NBER WORKING PAPER SERIES

TAXATION AND THE STOCK MARKET VALUATION OF CAPITAL GAINS AND DIVIDENDS: THEORY AND EMPIRICAL RESULTS

Roger H. Gordon

David F. Bradford

Working Paper No. 409

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

November 1979

The research reported here is part of the NBER's research project in the changing role of Debt and Equity Finance in U.S. Capital Formation in the United States, which is being financed by a grant from the American Council of Life Insurance. We would like to thank Stephen Williams for comments and extensive and careful computational assistance. An earlier version of this paper was presented at the NBER/SSRC conference on Econometric Studies in Public Finance, St. John's College, Cambridge, England, June 13-16, 1979. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Journal of Public Economics
Vol 14 No 2
Oct 1980
pp [illegible]-136

NBER Working Paper 409
November 1979

Taxation and the Stock Market Valuation of Capital Gains and Dividends: Theory and Empirical Results

ABSTRACT

Dividends seem to be more heavily taxed than capital gains. Why then do corporations pay dividends rather than repurchasing shares or retaining earnings? Either corporations are not acting in the interests of shareholders, or else shareholders desire dividends sufficiently for nontax reasons to offset the tax effect.

In this paper, we measure the relative valuation of dividends and capital gains in the stock market, using a variant of the capital asset pricing model. We find that dividends are not valued differently systematically from capital gains. This finding is consistent with share price maximization by firms but inconsistent with the fact that most shareholders pay a heavier tax on dividends.

We also show that the relative value of dividends provides an indirect measure of a marginal Tobin's q. The measured value of dividends relative to capital gains tends to be higher during prosperous periods, as is consistent with this interpretation. We hope that this time series on a marginal Tobin's q will prove to be useful in forecasting the rate of investment.

Roger H. Gordon
David F. Bradford
Department of Economics
Princeton University
Princeton, NJ 08544

609/452-4018
609/452-4804

TAXATION AND THE STOCK MARKET VALUATION OF
CAPITAL GAINS AND DIVIDENDS: THEORY AND EMPIRICAL RESULTS

Roger H. Gordon and David F. Bradford

Given the much more favorable tax treatment of capital gains than of dividends, why do U.S. corporations pay dividends when most shareholders ought to prefer stock repurchase? This is a question that has long puzzled those who have tried to use economic theory to predict the effect of taxes on the financial and other decisions of corporations. If investors evaluated securities the way they "ought to," firm managers attempting to maximize shareholder wealth should--because of the tax rules--be led to avoid dividends. Yet U.S. individual income taxpayers reported an aggregate of $ 25 billion dollars in dividends received in 1976, a figure that may be compared with an estimated aggregate corporation profit (after interest payments and corporate taxes) of $63 billion.

If economic theory fails by such a margin to predict the effect of existing taxes on individual and firm behavior, the confidence one can have in predictions derived on the same basis about alternative rules is seriously eroded. The research reported on here suggests that the explanation of the paradox of dividends is not to be found in non-maximizing behavior on the part of those controlling the financial policies of firms. Our conclusion is that the focus of attention, as far as this issue is concerned, is best directed toward explaining the preferences of the shareholding population, and our modelling of this side of the market for corporate securities suggests that here too the outcomes may be less at variance with maximizing behavior than usually thought, even taking into account only tax considerations.

We approach this question by developing and estimating a model of the relative value of dividends and capital gains in the U.S. stock market. The theory of portfolio choice we apply to the demand side of that market is an adaptation of the standard capital asset pricing model, particularly as modified by Brennan [1970]. Importantly if not solely for tax reasons, we expect to find in the taxpayer population a distribution of preferences between returns in the form of dividends and returns in the form of capital gains (increase in asset price). Given their financial policies, corporations' shares will differ in the division of their yields between the two forms, and these differences will be taken into account along with the risk properties of the shares in the portfolio decisions of investors.

As we show in the formal derivation, asset market equilibrium will generate a <u>single</u> rate of exchange between dividends and capital gains. While the capital gain equivalent to a dollar of dividends, which we denote by α, cannot be observed directly, it can be inferred from market data. Estimating α is the principal objective of our empirical work.

According to the theory the market α will be a weighted average of individual shareholder α's, that is, of individual valuations of dividends in terms of capital gains. Thus if, as is usually assumed, dividends are always taxed not less heavily than capital gains, and are usually taxed more heavily, tax considerations alone would imply a value of α less than one.

A finding otherwise should cause us to reassess the tax consequences at the shareholder level and to look for other reasons for valuing dividends.

On the supply side of the market for corporate shares, the theory of the firm implies that in equilibrium the two forms of returns to shareholders should be equally valued, that is, α equal to one. Since we expect the exchange market for shares to adjust very quickly to changing conditions, but firm policies to change slowly, this implies only a tendency for the value of α toward unity.

Our estimates of α indicate a cyclical pattern around one. Such a pattern is consistent with the view that the values of dividends and capital gains tend toward equality. That is, our empirical results do not cause us to question the usefulness of shareholder wealth maximization in predicting outcomes in the U.S. corporate sector. As we note in the next section, accepting this model allows us to draw conclusions from our estimates about the time path of the value in the stock market of an incremental dollar of real investment. In particular if corporate dividend policies are believed to adjust sufficiently rapidly to changing circumstances, Bradford-Gordon α can be taken as an estimate of a marginal Tobin's q. Potential uses of this are discussed in the concluding section of the paper.

Since the empirical results are consistent with the theory of the firm they are inconsistent with the view that dividends and capital gains are valued purely for their net of tax cash flow consequences, assuming dividends are relatively heavily taxed. Perhaps what is involved is an irrational preference for dividends, as has been suggested by Black [1976]. Other rational grounds seem worth investigating, such as the possibility, explored by Battacharya [1979], that dividends serve a signalling function concerning the future profitability of the firm. We do not pursue this investigation here, although we do note in the next section a frequently neglected class of shareholders for which dividends are less heavily taxed than capital gains.

The next section, Section I, presents certain institutional and theoretical background relevant to the interpretation of the empirical analysis. Section II describes the model of equilibrium share valuation, the estimation of which is the subject of Section III. Section IV contains the results of estimation and the paper concludes with a brief summing up and commentary in Section V.

I. Institutional and Theoretical Background

In this section we first briefly review the concepts of dividends and capital gains, and the tax treatment thereof, as these bear on the evaluation by demanders of common stock.[1] We then summarize the implications of the theory of the firm for the behavior of suppliers of common stock.

Dividends, Capital Gains, and Their Tax Consequences

The term "capital gain" as used in U.S. tax law is related to a transaction of sale or exchange of a "capital asset," as defined in the law. We use the term to refer simply to the increase in market value of an asset over a specified time period. It is thus unrelated to transactions. The equivalence between capital gains and dividends (which are cash distributions from a corporation to its shareholders)[2] is, however, dependent on transactions. The crucial point is that, in the absence of taxes and transactions costs, it is possible to produce precisely the same consequences by dividends and by share repurchase by the corporation. Both operations can be used to transfer funds out of the corporation. By entering the market to sell, shareholders can obtain the same cash flow in the repurchase case as would be provided by dividends. By entering the market to buy, recipients of dividends who prefer an increased ownership interest can reproduce the effect of declining to sell shares to the corporation in a disbursement of corporate funds by share repurchase. The ownership claim remaining in the shareholder after either transaction is the same. It is important to understand the equivalence of the two, since it explains the economist's conviction that the value of the corporation should be the same after disbursement of a given amount of funds by either form.[3]

Because there are transactions costs and because the tax consequences are very different, the expectation that future disbursements of corporate funds will take one or the other form should have a bearing on its current value.

The relevant tax rules provide in the case of individual (rather than corporate) shareholders (a) the first $100 ($200 for married couples filing jointly) of dividends are excluded from income tax; (b) further dividends are taxed as "ordinary" income (like interest receipts); (c) accruing capital gains induce no current tax liability, (d) the taxation of capital gains "realized" by sale depends upon the period over which the asset has been held, with short term gains taxed as ordinary income, but with only 40 percent of long term gains (asset held a year or more) subjected to tax as ordinary income.[4] These features imply that taxpayers with high marginal rates of tax on ordinary income and little use for current cash flow should strongly prefer accruing capital gains to dividend yield.[5] For taxpayers with zero or low marginal tax rates, transactions costs might be expected to play a more important role, with those wishing a steady flow of cash favoring dividends. In general one might expect the preferences of wealthy, highly taxed, individual shareholders for capital gains to be the dominant influence on individual valuation of corporate shares.

For shareholders other than individuals, an attitude ranging from neutrality between capital gains and dividends to a preference for dividends is implied by tax considerations.[6] Obviously, for tax exempt shareholders, such as pension funds, tax consequences are irrelevant, and matters of transactions costs and institutional features such as the rules limiting a university's cash draw on its endowment to "income" (often defined to exclude capital gains) may be the principal determinant of preferences. For shareholders that are themselves taxable corporations tax considerations are in favor of dividends. This is because 85 percent (100 percent in the case of a sufficient ownership interest) of dividends received are excluded from corporation income tax. This implies a tax rate in the typical case of .15 x .46 or 6.9 percent. Realized capital gains, on the other hand, are taxed at a flat 28 percent.[7] While the advantages of deferral of tax liability and the flexibility to choose the timing of cash flow, associated with capital gains, apply to corporate as well as individual shareholders, still taxable corporation shareholders should be expected typically to prefer dividends. Thus implications of tax rules for the preferences of potential shareholders between capital gains and dividends are not as unambiguously in favor of capital gains as is commonly believed. It appears possible that investors favoring dividends could be sufficiently influential to induce a temporary equilibrium value of α in excess of one.[8]

Implications of the Theory of the Firm for Dividend Policy

We turn next to the question of how the decisions of an individual firm will be related to its perception of the value of α implicit in the stock market. Specifically we consider three margins of choice: (1) that between retentions and dividends, (2) that between real investment and other uses of corporate funds, and (3) that between issuing (or retiring) debt or equity. In addressing these issues we assume the objective of the firm is to maximize the value of its shares.[9]

It follows that dividend policy will be set to bring about equality between α and the increase in per share market value consequent upon an extra dollar of retentions. If the market valued a flow of retentions at less than α per dollar, the value of the shares could be increased by converting the flow of retentions to a flow of dividends, while dividends would be reduced if the market valued retentions at more than α.

If the firm makes use of retained funds to maximize shareholder wealth, the stock market valuation of an incremental dollar of real investment must also be α (as long as real investment is among the best uses of retentions). Therefore, on the assumption that firms are setting dividend policy optimally, α represents a marginal Tobin's q (the market value of an extra unit of capital relative to cost).[10] Because it is a marginal rather than an average value (market value of the firm relative to replacement cost) as usually calculated, it ought to be more useful in analyzing firm investment behavior and financial structure.[11]

We have concluded so far that when the dividend payout rate is optimal α equals the value of a dollar of retentions, which also normally equals the value of a dollar of real investment. However the firm cannot be in long term equilibrium if α is different from one. A value of α greater than one implies the opportunity for profitable arbitrage between equity issue and the uses of new funds to retire debt (or purchase the debt of other firms) or to undertake real investment. If α is less than one the advantage shifts to debt finance, with the funds used to retire equity claims.[12]

II. Taxes, portfolio choice, and the relative valuation of dividends and capital gains.

As has been suggested in the preceding section, the relative valuation of dividends and capital gains in the market will be the net consequence of the portfolio choices of many separate wealth-holders, here collectively labeled "households." To study this we modify the standard analysis of portfolio choice to incorporate preferences between the two forms of returns. Thus, while we assume as usual that the household seeks to allocate its wealth among the available securities in the market to maximize a function $f(\mu, \sigma^2)$ of the mean μ and the variance σ^2 of the one-period real return on the portfolio, we take explicit cognizance of the fact that the return in question is a weighted sum of dividends and capital gains.

Individual Portfolio Optimality Conditions

Let d_{it} denote the dividends received on a dollar's worth of the i'th asset during period t, and r_{it} the increase in its market value over the period. The real after tax dollar value of the return on a share of equity in firm i to household h in period t is a weighted sum of r_{it} and d_{it}, less the rate of inflation, π_t, over the period:

$$(1) \quad R^h_{it} = \alpha^h_1 r_{it} + \alpha^h_2 d_{it} - \pi_t ,$$

where the α's capture tax and possibly other preference and signalling elements. Letting f^h represent this household's utility function of portfolio mean and variance and for the moment suppressing the time subscript, we can describe the household's problem as

(2) Choose x_i^h

to maximize $f^h(\mu_h, \sigma_h^2)$

subject to $\sum_i x_i^h = W^h$

$$\mu_h = \sum_i x_i^h \overline{R}_i^h$$

$$\sigma_h^2 = \sum_i \sum_j x_i^h x_j^h \operatorname{Cov}(R_i^h , R_j^h) ,$$

where a bar over a random variable denotes its subjective expected value. W^h denotes the household's wealth to be allocated, and the asset subscripts run over the set of securities available in the market. Note that we have not constrained x_i^h to be positive, implying the possibility of short sales.

For the next steps we shall focus on the single household, so where no ambiguity results we simplify notation by suppressing the identifying household index, h . First order conditions associated with a solution to problem (2) may then be written

$$(3) \quad f_1 \cdot \overline{R}_i + 2f_2 \cdot \sum_j x_j \operatorname{Cov}(R_i , R_j) = \lambda . \quad i=1,\ldots,n$$

where λ is the Lagrangian multiplier on the wealth constraint. These conditions imply in particular that for an asset z having zero covariance with the portfolio

$$(4) \quad f_1 \cdot \overline{R}_z = \lambda ,$$

and hence we may re-write (3) as

$$(5) \quad \overline{R}_i - \overline{R}_z = - \frac{2f_2}{f_1} \sum_j x_j \operatorname{Cov}(R_i , R_j) .$$

For the special case in which the asset i is the portfolio itself, indexed h , (5) implies

$$(6) \quad \bar{R}_h - \bar{R}_z = -\frac{2f_2}{f_1} \mathrm{Var}\,(R_h)\ .$$

Now (5) can be written

$$(7) \quad \bar{R}_i - \bar{R}_z = \frac{\bar{R}_h - \bar{R}_z}{\mathrm{Var}(R_h)} \sum_j x_j \,\mathrm{Cov}\,(R_i, R_j)$$

$$= \frac{\bar{R}_h - \bar{R}_z}{\mathrm{Var}(R_h)} \mathrm{Cov}\,(R_i , R_h)\ .$$

This together with (1) implies

$$(8) \quad \bar{r}_i + \alpha^h \bar{d}_i - \bar{g}_z^h = \beta_i^h \,(\bar{r}_h + \alpha^h \bar{d}_h - \bar{g}_z^h)$$

where $\alpha^h = \alpha_2^h / \alpha_1^h$

$$\beta_i^h = \frac{\mathrm{Cov}(R_i^h, R_h)}{\mathrm{Var}(R_h)} = \frac{\mathrm{Cov}(r_i + \alpha^h d_i - \pi/\alpha_1^h\ ,\ r_h + \alpha^h d_h - \pi/\alpha_1^h)}{\mathrm{Var}(r_h + \alpha^h d_h - \pi/\alpha_1^h)}$$

$$\bar{g}_z^h = \bar{r}_z + \alpha^h \bar{d}_z\ .$$

To remind us of potential aggregation problems, household indices have been included in equation (8), which can be thought of as a statement about the risk premium on different assets expressed in capital gains equivalents. For example, the relative weight, α^h , placed on dividends depends upon household preferences and circumstances. The optimizing portfolio, indexed h , is dependent upon the household, as is, therefore, the zero-covariance asset. In fact, even the expectations concerning the means and covariances of the returns might vary by household. Since we have no hope of gaining this detailed information for each household, we must make some assumptions about similarities of households so that we can derive an aggregate relationship among asset characteristics.

Market Equilibrium Derived from Individual Optimization

Black (1972) shows that equation (8) can be aggregated, yielding an equation relating the returns on each security with those of the market portfolio, under the following assumptions: 1) the relative weight, α^h, on dividends is the same across households, and 2) expectations concerning the means and covariances of returns are identical and rational, implying that the actual values of each correspond with expectations. To assume that all households share a common value of α^h, however, is to suppress an important aspect of the reality we are examining--the differences, especially due to different tax situtions, in evaluation of dividends and capital gains. Fortunately, we can relax this assumption. Equation (8) may be rewritten as

$$(9)\quad (\bar{r}_i + \alpha^h \bar{d}_i - \bar{g}_z^h)\ \Omega^h = \mathrm{Cov}\ (r_i + \alpha^h d_i - \pi/\alpha_1^h,\ r_h + \alpha^h d_h - \pi/\alpha_1^h)$$

where $$\Omega^h = \frac{\mathrm{Var}(r_h + \alpha^h d_h - \pi/\alpha_1^h)}{\bar{r}_h + \alpha^h d_h - \bar{g}_z^h} .$$

If dividends and inflation are both non-stochastic the right hand side of (9) becomes simply $\mathrm{Cov}\ (r_i, r_h)$. Consider the weighted sum of these covariances, where the weights are S^h, defined to be the ratio of the household's portfolio value, W^h, to the value of the market portfolio, W. Now use

$$(10)\quad \sum_h S^h r_h = r_m ,$$

to conclude

$$(11)\quad \sum_h S^h\ \mathrm{Cov}\ (r_i, r_h) = \mathrm{Cov}\ (r_i, r_m) .$$

Weighting each side of equation (9) by S_h and aggregating, using equation (11), then implies

$$(12) \quad \bar{r}_i + \alpha \bar{d}_i - \bar{g}_z = \gamma \, \mathrm{Cov}\,(r_i \, , r_m)$$

where
$$\gamma = \frac{1}{\sum_h S^h \Omega^h}$$

$$\alpha = \frac{\sum_h S^h \Omega^h \alpha^h}{\sum_h S^h \Omega^h}$$

$$\bar{g}_z = \frac{\sum_h S^h \bar{g}_z^h \Omega^h}{\sum_h S^h \Omega^h} \quad .$$

For the special case of the i'th asset being itself the market portfolio (12) implies

$$(13) \quad \bar{r}_m + \alpha \bar{d}_m - \bar{g}_z = \gamma \, \mathrm{Var}\,(r_m) \; .$$

This allows us to replace γ in (12) to obtain a relationship between expectations of returns on individual securities and expectations of returns on the market portfolio:

$$(14) \quad \bar{r}_i + \alpha \bar{d}_i - \bar{g}_z = \frac{\mathrm{Cov}(r_i \, , \, r_m)}{\mathrm{Var}(r_m)} \, (\bar{r}_m + \alpha \bar{d}_m - \bar{g}_z) \quad .$$

Even though equation (14) no longer depends on h, it still presents an estimation problem in that expectations concerning rates of return are not directly observable, only the actual outcomes. However, the definition of a covariance implies that were we to sample repeatedly from the subjective joint probability distribution for r_i and r_m, the resulting observations would satisfy

$$(15)\quad r_i - \bar{r}_i = \frac{\mathrm{Cov}(r_i, r_m)}{\mathrm{Var}(r_m)} (r_m - \bar{r}_m) + \epsilon_i \ ,$$

where ϵ_i is a random variable with zero expected value and uncorrelated with $r_m - \bar{r}_m$. With rational expectations, the actual outcomes will also satisfy (15). When the returns are joint normally distributed, as has implicitly been assumed in motivating (2), ϵ_i will be normally distributed.

Now use equation (14) to replace $\bar{r}_i$ in (15) by properties of the portfolio and zero-covariance asset:[13]

$$(16)\quad r_i + \alpha \bar{d}_i - \bar{g}_z = \beta_i (r_m + \alpha \bar{d}_m - \bar{g}_z) + \epsilon_i$$

$$\text{where} \quad \beta_i = \frac{\mathrm{Cov}(r_i, r_m)}{\mathrm{Var}(r_m)} \ .$$

The unobserved parameters in this equation are α, β_i, and $\bar{g}_z$. The entire derivation was for a given date. At other dates, the equation will have a similar form, but the parameters may all be different.

We then have (17) as a specification:

$$(17)\quad r_{it} + \alpha_t \bar{d}_{it} - \bar{g}_{zt} = \beta_{it}(r_{mt} + \alpha_t \bar{d}_{mt} - \bar{g}_{zt}) + \epsilon_{it} \ .$$

This is the first specification to be estimated in section IV.

Note that in this specification, α does not vary with i. According to the derivation, all firms will face the same relative value in the market for their dividends and capital gains, in spite of the fact that their shareholders may have very different characteristics. This result is contrary to that in Elton and Gruber (1970).

In addition, it follows from equation (6) that $\Omega^h = -f_1/f_2$. This implies that the weight given to each α^h in deriving α is not only proportional to the size of that investor's portfolio but only inversely proportional to the marginal degree of risk aversion of that investor. In particular, were any investor risk neutral, his α^h would receive infinite weight.

Refinements of the Estimating Equation

As has been noted in the derivation, specification (16) is implied when both dividends and inflation are nonstochastic. As we shall employ monthly observations in the estimations, the assumption that dividends are known for the period ahead does not seem extreme, but it is less plausible that the inflation rate can be forecast with confidence. If only dividends are nonstochastic, then the aggregation of individual equilibrium conditions produces a somewhat less neat specification. The right hand side of (9) now becomes $\mathrm{Cov}(r_i - \pi/\alpha_1^h,\ r_h - \pi/\alpha_1^h)$.

By a derivation analogous to that leading to equation (12) we now find

$$(18)\quad \bar{r}_i + \alpha\bar{d}_i - \bar{g}_z = \gamma\left[\mathrm{Cov}(r_i, r_m) + \delta_i\right],$$

where $\delta_i = -\sum_h S^h\left[\mathrm{Cov}(r_i, \pi/\alpha_1^h) + \mathrm{Cov}(\pi/\alpha_1^h, r_h) - \mathrm{Var}(\pi/\alpha_1^h)\right]$.

As in (13), use the special case of i=m to eliminate the parameter γ, giving

$$(19)\quad \bar{r}_i + \alpha\bar{d}_i - \bar{g}_z = \frac{(\mathrm{Cov}(r_i, r_m) + \delta_i)}{\mathrm{Var}(r_m) + \delta_m}(\bar{r}_m + \alpha\bar{d}_m - \bar{g}_z).$$

As in (15), use the definition of a covariance and (19) to produce

$$(20)\quad r_i + \alpha\bar{d}_i - \bar{g}_z = \beta_i\left(r_m + (\alpha\bar{d}_m - \bar{g}_z)\frac{\mathrm{Var}(r_m)}{\mathrm{Var}(r_m) + \delta_m} - \bar{r}_m\frac{\delta_m}{\mathrm{Var}(r_m) + \delta_m}\right)$$

$$+\ \delta_i\frac{(\bar{r}_m + \alpha\bar{d}_m - \bar{g}_z)}{\mathrm{Var}(r_m) + \delta_m} + \epsilon_i.$$

$$= \beta_i(r_m + \alpha^*\bar{d}_m - \bar{g}_z^*) + \delta_i^* + \epsilon_i.$$

where α^* can differ arbitrarily from α . This specification, with α, α^*, β_i, and δ_i^* all varying with time, is estimated in section IV. Note that when δ_i^* varies with time, neither $\bar{g}_{zt}$ nor $\bar{g}^*_{zt}$ can be estimated--only their combined effect with δ^*_{it} is identified.

If δ_m is small relative to $\mathrm{Var}(r_m)$, as seems plausible, then the only major difference between (17) and (20) is that in (20) there is a firm specific intercept, which might vary over time. For $\delta_m = 0$ this yields (21), which is also estimated in section IV.

$$(21)\quad r_{it} + \alpha_t \bar{d}_{it} - \bar{g}_{zt} = \beta_{it}(r_{mt} + \alpha_t \bar{d}_{mt} - \bar{g}_{zt}) + \delta_{it} + \epsilon_{it}$$

Again, $\bar{g}_{zt}$ is not separately identified when δ_{it} is also estimated, so is arbitrarily set to zero.

If we assume α_1^h is equal to one for all households, not a bad assumption except for corporate holders, then the right hand side of (18) becomes $\gamma \mathrm{Cov}(r_i - \pi, r_m - \pi)$. Using the same procedure as before, this implies

$$(22)\quad r_i - \pi + \alpha \bar{d}_i - (\bar{g}_z - \bar{\pi}) = \beta_i(r_m - \pi + \alpha \bar{d}_m - (\bar{g}_z - \bar{\pi})) + \epsilon_i \ .$$

This differs from (18) in that the capital gains are measured in real terms, and the zero beta rate, as estimated, is a real return rather than a nominal return. Equation (22) is also estimated in section IV.

Returning again to equation (9), what if dividends are also stochastic? We now find

$$(23)\quad \bar{r}_i + \alpha \bar{d}_i - \bar{g}_z = \gamma[\mathrm{Cov}(r_i, r_m) + \delta_i^*]$$

where $\delta_i^* = \sum_h S^h [\mathrm{Cov}(r_i, \alpha^h d_h - \pi/\alpha_1^h) + \mathrm{Cov}(\alpha^h d_i - \pi/\alpha_1^h, r_h) + \mathrm{Cov}(\alpha^h d_i - \pi/\alpha_1^h, \alpha^h d_h - \pi/\alpha_1^h)]$

Since equations (23) and (18) have the same structure, equation (23) would also imply the estimation specifications (20) and (21), with δ_i and δ_m suitably redefined.

If we again assume α_1^h equals one for all households, we now find using the same derivation

$$(24)\quad r_i - \pi + \alpha\bar{d}_i - (\bar{g}_z - \bar{\pi}) = \beta_i\left(r_m - \pi + (\alpha\bar{d}_m - \bar{g}_z)\frac{Var(r_m-\pi)}{Var(r_m-\pi)+\delta'_m} + \bar{\pi} - \bar{r}_m \frac{\delta'_m}{Var(r_m-\pi)+\delta'_m}\right)$$

$$+ \delta'_i \frac{\bar{r}_m + \bar{d}_m - \bar{g}_z}{Var(r_m-\pi) + \delta'_m} + \epsilon_i$$

where $\delta'_i = S^h[Cov(r_i-\pi, \alpha^h d_h) + Cov(\alpha^h d_i, r_h-\pi) + Cov(\alpha^h d_i, \alpha^h d_h)]$.

It seems very plausible that δ'_i is small relative to $Var(r_m-\pi)$, so the only real difference between equations (22) and (24) is the addition of a firm specific constant term. This equation with $\delta'_m=0$ is also estimated in section IV.

The derivation above made use of two strong assumptions: 1) there is no restriction on short sales and 2) expectations of households are identical. How sensitive is the specification to these assumptions?

Assume short sales are not possible, as an extreme alternative. It is not even clear that the constraint will be binding for many investors. When a risk-free asset exists and all households have the same α^h, they all will hold a proportional share of the market portfolio. A household's optimal portfolio will be a continuous function of α^h, so we expect that α^h will have to vary substantially from the "average" value if the household's desired holding of any security is to be negative. If the α^h do vary that substantially, it can then be shown that portfolio equilibrium for household h requires

$$(25) \quad \bar{r}_i + \alpha^h \bar{d}_i - \bar{g}_z^h \leq \beta_i^h (\bar{r}_h + \alpha^h \bar{d}_h - \bar{g}_z^h) \qquad i=1,\ldots,n$$

with an inequality only when the short sales constraint is binding. Let us rewrite (25) as (26), where $e_i^h \geq 0$.

$$(26) \quad e_i^h + \bar{r}_i + \alpha^h \bar{d}_i - \bar{g}_z^h = \beta_i^h (\bar{r}_h + \alpha^h \bar{d}_h - \bar{g}_z^h) \quad .$$

Were we to use equation (26) rather than equation (9) in the previous derivation, the only difference would be that there would be an additional term $e_i = \dfrac{\Sigma S^h \Omega^h e_i^h}{\Sigma S^h \Omega^h}$ on the left hand side. However, this term is just a firm specific intercept, which already exists in many of the previous specifications. We merely need to reinterpret it when a short sales constraint exists.

What if individuals have different expectations about $\bar{r}_i$, but not about $\mathrm{Cov}(r_i + \alpha^h d_i - \pi/\alpha_1^h , r_h + \alpha^h d_h - \pi/\alpha_1^h)$? Williams (1977) argues for the plausibility of this set of assumptions. In rederiving equation (12), $\bar{r}_i$ would be replaced by $\dfrac{\sum_h S^h \Omega^h \bar{r}_i^h}{\sum_h S^h \Omega^h}$. In order to be able to estimate the rederived equation, we would want to assume that this expression is consistent with actual outcomes. While previously we assumed that each $\bar{r}_i^h$ was consistent with actual outcomes, all we need is that this weighted average is consistent, a much weaker rationality assumption than before.

III. Estimation procedure.

In the previous section, we derived a series of specifications, for which equation (17) is representative. In this section, we first discuss further specific assumptions which must be made before the equation can be estimated. We then describe the data and the estimation procedure.

Further specification assumptions

In equation (17), the unobserved parameters are α_t, $\bar{g}_{zt}$ and β_{it}. Assumptions dealing with each of these will be described in turn.

$\underline{\alpha_{it}}$ Our derivation implies that α should not vary across assets but could vary across time, as the tax law and the wealth distribution change, as well as the size of any transactions costs and the importance of institutional constraints favoring dividends. To capture this smooth evolution of α over time, we normally assume that α_t is piecewise linear in t with break points every five years. For purposes of comparison with related work by Black and Scholes (1974) and Litzenberger and Ramaswamy (1979) (hereafter B-S and L-R), we also estimate a specification in which α is constant over time (except that for comparability with L-R we constrain α to equal one prior to the "normal" taxation of dividends in 1936).

$\underline{\bar{g}_{zt}}$ In L-R, $\bar{g}_{zt}$ is implicitly set equal to $a + \alpha r_{ft}$, where r_{ft} is an observed interest rate on short term high grade bonds. But because of inflation risk or default risk (when the series is for nongovernment bonds), the real yield on these assets is not variance free. Presumably the stochastic movement in these assets is positively correlated with that of the market portfolio, as both respond inversely to inflationary shocks. Therefore the zero beta rate ought to be less than αr_{ft}, and to a larger degree when the inflation rate is high on the assumption that the variance in the inflation rate is high when the inflation rate is high. Therefore, we set $\bar{g}_z = a + br_f$ where we expect $b < \alpha$. Since the relation between $\bar{g}_z$ and r_f may change over time, we also let "a" be a piecewise linear function of time.

$\underline{\beta_{it}}$ We assume that the β for a firm may drift smoothly over time, so treat the β for each firm as a piecewise linear function of time, with break points every five years. The coefficients of this linear function are estimated simultaneously with the others. In contrast, in B-S and L-R, as has been standard in papers estimating a capital asset pricing model, β is estimated from the previous five years or so of data, using the regression: $r_{it} + d_{it} - r_{ft} = \beta_i(r_m + d_{mt} - r_{ft}) + c + \epsilon_{it}$. Then $\hat{\beta}_i$ is used as an independent variable in the final regression for the other coefficients. There are at least three problems with this approach.

First, the specification of the regression for β_i assumes $\alpha = 1$ and $\bar{g}_{zt} = r_{ft}$. Neither is a maintained assumption in equation (17). Since the purpose of the paper is to estimate α, assuming it equal to one at an earlier step creates an internal inconsistency in the model. Simultaneous estimation of the parameters eliminates this inconsistency.

Second, the estimate of β_i in the previous work refers to the average β during the previous five years. Since β_i drifts over time, and people at the time likely observed this drift through daily observation or through specific knowledge about changes in the characteristics of the firm, the estimated β_i will have measurement error beyond that appearing in the standard error of β. L-R attempt to correct for the latter measurement error only. Conventional procedures for estimating capital asset pricing models have not worried about the first. This measurement error will cause bias in all coefficients.

In this paper we estimate β simultaneously with the other coefficients, so avoid problems with bias due to measurement error. Also, we allow for the drift in the value of β over time in a piecewise linear fashion. This procedure implies that the estimate of β for a year depends on subsequent as well as earlier data. We assume that individuals at the time knew much more about the firm than we can infer from monthly price data, so are not bothered by this implication. In effect, we assume rational expectations--that individuals know the parameters of the system, though not the stochastic element.

Third, the requirement that β be estimated on prior data results in the loss of five years of data from the sample, with clear efficiency costs.

$\underline{\epsilon_{it}}$ L-R assumed that $\text{var}(\epsilon_{it}) = \sigma_i^2$ and $\text{cov}(\epsilon_{it}, \epsilon_{j\tau}) = 0$, $i \neq j$ or $t \neq \tau$, in estimating the coefficients and in constructing the standard errors. We maintain L-R's stochastic assumption.[14]

$\underline{\delta_{it}, \alpha_t^*}$ When other coefficients are added, such as δ_{it} and α_t^*, each is assumed to be a piecewise linear function of time.

Data construction

Most of the data for this project come from the monthly returns file compiled at the Center for Research in Securities Prices (CRSP) at the University of Chicago. This data set provides monthly rates of price change and dividend yields on all securities traded on the New York Stock Exchange between 1926 and 1978.

beginning of the current period price, the rate of capital gains on that stock in the past year, r_{mt} , d_{mt} , r_{ft} , and a constant. The forecasts from the regressions are then used instead of d_{it} .

In L-R, when the amount of the dividend was not announced prior to the beginning of the period, the previous dividend payment was used as a proxy if the dividend was a recurring dividend, else zero was used as a proxy. If dividends tend to be increasing in dollar terms, then both of these proxies tend systematically to underestimate actual dividends. Such systematic errors will create biases in the coefficients. Also, L-R assume that the investor always knows in which month a periodic dividend will be paid, an assumption not imposed here.

$\underline{r_{ft}}$ A monthly time series for a high grade interest rate was kindly provided to us by Krishna Ramaswamy through 1977, so is the same series used in L-R. It consists of the interest rate on commercial paper prior to 1951 and the rate on Treasury bills with one month to maturity since then. For 1978 we use the Treasury bill rate as reported in the Federal Reserve Bulletin.

$\underline{r_{mt}, \bar{d}_{mt}}$ The rate of return on the market portfolio ought to be the value weighted average return on all assets, not just those traded on the New York Stock Exchange. However, as done in most all previous papers in the area, we use the value-weighted average capital gains rate and dividend yield for just NYSE securities in most of the specifications.[16] However, in one specification we try to improve on this. If we had the average rate of return on each _type_ of security, denominated in the capital gains equivalent, then the market return would be the value weighted average of these, or $R_m = \sum_{j=1} \Theta_j R_j$.

In addition to data on equity returns, we employed monthly time series on two other types of assets: corporate bond yields, r_{bt}, from Ibbotson and Sinquefield [1977],[17] and, as a proxy for the rate of return on real estate, As in L-R, but in contrast to many other recent papers, we use the firm as unit of observation rather than specially constructed portfolios of securities. The justification in B-S for using portfolios rather than individual securities is apparently 1) to minimize measurement error in β_i, and 2) to minimize correlation in the residuals across observations at any date. In our context, the first reason is moot, as β_i is a parameter, not a datum. Even given the second point, the coefficients ought to be estimated more efficiently by OLS on the individual data than by OLS on the portfolio data.[15]

Given the decision to use firm data, we next consider the problem of measuring each of the needed variables, r_{it}, π_t, $\bar{d}_{it}$, r_{ft}, r_{mt}, and $\bar{d}_{mt}$ appropriately.

$\underline{r_{it}}$ This series is directly available for all New York Stock Exchange securities on the CRSP monthly returns file.

$\underline{\pi_t}$ We used the inflation rate derived from the consumer price index, as reported in Ibbotson and Sinquefield [1977].

$\underline{\bar{d}_{it}}$ The dividend yield is also available monthly on the CRSP tape. When dividends are stochastic, however, we need expected, not actual, dividends. We therefore create an instrument for d_{it}, and use it even in the specificatinns where d_{it} is assumed to be nonstochastic, to minimize possible specification error.

In creating the instrument, we divide the sample into four subsamples. The classification of each observation depends on whether in the past year the firm had paid no dividends, one dividend, two dividends, or more than two dividends. Within each subsample, we regress actual dividend yield on the recent dollar dividends (corrected for stock splits) each divided by the

the inflation rate. The weighted average of these returns (with weights to be estimated) is substituted for the market return in equation (20) to obtain the additional specification[18]

$$(27)\quad r_{it} + \alpha_t \bar{d}_{it} - \bar{g}_{zt} = \beta_{it}(\theta_1(r_{mt} + \alpha_t d_{mt}) + \theta_2 r_{bt} + \theta_3 \pi_t - \bar{g}_{zt}) + \delta_{it} + \epsilon_{it} .$$

There is no unique optimal parameter set as the equation stands. However, if we make the arbitrary restrictions $\theta_1 = 1$ and $\bar{g}_{zt} = 0$, then there is a unique optimal set.

Estimation technique

For ease of exposition, we first discuss estimation of the simplest specification:

$$r_{it} = \beta_i(r_{mt} + \alpha d_{mt} - a - br_{ft}) + a + br_{ft} - \alpha d_{it} + \epsilon_{it} , \; E(\epsilon\epsilon') = \sigma^2 I .$$

This specification is nonlinear in the parameters β_i, α, a, and b. We use nonlinear least squares (equivalent to maximum likelihood under the assumption that the errors are distributed normally). This presents the immediate problem that a large number of parameters are involved. The procedure we follow is to estimate 1) β_i conditional on initial estimates of α, a, and b ($\alpha = b = 1$, $a=0$), using separate least squares regressions for each firm; 2) α, a, and b conditional on these β_i using a pooled regression, and then return to step 1) with the new estimates of α, a, and b and continue to iterate between steps 1) and 2) until convergence.

That the point of convergence minimizes the sum of squared residuals is immediate. In the first order conditions for this minimization, the function is minimized with respect to each parameter conditional on the other parameters. In the procedure described above, the set of parameters is subdivided into two groups, and the function is minimized with respect to each group conditional on the values of the other group. When the function is being jointly minimized, the entire set of first order conditions must be satisfied. It should be noted that the procedure used in L-R and B-S involves stopping this iterative procedure after one iteration rather than iterating until convergence.

One complication is that the standard errors of the parameters are not the standard errors reported by either regression. For the simplest model, where α did not depend on time, we report the maximum likelihood standard errors, as approximated by the square root of the diagonal elements in the inverse of the second derivative matrix of the log likelihood function with respect to the parameters. Since these estimates, as reported below, differed only slightly from the standard errors for the second regression, we report only the latter in the rest of our results.

Allowing each of the parameters to be piecewise linear in time presents no further complications. When δ_{it} and/or θ_j are estimated, as in equation (27), the only modification is that both β_{it} and δ_{it} are estimated in the first regression, while α_t and θ_j are estimated in the second regression. Assuming heteroskedasticity in the residuals across firms causes a few modifications. Since the residuals in each first stage regression for β_i remain homoskedastic and independent across regressions, this step is unchanged. The variance of the residual for each firm is estimated from the residuals in each of these regressions. The estimates of the variances are then used in constructing weighted least squares estimates for the second stage regression. The coefficients, and now variance estimates,are iterated until convergence.

IV. Coefficient estimates

The first step in the estimation process was to construct a forecasted value for dividends. As described before, the overall sample was divided into four subsamples, based on the number of dividend payments during the previous year. The resulting estimates for the four subsamples are reported in Table 1. The coefficients are all reasonably plausible.

These equations were then used to create forecasts for the dividend yield. As a result, the first year of data was dropped from later samples, as it was needed to construct fitted dividends. After dropping in addition observations with missing data and firms with too few observations to allow estimation of β_{it}, 614,150 observations were left. Initial estimation was done on this sample.

The first specification estimated corresponds to equation (17) with $\alpha_t = \alpha$ and $\overline{g}_{zt} = a + b\, r_{ft}$. This specification is similar to that estimated in L-R, though they assumed $b = \alpha$ and performed only one iteration, as described above. We estimate α to be .82, which is statistically significantly less than one, and corresponds to the estimate in L-R of .76. However, b is estimated to be less than zero and substantially less than α. This puts into question the use of r_f as an approximation to the risk-free rate. The problem, however, could well be the importance of firm specific intercepts, as rationalized above, which are captured here in the estimates of a and b.

The standard error estimates are found in columns 2 and 3. Here we find that the OLS standard errors, while they are systematic underestimates because they ignore that β is being estimated simultaneously, still are reasonably close to the maximum likelihood estimates.

In all other specifications estimated, we allowed α to be a piecewise linear function of time. The estimates for α at the breakpoints for all the other specifications are reported in Table 3. The implied value of α for any other date is calculated by taking an appropriately weighted average of the estimated values of α at the two nearest dates, e.g. the value in December 1953 in the first specification is .6 (1.28) + .4 (.96). The reported standard errors are those from the second stage OLS regression.

The time pattern of α, though varying somewhat among the different specifications, consistently follows the economic cycles very closely. It is lowest during the great depression through the end of World War II and is almost as low during the great recession of the early 70's. During the boom years of the 20's and the 50's through the mid-60's, it was above one virtually throughout. It is also very high in 1978, suggesting a favorable forecast for the future.

Estimates made assuming α constant throughout the period are very misleading, since the estimates will be very sensitive to the period chosen. B-S estimate α and its variance essentially by taking the mean and variance of annual estimates of α. It is apparent from these results why the standard error of their estimate was so large.

If one accepts the coefficients as they stand, they imply that dividends are not systematically undervalued. In fact they are often overvalued. That α tends to return to one, where dividends and capital gains are equally valued, following shocks is consistent with share price maximizing behavior by the firm. This result also implies, however, that in the weighted average of investor preferences dividends and capital gains are equally valued. Tax considerations alone seem to lead us to expect that dividends would be less valued. Our empirical finding might be explained either by nontax advantages of dividends to individuals (e.g. lower transactions costs or signalling implications), or by a sufficiently high relative weight placed on the preferences of institutional and corporate shareholders who would prefer dividends.

The cyclical pattern of the results is consistent with the marginal Tobin's q interpretation of α, where α represents the value in the market of an additional dollar of real investment in the firm. We find this value to be sharply procyclical as would be expected. Of particular interest in this connection is the very low value of α during the depression, when corporate investment at the margin was apparently almost valueless. Since α can in principle be measured arbitrarily close to the present, this suggests that the estimates could well prove to be a valuable forecaster for the investment rate.

Let us now examine the specifications individually. The estimates of equation (17), the equation most similar to those in previous studies, provides a basis for comparison. The implied estimates for $r_{zt} = a_t + br_{ft}$ for selected dates are reported in Table 4, column 1. As estimates of the risk free rate, they are not very plausible, fluctuating often between extreme values.[19] Our several justifications for including a firm specific intercept may provide a rationalization for this.

When we use real rather than nominal capital gains, corresponding to equation (22), the value of the log likelihood function falls substantially. The derivation of this specification required the additional assumption that $\alpha_1^h = 1$ for all h. We already noted that this assumption is poor for corporate holders of equity. The poorer statistical performance of this specification also suggests that this assumption may not be a good one. The estimates for r_z in Table 4 from this specification are normally lower than the previous ones in absolute value, as they ought to be, representing a real rather than a nominal interest rate. However, they are really no more plausible, again justifying including firm specific constants.

We therefore next estimated equation (21), where time varying firm specific intercepts are estimated simultaneously (though r_f is omitted). Here we allow for the possibility of stochastic inflation and dividends, or short sales constraints. The estimates of α_t, while changing rather little, seem to change most in those periods when our earlier estimates for r_z were least convincing. In those periods, we would expect the factors justifying the firm specific intercepts to be most important.

When we reestimate equation (21) using real rather than nominal capital gains, as reported in column 4 of Table 3, we again find that the log-likelihood falls substantially. The data reject the simplifying assumption $\alpha_1^h = 1$.

In the next specification, equation (27), we try to improve on our approximation to the return on the market portfolio by adding the corporate bond yield and the inflation rate, with estimated weights, to the market yield. The implied changes in our estimates of α, reported in Table 3, column 5, are minimal. The weights on these additional factors, reported in Table 4, are .11 and .17 respectively. Each coefficient ought to represent the market value of that type of security relative to equity multiplied by the "α" weight appropriate for that security, translating the actual returns into the equivalent capital gains returns. Under this interpretation, the coefficient of the corporate bond rate is quite plausible. While currently, the market value of debt is over half the market value of equity, during most of the sample period the value of debt was relatively much smaller. Were the average value for the period .15,[20] the implied value of "α" for bonds would be .73, a plausible value. While we have little evidence on the relative value of real estate, the estimated coefficient for the inflation rate does not seem grossly out of line.

In our final specification, equation (20), we no longer assume that the uncertainty in inflation or dividends is small relative to that on capital gains. Specifically, we estimate separate weights for market dividends and for individual firm dividends. The latter, which correspond to our previous estimates, cluster more tightly around one throughout the sample period though they continue to have the same cyclical time pattern. We argued in section I that these coefficients might be interpreted as estimates of the market value of a dollar of additional real investment. These final estimates of α give a very plausible estimate for the time pattern of this marginal Tobin's q, while our other estimates might seem a bit too erratic. In addition, the log-likelihood improves substantially here compared with that for equation (21).

The estimates for α^*, reported in Table 4, are less congruent with the predictions of the theory. According to our derivation, α^*_t/α_t represents $\mathrm{var}(r_m)/\mathrm{var}(r_m) + \delta_m$, an interpretation difficult to reconcile with the estimates. However, the estimates seem to fluctuate around plausible values, rather than being consistently implausible. Since surely our measure of the market rate of return is subject to error, when we allow α^* to be estimated separately, its coefficient will be biased upwards during periods when the true market yield is higher than the stock market yield and downwards when the true market yield is lower. We might therefore infer the nature of the true market yield relative to the stock market yield from our estimates of α^*.

To explore further this explanation for the fluctuations in α^*, we introduced the corporate bond rate and the inflation rate as components of r_m in equation (21). Unfortunately, the coefficients, not reported, changed little, and did not improve systematically. This suggests that there are other important factors omitted from our measure of the market rate of return.

While we tend to favor this last specification, fortunately our estimates of the general time pattern of α seem quite insensitive to the specification issues considered. Estimates derived from any of the specifications have the same economic implications, a comforting result if we are to use the estimates for forecasting purposes.

V. Conclusions

In this paper we have presented estimates of the relative value in the U.S. stock market of dividends and capital gains. We have concluded that over the sample period (1926 to 1978) the capital gain regarded by the market as equivalent to a dollar of dividends, denoted by α , has followed a cyclical path around one. The pattern of movement has roughly paralleled that of the business cycle.

The estimates are based on a modification of the capital asset pricing model. Successive estimates are presented, relaxing a priori restrictions. We are able to incorporate stochastic dividends and inflation and to allow for deviation between the returns on the stock market and that on the portfolio of all assets. In each case the empirical results are consistent with expectations. In particular, the most refined version has not only by far the highest value of the likelihood function, but also the most plausible path of α .

We interpret the tendency of α to one as consistent with a view of the firm as maximizing share value in making decisions about dividend policy, real investment and financial structure. An implication of this view is that α is an estimate of the value in the market of incremental real investment. The empirical results on the time pattern of α are consistent with this implication.

While the estimated value of α is consistent with maximizing behavior by the firm, it is less clearly consistent with maximizing behavior by investors. The market value of α is shown by the theory to be a weighted average of stockholder α's. The tax treatment of the two forms of returns implies that individuals should have α's less than one, and since the greater part of common stock is held by individuals one would expect the market equilibrium to be below one as long as a significant number of firms continue to pay dividends.

We note, though, that the presence in the market of tax exempt holders such as pension funds, and taxable corporations for which dividends are less heavily taxed that capital gains, clouds this conclusion. As we show, the weight attached to shareholders' α's is inversely related to their risk aversion. Hence it is not out of the question, though certainly not established, that an α of one is consistent with maximization generally, taking into account only tax issues.

However, it is less easy to reconcile the observed variation in α with this view. Our preferred estimate of α ranges from .04 to 1.37, and the range even in the non-depression years is .70 to 1.37. This degree of variation seems difficult to explain on the basis of differences in the distribution of tax circumstances in the investing population.

Naturally, even as public finance economists, we must acknowledge that taxes may not explain everything. The determinants of individual investors' valuation of dividends is but one of the questions remaining open, and we hope to follow up on some of these in further work. We see three related avenues for such work.

The first is a closer look at the relationship between tax rules and α . Over the sample period, the individual income tax evolved from a minor impost affecting few individuals to a mass tax with high rates.

The tax treatment of dividends, capital gains, corporate retained earnings, and retirement savings also varied substantially. These changes ought to have made a mark on α.

Second, the theory developed here implies a response of the dividend and other financial policies of individual firms to changes in α. Roughly speaking, high values of α should call forth a shift from debt to equity and decreases in payout rates, and low values should have the opposite effect. To deal properly with this relationship will require a careful treatment of the underlying reason for a change in α (e.g., change in tax rules vs. change in general outlook for profits), and of the determinants of the lag structure of responses (due, for example, to the difficulty of observing α).

The third idea of further work is the relationship between α and the rate of corporate real investment. Closely related to the investigation of the response of financial structure, this investigation will build on the implication of optimal firm behavior that α should be a measure of the value in the market of a dollar of additional investment, a marginal Tobin's q. Since the data are available almost immediately, α can be estimated very close to the present. This suggests that an estimated α may prove useful in forecasting the investment rate.

Table 1

Forecasting Equations for Dividends

	$\frac{D_{t-3}}{P_{t-1}}$	$\frac{D_{t-6}}{P_{t-1}}$	$\frac{D_{t-12}}{P_{t-1}}$	$\frac{P_{t-1}-P_{t-13}}{P_{t-13}}$	r_{mt}	d_{mt}	r_{ft}	constant
1. No dividends during previous year	--	--	--	.0001 (.00002)	-.0003 (.0002)	.1116 (.0078)	-.0584 (.0099)	.0001 (.00004)
# of observations 115,926								
2. One dividend during previous year			.146 (.0034)	.0012 (.0001)	.0038 (.0014)	.6394 (.0332)	-.2909 (.0516)	.0005 (.0002)
# of observations 26,770								
3. Two dividends during past year		.1324 (.0035)	.0939 (.0028)	.0014 (.0001)	.0005 (.0011)	.5896 (.0254)	-.2154 (.0361)	.0006 (.0002)
# of observations 29,675								
4. Three of more dividends during past year	.4159 (.0013)	.1543 (.0012)	.1553 (.0009)	.0007 (.00002)	-.0002 (.0002)	.2660 (.0034)	.0191 (.0044)	.0000 (.0000)
# of observations 441,375								

Table 2

Maximum Likelihood Estimates for Equation (17)

	Coefficients	Standard Errors OLS	Standard Errors Maximum Likelihood
α	.8238	.0191	.0207
a	.0007	.0003	.0004
b	-.6088	.0908	.1139

Table 3

Estimates of the Time Pattern of α

Date	(17)	(22)	(21)	(24)	(27)	(20)
	Equation Number					
1. Dec., 1925	1.50 (.27)	1.10 (.21)	.93 (.22)	1.09 (.23)	.87 (.22)	1.11 (.23)
2. Dec., 1930	-.11 (.09)	-.10 (.08)	.08 (.08)	.04 (.09)	.11 (.08)	.04 (.08)
3. Dec., 1935	.64 (.10)	.56 (.10)	.42 (.08)	.39 (.08)	.42 (.08)	.34 (.09)
4. Dec., 1940	.10 (.08)	.16 (.08)	.23 (.07)	.24 (.07)	.23 (.07)	.70 (.07)
5. Dec., 1945	.57 (.10)	.55 (.10)	.65 (.09)	.64 (.09)	.66 (.09)	.71 (.09)
6. Dec., 1950	.96 (.07)	.97 (.07)	.91 (.06)	.92 (.06)	.91 (.06)	1.00 (.06)
7. Dec., 1955	1.28 (.09)	1.27 (.09)	1.41 (.07)	1.37 (.06)	1.41 (.07)	1.37 (.07)
8. Dec., 1960	1.37 (.10)	1.30 (.10)	1.23 (.08)	1.26 (.07)	1.23 (.08)	1.20 (.09)
9. Dec., 1965	1.41 (.10)	1.46 (.10)	1.42 (.09)	1.40 (.09)	1.45 (.09)	1.21 (.09)
10. Dec., 1970	.75 (.09)	.75 (.09)	.54 (.08)	.55 (.08)	.56 (.08)	.72 (.08)
11. Dec., 1974	.45 (.06)	.45 (.06)	.46 (.05)	.43 (.05)	.48 (.05)	.88 (.06)
12. Dec., 1978	1.59 (.09)	1.65 (.08)	1.68 (.01)	1.79 (.02)	1.71 (.01)	1.09 (.05)
Log-Likelihood	652304	651704	658821	658363	658900	659500

Notes: Standard errors, taken from the final OLS regression determining α, are reported in parentheses.

Table 4

Estimates of Supplementary Parameters

Date	Equation Number (17) r_z	(22) r_z	(20) α *	(27)	
1. Dec., 1925	.25	.14	-2.44 (.32)	θ_2	.11 (.005)
2. Dec., 1930	-.33	-.28	1.40 (.12)	θ_3	.17 (.006)
3. Dec., 1935	.03	.005	.74 (.10)		
4. Dec., 1940	-.15	-.18	-1.82 (.08)		
5. Dec., 1945	-.02	-.11	.05 (.10)		
6. Dec., 1950	.11	.08	.55 (.07)		
7. Dec., 1955	.07	.07	1.42 (.10)		
8. Dec., 1960	.29	.21	1.38 (.10)		
9. Dec., 1965	-.18	-.16	2.56 (.11)		
10. Dec., 1970	.20	.08	-1.12 (.11)		
11. Dec., 1974	.00	-.08	-2.56 (.08)		
12. Dec., 1978	-.06	-.10	2.04 (.01)		

Notes: The values for r_{2t} equal $a_t + br_{ft}$, where a_t and b are both estimated.

FOOTNOTES

[1]The tax rules described below are those currently in effect. They are representative in a qualitative sense of the rules that were in effect over the recent decades of the sample period of the empirical work, although there have been important changes in the treatment of capital gains in the individual income tax. The important fact about the early decades of the sample is that the income tax affected very few people.

[2]For simplicity we neglect here distributions of property other than cash. Stock dividends are treated as a simple redefinition of the units of ownership, although in fact they do have some implications for the corporation's books.

[3] By buying for cash equity claims of <u>other</u> corporations, a corporation can accomplish much the same effect as repurchasing its own shares. Because of the rules for taxation of intercorporate dividends (discussed in the text), this continues to be roughly true even in the presence of taxes.

[4] Under some circumstances capital gains are subject to a special "alternative minimum tax," with but trivial consequences for the effective marginal rate. More important is the fact that if an asset is held to the owner's death capital gain to that date goes free of income tax.

[5]To analyze fully the tax treatment of capital gains it is necessary to take account of future tax consequences via changes in "basis" of the shares owned by the taxpayer. For details see Bradford [1980].

[6] This does not encompass mutual funds, which serve as conduits for individual shareholders.

[7] The 28 percent rate is an alternative tax. Since the corporation income tax is assessed on a graduated schedule, it will be advantageous for the firm to treat capital gains as ordinary income when total taxable profits are low enough. Note also that under rare conditions corporate capital gains are subject to an additional "minimum tax."

[8] As we shall see below the influence of an agent on the market α is negatively associated with his risk aversion. Arguably, large institutional investors, with relatively high α's are also relatively risk neutral.

[9] We are inclined to this even though, as King [1977] shows, maximization of market value will not in general be the preferred objective of all or even the majority of stockholders. Without claiming to make the point precisely, we conjecture that the deviation of firm behavior from that implied by wealth maximization will be small in a large system such as the U.S. economy. As we have noted, our empirical results seem consistent with this view.

[10] Cf. Tobin [1969], Ciccolo [1975], or von Furstenberg [1977].

[11] See Gordon [1979] for further discussion.

[12] Note that much the same effect as stock repurchase can be accomplished by the take-over of other corporations in a purchase involving cash, or even by the simple purchase of stock in the market, taking advantage of the 85 percent dividend received deduction. If such repurchase of equity is ruled out, however, as in Bradford (1977), then α may remain below one even in equilibrium.

[13] That (16) follows directly from (14) and rational expectations seems to have been overlooked in the finance literature, e.g. Fama [1968].

[14] When we measured the covariance in the residuals at any date among firms in the same industry, and among firms in different industries, we found that the size of the covariance was normally about ten per cent of the variance of the residual, implying little bias in the reported standard errors.

[15] As a simple case, assume N firms per portfolio and I portfolios at each of T dates. Assume that the residuals all have variance σ^2, while residuals for firms in the same portfolio at any date have covariance ρ. Assume one independent variable for notational simplicity whose average squared value is x^2, regardless of firm or date. For any portfolio, where the average of the independent variables for the component firms becomes the new independent variable, assume its average squared value to be $p\,x^2$, where $0 < p < 1$ due to the averaging out of individual variation among firms in a portfolio. By solving explicitly for the variance of the coefficient estimate under each procedure, one can show after some messy algebra that the variance of the estimate on portfolio data is

$$\frac{\sigma^2 + (N-1)\rho}{Np\,ITx^2}$$ while that on individual data is

$$\frac{\sigma^2 + (N-1)\rho}{NITx^2} \cdot \frac{\sigma^2 - \rho}{\sigma^2 + \rho(N(1-p)-1)}$$. It is straightforward to show that the latter is necessarily smaller.

[16] The dividend yield was used directly, without an instrument, since it was very stable over time.

[17] We have no breakdown here into interest payments and capital gains. Therefore, the average "α" for bonds will be incorporated into the θ weight on bonds.

[18] An alternative rationalization for this specification, suggested to us by Stephen Ross, involves assuming three underlying stochastic factors, f_1, f_2, and f_3 which jointly determine the return on all types of securities, that is $R_j = \gamma_j \; f_1 + \gamma_{j2} f_2 + \gamma_{j3} f_3$. $R_m = \gamma_{m1} f_1 + \gamma_{m2} f_2 + \gamma_{m3} f_3$. Then, except in degenerate cases, there exists a set of weights θ_j such that $R_m = \sum_{j=1}^{3} \theta_j R_j$.

[19] Here, b is estimated to be -4.3. While less than α , as expected, it is implaus8bly low, again rationalizing the use of firm specific intercepts.

[20] Gordon and Malkiel [1979] find that since the late 1950's, the debt equity ratio has grown steadily from a value in the late 1950's of about .18.

REFERENCES

Bhattacharya, Sudipto, "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," The Bell Journal of Economics, 10 (Spring, 1979), pp. 259-270.

Black, Fischer, "Capital Market Equilibrium with Restricted Borrowing," Journal of Business, 45 (1972), pp. 444-455.

Black, Fischer, "The Dividend Puzzle," Journal of Portfolio Management, 2 (Winter 1976), pp. 5-8.

Black, Fischer and Myron Scholes, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," Journal of Financial Economics, 1 (1974): pp. 1-22.

Brennan, Michael J., "Taxes, Market Valuation and Corporate Financial Policy," National Tax Journal, 23 (1970): pp. 417-27.

Bradford, David F., "The Incidence and Allocation Effects of a Tax on Corporate Distributions," Discussion Paper #7738, Center for Operations Research and Econometrics, Universite Catholique de Louvain, Belgium, August, 1977. For a revised version, see N.B.E.R. Working Paper #349, May, 1979.

Bradford, David F., "The Economics of Tax Policy Towards Savings," in George M. Von Furstenberg, ed., The Government and Capital Formation, Cambridge, MA: Ballinger, 1980.

Ciccolo, John H., "Four Essays on Monetary Policy," Unpublished doctoral dissertation, Yale University, 1975.

Elton, Edwin J. and Martin J. Gruber, "Marginal Stockholder Tax Rates and the Clientele Effect," Review of Economics and Statistics, 52 (Feb., 1970), pp. 68-74.

Fama, Eugene F., "Risk, Return and Equilibrium: Some Clarifying Comments," The Journal of Finance, 23 (1968): pp. 29-40.

Gordon, Roger H., "Tobin's q and the Rate of Investment: A Reappraisal in a World with a Flexible Debt-Equity Ratio," mimeo, May, 1979.

Gordon, Roger H. and Burton G. Malkiel, "Taxation and Corporation Finance," mimeo, September, 1979.

Ibbotson, Roger G. and Rex A. Sinquefield, Stocks, Bonds, Bills and Inflation: The Past (1926-1976) and the Future (1977-2000) (Charlottesville, Va.: The Financial Analysis Research Foundation, 1977).

King, Mervyn A., Public Policy and the Corporation (London: Chapman and Hall, 1977).

Litzenberger, Robert H. and Krishna Ramaswamy, "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence," mimeo, 1978.

Rosenberg, Barr and V. Marathe, "Tests of Capital Asset Pricing Hypotheses," Journal of Financial Research, forthcoming.

von Furstenberg, George M., "Corporate Investment: Does Market Valuation Matter in the Aggregate," Brookings Papers on Economic Activity, 2 (1977): pp. 347-397.

Williams, Joseph T., "Capital Asset Prices and Heterogeneous Beliefs," Journal of Financial Economics, 5 (Nov., 1977), pp. 219-240.