This paper reexamines some recent tests of whether holders of shares with higher dividend yields receive higher risk-adjusted rates of return to compensate for the heavier taxes on dividend payments than on long-term capital gains. Our particular concern is with tests using short-run measures of dividend yield—that is, measures that seek to deduce the differential tax burden on dividends over long-term capital gains from differences in rates of return on shares that do and shares that do not pay a cash dividend during the return interval. We show that such measures are inappropriate for that purpose. Any yield-related effects associated with such measures must arise from sources other than the long-term tax differential. For the short-run measures considered here, the yield-related effects found in some tests are traced to biases, one of a fairly subtle kind, introduced by dividend announcement effects.

I. Introduction

The publication of Black and Scholes's (1974) study came after nearly two decades of intense academic controversy over the effects of dividends on stock prices. Using the large Center for Research in Security Prices (CRSP) data base and employing new econometric methods...
that avoided many of the difficulties that had hampered earlier testing efforts, Black and Scholes found no significant relation between stock returns and dividend yield or dividend payout. Their results thus lent support to neither of two contending hypotheses about dividend effects—the conventional view that the market prefers to obtain the income from stock as dividends and the contrary view, widely held among academics, that the market demands higher returns on dividend-paying shares to compensate for tax penalties on dividend income.

Although academic researchers continued to speculate about why so seemingly important a yield-related effect as the tax penalty on dividends should have left so small a trace in the data, further empirical research seemed unpromising, barring new sources of data (perhaps from foreign countries) or new and more powerful methods of data analysis. Within the last few years, studies claiming just such improvements in methodology or data have appeared (e.g., Long 1978; Litzenberger and Ramaswamy 1979, 1980, 1981; Rosenberg and Marathe 1979; Stone and Bartter 1979; Banz 1980; Blume 1980; Gordon and Bradford 1980; Hess 1980, 1982; Morgan 1980a, 1980b) with some, but by no means all, reporting significant, yield-related tax effects.

We will not here review these many studies in detail. Our purpose rather is to warn against accepting as tax effects the yield-related effects reported in those studies using short-run definitions of dividend yield—that is, in tests seeking to deduce the differential tax burden on dividends over long-term capital gains from differences in rates of return on shares that do and shares that do not pay a cash dividend during the return interval. Tests employing such short-run definitions of dividend yield are inappropriate for that purpose. Ex-dividend day returns that reflected the long-term tax differential would imply substantial profit opportunities in short-term trading around ex-dividend days, particularly for brokers and dealers in securities. The cum-ex price differentials that maintain market equilibrium and keep such profit opportunities from arising obliterate the traces of the long-run tax differential that the tests with short-run yield definitions seek to measure.

Although tests with short-run measures of expected dividend yield cannot reliably calibrate the effects of differences in the tax treatment of dividends and long-term capital gains, they can, and frequently do, turn up what appear to be significant yield-related differences in rates of return. Researchers reporting such differences, however, must recognize that they have a problem, not a solution. The challenge is to account for them. For the particular short-run measures considered here, the yield-related effects found in some tests are traced to biases,
one, of a fairly subtle kind, introduced by dividend announcement effects.

The plan of the paper is as follows. Section II describes our method of testing for yield-related tax effects—essentially tests of an after-tax capital-asset pricing model (CAPM) using the familiar Fama-MacBeth (1973) method of time-series pooling of cross-sectional coefficients (but applied to individual stocks rather than portfolios). The problems of defining the dividend-yield variable are then briefly explored along with some discussion of the properties of estimates based on short-run measures of yield. Section III presents the estimates of yield-related tax effects for several alternative short-run yield measures. The estimates, most of which seem to imply substantial tax effects, are sensitive to the choice of dividend variable, largely, we argue, because the short-run measures are distorted to different degrees by dividend information effects. Purging these measures of dividend yield of information effects gives estimates of yield-related tax effects that are both statistically and economically insignificant. The next two sections show that this failure to find significant tax effects with our purged yield measures cannot plausibly be attributed to the inadequacies in our controls for risk (Sec. IV) nor to our neglect of nonlinear clientele effects (Sec. V). Nor should this failure be regarded as surprising. We show, in our concluding Section VI, why tests relying on short-run responses to dividend payments cannot be expected to provide reliable estimates of the effects of a tax penalty on dividends over long-term capital gains. Whether such a tax effect, if it exists, can be detected with yield measures reflecting long-run dividend policies rather than short-run payments remains an open question.

II. The Dividend Coefficient and the Dividend Variable

We seek to estimate the dividend coefficient $a_3$ in the regression equation:

$$\hat{R}_{it} - R_R = a_1 + a_2\hat{b}_{it} + a_3(\hat{d}_{it} - R_R) + \hat{\epsilon}_{it},$$

(1)

where $\hat{R}_{it}$ is the rate of return on share $i$ during period $t$, $R_R$ is the riskless rate of interest during period $t$, $\hat{b}_{it}$ is the estimated beta or systematic risk coefficient for stock $i$ for period $t$, and $\hat{d}_{it}$ is an estimate of the dividend yield of stock $i$ in period $t$. We estimate equation (1) by the now familiar three-step, pooled cross-section, and time-series approach of Fama and MacBeth (1973) and Fama (1976). First, the risk coefficient, beta, is estimated from a market model regression of the form
over the 60 months previous to the test month $t$. For month $t$, the risk coefficient, $\hat{b}_t$, and an estimate of the dividend yield for each company are then treated as independent variables in a cross-sectional multiple regression of the form:

$$\tilde{R}_{it} - R_{ft} = a_{it} + b_i(\tilde{R}_{mt} - R_{ft}) + \tilde{e}_{it}, T = t - 60, \ldots, t - 1$$

(2)

This step is repeated month by month, with the estimated risk and dividend-yield variables updated each time. In the final step, the coefficient $a_3$ is estimated as the sample mean, $\bar{a}_3$, of the monthly cross-section regression coefficients $\hat{a}_{3t}$. The standard error of the estimate is computed as $\sigma_{a_3}/\sqrt{t'}$, where $\sigma$ is the standard deviation of the time series of $\hat{a}_{3t}$, and $t'$ is the number of months in the sample.

Unlike Fama and MacBeth (1973), Black and Scholes (1974), or Fama (1976), we apply the three-step method using individual company data instead of portfolios or grouped data. We do this to simplify comparison of various dividend measures and not from any belief that tests with ungrouped data are necessarily more efficient than tests with grouped data.

Readers are warned that the distributional assumptions justifying the three-step procedure are not always well approximated. The coefficients $\hat{a}_{3t}$, for example, are not generally distributed independently and identically over the sample period, and their distributions are sometimes asymmetric and fat-tailed. Nor does the sequence of coefficients $\hat{a}_{1t}$, which can be interpreted as the returns on a “zero beta” portfolio, always have a zero beta. When the sequence of coefficients $\hat{a}_{3t}$ varies systematically with the market, Black and Scholes (1974) have suggested a fourth step in which the sequence of coefficients $\hat{a}_{3t}$ is regressed on the market excess rate of return. At appropriate points in the discussion, we will take this fourth step as well. In the first part of the analysis, however, we stipulate the statistical assumptions, as the lawyers might put it, and concentrate on the dividend variable.

The dividend coefficient $\bar{a}_3$, if positive and within a reasonable range (positive but less than .6), is often interpreted as an implicit “tax bracket” or “tax differential.” This interpretation can be justified rigorously in a CAPM framework by assuming investors maximize expected after-tax returns subject to the standard constraints. Brennan (1970) formulated such an after-tax CAPM relating expected before-tax returns to risk, $\beta_i$, and dividend yield of the following form:

$$E(R_i) - R_f = b\beta_i + \tau(d_i - R_f),$$

(4)
where $\tau$, the coefficient of the dividend term, is a weighted-average marginal tax differential of dividends over capital gains. The Brennan model (and Litzenberger and Ramaswamy's [1979] generalization of it) implies that the dividend coefficient has the same value for all shares. This prediction was tested by Hess (1980), who finds it not descriptive of the data. Despite this evidence of misspecification of after-tax models such as (4) and its variants, the conventional reference to $a_3$ as an implicit tax bracket or a tax differential will be maintained, at least until we present alternative interpretations.

**Problems in the Definition of Dividend Yields**

The appropriate measure of dividend yield in tests for tax or other yield effects is by no means clear. The underlying valuation models call for a measure of the market's expectation of future dividend yield. But over what horizon is that expectation to be measured? Is it a one-step-ahead forecast? Or is it a forecast of the average dividend yield that might be realized from holding the share over a considerable period of time? And if the latter, how long?

Remember, also, that the tax differential between dividends and capital gains is itself a function of the expected length of the holding period. Does the market accrue the tax differential, adjusting returns more or less evenly over time? Or is the adjustment for tax effects concentrated mainly at times when dividends are paid and tax liabilities incurred? If the former, a long-run measure of expected dividend yield is the appropriate dividend variable. Black and Scholes (1974), for example, took as their yield measure the realized dividend yield of portfolios selected by ranking securities by the sum of dividends per share paid during the previous year divided by the price per share at the end of the year. Their ranking variable is thus a way, albeit a simple one, to approximate the average annual dividend yield expected by someone who bought one of their portfolios at the start of the year and planned to hold it for a year or more.

Black and Scholes's failure to find significant yield-related tax effects with their long-run variable led researchers to try the short-term approach by focusing on returns in and around the actual ex-dividend dates. The appropriate expected yield for the test equations

---

1 The weights are the risk tolerance of each investor relative to the total risk tolerance. If individuals are assumed to have utility functions with constant relative risk aversion, then the weights become proportional to market holdings.

2 Models of equilibrium valuation that do not imply the same tax differential for all shares have recently been developed by Constantinides (1980) and by Litzenberger and Ramaswamy (1980).
is, then, not some long-run average but only the dividend yield, if any, expected by the market during the next return interval.

If the return interval is a month, as here and in most of the cited studies of yield effects, the expected dividend yield for about two-thirds of the firms in the sample will be zero, because the vast bulk of dividend-paying firms follow a quarterly payment cycle. Each monthly slope coefficient $a_{t}$ in the second-pass regression will combine two sources of variation in monthly dividend yields. The first is the cross-sectional variation in dividend yields among those firms expected to go ex dividend during the month. If a tax effect exists, the conditional mean returns of the high-yield ex-dividend firms will be higher than for the lower-yield firms. A scatter plot for these ex-dividend firms would then be upward sloping. The second source of variation in yields is that between the ex-dividend firms as a group and the non-ex firms. The returns of the non-ex firms will lie along a vertical line through the origin (or more precisely through the negative of the riskless rate since most tests are run with the variables in premium form). The location of the mean return of the non-ex firms (about two-thirds of the sample each month) will thus have substantial weight in determining the intercept of each monthly regression (and hence, of course, also indirectly the slope coefficients $a_{t}$).

Including both groups rather than only those expected to go ex dividend increases the effective range of the critical yield variable and thereby presumably also the efficiency of the estimates of any yield-related tax effects. Whether these hopes for greater efficiency can in fact be realized by the short-run approach, or whether they are thwarted by other disadvantages, will be our principal concern in the sections to follow.

III. Estimates of Yield-related Tax Effects under Alternative Short-Run Definitions of Yield

Even if we accept the logic of the short-run approach, the market’s expectations of the cash dividends to be paid in month $t$ must still be specified. An obvious first approximation is the actual dividend payment in month $t$—that is, assume the actual dividend payment

---

3 Some dividend studies use daily data, though in a test format somewhat different from the three- or four-pass procedure used here (see, e.g., Elton and Gruber 1970; Black and Scholes 1973; Kalay 1977). One (Blume 1980) uses a quarterly return interval (and a long-run dividend measure) precisely to smear over the cum-ex differentials that motivate the short-term yield approach.

4 A scatter plot of monthly excess returns against dividend yields for a combined sample of ex-dividend and non-ex-dividend firms is shown in fig. 1. Readers are warned, however, that the picture is intended at this point only to portray the two sources of variation; analysis of that scatter must be deferred.
equaled the expected dividend payment. Such an assumption will not be universally true, of course, for dividend surprises do occur; but a month is a short forecasting interval after all, and as Hess (1980) points out, the perfect-foresight estimate can at least be considered one bound on the set of permissible approximations. Certainly if no dividend effect turned up with this approximation, there would be little point in going further.

The Perfect-Foresight Definition of Dividend Yield

The estimated tax effect coefficients obtained under this definition of dividend yield are shown in the first row of table 1. For the period 1940–78—which, for reasons noted later, is the longest benchmark period over which the effects of the different dividend variables can be compared—the coefficient \( \bar{a}_3 \) turns out to be .317 with a \( t \)-ratio of no less than 10.2. Thus at first sight the case for the short-run approach seems amply vindicated. The coefficients are not only in the plausible range for a tax effect but appear estimated with great precision.

These appearances may be deceptive, however, because forces at work in the data impart an upward bias to the dividend coefficient. Between 30 and 40 percent of the firms going ex dividend in month \( t \)...

### TABLE 1

<table>
<thead>
<tr>
<th>Definition of Expected Dividend Yield</th>
<th>( \bar{a}_1 )</th>
<th>( \bar{a}_2 )</th>
<th>( \bar{a}_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual dividend yield</td>
<td>.0059</td>
<td>.0024</td>
<td>.3173</td>
</tr>
<tr>
<td>(4.5)</td>
<td>(1.6)</td>
<td>(10.2)</td>
<td></td>
</tr>
<tr>
<td>Level-revised (LR) monthly dividend yield</td>
<td>.0065</td>
<td>.0022</td>
<td>.1794</td>
</tr>
<tr>
<td>(4.9)</td>
<td>(1.4)</td>
<td>(6.1)</td>
<td></td>
</tr>
<tr>
<td>Dividend yield of 12 months ago (no regard to ex-dividend month)</td>
<td>.0038</td>
<td>.0019</td>
<td>.0376</td>
</tr>
<tr>
<td>(2.9)</td>
<td>(1.3)</td>
<td>(1.3)</td>
<td></td>
</tr>
<tr>
<td>Positive LR dividend yields set to dividend yields of 36 months ago</td>
<td>.0041</td>
<td>.0021</td>
<td>.1544</td>
</tr>
<tr>
<td>(3.2)</td>
<td>(1.4)</td>
<td>(4.8)</td>
<td></td>
</tr>
<tr>
<td>LR dividend yield set to 0 if positive (ex-dividend month), otherwise 0</td>
<td>.0054</td>
<td>.0022</td>
<td>.0026</td>
</tr>
<tr>
<td>(4.0)</td>
<td>(1.4)</td>
<td>(5.9)</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** — T-values in parentheses. Definition used is:

\[
R_d - R_p = a_1 + a_2 \bar{a} + a_3 (\bar{d}_u - R_p) + \bar{e}_u. \tag{1}
\]

**Source.** — All price and dividend data are from the Center for Research in Security Prices (CRSP) monthly tapes.
also declared during the same month \( t \). Consider two such stocks and suppose that at the end of the previous month the market's expected dividend was $1.00 for each. On the basis of this expectation, assume the price of each has been set at $50 per share. Suppose one firm increases its dividend to $1.50 a share and the other cuts its dividend to $0.50 a share. The first firm will now be recorded as having a dividend yield of 3 percent and the second of only 1 percent. But the high-yield firm will also have a higher realized rate of return for the month if, as is often the case, the market interprets the unexpected increase in dividends as signaling improved earnings prospects for the company. By the same token, the company with the disappointing cut in dividends may find its return for the month dragged down by a fall in the end-of-month price that reflects the market's downward revision of the company's future earnings. Errors in dividend anticipations and realized rates of return will thus tend to be positively correlated.

**The Litzenberger-Ramaswamy Variable**

Litzenberger and Ramaswamy (1979) sought to eliminate this information bias by exploiting an additional piece of information on the CRSP tape, to wit, the declaration date. On the basis of the reported dividend declaration date, they defined a revised dividend variable as follows: (1) If a firm declared prior to month \( t \) and went ex dividend in month \( t \) (about 60 percent of those going ex dividend in \( t \)), then the expected dividend yield, \( \delta_{\text{ur}} \), was computed using the actual dividend paid in \( t \), \( D_{\text{ur}} \), divided by the price at the end of month \( t - 1 \). (2) If the firm both declared and went ex dividend in the same month \( t \), then \( \delta_{\text{ur}} \) was computed using the last previous regular dividend (looking back up to one year). If the backward search yields no such regular dividend or if the dividend was an extra dividend, then \( \delta_{\text{ur}} \) is set equal to zero (about 10 percent of those declaring and paying in \( t \)). We call this definition of the expected dividend yield the level-revised (LR) dividend variable.

The results obtained with this dividend variable are shown in line 2 of table 1. Note the large response to this seemingly minor technical adjustment. The coefficient drops substantially from .317 to .179 or by nearly 45 percent, though it remains highly significant and still within the plausible range for a tax effect.\(^3\)

\(^3\) Our estimate is somewhat lower than that reported by Litzenberger and Ramaswamy (1979), whose sample period is somewhat different (1936–77 rather than 1940–78). We start from 1940 because the CRSP tape lists few declaration dates prior to that year. Hence corrections cannot be made (either by the LR method or others to be considered below) for most of the cases in which the declaration date and ex date are in
Exploiting the declaration date in this fashion reduces the upward bias in $\bar{a}_t$ but does not eliminate it. One group of firms still remains for which the LR variable is the actual dividend yield even though the firms have declared and paid in the same month. These are the firms that might have paid a dividend in month $t$ on their regular quarterly schedule, but whose directors, at some time during the month, have voted to omit dividends. The CRSP tape will report, correctly, no dividends paid during the month and no date on which a dividend was declared. But absence of a dividend declaration during the month is information not available to the market at the beginning of the month. And as the old story goes, there may be an important clue in knowing that a dog did not bark!

To see how overlooking this clue can bias regressions employing the LR dividend variable, consider again the two similar firms for which the market’s expected dividend was $1.00 per share. Suppose that for each firm separately the market’s consensus probability was .5 that the firm would declare a dividend of $2.00 and .5 that the firm would omit the dividend. Suppose further that the price of the shares post announcement will double if the dividend announced is $2 and will halve if the dividend is omitted. If the initial price for each firm were, say, $10, the realized rate of return would be 120 percent when a $2.00 dividend is announced and −50 percent when the dividend is set to zero, which implies an ex ante (and, of course, also average ex post) rate of return of 35 percent. The ex ante expected dividend yield for both shares is 10 percent. Regressions of realized returns on expected dividend yields for such ex ante matched pairs would show no relation between the two variables.

Suppose, however, that in those regressions the LR variable had been used instead of the market’s expected dividend yield. And suppose, for concreteness, that the most recent regular dividend within the last 12 months had been $2.00 for each firm. A firm now announcing a dividend of $2.00 in month $t$ would be recorded as having a dividend yield of 20 percent and a realized return of 120 percent. A firm passing its dividend would have a realized return of −50 percent but a dividend yield of zero under the LR definition. Thus the regressions with the LR variable would show what appears to be a positive association between returns and dividend yields even though returns are actually independent of correctly measured ex ante dividend yields.  

The bias introduced by zero-dividend firms under the LR definition can persist for up to three additional quarters after the first regular dividend is passed. Those firms that resume regular dividends one, two, or three quarters later will have higher returns the same month. Tests that include the pre-1940 years will to that extent still be subject to the information-effect bias described above.
Information Effects or Tax Effects?

That valuable information may be contained merely in the knowledge that a dividend was or was not paid during a month is strongly suggested by some of the other tests reported in table 1. Note, for example, that when we use not the firm’s actual dividend yield but its dividend yield 12 months ago, if any, no significant coefficient emerges. Yet the cross-sectional variation among the firms is just as large and the yields show substantial stability over time. But when we use what will presumably be an even more out-of-date predictor, namely, the dividend yield of 3 years ago, we get a highly significant positive coefficient, provided we restrict the nonzero entries to those firms for which the level-revised yield was nonzero in \( t \). Even more striking is the fact that a significant coefficient is produced without reference to any specific dollar value for the dividends, but merely a dummy variable set equal to one if the LR dividend variable was nonzero.

Although these tests are suggestive, they cannot by themselves tell whether the ex-dividend month shows up because of information effects or tax effects. But there is a way to find out.

Yield Variables Using Only Dividends Declared in Advance

Table 2 presents estimates of \( \bar{a}_3 \) for several different yield measures, in each of which the nonzero elements of the vector include dividends paid in month \( t \), but only if declared in a month prior to \( t \). For these firms, at least, unwanted information effects of the kind found in the previous definitions are absent. Yet the logic underlying the short-run approach is maintained because these firms did go ex dividend in month \( t \).

In the first panel of part A, the entry in the yield vector for any firm that both declared and paid its dividend in month \( t \) is set to zero. This approach seeks to offset the downward pull on the intercept by the “dogs that did not bark” with the upward pull from those declaring and paying nonzero dividends during the same month. Announce-

and higher LR dividend yields than those otherwise comparable firms that disappointed the market by not resuming their regular dividends during the 12-month interval following the cut.

Table 2 presents estimates both for the entire sample period 1940–78 and for four subperiods. Great caution should be taken in interpreting the subperiod results, which are presented only to give some idea of the substantial variability in the dividend coefficients over time. The patterns of subperiod coefficients appear unrelated to changes in the tax law or to the downward drift in the weighted-average marginal tax bracket (see n. 1) that would presumably have accompanied the steady increase in the proportion of stocks held by tax-exempt institutions over the sample period.
ment effects should net to zero if all announcements are included. The disadvantage of the approach is that setting a nonzero dividend to zero creates an errors-in-variables problem that attenuates the estimate of $\overline{a}_3$. But the degree of attenuation from this source is known—the bias toward zero can be shown to be approximately the proportion of firms whose yields are changed (about 40 percent over the sample period)—and can be taken into account before conclusions are drawn.

As can be seen from the table, however, interpretations are unlikely to be much affected by adjustments for attenuation. The dividend coefficient for the entire sample period 1940–78 is close to zero (−0.0268) and statistically insignificant.

The tests in the second panel differ from those just mentioned by eliminating firms both declaring and paying dividends in month $t$.
TABLE 2 (Continued)

<table>
<thead>
<tr>
<th>Years</th>
<th>$\tilde{a}_1$</th>
<th>$\tilde{a}_2$</th>
<th>$\tilde{a}_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940–78</td>
<td>.0043</td>
<td>.0035</td>
<td>.0135</td>
</tr>
<tr>
<td></td>
<td>(2.5)</td>
<td>(2.2)</td>
<td>(.1)</td>
</tr>
<tr>
<td>1940–48</td>
<td>.0080</td>
<td>.0025</td>
<td>.0306</td>
</tr>
<tr>
<td></td>
<td>(2.1)</td>
<td>(3.8)</td>
<td>(.0)</td>
</tr>
<tr>
<td>1949–58</td>
<td>.0095</td>
<td>.0049</td>
<td>-.2694</td>
</tr>
<tr>
<td></td>
<td>(4.0)</td>
<td>(1.7)</td>
<td>(-2.4)</td>
</tr>
<tr>
<td>1959–68</td>
<td>.0042</td>
<td>.0034</td>
<td>.1557</td>
</tr>
<tr>
<td></td>
<td>(1.2)</td>
<td>(1.1)</td>
<td>(.6)</td>
</tr>
<tr>
<td>1969–78</td>
<td>-.0043</td>
<td>.0033</td>
<td>.1635</td>
</tr>
<tr>
<td></td>
<td>(-1.1)</td>
<td>(.9)</td>
<td>(.8)</td>
</tr>
</tbody>
</table>

B. Definition (1d):  
DIVIDEND YIELD SLOPE DUMMY ADDED: DUMMY K = 1  
IF DIVIDEND DECLARED IN ADVANCE; OTHERWISE ZERO ($\tilde{d}_a = \tilde{d}_a''''$)

<table>
<thead>
<tr>
<th>Years</th>
<th>$\tilde{a}_1$</th>
<th>$\tilde{a}_2$</th>
<th>$\tilde{a}_3$</th>
<th>$\tilde{a}_4$</th>
<th>$\tilde{a}_3 + \tilde{a}_4 = \tilde{a}_5$</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940–78</td>
<td>.0071</td>
<td>.0022</td>
<td>.3702</td>
<td>-.4585</td>
<td>-.0883</td>
<td>-.0417</td>
</tr>
<tr>
<td></td>
<td>(5.4)</td>
<td>(1.4)</td>
<td>(10.7)</td>
<td>(-7.8)</td>
<td>(-1.7)</td>
<td>(-.8)</td>
</tr>
<tr>
<td>1940–48</td>
<td>.0073</td>
<td>.0048</td>
<td>.3107</td>
<td>-.3294</td>
<td>-.0186</td>
<td>.0290</td>
</tr>
<tr>
<td></td>
<td>(2.1)</td>
<td>(1.3)</td>
<td>(4.4)</td>
<td>(-3.7)</td>
<td>(-.3)</td>
<td>(.4)</td>
</tr>
<tr>
<td>1949–58</td>
<td>.0108</td>
<td>.0023</td>
<td>.2903</td>
<td>-.4788</td>
<td>-.1886</td>
<td>-.1393</td>
</tr>
<tr>
<td></td>
<td>(6.5)</td>
<td>(9.0)</td>
<td>(5.9)</td>
<td>(-7.5)</td>
<td>(-3.8)</td>
<td>(-2.7)</td>
</tr>
<tr>
<td>1959–68</td>
<td>.0077</td>
<td>.0018</td>
<td>.2899</td>
<td>-.3563</td>
<td>-.0664</td>
<td>-.0100</td>
</tr>
<tr>
<td></td>
<td>(3.4)</td>
<td>(7.0)</td>
<td>(3.5)</td>
<td>(-2.3)</td>
<td>(-.5)</td>
<td>(-.1)</td>
</tr>
<tr>
<td>1969–78</td>
<td>.0027</td>
<td>.0000</td>
<td>.5840</td>
<td>-.6567</td>
<td>-.0727</td>
<td>-.0964</td>
</tr>
<tr>
<td></td>
<td>(0.9)</td>
<td>(0.0)</td>
<td>(8.9)</td>
<td>(-5.1)</td>
<td>(-.6)</td>
<td>(-.8)</td>
</tr>
</tbody>
</table>

Note. —t-ratios in parentheses. For definitions (1a)–(1b) regression eq. is:  
$\tilde{R}_a - R_0 = a_1 + a_2 \tilde{d}_a + a_3 d_a - R_0$.

For definition (1d), regression eq. is:  
$\tilde{R}_a - R_0 = a_1 + a_2 \tilde{d}_a + a_3 d_a - R_0 + a_4 \tilde{d}_a - R_0 K + \tilde{e}_n$.

* After regression of $\tilde{a}_3$ on excess market return.

rather than shifting them to the zero dividend group. The attenuation bias of (1a) is thereby reduced in (1b), but at the cost of removing the offset to the net negative information effects in the zero dividend group. As expected, the estimated $\tilde{a}_3$ for the overall period, .0368, is indeed higher than in the first panel, but still far too small to be considered a significant tax effect.

The third panel of table 2 goes a step further and drops from the sample all firms except those that both paid dividends in $t$ and announced them in advance. This definition avoids the biases in the first two sets of tests but sacrifices any contribution to the precision of
estimating the tax effect that might come from including both ex-dividend and non-ex-dividend shares in the same sample.

A glance at this third panel shows that when the known biases are removed in this fashion, no relation remains between excess returns and dividend yields for the ex-dividend firms. The point estimate of the coefficient $a_3$ falls to an economically negligible and statistically insignificant value of .0135. Thus if yield-related effects do exist, they clearly cannot be attributed to differences in the dividend yields of the firms actually paying dividends in the month.

Part B of table 2 presents a test that combines advantages of some of the previous approaches, but in a more efficient way. Rather than throw away observations (as with [1b] and [1c]) or change their values (as with [1a]), the test uses a slope dummy to estimate the dividend coefficient separately for those ex-dividend firms that did and that did not declare in advance. The dummy variable is set to unity if the dividend is declared in advance and zero otherwise. Hence the sum of the two coefficients measures the effect of dividend yields on returns net of direct announcement effects. For the overall sample period the coefficient is negative and not statistically significant, though somewhat larger in absolute value than in part A(1a). The last column of part B is an adjustment (see the discussion in Black and Scholes [1974]) to allow for the correlation between the monthly dividend coefficients and the market return coefficients, $a_{2t}$. This correction further reduces the size and significance of the dividend coefficient.

**Summary**

Summing up to this point, we find the relation between returns and dividend yields to be sensitive to the definition of dividend yield. The
differences in estimated yield effects appear to reflect differences in the degree to which the various short-run measures of expected dividend yield introduce unwanted information effects. After correcting those measures for their information effects we find no significant remaining relation between returns and expected dividend yields—certainly nothing that could be considered a yield-related tax effect of the classic kind. Before we seek to explain this finding, prudence suggests a check for other deficiencies in the testing procedure that may be obscuring a tax effect.

IV. Possible Biases from Inadequate Risk Corrections

The tests of table 2 have been designed to remove information effects, but other biases affecting the dividend coefficient may still remain. Dividend yields, for example, may be proxying for risk. If dividend yields tend to be lower for high-risk firms, as seems likely, and if risk is measured with error, as seems even more likely, the dividend coefficient in a cross-sectional, multiple regression of the kind in tables 1 and 2 will be given a negative twist (at least over those sample periods in which the market return is positive).

A twist of the opposite kind occurs, however, to the extent that firms adjust their dividends slowly to changes in their earnings prospects. In the months following the announcement of good news, a firm’s price per share will tend to be higher and its “true” beta tend to be lower (because of leverage and option effects, if nothing else) than in the months before the announcement. If, at the same time, the firm maintains its dividend or increases it only gradually, its dividend yield will be smaller than before. For these good news firms, low dividend yields will appear to be associated with abnormally low returns in the months after the announcement because the measured beta, based mainly on preannouncement months, will be too high; while for firms maintaining their dividends in the face of bad news, high dividend yields will be associated with high abnormal returns.

Correcting for biases of these kinds is more difficult than for the information effects of the previous section. Since these biases arise, in part, out of nonstationarities and misspecifications in the underlying

---

10 Many of the issues considered in this section are treated in greater detail in Hess (1980).

11 Bias due to correlation of dividend yields with measurement errors in beta cannot be eliminated merely by taking pains to assure that the measures of expected dividend yields in the tests use only past data. In fact, extrapolative measures of expected dividends (of the kind proposed, say, in Morgan [1980a, 1980b] or Litzenberger and Ramaswamy [1981, pp. 7–8]) are likely to increase the bias. Extrapolative measures, by their very nature, will give the appearance of “dividend stabilization” of the kind described even when the firm has actually adapted its dividend to the change in its circumstances.
processes describing returns, we cannot rely on ordinary errors-in-variables approaches (such as those proposed, e.g., in Litzenberger and Ramaswamy [1979]). But we can at least check the sensitivity of the dividend coefficient $\tilde{a}_3$ to variations in the risk measures.

Part A of table 3 shows the effects of a risk measure even worse than the standard, first-pass 5-year $\tilde{b}_{it}$ used in tables 1 and 2, to wit, none at all. (For simplicity, the results are presented only for the dummy variable test corresponding to part B of table 2, since that test is the most efficient.) Note that for the overall sample period, the value of the dividend coefficient is little changed.

Finding a better risk proxy is not as simple as finding a worse one. But the search for risk proxies need not be restricted to single-variable measures such as $\tilde{b}_{it}$. The discussion of the biases above suggests that the current end-of-previous-month price $p_{i,t-1}$ may

### Table 3

**A. Definition (1d) with $\tilde{b}_{it}$ Eliminated**

<table>
<thead>
<tr>
<th>Years</th>
<th>$\tilde{a}_1$</th>
<th>$\tilde{a}_2$</th>
<th>$\tilde{a}_3$</th>
<th>$\tilde{a}_4$</th>
<th>$\tilde{a}_3 + \tilde{a}_4 = \tilde{a}_5$</th>
<th>Adjusted†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940–78</td>
<td>.0101</td>
<td>3682</td>
<td>.4542</td>
<td>-.0859</td>
<td>-.0045</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.3)</td>
<td>(9.0)</td>
<td>(-7.60)</td>
<td>(-1.4)</td>
<td>(-0.1)</td>
<td></td>
</tr>
<tr>
<td>1940–48</td>
<td>.0140</td>
<td>3071</td>
<td>-.3532</td>
<td>.0281</td>
<td>.0518</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.3)</td>
<td>(3.8)</td>
<td>(-3.5)</td>
<td>(-0.3)</td>
<td>(0.7)</td>
<td></td>
</tr>
<tr>
<td>1949–58</td>
<td>.0137</td>
<td>2836</td>
<td>-.4739</td>
<td>-.1903</td>
<td>-.1426</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.2)</td>
<td>(5.7)</td>
<td>(-7.2)</td>
<td>(-3.6)</td>
<td>(-2.7)</td>
<td></td>
</tr>
<tr>
<td>1959–68</td>
<td>.0101</td>
<td>2749</td>
<td>-.3904</td>
<td>-.1155</td>
<td>.0179</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.9)</td>
<td>(2.8)</td>
<td>(-2.5)</td>
<td>(-0.8)</td>
<td>(0.1)</td>
<td></td>
</tr>
<tr>
<td>1969–78</td>
<td>.0028</td>
<td>6009</td>
<td>-.6055</td>
<td>-.0046</td>
<td>-.0478</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(6.8)</td>
<td>(-4.6)</td>
<td>(-0.0)</td>
<td>(-0.4)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3

**B. Definition (1d) with $1/(p_{i,t-1})$ Added‡**

<table>
<thead>
<tr>
<th>Years</th>
<th>$\tilde{a}_1$</th>
<th>$\tilde{a}_2$</th>
<th>$\tilde{a}_3$</th>
<th>$\tilde{a}_4$</th>
<th>$\tilde{a}_6$</th>
<th>$\tilde{a}_3 + \tilde{a}_4 = \tilde{a}_5$</th>
<th>Adjusted†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940–78</td>
<td>.0067</td>
<td>.0006</td>
<td>.3914</td>
<td>-.4588</td>
<td>.0357</td>
<td>-.0674</td>
<td>-.0223</td>
</tr>
<tr>
<td></td>
<td>(5.1)</td>
<td>(.5)</td>
<td>(12.0)</td>
<td>(-8.4)</td>
<td>(3.0)</td>
<td>(-1.37)</td>
<td>(-.5)</td>
</tr>
<tr>
<td>1940–48</td>
<td>.0085</td>
<td>.0003</td>
<td>.3403</td>
<td>-.2884</td>
<td>.0445</td>
<td>.0519</td>
<td>.0959</td>
</tr>
<tr>
<td></td>
<td>(2.6)</td>
<td>(1.)</td>
<td>(5.2)</td>
<td>(-3.2)</td>
<td>(1.8)</td>
<td>(.8)</td>
<td>(1.56)</td>
</tr>
<tr>
<td>1949–58</td>
<td>.0108</td>
<td>.0022</td>
<td>.2947</td>
<td>-.4811</td>
<td>.0008</td>
<td>.1865</td>
<td>-.1132</td>
</tr>
<tr>
<td></td>
<td>(7.0)</td>
<td>(9.)</td>
<td>(6.1)</td>
<td>(-7.4)</td>
<td>(1.)</td>
<td>(-3.7)</td>
<td>(-2.1)</td>
</tr>
<tr>
<td>1959–68</td>
<td>.0048</td>
<td>.0016</td>
<td>.3258</td>
<td>-.4358</td>
<td>.0931</td>
<td>-.1100</td>
<td>-.0563</td>
</tr>
<tr>
<td></td>
<td>(2.0)</td>
<td>(.6)</td>
<td>(4.1)</td>
<td>(-3.0)</td>
<td>(3.4)</td>
<td>(.9)</td>
<td>(-.5)</td>
</tr>
<tr>
<td>1969–78</td>
<td>.0030</td>
<td>-.0016</td>
<td>.5997</td>
<td>-.6130</td>
<td>.0054</td>
<td>-.0132</td>
<td>-.0368</td>
</tr>
<tr>
<td></td>
<td>(1.0)</td>
<td>(-.5)</td>
<td>(9.9)</td>
<td>(-5.2)</td>
<td>(2.)</td>
<td>(-1.)</td>
<td>(-.3)</td>
</tr>
</tbody>
</table>

**Note:** t-ratios in parentheses.

* $R_{it} - R_{it} = a_{1} + a_{2}d_{it} + a_{3}(d_{it} - R_{it}) \cdot K + \epsilon_{it}$.

† After regression of $\tilde{a}_w$ on excess market return.

‡ $R_{it} - R_{it} = a_{1} + a_{2}d_{it} + a_{3}(d_{it} - R_{it}) + a_{4}(d_{it} - R_{it}) \cdot K + a_{5}(1/p_{i,t-1}) + \epsilon_{it}$. 

---

**Finding a better risk proxy is not as simple as finding a worse one. But the search for risk proxies need not be restricted to single-variable measures such as $\tilde{b}_{it}$. The discussion of the biases above suggests that the current end-of-previous-month price $p_{i,t-1}$ may**
contain additional and more recent information about the firm’s risk and prospects than the 5-year \( \tilde{b}_u \). Part B of table 3 shows the effect of adding that variable—actually its reciprocal \( 1/p_{it-1} \)—to make its scaling comparable to that of the dividend-yield variable. Note that it does indeed contribute significantly (more so, in fact, than \( b_u \) itself). But the key dividend coefficient remains as small and insignificant as before.\(^{12}\)

V. The Results for Subgroups

Significant tax effects in the aggregate sample may perhaps have been obscured by nonlinear clientele effects of the kind discussed in Litzenberger and Ramaswamy (1980). As a check we present tests similar to theirs in which \( \bar{a}_3 \) is estimated separately under various short-run

\(^{12}\) The variable \( 1/p_{it-1} \) can also provide a striking illustration of our warning (see nn. 8, 11) that short-run measures of expected dividend yield, even when purged of announcement effects, may still lead to biased estimates of yield-related tax effects. That variable, after all, can be considered a measure of expected dividend yield in which the forecasted dividend of every firm next period is taken to be $1. Such a forecast is crude but is at least based entirely on past information and hence is free from announcement effects of the kind considered earlier. For all its crudeness as a forecast, however, \( 1/p_{it-1} \) has a positive and significant coefficient when used as the measure of expected dividend yield in tests for tax effects. Since the numerator is a constant, the significant positive value of the coefficient can come only from the information about the firm’s future prospects, additional to that in \( \tilde{b}_u \), conveyed by the current price in the denominator. Any other measure of dividend yield with \( p_{it-1} \) in the denominator, such as those in Litzenberger and Ramaswamy (1981) or in Auerbach (1981, esp. pp. 16–18), will, of course, be subject to the same upward bias.
definitions of dividend yield for subgroups of firms ranked by past average dividend yield. In the tests shown in table 4, the ranking is by the Black-Scholes yield-ranking variable (previous year dividends divided by end-of-year price) updated monthly. Group I contains the 20 percent of firms ranked lowest by this measure, group II the next 20 percent, and so on to group V, the highest.

The estimates of $\tilde{a}_3$ for the subgroups in table 4 appear to be no less sensitive to the definition of dividend yield than the overall sample estimates considered earlier. But note that for these tests, correcting for information bias (last column) by the method in A(1a) of table 2 now does not reduce all the dividend coefficients to insignificance. The coefficient for the lowest yield group remains positive and significant even when the nonzero components of the yield variable include only dividends declared in advance.

A closer look at the underlying data, however, raises doubts about the reliability of the estimate for the low-yield group. The cumulants, fractiles, and other relevant sample statistics of the 468 monthly cross-sectional regression coefficients $\tilde{a}_{3t}$ for the lowest dividend-yield group are given in table 5. Note that the observations cluster very tightly around the modal value of zero. A number of extreme outliers are present, however, at both ends of the scale. The highest estimated "tax bracket" is the "1,555" percent reached in October 1968, and the lowest is the "−1,041" percent in January 1971. Not only are the positive outliers larger, but they are more numerous. The right skewness in the distribution is apparent in the fractiles reported in table 5. Note in particular how far the median is below the mean—.098 as compared with .373.

Ample justification exists for throwing out these extreme observations in computing an average tax effect for the period as a whole. But the case becomes more compelling given the underlying observations on yields and returns that produced the extreme coefficients. Figure 1, for example, shows the scatter plot of the relation between returns and yields for the lowest yield group in the month April 1967—another month with estimated $\tilde{a}_{3t} > 15$. Table 6 shows the numerical values for the first 40 observations ranked by size of dividend yield. Note first that only 16 of the 178 firms in the group had nonzero dividend yields—zero yield being entered not at zero but at −.0042 after subtracting the estimated riskless rate of interest. Thus in the scatter plot of figure 1, 91 percent of the observations lie along the vertical straight line through the −.0042 point on the dividend-yield axis. Of the 16 firms with nonzero dividends, two happened to have extremely large increases in price during the month—one with a return of over 60 percent and one with a return of 22.8 percent. These two points alone account for the sign, size, and apparent
in order of magnitude or more beyond permissible tax effects.

Monmouth's such as those pictured in Figure 1, in which one or two
months such as those pictured in Figure 1,

The solid line in Figure 1

...comparing the dotted line with

...group. With these points detected, the slope drops dramatically (be-

...of the cross-sectional regression coefficient for the entire

<table>
<thead>
<tr>
<th>Class Width (Cumulative)</th>
<th>Cumulative Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>0.20</td>
<td>0.30</td>
</tr>
<tr>
<td>0.30</td>
<td>0.60</td>
</tr>
<tr>
<td>0.40</td>
<td>0.80</td>
</tr>
<tr>
<td>0.50</td>
<td>0.90</td>
</tr>
<tr>
<td>0.60</td>
<td>1.00</td>
</tr>
</tbody>
</table>

For lowest division—whole group, 1940–78

Cumulants, Fractiles, and Summary Statistics of \( \hat{g} \)

TABLE 5

DIVIDENDS AND TAXES

1137
turn out to be largely responsible for the seemingly significant value of $\bar{a}_3$ for the lowest yield group. The sensitivity of the sample mean to the extreme values of $\hat{a}_3$ is shown in table 7. When we chop off the 33 cases or 7 percent with cross-sectional regression coefficients greater than 5 in absolute value, the sample mean falls from .373 with a $t$-value of 2.8 to .098 with a $t$-value of only 1.0. If we chop a further 32 whose coefficients fall outside a range of plus or minus 4, the estimated implicit tax bracket falls to .036 with a $t$-value of 0.44.

In principle, we might go on to examine the observations in the other four subgroups. But why bother? In table 4 and some other recent studies are too inefficient and unreliable, even apart from their vulnerability to outliers, to throw

---

---

13 Actually, of course, we did examine the other groups and found, as might be expected, that the outlier problem was less severe because more dividend entries were nonzero in those higher-yield groups in the typical month. The coefficients, however, were still sensitive to trimming.
any light on the issues in dispute. The within-group estimates of tax brackets reflect only what little variation in short-run dividend yields remains after the substantial between-group variation in average yields has been removed. In fact, if the preclassification were completely successful, the within-group variation would be mostly noise!14

14 The substantial negative coefficient that emerges for the highest yield portfolio in table 4 after announcement effects are eliminated remains a puzzle. Cum-ex trading by
An instrument so unreliable permits no firm conclusions about the existence or absence of tax clienteles. But we can say, at least, that the subgroup tests presented provide no grounds for suspecting that important, yield-related tax effects in our aggregate sample are obscured by fitting a straight line to a nonlinear relation.\textsuperscript{15}

VI. Why Tax Effects Should Not Be Expected in Tests Using Short-Run Measures of Expected Dividend Yield

We recognize, as did Black and Scholes (1974), that our failure to detect yield-related tax effects leaves something of a puzzle. Why does a tax effect that seems so plausible to so many on a priori grounds appear to leave no detectable track in the data? In the case of the short-run dividend measures, at least, we believe the puzzle may have a relatively simple solution. Recall that the presumed tax effect with that yield variable is essentially the average difference in rates of return on those shares that go ex dividend in a given month and those that do not. Some or all of this difference is supposedly the equalizing differential necessary to make those owning the shares at the start of the month indifferent between continuing to hold the shares (and paying full tax on the forthcoming dividend) and selling the shares corporate investors might conceivably be responsible. But a more likely candidate is the proxy bias discussed earlier. Fortunately, no great urgency attaches to resolving the puzzle. As we show in the next section, tests of the kind in table 4 could not shed any reliable light on tax-related clienteles even if their econometric deficiencies were repaired.

\textsuperscript{15} A highly nonlinear, in fact U-shaped, relation between returns and yields is reported by Blume (1980), although in tests using a long-run measure of dividend yield rather than the short-run measures considered here. Later work by Keim (1982), however, suggests that Blume's yield effects are confounded with the so-called small-firm effect. After one controls for size of firm, the nonlinear yield effects largely vanish.
cum dividend (and paying tax on the implicit dividend at the long-
term capital gains tax rate). As Elton and Gruber (1970) have shown,
if the capital gains tax rate is less than the rate on dividends, the price
fall from the last cum-dividend day to the first ex-dividend day will be
less than the dividend paid. The (risk-adjusted) rate of return on the
ex-dividend day (and, by extension, on the months containing the
ex-dividend days) will thus be higher than on non-ex days and
months.

This plausible and oft-invoked explanation has, however, a fatal
flaw. If the price falloff on the ex-dividend day were really less than
the dividend (after correcting for risk), then short-term traders, buy-
ing cum-dividend shares and selling them ex dividend, would earn
above-normal profits. Remember that for such in-and-out traders
(and also for tax-exempt institutions) the tax rate on dividends and on
capital gains—in this transaction actually capital losses—is precisely
the same.16

Short-term traders can be expected to dominate the short-term
equilibrium and hence to eliminate the presumed compensating
cum-ex differential. In principle, every investor, taxable or not, can
exploit that profit opportunity from short-term trading, whereas the
potential sellers are only those taxable investors who happen to own
the shares (and are not locked in by holding period restrictions or by
potential taxes on past unrealized gains). Some individual taxpayers
tempted to exploit the profit opportunities in short-term cum-ex
trading may find themselves constrained by the capital loss limitations
and the wash-sale rules. But these provisions do not apply to brokers
or dealers in securities. For such brokers and dealers organized as
corporations, moreover, and for corporate traders generally (includ-
ing casualty insurance companies) the profit potential in short-term
cum-ex trading is further enlarged by the exclusion of 85 percent of
dividends received from any taxable U.S. corporation from taxable
corporate profits.

Transactions costs reinforce the dominance of short-term buyers in
setting the equilibrium cum-ex differential. The round-trip costs
faced by potential sellers among the taxable investors are substantially
larger than those incurred by the brokers and dealers standing ready
capitalize on deviations from the short-run trading equilibrium.

The presence of transactions costs, even the comparatively small
inside market costs of the brokers and dealers, may well keep the
ex-dividend price from always falling by the full amount of the divi-

16 A number of writers, notably Kalay (1977), have pointed out this flaw in the
Elton-Gruber reasoning. In fact the point appears already to have reached the textbook
level—witness the discussion in Copeland and Weston (1979, p. 353).
dend. And since transactions costs are roughly proportional to price, the observed price falloff may well be smaller relative to the dividend yield for low- than for high-yield shares, exactly as the tax-clientele model seems to predict. But such yield-related effects are not tax effects, or at least not tax effects reflecting the presumed tax penalty on dividends over long-term capital gains that the shooting is all about.17

Solving the puzzle for the short-run measures of dividend yield still leaves unanswered why yield-related tax effects have not been convincingly demonstrated in tests using long-run measures, like those of Black and Scholes (1974). Perhaps the explanation is that yield-related tax effects do not accrue month by month as assumed implicitly in many standard after-tax models of asset pricing. They can show up in principle only in ex months, and there they are eliminated by short-run tax trading. Or perhaps month-by-month accruals of tax-related differences in returns could arise but are eliminated by supply adjustments as described in Black and Scholes (1974) or by tax sheltering as in Miller and Scholes (1978). Or perhaps the tax effect does accrue month by month, but the data are so noisy that the tax effect has escaped detection by existing instruments. Which of these explanations is correct remains to be seen. But we hope that the search can proceed more effectively now that we know at least where not to look.

References


17 Tests rejecting the tax interpretation of ex-dividend day returns are presented in two recent studies, one on U.S. data by Hess (1982) and one on Canadian data by Lakonishok and Vermaelen (1981).


