average linear effect in the literature is misplaced. Nevertheless, in the interest of providing a comparison to prior studies and of exploring the possible information effect in tests of the After-tax CAPM, the next section examines the results of tests of a linear relationship.

III.3 Results of Tests of After Tax CAPM

In this section the results of the After Tax CAPM which predicts a linear relationship between expected returns and yield are presented. The results in Panels A, B, C of Table 6 correspond to the results in Table 2, Table 3, and Table 5 respectively. Panel A reports results that extend those in Litzenberger and Ramaswamy (1979) through 1980: the dividend yield variable d_{iit} is employed here. Panel B reports results with the predicted yield d_{2it} : the average coefficient $\hat{\gamma}_2$ is 0.151 and statistically significant at the 0.05 level. The drop in the coefficient on yield between Panels A and B is approximately 8 basis points. Panel C reports results with the subsample of securities that have either announced a dividend prior to the ex-month, or have just paid a regular dividend. The coefficient $\hat{\gamma}_3$ here is 0.135 and this is statistically significant as well.

The drop in the coefficient on yield in Panel B from its value in Panel A could be due to information effects. Alternatively, decline could be attributed to the tax effect occurring over a period prior to the ex-date, as suggested by Green (1980).

Table 6
Pooled Time Series and Cross Section Test of After Tax CAPM, 1940-80

$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1\beta_i + \gamma_2(d_i - r_f), \forall i$			
	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
PANEL A			
Dividend variable d_{1i} used	0.00313 (1.81)	0.00484 (2.15)	0.233 (8.79)
PANEL B			
Dividend variable d_{2i} based on prediction rule, used	0.00337 (1.95)	0.00470 (2.08)	0.151 (5.39)
PANEL C			
Subsample dividend variable d_{3i} used	0.00097 (0.52)	0.00527 (2.33)	0.135 (4.38)

Note: *t*-values are in parentheses under each coefficient.

Table 7
Pooled Time Series and Cross-Section Test of After Tax CAPM, 1940-80 With Scaled Yield Variables

$E(\tilde{R}_i) - r_f = \gamma_0 + \gamma_1\beta_i + \gamma_2(s_i * d_i - r_f), \forall i$			
	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
PANEL A			
Dividend yield variable d_{1i} used	0.00318 (1.85)	0.00499 (2.21)	0.401 (14.29)
PANEL B			
Subsample dividend yield variable d_{3it} used	0.00107 (0.57)	0.00529 (2.34)	0.297 (7.19)

The time from the beginning of the month to the ex-date is frequently less than 2 weeks for stocks that announce prior to the ex-month. Assuming the tax effect occurs uniformly over the two week period prior to the ex-date, the dividend variable for stocks that went ex-dividend within the first two weeks of the month may be scaled by the ratio of the number of days until the ex-date to two weeks. If this explanation of the decline in the coefficient is valid, the coefficients of the scaled variables for the subsample and for the total sample should be of the same magnitude.

Let ND_i be the number of days from the beginning of the month to the ex-date for the i^{th} stock. Then the scale factor that was employed is given by $s_i = \text{Min}\{ND_i/15., 1.\}$. Panels A and B of Table 7 report results of tests identical to those in Panels A and C of Table 6 respectively; except that the yield variables d_{1it} in Table 5 and d_{3it} in Table 7 have both been scaled by s_i . As is evident, the coefficients are not of the same order of magnitude, lending little support to the hypothesis that the tax effect is spread over a two week period prior to the ex-date. It is possible, of course that the scale factor applied is incorrect.

The lack of significance of the dividend yield coefficient in the Black-Scholes study has been subject to alternative interpretations. Rosenberg and Marathé (1979) have attributed this to the inefficiency of the OLS estimating procedure

Table 8
Pooled Time Series and Cross Section
Test of After Tax CAPM, 1936-78
Twenty-five Portfolios* in Cross
Sections

$$E(\tilde{R}_p) - r_f = \gamma_0 + \gamma_1 \beta_p + \gamma_2 (d_{1p} - r_f),$$

$$p = 1, 2, \dots, 25$$

$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
0.00466	0.00592	0.125
(2.73)	(1.98)	(0.21)

* Portfolios ranked first by annual yield, last 12 months and within 5 portfolios so formed, ranked by beta to construct 5 portfolios. Equal numbers of stocks in each portfolio.

Note: *t*-values in parentheses under each coefficient; *t* values in Table 8 are not corrected for first order autocorrelation in coefficients.

and to the loss in efficiency which arises from grouping stocks into portfolios. Miller and Scholes (1981) argue that the Black-Scholes study points up the absence of a long-run dividend yield effect by virtue of its use of an expected annual yield variable from past data, and that the use of a short-run variable (such as d_{1it}) is beset with potential information effects. The power of the Black-Scholes procedure was examined by replicating the Black-Scholes study, but using the dividend yield variable, d_{1it} . If the procedure is sufficiently powerful this should result in a statistically significant yield coefficient, since the dividend yield variable which impounds knowledge of the ex-dividend months is used.

There are some differences in the replication which must be noted. Twenty-five portfolios of stocks, with equal numbers of stocks in each group, were formed by ranking *every month* first by yield (defined as the sum of all dividends, adjusted for splits, etc., divided by the end of the previous month price), forming five portfolios, and then ranking stocks in each of the five portfolios by beta. Thus the composition of the portfolios varies from month to month. For n_t stocks in *each* of the twenty-five portfolios in month t , the dividend yield of the portfolio is computed as:

$$d_{1pt} = \sum_{i=1}^{n_t} d_{1it}/n_t. \quad (13)$$

The value weighted index \tilde{R}_m is used to compute betas. In addition, the yield variable employed on the RHS of the cross sectional regression is $(d_{1pt} - r_{ft})$, and not $(d_{1pt} - d_{mt})/d_{mt}$ as used by Black and Scholes.

The results are reported in Table 8. The coefficient on dividend yield is insignificant, implying that the Black-Scholes procedure as replicated here is not sufficiently powerful to pick up potential information effects.

IV. Conclusion

This study has presented empirical evidence consistent with the Tax-Clientele CAPM: the data indicate that there is a positive but non-linear association between common stock returns and dividend yields. The prediction rule for the

expected dividend yield is based solely on information that would have been available to the investor *ex-ante*, and hence is free from potential information effects that are contained in dividend yield variables that anticipate the occurrence (or lack thereof) of a dividend. Nevertheless, the results here are similar to those obtained earlier.

Whether the effect of dividend yields on common stock returns (as indicated by the data) can be attributed to taxes or is due to some omitted variable(s) remains an open question. The conclusion of the present study is that these significant yield effects cannot be pinned to the information content in the prior knowledge that the firm will declare a dividend of unknown magnitude.

REFERENCES

- Bhattacharya, S., 1979, Imperfect information, dividend policy and the bird-in-the-hand fallacy, *The Bell Journal of Economics*, 10, 259-270.
- Black, F. and M. Scholes, 1974, The effects of dividend yield and dividend policy on common stock prices and returns, *Journal of Financial Economics*, 1, 1-22.
- Blume, M. E., 1980, Stock returns and dividend yields: some more evidence, *Review of Economics and Statistics*, 62, 567-577.
- Bradford, D. F. and R. H. Gordon, 1980, Taxation and the stock market valuation of capital gains and dividends, *Journal of Public Economics*, 14, 109-136.
- Brennan, M. J., 1970, Investor taxes, market equilibrium and corporation finance, Unpublished Ph.D. Dissertation (Massachusetts Institute of Technology, Cambridge, Mass.).
- Charest, G., 1978, Dividend information, stock returns and market efficiency—II, *Journal of Financial Economics*, 5.
- Elton, E. and M. Gruber, 1970, Marginal stockholder tax rates and the clientele effect, *Review of Economics and Statistics*, 52, 68-74.
- Fama, E. and H. Blasiak, 1968, Dividend policy: an empirical analysis, *Journal of the American Statistical Association*, 63, 1132-61.
- Green, J., 1980, Taxation and the ex-dividend day behavior of common stock prices, *National Bureau of Economic Research*, W. P. 496, Cambridge, Mass.
- Hess, P., 1979, The empirical relationship between dividend yields and stock returns: tax effects or non-stationarities in expected returns, mimeo, The Ohio State University, Columbus, Ohio.
- Lintner, J., 1956, Distribution of incomes of corporations among dividends, retained earnings and taxes, *American Economic Review*, 46, 97-113.
- Litzenberger, R. H. and K. Ramaswamy, 1979, The effects of personal taxes and dividends on capital asset prices: theory and empirical evidence, *Journal of Financial Economics*, 7, 163-195.
- Litzenberger, R. H. and K. Ramaswamy, 1980, Dividends, short-selling restrictions, tax-induced investor clienteles and market equilibrium, *Journal of Finance*, 35, 469-482.
- Miller, M. and M. Scholes, 1978, Dividends and taxes, *Journal of Financial Economics*, 6, 333-364.
- Miller, M. and M. Scholes, 1981, Dividends and taxes: some empirical evidence, W. P. 55, Center for Research in Security Prices, Graduate School of Business, University of Chicago, Chicago, Illinois.
- Morgan, I. G., 1980, Dividends and capital asset prices, (Unpublished mimeo, School of Business, Queens University, Kingston, Ontario).
- Rosenberg, B. and V. Marathé, 1979, Tests of capital asset pricing hypotheses, *Research in Finance*, 1, 115-223.
- Ross, S. A., 1977, The determination of financial structure: the incentive signalling approach. *The Bell Journal of Economics*, 8, 23-40.
- Sharpe, W. F., 1981, Some factors in New York Stock Exchange security returns, 1931-1979, Unpublished mimeo (August 1981), Graduate School of Business, Stanford University, Stanford, California.
- Stern, J., The dividend question, Opinion Column, *Wall Street Journal*, July 15, 1979.
- Stone, B. K. and B. J. Bartter, 1979, The effect of dividend yield on stock returns: empirical evidence on the relevance of dividends, W. P. E-76-8, Georgia Institute of Technology, Atlanta, Georgia.

Copyright of *Journal of Finance* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.