Patrick J. Hess
Ohio State University

Test for Tax Effects in the Pricing of Financial Assets*

This paper tests four models describing the impact of taxes on financial asset returns: no-tax effects; Miller's equilibrium; the two-parameter-pricing models of Brennan and Litzenberger and Ramaswamy; and tax-induced-dividend clientele effects. These tests are carried out by casting the hypotheses in terms of systems of equations and identifying the parameter restrictions implied by each of the hypotheses. The systems approach is particularly useful for testing for clientele effects because we are able to identify within- and between-equation restrictions without specifying the marginal tax rate of the clientele which holds a particular security. My results suggest that none of tax-motivated hypotheses is descriptive of the data.

This paper tests four competing hypotheses regarding taxation and the pricing of financial assets: no differential tax effects in the pricing of stocks or bonds; the equilibrium condition described by Miller (1977); the two-parameter after-tax pricing equations of Brennan (1970) and Litzenberger and Ramaswamy (1979); and dividend-related clientele effects described by Elton and Gruber (1970) and Litzenberger and Ramaswamy (1980). These hypotheses make substantially different predictions about the impact of taxes on financial decisions; however, when stated in terms of returns the hypotheses imply parameter restrictions with the same functional form. The commonality of the parameter restrictions allows simultaneous testing of all four models by testing the form of the parameter restrictions. This approach differs from that of previous researchers who have by and large implicitly assumed that the parameter restrictions implied by the respective models hold (see Elton and Gruber 1970; Black and Scholes 1973, 1974; Litzenberger and Ramaswamy 1979, 1980; Miller and Scholes 1982). Conditional on this assumption, parameter estimates have been made and

* This paper is a distillation of several research projects I have undertaken in the last two years. Special thanks are in order to G. Constantinides, E. Fama, J. Ingersoll, A. Madansky, and M. Scholes. I owe a particularly important debt to Merton Miller for help and encouragement. Finally, I would like to thank Steve Buser and E. Han Kim.

(*Journal of Business*, 1983, vol. 56, no. 4)

0021-9398/83/5604-0006$01.50

the debate has centered on the value of the parameters.[1] Past research has focused on which hypotheses should be accepted, ignoring the possibility that none of the hypotheses are descriptive of the data. My tests address this crucial question.

The next section of this paper lays out two simple statistical models which are the basis of my empirical tests. I show that the competing hypotheses are nested within this model and what parameter restrictions may be extracted from the hypotheses. Following this discussion I outline the estimation and testing procedure, present the empirical results, and finally state my conclusions. Interestingly, the results turn out to be inconsistent with all four hypotheses.

Statistical Models

The simplest model used in subsequent empirical tests is

$$\tilde{R}_{it} = \gamma_{0i} + \gamma_{1i}d_{it} + \gamma_{2i}r_{ft} + \tilde{\mu}_{it}, \qquad i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T, \tag{1}$$

where $\tilde{R}_{it}$ is the return of security i in period t, d_{it} is security i's dividend yield in period t, r_{ft} is the taxable riskless rate of interest set at the beginning of period t, $\tilde{\mu}_{it}$ is a disturbance, and N is the number of securities. It is assumed that the disturbance vector, $\mu_t' = (\mu_{1t}, \mu_{2t}, \ldots, \mu_{Nt})$, has the distribution $NID(0, \Sigma)$ where 0 is the $N \times 1$ null vector and Σ is an unrestricted covariance matrix. Because (1) holds for all assets, it also holds for any portfolio; hence,

$$\tilde{R}_{mt} = \gamma_{0m} + \gamma_{1m}d_{mt} + \gamma_{2m}r_{ft} + \mu_{mt}, \qquad t = 1, 2, \ldots, T, \tag{2}$$

where $\tilde{R}_{mt}$ is the return on an arbitrary portfolio m. Combining (1) and (2),[2]

$$\tilde{R}_{it} = \alpha_{0i} + \alpha_{1i}\tilde{R}_{mt} + \alpha_{2i}d_{mt} + \alpha_{3i}r_{ft} + \alpha_{4i}d_{it} + \tilde{\nu}_{it} \quad i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T, \tag{3}$$

where

$$\alpha_{0i} = \gamma_{0i} - \beta_i\gamma_{0m}$$

$$\alpha_{1i} = \beta_i$$

$$\alpha_{2i} = -\beta_i\gamma_{1m}$$

$$\alpha_{3i} = \gamma_{2i} - \beta_i\gamma_{2m}$$

1. An example of this is provided by Litzenberger and Ramaswamy (1979, 1982) and Miller and Scholes (1982).

2. All variables except for security returns are predetermined.

$$\alpha_{4i} = \gamma_{1i}$$

$$\tilde{\nu}_{it} = \tilde{\mu}_{it} - \beta_i \mu_{mt}$$

$$\beta_i = \text{cov}(\tilde{R}_i, \tilde{R}_m)/\text{var}(\tilde{R}_m).$$

The models shown in (1) and (3) are devoid of theoretical content. To give the models content, we need to identify parameter restrictions implied by alternative theoretical models. Four such tax-motivated models will be considered. The first is simply no differential tax effects in the pricing of stocks or bonds, as in a world in which the marginal tax rate on investment income is zero, or if nonzero is the same for all investment income regardless of kind or source. The second alternative is derived from the equilibrium conditions on corporate capital structures described by Miller (1977). The third comprises the after-tax, two-parameter models of Brennan (1970) and Litzenberger and Ramaswamy (1979); and the fourth is the dividend-related clientele models proposed by Elton and Gruber (1970) and Litzenberger and Ramaswamy (1980).

No-Tax-Effect Hypothesis

With equal taxation of all investment income, pretax returns of stocks will not adjust to reflect differential taxation of dividends over capital gains. Ignoring imperfections in capital markets, security returns will be independent of dividend yields;[3] the dividend yield coefficients in (1), γ_{1i}, will all equal zero. The impact of the taxable-riskless rate depends on the behavior of securities' risk premiums over time. If we assume the risk premiums are constant, equal taxation of stocks and bonds implies that expected returns will adjust one for one with changes in the taxable-riskless rate. Thus the no-tax-effect hypothesis implies the parameter restrictions, $\gamma_{1i} = 0$, $\gamma_{2i} = 1$, $i = 1, 2, \ldots, N$. Applying these restrictions to the system (3), $\alpha_{2i} = \alpha_{4i} = 0$, $\alpha_{3i} = 1 - \beta_i$, $i = 1, 2, \ldots, N$.

Miller's Equilibrium

Miller (1977) describes an equilibrium embedding corporate and personal taxes in which firms' capital-structure decisions are irrelevant. In its simplest form, Miller's equilibrium assumes that the returns to corporate bondholders are fully taxable while the returns to stockholders are effectively tax exempt. The irrelevance of the corporate-capital-structure decision implies that the marginal bondholder has a tax rate equal to the corporate tax rate. At the margin, the reduction in corporate taxes from an additional dollar of interest deductions is exactly

3. Miller and Modigliani (1961) demonstrate the irrelevancy of dividend policy in a perfect capital market with equal taxation of dividends and capital gains.

offset by increased taxes for bondholders. The tax-exempt status of stocks and the equilibrium condition implicit in the pricing of bonds restrict the γ_{1i} and γ_{2i} coefficients of (1). Because stock returns are tax exempt, there is no distinction between dividends and capital gains; once again γ_{1i}'s all equal zero. However, the heavier taxation of bonds relative to stocks means that expected returns of stocks will not adjust one for one with changes in the taxable-riskless rate. Instead, the adjustment will only be at one minus the corporate tax rate, τ_c, or approximately half as fast. In summary, Miller's equilibrium imposes the restrictions, $\gamma_{1i} = 0, \gamma_{2i} = 1 - \tau_c, i = 1, 2, \ldots, N$. For the system (3), we have $\alpha_{2i} = \alpha_{4i} = 0, \alpha_{3i} = (1 - \beta_i)(1 - \tau_c), i = 1, 2, \ldots, N$.

Two-Parameter After-Tax Pricing Equation

Brennan (1970) and Litzenberger and Ramaswamy (1979) develop two-parameter pricing models incorporating differential taxation of dividends over capital gains. Both these models assume that dividends and interest are taxed as ordinary income and capital gains are taxed at more favorable rates. Brennan's model is

$$\begin{aligned} E(\tilde{R}_i) - r_f &= b^*\beta_i + \tau(d_i - r_f) \\ b^* &= \mathrm{E}(\tilde{R}_m) - r_f - \tau(d_m - r_f), \end{aligned} \tag{4}$$

where $E(\tilde{R}_m)$ is the expected return on the market portfolio of all risky assets, τ represents a tax differential between dividends and capital gains, d_m is the dividend yield of the market portfolio, and β_i is the covariance of asset i's return with the market divided by the variance of the market. Litzenberger and Ramaswamy extend this model by assuming both margin and income constraints on borrowing. Their model is

$$\begin{aligned} E(\tilde{R}_i) - r_f &= a + b\beta_i + c(d_i - r_f) \\ a &= E(\tilde{R}_{Z^*}) - r_f \\ b &= E(\tilde{R}_m) - E(\tilde{R}_{Z^*}) - c(d_m - r_f), \end{aligned} \tag{5}$$

where c is a tax differential reduced by a shadow price that reflects increases in investors' ability to borrow from an additional dollar of dividends and $E(\tilde{R}_{Z^*})$ is the expected return on a zero-beta portfolio with a dividend yield equal to the taxable riskless rate. These models are the standard two-parameter pricing models adjusted for differential taxation of dividends and interest income relative to capital gains.

If the models (2) and (3) hold period by period, the stochastic versions are

$$\tilde{R}_{it} - r_{ft} = b_t^*\beta_{it} + \tau_t(d_{it} - r_{ft}) + \tilde{\mu}_{it}, \qquad i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T, \tag{4a}$$

and

$$\tilde{R}_{it} - r_{ft} = a_t + b_t\beta_{it} + c_t(d_{it} - r_{ft}) + \tilde{\mu}_{it}, \quad i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T. \tag{5a}$$

For purposes of estimation the parameters of (4a) and (5a) are assumed to be constant. In terms of equation (1), the parameter values for Brennan's model are $\gamma_{0i} = b^*\beta_i$, $\gamma_{1i} = 1 - \gamma_{2i} = \tau$, $i = 1, 2, \ldots, N$. For Litzenberger and Ramaswamy the parameter values are $\gamma_{0i} = a + b\beta_i$, $\gamma_{1i} = 1 - \gamma_{2i} = c$, $i = 1, 2, \ldots, N$. In principle, the parameter values of these models may differ; however, the parameter restrictions of the models are identical. In both cases the dividend yield effect is the same across securities and the dividend yield and risk-free rate coefficients sum to one. The second restriction reflects the fact that both models assume that dividends and interest income are taxed at the same rate.

Conditional on identifying the true market portfolio of all risky assets, Brennan's model restricts the system (3), as $\alpha_{0i} = 0$, $\alpha_{2i} = -\beta_i\tau$, $\alpha_{3i} = (1 - \beta_i)(1 - \tau)$, $\alpha_{4i} = \tau$, $i = 1, 2, \ldots, N$. If m is any arbitrary portfolio, the parameter restrictions of (3) are identical with the exception of the intercepts. Without knowledge of the relative betas of the securities and portfolio m the intercepts of (3) are unrestricted.

The respective parameter restrictions for Litzenberger and Ramaswamy model also depend on the ability to identify the true market portfolio of all risky assets. For the market portfolio, the restrictions are $\alpha_{0i} = E(\tilde{R}_{Z^*})(1 - \beta_i)$, $\alpha_{2i} = -\beta_i c$, $\alpha_{3i} = (1 - \beta_i)c$, $\alpha_{4i} = c$, $i = 1, 2, \ldots, N$. For any arbitrary portfolio m, there are no restrictions on the intercepts of (3); otherwise the restrictions are identical to those of Brennan's model.

Tax-Clientele Effects

The final model we consider is in the spirit of the dividend-clientele models proposed by Elton and Gruber (1970) and Litzenberger and Ramaswamy (1980). In particular, assume that each security potentially has its own unique tax parameter. In terms of the system (1) this implies the parameter restrictions $\gamma_{1i} = \tau_i$, $\gamma_{2i} = 1 - \tau_i$, $i = 1, 2, \ldots, N$. The parameter restrictions for the system (3) are $\alpha_{2i} = -\beta_i\tau_m$, $\alpha_{3i} = 1 - \tau_i - \beta_i(1 - \tau_m)$, $\alpha_{4i} = 1 - \tau_i$. For the system (1), tax clienteles impose within-equation restrictions; for (3), tax clienteles impose both within- and between-equation restrictions.

Summary

Each of the theoretical models analyzed imposes restrictions on the parameters of (1) and (3), which in turn provide a basis for testing the respective models. It is interesting to note that though the parameter values of the three models differ, the models often share parameter restrictions.

It is appropriate to note also that the approach taken here differs from that of most previous researchers on the subject. A common procedure of past research is estimation of the cross-sectional regression (see, e.g., Litzenberger and Ramaswamy 1979, 1980; Miller and Scholes 1982).

$$\tilde{R}_{it} - r_{ft} = a_t + b_t\hat{\beta}_{it} + c_t(\hat{d}_{it} - r_{ft}) + \hat{v}_{it}, \quad i = 1, 2, \ldots, N, \qquad (5)$$

where $\tilde{R}_{it}$ is the return of security i in month t, r_{ft} is the 1-month T-bill rate set at the beginning of t, $\hat{\beta}_{it}$ is estimated from month $t - 60$ to month $t - 1$, and $\hat{d}_{it}$ is a measure of the anticipated dividend yield of i in month t. This regression is estimated monthly and the estimates are pooled over time as a simple average. While a number of questions might be raised regarding this estimation procedure, the most significant feature, from our point of view, is the implicit assumption that the marginal tax differential, as measured by the average dividend-yield coefficient $\bar{c}$, is the same across securities and that it equals one minus the risk-free-rate coefficient. In this respect the research here asks a more fundamental question: Can *any* of the proposed differential-tax models explain the structure of security returns?

Estimating and Testing the Models

With equations (1) and (3), capital markets are described by a system of time-series regressions: one equation for each security.[4] To estimate such a system and test the respective parameter restrictions, the cross-sectional dependence in security returns must be taken into account. Zellner (1971) has called systems like (1) and (3) "seemingly unrelated regressions" (SUR). The regressions have different independent variables but are related by the contemporaneous correlation of disturbances. While the disturbances are presumed to be contemporaneously correlated across equations, the noncontemporaneous covariances and autocovariance are here all assumed to be zero. If the disturbances of the system are normally, independently, and identically distributed over time, the SUR estimator is the maximum likelihood estimator. The feasible estimator involves an estimated covariance matrix rather than the true one. In general, the feasible estimator has the same asymptotic distribution as the maximum likelihood estimator (Schmidt 1976, p. 85). Unlike the two-pass, cross-sectional regression of many previous authors (e.g., Litzenberger-Ramaswamy, Miller-Scholes), the

4. Gordon and Bradford (1980) estimate a model similar to (3), but they did not test any of the restrictions implied by their model. Gibbons (1982) tests a class of asset pricing models by examining parameter restrictions.

SUR systems avoid some of the problems posed by errors of measurement in the risk parameters, β_i, as well as allow for explicit testing of the parameter restrictions implied by the various models.[5]

Testing the Parameter Restrictions

For all four of the hypotheses, the system (1) is linear in the parameter. To test the hypotheses with (1), I estimate the unrestricted system, and, conditional on the estimated covariance matrix, test the individual restrictions. For a known covariance matrix the reported test statistic is distributed as F with degrees of freedom equal to the number of equations times the degrees of freedom per equation (see Theil 1971, p. 313). Because the covariance matrix of the unrestricted system is unknown, the asymptotic distribution is χ^2 divided by its degree of freedom (the number of restrictions imposed) (Theil 1971, p. 402). The system (1) tests three restrictions: the coefficients on the dividend yields are the same across equations, the coefficients on the riskless rate are the same across equations, and the sum of the riskless rate coefficients and the dividend yield coefficients add up to one. The no-tax hypothesis and the hypothesis of Brennan and Litzenberger and Ramaswamy impose all three of these restrictions. Miller's hypothesis imposes only the first two, and only the third restriction applies to the tax-clientele-effect hypothesis. To get an idea of what kind of model might best fit the data, each of the restrictions is tested independently.

For all but the no-tax-differential hypothesis, the system (3) is nonlinear in its parameters. These systems are estimated with a Gauss-Newton iterative procedure. For a normally, independently, and identically distributed disturbance vector, the Gauss-Newton procedure is asymptotically maximum likelihood (see Judge et al. 1980, p. 735). At each iteration the estimation procedure linearizes the likelihood function, and, of course, the linearizations depend on the parameter space. Because of this, it is not possible to estimate the unrestricted model and then independently test each of the restrictions; instead, each of the four hypotheses is tested with a likelihood ratio test. The test statistic is $T(\ln|\hat{\Sigma}_r| - \ln|\hat{\Sigma}_u|)$, where T is the number of observations per equation, $|\hat{\Sigma}_r|$ is the determinant of the estimated covariance matrix for the restricted system, and $|\hat{\Sigma}_u|$ is the determinant of the estimated covariance matrix for the unrestricted system. This test statistic is distributed asymptotically as χ^2 with degrees of freedom equal to the number of restrictions (the reduction in the number of parameters) (see Theil 1971, p. 396).

5. Litzenberger and Ramaswamy (1979, 1980) have suggested an estimator adjusted for errors in beta. Hess (1980) shows that without prior knowledge about the *true* covariance matrix of estimation errors, this adjusted estimator is not consistent.

The Sample of Securities

In the system (1) and (3) we have assumed that the dividends paid in any period t are known at the end of period $t - 1$; hence dividend yields are predetermined variables, and they affect only expected returns. However, (1) and (3) are estimated with monthly returns, and firms often announce and pay dividends in the same month. To the extent that dividend announcements convey information to the market, the coincident payment and announcement of dividends will contaminate our tests for tax effects.[6] In principle, this problem could be solved by modeling the expected dividends of securities and appending these equations to the systems (1) and (3). Such a procedure would result in nonlinear systems of $2N$ equations with across-equation rational expectation restrictions of the type discussed by Mishkin (1982).[7] Unfortunately, because dividends are only observable quarterly, we can at most use quarterly returns to estimate the systems. With an unrestricted covariance matrix, a system of $2N$ equations and k parameters per equation requires a minimum of $2N + k$ observations per equation. Estimation of a moderate-size system of 30 securities would require over 15 years of quarterly returns, and it seems unreasonable to assume that the parameters of the return distribution are constant for such long periods.

While the rational expectation definitions of dividend yields impose no a priori limits on our sample, the assumption that the parameters of (1) and (3) are constant during each sample period severely limits the size of systems that can be estimated with the rational-expectations definition. To avoid these problems a sample of securities is identified which always paid regular quarterly dividends and announced all dividends before the month of payment.[8] For these securities, dividend yields truly are predetermined variables.[9]

Five years of monthly returns are used to estimate the systems (1) and (3), and these systems are limited to 30 securities. Before 1951 only a few securities meet the requirements on the announcement and payment of dividends. For each of the two 5-year periods between January 1951 and December 1960 it is possible to estimate a system of 30 securities. For the three 5-year periods between January 1961 and December 1975 three systems of 30 could be estimated; and finally, for the 5-year period January 1976–December 1980 four systems are es-

6. See Miller and Scholes (1982) for a discussion of the problems of defining expected dividend yields in the presence of announcement effects.

7. Although Litzenberger and Ramaswamy (1982) and Morgan (1982) use forecasted dividend yields in their tests, they ignore the additional restrictions implied by this procedure.

8. For present purposes, the date of payment is the ex date.

9. Obviously this sample is not randomly selected; however, it is unclear why the selection criteria would introduce a bias into the tests.

timated. While this sample does not include all securities, the sample is large and does effectively eliminate announcement effects.

Data

Monthly returns are taken from the Center for Research in Security Prices (CRSP) monthly return file. Yields on 1-month T-bills are used to proxy for the riskless rate of interest. The T-bill yields are taken from Ibbotson and Sinquefield (1982). The CRSP value-weighted index of all New York Stock Exchange common stock is portfolio m in system (3). When the intercepts of (3) are restricted, the value-weighted index is proxying for the market portfolio of all risky assets; in all other tests it is simply an arbitrary portfolio. The amount of the dividends, their announcement dates and payment dates (ex-date) are taken from the CRSP monthly master file.

Empirical Results

Tests without a Market Index: System (1)

The results of estimating system (1) are reported in table 1. For each system I report the F-statistic for each of the restrictions, degrees of freedom, probability of observing the sample given each of the restrictions along with restricted estimates, and their asymptotic standard errors. The restriction that the dividend yield coefficient is the same across securities is inconsistent with the results in table 1. The restriction is typically rejected at the .0001 significance level (one chance in 10,000).[10] Given that the restriction on dividend-yield coefficients fails to hold, economic interpretation of the restricted estimates is suspect. Nevertheless, it is interesting to note that in 11 of the 15 systems the restricted estimates are negative. The restrictions on the risk-free rate coefficients are not as easily rejected, but there is only scant evidence of a common effect. In six of the 15 systems the restriction of a common coefficient on the risk-free rate is rejected at the .0001 level, and in six other systems it is rejected at the .01 level. On balance it seems reasonable to conclude that we can reject both the restriction of a common dividend-yield coefficient and a common risk-free rate coefficient. This evidence conflicts with the no-tax-differential hypothesis, Miller's simple capital structure hypothesis, and the two-parameter pricing equations of Brennan and Litzenberger and Ramaswamy.

In system (1) the restriction that the sum of the dividend-yield

10. Inspection of the unrestricted estimated coefficients reveals no clear pattern. Many of the estimates are greater than $+1$ and many are less than -1 while the remaining estimates appear to be approximately uniformly distributed over the interval -1 to $+1$.

TABLE 1 Estimates and Tests without a Market Index: System (1)
Equation (1): $\bar{R}_{it} = \gamma_{0i} + \gamma_{1i} d_{it} + \gamma_{2i} r_{ft} + \bar{\epsilon}_{it}$

	Restrictions								
	January 1951–December 1955			January 1956–December 1960			January 1961–December 1965		
	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$
System (1):									
F-statistics	3.350	2.360	3.046	1.740	1.956	2.506	4.184	1.884	1.752
Probability level	.0001	.0001	.0001	.0088	.0018	.0001	.0001	.0031	.0073
Restricted estimates	−.086	14.666	. . .	−.727	9.195	. . .	−.795	.091	. . .
Standard errors	.084	3.083	. . .	.151	2.422	. . .	2.225	9.1083	. . .
System (2):									
F-statistics	. . .	. . .	. . .	. . .	. . .	. . .	2.452	2.189	2.146
Probability level	. . .	. . .	. . .	. . .	. . .	. . .	.0001	.0003	.0001
Restricted estimates	. . .	. . .	. . .	. . .	. . .	. . .	−.784	−16.161	. . .
Standard errors	. . .	. . .	. . .	. . .	. . .	. . .	.177	7.826	. . .
System (3):									
F-statistics	. . .	. . .	. . .	. . .	. . .	. . .	2.150	1.387	1.227
Probability level	. . .	. . .	. . .	. . .	. . .	. . .	.0004	.0825	.1857
Restricted estimates	. . .	. . .	. . .	. . .	. . .	. . .	.836	−.409	. . .
Standard errors	. . .	. . .	. . .	. . .	. . .	. . .	.243	9.098	. . .

	January 1966–December 1970			January 1971–December 1975			January 1976–December 1980		
	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$	$\gamma_{1i} = \gamma_{1j}$	$\gamma_{2i} = \gamma_{2j}$	$\gamma_{1i} + \gamma_{2i} = 1$
System (1):									
F-statistics	2.033	3.428	3.062	2.784	2.088	1.352	4.700	1.318	2.420
Probability level	.0010	.0001	.0001	.0001	.0006	.0972	.0001	.1202	.0001
Restricted estimates	−.242	2.565	. . .	−.175	−4.584	. . .	−.019	−.692	. . .
Standard errors	.238	3.219	. . .	.145	2.119	. . .	.0837	.637	. . .
System (2):									
F-statistics	3.091	1.919	1.778	2.340	2.879	3.091	5.651	3.067	2.996
Probability level	.0001	.0024	.0060	.0001	.0001	.0001	.0001	.0001	.0001
Restricted estimates	.571	3.184	. . .	−.433	3.993	. . .	−.174	−2.147	. . .
Standard errors	.210	4.117	. . .	.182	2.194	. . .	.130	.545	. . .
System (3):									
F-statistics	2.614	1.279	1.207	4.013	3.904	3.787	5.020	1.647	2.137
Probability level	.0001	.1468	.2037	.0001	.0001	.0001	.0001	.0160	.0003
Restricted estimates	−.575	−2.563	. . .	.389	−.494	. . .	.237	.196	. . .
Standard errors	.194	3.254	. . .	.210	2.811	. . .	.100	.563	. . .
System (4):									
F-statistics	. . .	. . .	. . .	. . .	. . .	. . .	2.839	3.879	1.927
Probability level	. . .	. . .	. . .	. . .	. . .	. . .	.0001	.0001	.0001
Restricted estimates	. . .	. . .	. . .	. . .	. . .	. . .	−.040	1.205	. . .
Standard errors	. . .	. . .	. . .	. . .	. . .	. . .	.132	.280	. . .

NOTE.—Degrees of freedom are 29 and 1,710 for the restrictions $\gamma_{1i} = \gamma_{1j}$ and $\gamma_{2i} = \gamma_{2j}$. For the restriction that $\gamma_{1i} + \gamma_{2i} = 1$, the degrees of freedom are 30 and 1,710. The minimum probability level reported is .0001.

coefficients and the risk-free rate coefficient add up to one is the identifying restriction for the clientele-effect hypothesis. This restriction fares only slightly better than the common coefficient restrictions of the other hypotheses. In nine out of 15 systems the restriction that the sum of the coefficients equals one can be rejected at the .0001 level and at the .01 level for two systems. In short, even the relatively weak within-equation restrictions implied by the dividend-induced clientele model fail to conform to the data.

On balance our tests with system (1) reveal no apparent pattern in the relations between security returns, dividend yields, and the riskless rate of interest. None of the competing hypotheses provides an adequate description of the relations.

Tests with a Modified Market Model: System (3)

Tests with system (3) are reported in table 2. For each hypothesis and system the table reports a χ^2 statistic and approximate probability level that the parameter reduction implied by the hypothesis conforms to the data. For the no-tax-differential effect hypothesis there are no tax parameters to estimate; however, for the other systems there are one or more parameters to be estimated. From inspection of the probability levels in table 2, it is clear that none of the hypotheses fits the data very well: the two-parameter models appear to do the worst and the clientele hypothesis the best. The dividend-yield coefficients in the two-parameter model are often negative and generally close to zero. For the clientele model, the estimated tax rate on the CRSP value-weighted index is often negative and large in absolute value. As is true with the system (1), the results for system (3) suggest that none of the hypotheses conform to the data.

Dividend Yields and Shifts in Expected Returns

Tables 1 and 2 reveal a complex relation between expected returns and dividend yields. In many respects, the results are unsettling because they reject the simplest models of the impact of dividends: the data tell us that all four hypotheses are bad approximations, but they do not tell us why.

One possible explanation is that dividend yields are proxying for other factors that affect returns. Imagine, for example, a world where dividends are "steady" in the sense of not being adjusted immediately by firms to maintain constant dividend yields. A decrease in the price of a firm's shares, on average, will thus imply a higher dividend yield and an increase in price will, on average, imply a lower dividend yield. For leveraged firms, changes in stock prices may also result in changes in the riskiness of the firms' stock. Unless firms make compensating adjustments in leverage, a decrease in the price of a common stock implies a higher leverage ratio and an increase a lower leverage ratio. As a result, both dividend yields and expected returns may be increasing or decreasing simultaneously. Stated differently, equations have been omitted, namely, the ones determining expected returns, and dividend yields are proxying for the omitted equations.

To see the possible impact of this omission, consider equation (1). The intercept of equation (1) may be expressed as $\gamma_{0it} = \bar{\gamma}_{0i} + \tilde{v}_{it}$, where $\tilde{v}_{it}$ is the difference between the expected return of security i in period t and its average during the sample period (ignoring taxes). Substituting this expression in equation (1),

$$\tilde{R}_{it} = \bar{\gamma}_{0i} + \gamma_{1i}d_{it} + \gamma_{2i}r_{ft} + \tilde{\epsilon}_{it} + \tilde{v}_{it}. \tag{1a}$$

The disturbance term in equation (1a) consists of the true disturbance plus the deviation of the expected return in period t from its average during the sample period.

Our previous discussion suggests that, on average, $\tilde{v}_{it}$ and d_{it} are both negatively related to stock prices and therefore are positively related to each other. Further, the relation is not easily modeled. The best that can be done is to present evidence that establishes the plausibility of the connection; however, the evidence has important implications for interpreting tests of the relation between dividend yields and expected returns (and perhaps other anomalies of asset pricing).

To test for this relation, equation (1) can be generalized to

$$\tilde{R}_{it} = \gamma_{0i} + \gamma_{1i}r_{ft} + \gamma_{2i}d_{it} + \gamma_{3i}d_{it-1} + \gamma_{4i}d_{it-2} + \tilde{\epsilon}_{it}, \tag{1b}$$

where d_{it} is the dividend yield in month t. In effect, (1b) includes the dividend yield terms of (1) in both the cum and ex months. For example if a stock goes ex in December, (1b) includes December's yield in January (γ_{3i}) and in February (γ_{4i}). Because of the characteristics of our sample, there can be no information effects in the coefficients γ_{2i}, γ_{3i}, or γ_{4i}. Clearly, there are no tax effects of the kind hypothesized by Brennan and Litzenberger and Ramaswamy in the coefficients γ_{3i} and γ_{4i}, since the dividend has already been paid. The cum-month effects are purely statistical artifacts from their point of view. The tests conducted with (1b) are tests of no effect of current or lagged dividend yields. If both lagged and current dividend yields turn out to be nonzero, there is reason to believe yields are proxying for shifts in expected returns. The results of these tests are reported in table 3.

As can be seen, both the lagged and current dividend yields appear to be nonzero in all time periods. It would be difficult, on the basis of these results, to conclude either that the current values dominate the lagged value or vice versa. The current and lagged values both contribute to the explanation of expected returns, thus warning again of the

TABLE 2 Estimates and Tests with a Modified Market Model: System (3)

		No Tax Effect		Litzenberger and Ramaswamy				Brennan		
Period	System No.	χ^2	Probability Levels	χ^2	Probability Levels	$E(\tilde{R}_{Z^*})$	C	χ^2	Probability Levels	τ
January 1951–December 1955	(1)	152.6	.0001	185.4	.0001	.008	−.197	194.5	.0001	−.075
						(.001)	(.080)			(.077)
January 1956–December 1960	(1)	169.1	.0001	194.5	.0001	.015	−.643	207.7	.0001	−.478
						(.002)	(.146)			(.148)
January 1961–December 1965	(1)	189.4	.0001	229.1	.0001	.006	−.462	228.6	.0001	−.413
						(.004)	(.197)			(.196)
	(2)	180.0	.0001	215.2	.0001	.004	−.536	215.5	.0001	−.546
						(.003)	(.158)			(.158)
	(3)	188.7	.0001	222.5	.0001	.010	.390	221.3	.0001	.467
						(.004)	(.194)			(.190)
January 1966–December 1970	(1)	159.9	.0001	209.0	.0001	.008	−.248	206.9	.0001	−.368
						(.003)	(.227)			(.222)
	(2)	145.4	.0002	178.0	.0003	−.017	.471	189.1	.0001	.380
						(.004)	(.199)			(.201)
	(3)	136.9	.0011	173.0	.0007	.002	−.565	175.3	.0006	−.635
						(.004)	(.171)			(.170)
January 1971–December 1975	(1)	183.3	.0001	223.5	.0001	.003	−.116	223.3	.0001	−.108
						(.004)	(.141)			(.141)
	(2)	168.3	.0001	191.1	.0001	.007	−.245	191.0	.0001	−.254
						(.005)	(.157)			(.157)
	(3)	182.1	.0001	199.9	.0001	.007	.634	202.0	.0001	.542
						(.004)	(.204)			(.204)
January 1976–December 1980	(1)	172.3	.0001	205.0	.0001	.002	−.098	205.7	.0001	−.154
						(.004)	(.086)			(.088)
	(2)	182.7	.0001	228.2	.0001	.0003	−.126	233.5	.0001	−.095
						(.004)	(.125)			(.126)
	(3)	184.0	.0001	228.7	.0001	.0004	.044	227.1	.0001	−.065
						(.003)	(.103)			(.100)
	(4)	160.9	.0001	210.5	.0001	−.003	.251	208.7	.0001	.114
						(.004)	(.120)			(.117)

		Free Intercept			Clientele			Miller		
		χ^2	Probability Levels	τ	χ^2	Probability Levels	τ_m	χ^2	Probability Levels	τ_b
January 1951–December 1955	(1)	154.6	.0001	−.188 (.081)	88.6	.0076	−1.054 (.496)	151.5	.0001	7.09 (3.35)
January 1956–December 1960	(1)	162.8	.0001	−.641 (.147)	113.1	.0001	−.016 (.78)	165.9	.0001	6.565 (2.69)
January 1961–December 1965	(1)	191.4	.0001	−.421 (.200)	123.9	.0001	−2.69 (.99)	184.4	.0001	61.3 (13.1)
	(2)	174.4	.0001	−.454 (.158)	119.6	.0001	1.640 (.943)	179.5	.0001	4.831 (5.79)
	(3)	185.8	.0001	.448 (.194)	124.3	.0001	−1.928 (.798)	184.7	.0001	38.54 (7.27)
January 1966–December 1970	(1)	158.2	.0001	−.326 (.229)	111.1	.0001	−.052 (.806)	143.1	.0002	−16.87 (2.97)
	(2)	142.8	.0003	.315 (.200)	89.1	.0069	.577 (.798)	138.6	.0006	−15.59 (4.22)
	(3)	128.6	.0039	−.644 (.173)	83.2	.0209	1.316 (1.026)	135.4	.0011	−7.50 (4.07)
January 1971–December 1975	(1)	182.5	.0001	−.159 (.143)	132.5	.0001	.698 (.926)	182.1	.0001	5.01 (2.28)
	(2)	166.2	.0001	−.256 (.158)	113.6	.0001	2.893 (.933)	168.4	.0001	1.11 (3.19)
	(3)	178.3	.0001	.520 (.206)	98.8	.0009	1.698 (.855)	177.3	.0001	−11.716 (3.74)
January 1976–December 1980	(1)	172.3	.0001	−.138 (.089)	112.0	.0001	−.458 (.675)	172.2	.0001	.261 (1.491)
	(2)	182.2	.0001	−.052 (.126)	96.7	.0014	−.011 (.461)	177.8	.0001	−1.917 (1.37)
	(3)	184.6	.0001	−.083 (.098)	113.3	.0001	.234 (.636)	183.9	.0001	−2.060 (1.300)
	(4)	161.2	.0001	.079 (.116)	112.4	.0001	−.287 (.773)	158.2	.0001	−3.760 (1.52)

NOTES.—For the no-differential-tax-effect hypothesis, the degrees of freedom are 90. For Litzenberger and Ramaswamy, $E(\tilde{R}_{Z^*})$ is the expected return on a zero-beta portfolio with a dividend yield equal to the taxable riskless rate and c is the coefficient on dividend yields and the taxable riskless rate. The degrees of freedom for Litzenberger and Ramaswamy are 118. For Brennan τ is the tax rate on dividends and interest income and the degrees of freedom are 119. With a free intercept τ corresponds to c in Litzenberger and Ramaswamy's model or the tax rate on dividends and interest income for Brennan. The degrees of freedom for the free intercept are 89. For clientele effects τ_m is the tax rate of the clientele which would own the CRSP value-weighted index. The degrees of freedom are 59. For Miller τ_b is the marginal bondholders' tax rate and the degrees of freedom are 89. The minimum probability level reported is .0001.

TABLE 3 Test of Cum- and Ex-Month Effects

Equation (1b): $\hat{R}_{it} = \gamma_{0i} + \gamma_{1i} r_{ft} + \gamma_{2i} d_{it} + \gamma_{3i} d_{it-1} + \gamma_{4i} d_{it-2} + \hat{\epsilon}_{it}$

	January 1951–December 1955			January 1956–December 1960			January 1961–December 1965		
	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$
System (1):									
F-statistics	2.921	2.518	2.112	3.278	2.837	3.465	3.687	3.618	3.351
Probability level	.0001	.0001	.0004	.0001	.0001	.0001	.0001	.0001	.0001
System (2):									
F-statistics	...	...	...	...	...	...	3.468	3.550	4.240
Probability level	...	...	...	...	...	...	.0001	.0001	.0001
System (3):									
F-statistics	...	...	...	...	...	...	4.670	5.113	4.197
Probability level	...	...	...	...	...	...	.0001	.0001	.0001

	January 1966–December 1970			January 1972–December 1975			January 1976–December 1980		
	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$	$\gamma_{2i} = 0$	$\gamma_{3i} = 0$	$\gamma_{4i} = 0$
System (1):									
F-statistics	2.531	2.416	2.417	2.768	2.945	3.247	4.509	4.155	3.468
Probability level	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
System (2):									
F-statistics	4.034	3.879	3.455	4.082	5.042	5.128	5.088	4.077	3.790
Probability level	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
System (3):									
F-statistics	5.431	3.817	3.322	4.617	4.705	4.289	5.010	3.097	3.003
Probability level	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
System (4):									
F-statistics	...	...	...	...	...	...	5.004	3.951	4.021
Probability level	...	...	...	...	...	...	.0001	.0001	.0001

NOTE.—In all cases degrees of freedom are 30 and 1,650. Minimum probability level reported is .0001.

perils of attaching too much economic significance to the observed relations between expected returns and dividends yields.[11]

Conclusions

This paper has tested four hypotheses about the impact of taxation on the pricing of financial assets. The tests employ systems of time-series regressions with the competing hypotheses taking the form of cross-equation and within-equation parameter restrictions.

The major finding is that none of the proposed hypotheses is consistent with the data. In almost all cases examined, the hypothesis that dividends and the taxable riskless rate have a common effect on expected returns can be decisively rejected. This conclusion is inconsistent with the no-tax-differential hypothesis; Miller's simple, capital structure equilibrium model; and the two-parameter models of after-tax asset prices of Brennan and Litzenberger and Ramaswamy. Further, the parameter restrictions implied by dividend-induced tax clientele models such as those of Litzenberger-Ramaswamy or Elton-Gruber are easily rejected. The tests do show a statistically significant relation between yields and returns, but one that is not well described by any of the models. The unresolved question is, What does explain the results? The observed structure of dividend payments suggests that dividend yields may be proxying for changes in the expected returns of securities over time. That possibility was tested by including the lagged dividend yield in cum months. The cum-month dividend yields have no impact on investors' tax liabilities and have no informational content since they are lagged values. Nevertheless, the cum-month dividend yields appear to be as important statistically as the ex-month dividend yields. The cum-month results emphasize the need for great caution in interpreting the dividend tests here and more generally. They should especially discourage attempts to appeal to existing empirical tests for justifying either the dividend policies of particular firms or the security selection policies of portfolio managers.

References

Black, F., and Scholes, M. 1973. The behavior of security returns around ex-dividend days. Unpublished manuscript, August.

Black, F., and Scholes, M. 1974. The effects of dividend yield and dividend policy on common stock price and returns. *Journal of Financial Economics* (May), pp. 1–22.

Brennan, M. 1970. Taxes, market valuation and corporate financial policy. *National Tax Journal* 23 (December): 417–27.

Elton, E., and Gruber, M. 1970. Marginal stockholder tax rates and the clientele effect. *Review of Economics and Statistics* 52 (February): 68–74.

11. This sample was also used to test for common cum-month effects and common differences between cum- and ex-month effects. In general these restrictions were rejected at high probability levels.

Gibbons, M. R. 1982. Multivariate texts of financial models: A new approach. *Journal of Financial Economics* 10 (October): 3–27.

Gordon, R., and Bradford, D. 1980. Taxation and the stock market valuation of capital gains and dividends: Theory and empirical results. *Journal of Public Economics* 14 (October): 109–36.

Hess, P. 1980. Dividend yields and stock returns: A test for tax effects. Ph.D. dissertation, University of Chicago.

Judge, G. G.; Griffiths, W. E.; Hill, R. C.; and Lee, T. C. 1980. *The Theory and Practice of Econometrics*. New York: Wiley.

Litzenberger, R., and Ramaswamy, K. 1979. The effects of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics* 7 (June): 163–95.

Litzenberger, R., and Ramaswamy, K. 1980. Dividends, short selling restrictions, tax-induced investor clienteles, and market equilibrium. *Journal of Finance* 35 (May): 469–82.

Litzenberger, R., and Ramaswamy, K. 1982. The effects of dividends on common stock prices tax effects and information effects. *Journal of Finance* 37 (May): 429–44.

Miller, M. 1977. Debt and taxes. *Journal of Finance* 32 (May): 261–76.

Miller, M., and Modigliani, F. 1961. Dividend policy, growth and the valuation of shares. *Journal of Business* 4 (October): 411–33.

Miller, M., and Scholes, M. 1978. Dividends and taxes. *Journal of Financial Economics* 6 (December): 333–64.

Miller, M., and Scholes, M. 1982. Dividends and taxes: Empirical evidence. *Journal of Political Economy* 90 (December): 1118–41.

Mishkin, F. 1982. Does anticipated monetary policy matter? *Journal of Political Economy* 90 (February): 22–51.

Morgan, I. G. 1982. Dividends and capital asset prices. *Journal of Finance* 37 (September): 1071–86.

Schmidt, P. 1976. *Econometrics*. New York: Dekker.

Theil, H. 1971. *Principles of Econometrics*. New York: Wiley.

Zellner, A. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.