Volume Title: Determinants of Investment Behavior

Volume Author/Editor: Robert Ferber, ed.

Volume Publisher: UMI

Volume ISBN: 0-87014-309-3

Volume URL: http://www.nber.org/books/ferb67-1

Publication Date: 1967

Chapter Title: Estimates of the Cost of Capital Relevant for Investment Decisions Under Uncertainty

Chapter Author: Merton Miller, Franco Modigliani

Chapter URL: http://www.nber.org/chapters/c1236

Chapter pages in book: (p. 189 - 214)

certainty case (see [9], footnote 15) the current market value of the firm can be expressed as

$$V = \frac{1}{\rho}\bar{X}(1-\tau) + \tau D + k\bar{X}(1-\tau)\left[\frac{\rho^* - C}{C(1+C)}\right]T, \tag{8}$$

where the first two terms, as before, represent the capitalized value of the current tax-adjusted earning power plus the tax benefits on debt, and the last term is the contribution to value of the future growth potential.

Despite the heroic simplifications invoked in its derivation, the above expression for growth potential is still by no means a simple one. It is the *product* of three separate elements: the profitability of the future opportunities as measured by the difference between ρ^* and $C(L)$; the size of these opportunities $(k\bar{X}(1-\tau))$; and how long they are expected to last (T). None of these component terms is directly observable, though some such as $k\bar{X}(1-\tau)$ (and possibly ρ^*) might be approximated by extrapolating recent past experience. In this paper, we take the simplest way out by focusing on the most tractable component $k\bar{X}(1-\tau)$, the level of investment opportunities, and impounding the others in its regression coefficient.

As an empirical estimate of investment opportunities, we have used in the subsequent estimating equations the quantity $(1/5(A_t - A_{t-5})/A_{t-5}) \cdot A_t$. That is, we have used a linear five-year average growth rate of total assets times current assets denoted for simplicity hereafter as $\Delta\bar{A}$. This particular form of average, for reasons still not entirely clear to us, yields consistently higher gross and net correlations with total value than other simple averages we have tried. But the differences are not large and the estimates of the other coefficients are not sensitive to the specific measure used.

5. DIVIDEND POLICY, VALUATION AND THE COST OF CAPITAL

Under ideal conditions of perfect capital markets, rational investor behavior, and no tax discrimination between sources of income, dividend policy would present no particular problem. In such a setting, we have shown [9] that, given a firm's investment policy, its dividend policy will have no effect whatever on the current market value of its shares or on its cost of capital; and that despite the impressions of some writers to the contrary (see, e.g., Lintner [7]), this conclusion is equally valid whether one is considering a world of certainty or uncertainty. Dividend policy serves to determine only the division of the stockholders' return between current cash receipts and capital appreciation, and the division of the firm's equity financing between retained earnings and external flotations.

The picture becomes considerably more complicated, however, as soon as we weaken the assumptions to allow for the present tax subsidy to capital gains and for the existence of brokerage fees and flotation costs. Under these conditions, a firm's dividend policy can, in general, be expected to have some effect on its market value though the precise amount of the effect is impossible to determine a priori.

Given this uncertainty as to the size and, to some extent, even the direction of the dividend effect, the indicated course might seem to be simply to add a dividend term with an unspecified coefficient to the structural equation (8) and let the sample determine its value. From such a valuation equation we could, of course, also go on to derive an extension of the cost of capital formula (7) running in terms of dividend policy as well as debt policy.[8]

The trouble with such an approach, however, is that if it is applied in a straightforward fashion, as in Gordon [3] or Durand [2], the resulting estimate of the dividend coefficient will inevitably be strongly biased upward (and the key earnings coefficient correspondingly biased downward). Since the precise mechanism generating this bias has been described at length in our [11] and will be further referred to below, we need not dwell on the matter further at this point beyond observing that the difficulty arises from the widespread practice of dividend stabilization. With current dividends based in large part on management's expectations of long-run future earnings, the dividend coefficient in the regression equation will reflect this substantial informational content about $\bar{X}(1-\tau)$ along with the true effect, if any, of dividends per se on valuation.

Because of this confounding of the earnings and dividend coefficients, our approach here will be to omit the dividend variable entirely and to focus on the problem of estimating the earnings coefficient (which is, of course, to be interpreted as the capitalization rate for earnings for companies following the sample average dividend policy). As it turns out, tests of the dividend effect (presented in detail in the unabridged

[8] Although the procedure for deriving the marginal and average costs of capital in the dividend case is analogous to that for the leverage case, the derivation is considerably more complicated. Further difficulties arise from the fact that, in such a setting, maximizing market value is no longer always equivalent to maximizing the economic welfare of the owners. Since these and related problems are largely peripheral to the main concerns of this paper, further discussion of them will be deferred to a separate paper.

version) indicate that it is quite small in this industry for the years under study and can safely be neglected for our main purposes here.

6. SIZE AND VALUATION

All the valuation equations considered so far have been written as linear homogeneous functions of the independent variables, implying, among other things, that a given proportionate change in the values of all the independent variables leads to an equal proportionate change in the market value of the firm. The results of previous valuation studies (see, e.g., Gordon [4]) suggest, however, that the true market capitalization rate for the expected earnings of large firms may tend to be larger than that of small firms in the same industry.

As was true of the growth effect, there are a number of possible ways of incorporating this size or scale effect into the model. By far the simplest is merely to add a constant term to the valuation equation. The resulting nonhomogeneous equation must then be interpreted as the linear approximation over the sample range to the underlying nonlinear relation, and the coefficient of the earnings variable as the (constant) marginal capitalization rate in the industry. The magnitude and direction of the scale effect would be indicated by the size and sign of the constant term. A negative constant term would confirm that the average capitalization rate is less than the marginal rate and hence that the average capitalization rate tends to rise with increasing size of firm. A positive value for the constant term, on the other hand, would imply decreasing returns to scale in valuation.

7. THE ECONOMETRIC MODEL AND THE METHOD OF ESTIMATION

Our analysis of the theory of valuation thus leads to the following structural equation:

$$(V - \tau D) = a_0 + a_1\overline{X}(1 - \tau) + a_2\overline{\Delta A} + U, \tag{9}$$

where a_1 is the marginal capitalization rate for pure equity streams in the class and hence the key parameter for deriving the cost of capital, a_0 is an intercept term whose size and sign will measure any effects of scale on valuation, a_2 is a measure of the effects of growth potential on value, and U is a random disturbance term. Note that since the theory implies that the coefficient of the leverage variable D is equal to the marginal corporate tax rate τ, we shall, to increase the efficiency of estimation, so constrain it by incorporating it with the dependent variable.

Least-squares estimates of the coefficients of (9) will be efficient and unbiased only if, among other things, the variance of the disturbance term U is a constant, independent of the size of the firm, and the disturbances are not correlated with the independent variables. Unfortunately, neither of these conditions can reasonably be expected to hold in our sample.

As for the variance of the disturbances, one would certainly suppose that the errors in a valuation equation, including errors in measuring $(V - \tau D)$, are of the multiplicative rather than the additive variety. And indeed, check of the simple scatter of value on measured earnings suggests that the error term is approximately proportional to the size of firm. Any attempt to fit (9) directly, therefore, would be highly inefficient and in our sample (where the largest firm is on the order of 100 times the smallest) the results would be completely dominated by a handful of giant companies.

In the present context, there are at least two approaches worth considering as possible solutions for this problem of heteroscedasticity: (i) dividing (9) through by $(V - \tau D)$ and re-expressing the structural relation in so-called "yield" form; or (ii) weighting each observation in inverse proportion to the size of the firm and hence to the size of the standard deviation of the error. The former leads to the estimating equation

$$\frac{\overline{X}(1 - \tau)}{V - \tau D} = a'_1 + a'_0 \frac{1}{V - \tau D} + a'_2 \frac{\overline{\Delta A}}{V - \tau D} + u', \qquad (10)$$

where $a'_1 = \rho =$ the reciprocal of the capitalization rate for pure equity streams (or, equivalently, the "marginal cost of equity capital"), $a'_0 = a_0\rho$, $a'_2 = -a_2\rho$, and $u' = -\rho(U/V - \tau D)$, with Var (u') approximately a constant for all firms.

While an approach of this kind has the virtue of simplicity, it suffers from the fact that the variable $(V - \tau D)$ enters into the denominator of the ratios on both sides of the equation. This is not only somewhat unesthetic—since we are, in effect, using V to explain V—but will lead to biased estimates to the extent that $(V - \tau D)$ contains stochastic elements independent of those in the numerator of the ratios. In the present case, this will mean that the coefficients of the growth and size variables will be too high (i.e., less negative) and that the estimate of the cost of capital (from the intercept term a_1) will be correspondingly too low. Since $(V - \tau D)$ certainly does have a stochastic component—impounded in the term U in (9)—and since we have, at this stage, no basis for judging how large the resulting bias really is, we obviously cannot afford to rely on estimating equations of this form. We shall therefore rely primarily on the weighted regression approach.

Assuming that the standard deviation of the error term in (9) is roughly proportional to size of firm, the required weighting can be effected by the relatively simple expedient of deflating each of the variables by the book value of total assets, denoted by A. Our reason for using total assets as a deflator rather than, say, total sales (as, e.g., in Neilsen [13]) is mainly that in the utility industry at least such deflated terms as V/A, D/A, or $\bar{X}(1-\tau)/A$ have natural and useful economic interpretations in their own right. The equation to be fitted, then, will be of the form

$$\frac{V-\tau D}{A} = a_0\frac{1}{A} + a_1\frac{\bar{X}(1-\tau)}{A} + a_2\frac{\overline{\Delta A}}{A} + u, \qquad (11)$$

with $u = U/A$ and Var (u) = a constant.

One question that immediately arises in connection with (11) is the status of the constant term. Recall that we are interpreting the basic valuation equation (9) in the original, undeflated variables as a linear approximation over the sample range, with its constant term a_0 serving as a measure of the effect of scale on valuation. To preserve this interpretation, we must, therefore, regard the derived deflated regression (11) as homogeneous, that is, as being fitted with no constant term and with the coefficient of the variable $1/A$ now measuring the size effect.

A potentially much more serious problem than heteroscedasticity is that posed by the lack of independence between the disturbance term in (11) and the independent variables, particularly the key earnings variable $\bar{X}(1-\tau)/A$. That variable is defined, it will be recalled, as the market's expectation of the long-run, future earning power of the assets currently held by the firm. Since it is an expectation, it is not directly observable or measurable and the best that can normally be done is somehow to approximate it from the firm's published accounting statements. This best, unfortunately, is likely to be none too good even in an industry, such as the electric utility industry, where there is substantial uniformity of accounting conventions among firms, where there are (at least in our sample period) no firms suffering net losses, and where large, year-to-year random fluctuations in reported earnings seem to be relatively rare.

The implications of these inevitable errors in the measurement of earnings for the problems at hand are perhaps most easily seen by expressing the underlying structure as the following *system* of equations (where, to simplify the notation, we let $V^* = (V-\tau D)/A$, $X^* =$

$\bar{X}(1-\tau)/A$ = the "true" unobservable expected earnings, X = deflated earnings as measured from the accounting statements, and Z_i, $i = 1 \ldots m$ stand for all other relevant variables (including constants, where appropriate):

$$V^* = \alpha X^* + \sum_{i=1}^{m} \beta_i Z_i + u \tag{12a}$$

$$X = X^* + v \tag{12b}$$

$$X^* = \sum_{i=1}^{m} \gamma_i Z_i + w \tag{12c}$$

where some γ_i and β_i may be zero, and with the error terms assumed to be independent of each other and to have mean zero and (constant) variances σ_u^2, σ_v^2, and σ_w^2 respectively. In other words, the value of the firm depends on expected earnings and certain additional explanatory variables; measured earnings are merely in approximation to true expected earnings, the error of measurement being v; and lastly, at least some of the explanatory variables are also correlated with (and hence convey information about) the true but unobservable X^*.

The equations above thus constitute a simultaneous system in which V^* and X are, in effect, the endogenous variables, and the Z_i are the exogenous variables. It follows, then, that if we attempt to fit by direct least squares the single equation.

$$V^* = aX + \sum_{i=1}^{m} b_i Z_i + u' \tag{13}$$

in which V^* is regressed on the Z_i and the endogenous, measured earnings X, the error term u' will not be independent of X and the coefficients of (13) will be biased. More concretely, it can readily be shown (see, e.g., Chow [1], esp. pp. 94–98) that, in the limit for large samples, the coefficients of X will be given by

$$a = \alpha \frac{\sigma_w^2}{\sigma_w^2 + \sigma_v^2}$$

which is less than the true value α and the more so the larger the variance of the error of measurement σ_v^2 and the better proxies the included exogenous variables are for earnings (i.e., the smaller the value of σ_w^2).[9] As for the other variables, the coefficients will be given by

$$b_i = \beta_i + \gamma_i \alpha \frac{\sigma_v^2}{\sigma_w^2 + \sigma_v^2} = \beta_i + \gamma_i[\alpha - a]$$

and thus may be larger or smaller than their true values (β_i) depending on the direction of correlation with X^* (i.e., on the sign of γ_i).[10]

Recasting the original structure in the form of (12) not only serves to clarify the nature of the biases introduced by errors of measurement, but also suggests a remedy, namely, an instrumental variable approach. For reasons of computational simplicity as well as ease of interpretation in the present context, we shall implement this approach by means of a two-stage procedure formally equivalent to the two-stage least-squares method of Theil [14]. (See also Madansky [8].) Operationally, this means first regressing the endogenous variable X on all the instrumental variables Z_i, thereby obtaining estimates g_i of the coefficients γ_i in (12c). From these estimates, a new variable $\bar{X}$ is formed, defined as $\sum_{i=1}^{m} g_i Z_i$, and thus constituting an estimate of X^* from which, if our assumptions are correct, the error of measurement v will have been purged. If $\bar{X}$ is then used in the second stage as the earnings variable in (13) in place of X (and if the conditions for identification are met), the resulting estimates of a and the b_i can be shown to be consistent estimates of α and the β_i in the basic structural equation (12a).

As for the specific exogenous or instrumental variables to be used, we have already considered two, growth and size. In addition—for reasons discussed in detail in the unabridged version—we shall use total assets, two capital structure variables (D/A and P/A), and total dividends paid.

[9] The above expression for the bias in the earnings coefficient was derived on the assumption that (13) was fitted *with* a constant term. If the equation were fitted without a constant term (and if, as we have assumed, there is no constant term in the true specification), then the apparent bias will be considerably smaller. The reason is, of course, that the bias, by flattening the slope of the regression, tends to produce a positive intercept even where none really belongs. Hence forcing the regression through the origin and eliminating the artificial intercept offsets some of the distortion. The offset is only partial, however, and forcing the regression through the origin cannot be regarded as a satisfactory substitute for the more elaborate methods for eliminating the bias to be introduced below.

[10] Note that even if some $\beta_i = 0$ (implying that the corresponding Z_i really has no effect on market value) its estimate b_i might still be positive if $\gamma_i > 0$. And the b_i might be quite large if γ_i is large and if the measurement error in X is substantial (so that a is considerably smaller than α). This is, of course, precisely the "information effect" or proxy variable bias we were concerned about in connection with the dividend variable (see section II 5).

III. The Results

1. THE VALUATION EQUATION

The two-stage least-squares estimates of (11) for the three sample years are presented in Table 1. Since our concern here is primarily with

TABLE 1

Two-Stage Least-Squares Estimates of the Basic Valuation Model (Dependent Variable: $(V - \tau D)/A$)

	Coefficients of					
Year	Earnings $(\overline{X}^{\tau} - \tau\overline{R})/A$	Size $1/A \cdot 10^7$	Growth $\overline{\Delta A}/A$	Mult. R	Adjusted Standard Error	Ratio of Adjusted Standard Error to Mean, V/A
1957	16.1 (.46)	−.280 (.08)	1.36 (.23)	.88	.057	.052
1956	16.7 (.40)	−.114 (.07)	.896 (.21)	.87	.057	.051
1954	19.7 (.45)	−.244 (.07)	.299 (.18)	.73	.063	.053

the cost of capital rather than valuation per se, we shall not comment on these estimates in any detail. Suffice it to say that the results compare quite favorably in terms of explanatory power—as measured, for example, by the ratio of the standard error of the regression to the mean value of V/A—with those that have been obtained in other valuation studies using very different (and to us, at least, very unsatisfactory) specifications. Of particular interest, of course, is the behavior of the growth variable. Future growth potential (though small relative to current earning power and the tax subsidy to debt in terms of contribution to total market value) apparently increased steadily in absolute and relative importance over this period and by 1957 accounted for something over 10 per cent of total market value for the average firm in the sample.

For further reference and comparison, we present in Table 2 two alternative sets of estimates of the valuation equation. The first is the direct least-squares regression of $(V - \tau D)/A$ on measured earnings, with the constant term suppressed. As can be seen, the differences from the two-stage estimates are generally quite small, a result not really

TABLE 2

Direct Least-Squares Estimates with Measured Earnings

Part A: Value Form (Dependent Variable: $(V - \tau D)/A$)

Year	Coefficients of: Earnings $(\bar{X}^\tau - \tau\bar{R})/A$	Size $1/A \cdot 10^7$	Growth $\overline{\Delta A}/A$	Mult. R	Adjusted Standard Error	Ratio of Adjusted Standard Error to Mean, V/A
1957	16.0 (.44)	−.277 (.08)	1.39 (.23)	.88	.057	.052
1956	16.6 (.39)	−.111 (.07)	.926 (.21)	.87	.057	.051
1954	19.2 (.43)	−.205 (.07)	.466 (.17)	.75	.063	.053

Part B: Yield Form (Dependent Variable: $(\bar{X}^\tau - \tau\bar{R})/V - \tau D$)

Year	Coefficients of: Constant	Size $(1/V - \tau D) \cdot 10^6$	Growth $\overline{\Delta A}/V - \tau D$	Mult. R	Reciprocal of Constant Term and Its Implied Standard Error	Ratio of Standard Error of Regression to Mean of Dependent Variable
1957	.0592 (.002)	.166 (.04)	−.0516 (.02)	.58	16.8 (.44)	.07
1956	.0582 (.001)	.066 (.04)	−.0325 (.01)	.39	17.4 (.38)	.07
1954	.0506 (.001)	.121 (.04)	−.0124 (.01)	.45	19.8 (.45)	.07

very surprising in the light of the suppression of the constant term.[11] The differences turn out to be considerably larger (particularly in 1954) when the constant term is not suppressed and, indeed, it is only in the case of the two-stage estimates that the constant term really does approach zero. The differences become larger still when other tests of the basic specification (notably those concerned with the leverage and dividend variables) are considered.

The second panel shows the estimates obtained by direct least-squares regressions in the "yield form" of equation (10) above. The main drawback of this approach, it will be recalled, comes from the presence of $V - \tau D$ in the denominators of variables on both sides of the equation, which imparts an upward bias to the coefficients of the independent variables and a consequent downward bias to the crucial constant term. Since the direction of the bias is known, however, we can use equations of this form to provide at least a rough check on the reasonableness of the estimates obtained by the more roundabout, two-stage approach.

To facilitate comparison with the estimates in Table 1, a column has been added showing the reciprocal of the constant term, which is the estimate of the capitalization factor for earnings implied by the observed constant terms in the yield equations. As predicted, the capitalization factors obtained via the yield equations are indeed all higher than those obtained via the two-stage approach. The gap between the two sets of estimates tends to widen somewhat over time, but the differences are never very large. This close agreement should remove any lingering fears that major distortions in the estimates may somehow have been introduced in the two-stage approach. At the same time, it suggests that the simpler yield equations may still have a useful role to play in valuation studies, particularly where the interest is mainly in determining the direction of changes in the cost of capital over time rather than developing precise estimates or testing the basic specification as developed here.

2. THE COST OF EQUITY CAPITAL

Turning now from valuation to the other side of the coin, the cost of capital, we show in Table 3 the estimates of the cost of equity capital implied by the earnings coefficients of Table 1. For comparison, the table also shows two other measures of the cost or "ease of acquisition" of equity capital frequently used by economists in investment studies, namely, the average earnings-to-price ratio and the reciprocal of the

[11] See footnote 9.

TABLE 3

Estimated Cost of Equity Capital and Some Alternative Measures of Equity Costs

Year	Estimated Cost of Equity Capital (ρ)		Average Earnings Yield on Shares ($\bar{\pi}^{\tau}/S$)		Reciprocal of Ratio of Price to Book Value (B/S)		Average Tax and Leverage Adjusted Total Earnings Yield, $(\bar{X}^{\tau} - \tau\bar{R})/(V - \tau D)$	
	Amt. (1)	As Per Cent of 1954 (2)	Amt. (3)	As Per Cent of 1954 (4)	Amt. (5)	As Per Cent of 1954 (6)	Amt. (7)	As Per Cent of 1954 (8)
1957	.062 (.002)	122	.070	106	.64	105	.056	110
1956	.060 (.001)	118	.070	106	.63	103	.056	110
1954	.051 (.001)	100	.066	100	.61	100	.050	100

ratio of the average price to the book value of the shares. Notice that all three measures indicate a rise in the cost of equity capital between 1954 and 1957, but our measure indicates a steeper and more substantial increase over the interval. The causes and implications of this apparently lesser responsiveness of the standard measures will become clear in subsequent discussion.

Insofar as levels are concerned, notice that the average earnings yield happens to be consistently higher than our estimate of the cost of equity capital. We say happens to be to emphasize that, under our model of valuation, there is no "normal" or even simple relation to be expected between the two concepts. The earnings yield for any company is not a given fixed number for each member of the class, but rather a function whose arguments include the cost of equity capital for the class, the firm's growth potential, its leverage policy, and its size. The sample mean earnings yield shows only the combined effect of these different and, to some extent, offsetting influences.

3. VALUATION, GROWTH, AND THE COST OF EQUITY CAPITAL

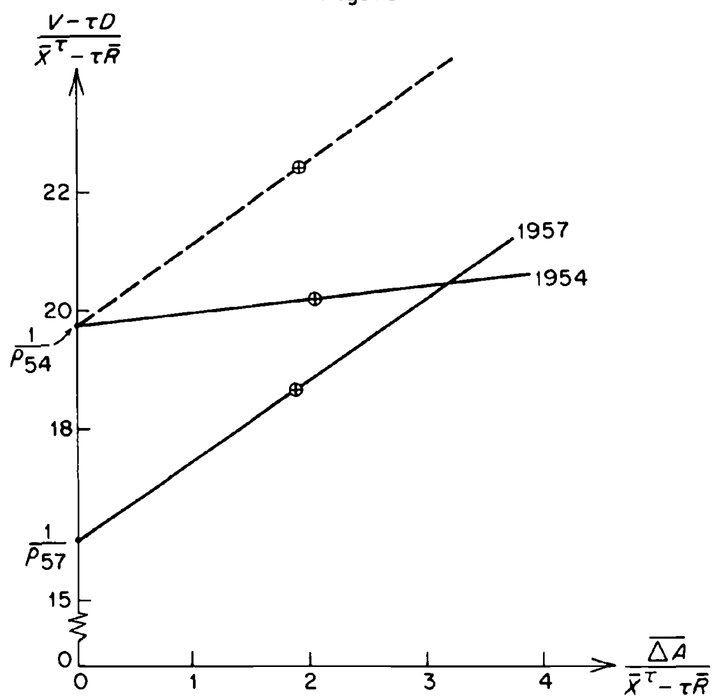
Because of the distortions resulting both from leverage effects and from the fact that the market value of shares incorporates the capitalized value of growth opportunities, we must conclude that the earnings-to-price ratio—which together with long-term interest rates is the most widely used measure of the cost of capital in investment studies—is unlikely to provide an adequate approximation.

The first of these distortions—that arising from leverage—could be handled by falling back on a measure of market yield somewhat different from the earnings-to-price ratio, though related to it. Let us suppose first that there were no corporate income taxes. Then we know that, in the absence of valuable growth opportunities, the fundamental valuation equation, expressed in yield form, would be simply $\bar{X}/V = \rho$. The ratio of expected total earnings to total market value—which may be thought of as a "leverage-corrected" yield—would thus provide a direct estimate of ρ. In principle, any firm could so approximate its cost of equity capital from its own company data, although, of course, as a practical matter, a better estimate would be obtained by averaging over a large group of similar firms so as to wash out any random noise in $\bar{X}$ or V. When we allow for taxes and the consequent tax subsidy on debt the picture becomes slightly more complicated, but a direct approximation of ρ still exists. The appropriate yield—see (5)—now becomes $(\bar{X}^{\tau} - \tau\bar{R})/(V - \tau D)$, the ratio of total tax-adjusted earnings to total market value minus the value of the tax subsidy.

This method of direct approximation breaks down, however, in the presence of growth. The leverage adjusted yield will be systematically too low as an estimate of ρ for any company with growth potential as will be the group average yield for any sample that contains significant numbers of growth companies. Nor will the movements of the yield over time conform well with changes in ρ to the extent that the market's evaluation of future growth potential changes over time (and, of course, much of the short-term variation visible in share prices stems precisely from this source). Some idea of how sizable the distortions of level and movement of the yield relative to ρ can be—even in such a low-growth industry as our electric utilities and even over such a short span of time —is provided by a comparison of our estimates of the cost of equity capital in the first column of Table 3 with those of the tax and leverage adjusted yield, $[(\bar{X}^\tau - \tau\bar{R})/(V - \tau D)]$, in the last column of that table.

One somewhat surprising aspect of this comparison, already noted above, is the relative stability of the leverage-adjusted yield series over this period. Because of the many uncertainties surrounding estimates of

Figure 1



future growth potential and because of the sensitivity of current market values to even small changes in projected growth rates, one would expect the growth component in the denominator of the yield ratio to be quite volatile, and hence that the market yield would tend to swing quite substantially in response to these continuing re-evaluations. Some idea of why this "normal" pattern did not obtain during our sample period can be gained from Figure 1. The solid-line functions plotted there are the basic value regressions of Table 3 for the beginning and ending years, 1954 and 1957, expressed in ratio form as

$$\frac{V - \tau D}{\overline{X}^{\tau} - \tau \overline{R}} = a_1 + a_2 \frac{\overline{\Delta A}}{\overline{X}^{\tau} - \tau \overline{R}}$$

(and hence ignoring the minor size effect). The dependent variable is thus the reciprocal of the tax- and leverage-adjusted yield; the intercept a_1 is our estimate of $1/\rho$; and the slope a_2 is the coefficient of growth in the V equation.[12]

Notice that in 1954, at the beginning of the period, the market's estimate of the growth potential of the industry was quite low. Because the slope was so flat, the approximate sample mean value of $(V - \tau D)/(\overline{X}^{\tau} - \tau \overline{R})$—indicated by the circled cross—differed only very slightly from the estimate of $1/\rho$ implied by the intercept. By 1957, however, a striking increase had taken place in the market's valuation of the prospects for continuing profitable growth in the industry. As can be seen from the broken line—which has been plotted with 1954 intercept and 1957 slope—this large revaluation would have pushed the average value of $(V - \tau D)/(\overline{X}^{\tau} - \tau \overline{R})$ up by nearly 15 per cent to 22.4 (equivalent to a yield of about .044) if no other changes had occurred. But instead of this "pivoting" around a stable intercept of $1/\rho_{54}$, our estimates indicate that a simultaneous and quite substantial drop in the intercept took place (i.e., rise in the cost of equity capital). So substantial was this drop, in fact (when combined with the slight fall in the mean value of the growth variable itself), that the upward push of the revaluation of growth was more than offset; and the mean value of $(V - \tau D)/(\overline{X}^{\tau} - \tau \overline{R})$ actually fell by about 10 per cent.

Although these compensating movements in ρ and the market's evaluation of growth "explain" the relative stability of the tax- and leverage-adjusted yield during the sample period, the explanation may

[12] The use of the reciprocal of the yield rather than the yield itself is simply a matter of convenience since the presence of growth impounded in V would lead to a nonlinear relation between $(\overline{X}^{\tau} - \tau \overline{R})/(V - \tau D)$ and growth as measured by $\overline{\Delta A}/(V - \tau D)$.

strike the reader as somewhat paradoxical. Growth potential, after all, is the opportunity to invest in the future in projects whose rates of return exceed the cost of capital. One would expect, therefore, that a rise in the cost of capital would normally be associated with a *fall* in growth potential. There are a number of possible explanations for the opposite behavior in the present instance, but discussion of them is perhaps best postponed until we have first provided estimates of the average cost of capital relevant for investment decisions.

4. THE REQUIRED YIELD OR AVERAGE COST OF CAPITAL

As emphasized earlier, the relevant cost of capital for investment decisions at the level of the firm is the average cost of capital, $\rho(1 - \tau L)$, where L measures the "target" proportion of debt in future financing. The average cost is thus not a fixed number, but a schedule or function whose arguments are ρ (which is an "external" property of the class or industry determined by the market) and L (which is a matter of "internal" company policy). Although the average cost of capital, unlike the cost of equity capital, is thus in principle different for each firm in the industry, we can get some idea of its value and behavior for the typical electric utility by using a typical or average value for L. The obvious candidate, of course, is the actual sample average of D/A for each year, since D/A measures the average proportion of debt in past financing and this proportion is likely to be quite stable (particularly when averaged over the industry). Estimates with these values for L are shown in Table 4, columns 1 and 2. (For notational convenience, we shall hereafter refer to these estimates as $C(\bar{D}/\bar{A})$, using $C(L)$ to mean the function itself and $C(D/A)$, i.e., without the bar on D/A, to refer to the function evaluated not at the industry mean, but at a particular company's value of D/A.)

These estimates $C(\bar{D}/\bar{A})$ of the average cost of capital are, of course, always below the corresponding estimates of ρ (Table 3, column 1); but the movements over time of the two series are closely similar since, as expected, the sample mean value of $\bar{D}/\bar{A}$ is quite stable. Notice also that the estimates of ρ and hence of $C(\bar{D}/\bar{A})$ conform quite closely in their movements with the average yield on AAA bonds in the industry (Table 4, columns 3 and 4)—probably the most popular surrogate for the cost of capital in investment studies. This conformity is particularly interesting since the rate of interest on bonds enters only very indirectly into our calculations, of ρ and $C(\bar{D}/\bar{A})$ and, as can be seen from columns 5 and 6, the implied average rate of interest in our sample does not even seem to conform well with the AAA series. From the

TABLE 4

Average Cost of Capital and Some Comparison Series

Year	Average Cost of Capital, $C(\overline{D}/\overline{A})$		Yields on AAA Public Utility Bonds[a]		Sample Mean Yield on Debt (R/D)		Sample Mean Yield on Preferred Stock (Pdv/P)		Weighted Average Yields: With Book Value Weights $\overline{C}_B(\overline{D/A})$		Weighted Average Yields: With Market Value Weights $\overline{C}_M(\overline{D/V})$		Growth-Adjusted Average Yield[b] $(\overline{X}^{\tau}-\tau\overline{R})/V-G$	
	Amt. (1)	As Per Cent of 1954 (2)	Amt. (3)	As Per Cent of 1954 (4)	Amt. (5)	As Per Cent of 1954 (6)	Amt. (7)	As Per Cent of 1954 (8)	Amt. (9)	As Per Cent of 1954 (10)	Amt. (11)	As Per Cent of 1954 (12)	Amt. (13)	As Per Cent of 1954 (14)
1957	.046	128	.043	137	.029	112	.047	115	.034	104	.042	108	.047	122
1956	.045	125	.039	126	.029	112	.046	112	.035	106	.043	110	.045	115
1954	.036	100	.031	100	.026	100	.041	100	.033	100	.039	100	.039	100

[a] Monthly average for December (from *Federal Reserve Bulletin*).

[b] With G equal to growth coefficient in Table 1 times sample mean value of $\overline{\Delta A/A}$.

economic point of view, this parallelism between movements in ρ and the AAA yields would seem to suggest that, over this short interval at least, the movements of both series were dominated by factors affecting the supply of and demand for capital generally. Changes, if any, in investors' tastes for risk-bearing or in their evaluation of the riskiness of this industry in relation to others were apparently not large enough (except possibly in 1957) to cause any significant divergence of movement in the period under study.

It is also instructive to contrast our estimates of the average cost of capital with those that would be obtained by following the prescriptions laid down in much of the traditional literature of corporation finance. Essentially, these call for computing the weighted sum of the market yields of each type of security, the weights being the "target" proportions of each security in the capital structure. That is, if we let i equal the earnings-to-price ratio (our $\bar{\pi}^{\tau}/S$), p the preferred yield (our $\overline{Pdv}/P$), r the average rate of interest on bonds (our $\bar{R}/D$), l the target debt ratio, and l' the target preferred ratio, then the weighted average cost of capital function under the traditional view can be expressed as $i(1-l-l')+p(l')+r(1-\tau)(l)$. Where the target weights l and l' are computed at book value as is usually recommended (i.e., with $l=D/A$ and $l'=P/A$ and $(1-l-l')=B/A$ in our notation), we shall refer to the resulting average as $\bar{C}_B(D/A)$; where they are taken at market value (i.e., with $l=D/V$ and $l'=P/V$ and $1-l-l'=S/V$), we shall refer to the average as $\bar{C}_M(D/V)$, with unbarred values of the argument standing as before for a single company value and barred values for industry means. Estimates of both $\bar{C}_B(\overline{D/A})$ and $\bar{C}_M(\overline{D/V})$ for the typical firm in the sample, using actual sample mean values of i, p, and r, as well as of the book and market value measures of l and l' in each case, are shown in Table 4, columns 9–12.

As can be seen from Table 4, both the levels and the time paths of $\bar{C}_B(\overline{D/A})$ and $\bar{C}_M(\overline{D/V})$ differ significantly from those of $C(\overline{D/A})$. The largest discrepancies arise in the case of the widely advocated $\bar{C}_B(\overline{D/A})$ measure, which is substantially below $C(\overline{D/A})$ in all three years and which shows only a very slight rise over the period. The market value estimates, $\bar{C}_M(\overline{D/V})$, are considerably closer to those of $C(\overline{D/A})$, but they too fail to indicate the sizable increase in the cost of capital which seems to have occurred during this period.

5. RECONCILIATION WITH CONVENTIONAL AVERAGES

To understand precisely why these three methods of estimating the average cost of capital gave such different answers for the years under study (and why they are likely to continue to diverge for other years and other industries), it is helpful to begin by showing how these estimates would relate to each other in a much simpler world in which no growth potential ever existed. In such a world, we have seen that the ratio $(\bar{X}^\tau - \tau\bar{R})/(V - \tau D)$—the tax- and leverage-adjusted yield of the previous section—would be a measure of our ρ, the cost of equity capital. Hence, from the standpoint of any individual firm, the $C(L)$ function can be expressed as

$$C(L) = \frac{\bar{X}^\tau - \tau\bar{R}}{V - \tau D}(1 - \tau L) = \frac{\bar{X}^\tau - \tau\bar{R}}{V} \cdot \frac{1 - \tau L}{1 - \tau D/V}. \tag{14}$$

The weighted average cost of capital function with market value weights is

$$\bar{C}_M(D/V) = \frac{\bar{\pi}^\tau}{S} \cdot \frac{S}{V} + \frac{\overline{Pdv}}{P} \cdot \frac{P}{V} + \frac{\bar{R}(1 - \tau)}{D} \cdot \frac{D}{V} = \frac{\bar{X}^\tau - \tau\bar{R}}{V} \tag{15}$$

and with the book value weights

$$\bar{C}_B(D/A) = \frac{\bar{\pi}^\tau}{S} \cdot \frac{B}{A} + \frac{\overline{Pdv}}{P} \cdot \frac{P}{A} + \frac{\bar{R}(1 - \tau)}{D} \cdot \frac{D}{A}. \tag{16}$$

Notice first that for the special case of $L = D/V$—i.e., when the target leverage coincides with current leverage at market value, D/V, the function $C(L)$ takes the value $C(D/V) = (\bar{X}^\tau - \tau\bar{R})/V = \bar{C}_M(D/V)$. In other words, if a firm's current and future target leverage is D/V, it will get precisely the same estimate for its average cost of capital regardless of whether it chooses to multiply its current tax- and leverage-adjusted yield by $(1 - \tau D/V)$, to compute the weighted average of the current yields of its outstanding securities with market value weights for each, or to simply use the ratio of expected tax-adjusted earnings to total market value. A similar equivalence of estimates (at least to a very close degree of approximation) would also hold, of course, for economists concerned with "typical" values for the industry and using industry mean values of D/V and of the various yields, i.e., $C(\overline{D/V}) \cong \bar{C}_M(\overline{D/V})$.[13]

[13] Although the equivalence holds for individual company data and for industry averages, there is one important case in which the equivalence very definitely does *not* hold. This is the common case of the firm following the weighted average approach of (15) or (16) with current (or prospective future) company weights, but using industry-wide averages of the component yields so as to obtain more reliable estimates. The trouble here is that the market yield on shares (and to some extent the yields on preferred and bonds as well) are increasing functions of leverage. Hence, for a firm whose target leverage is greater (smaller) than the average for the industry mean yield will be an underestimate (overestimate) of its own yield and the resulting average cost of capital will be too low (high). This problem does not arise under our (14), of course, since $(\bar{X}^\tau - \tau R)/(V - \tau D)$ is not a function of firm policy (as is $\bar{\pi}\ /S$) but an estimate of the external, market-given parameter, ρ.

Note also that if V equals A—which would tend to be the case if there were no growth past or future—then $\bar{C}_B(\bar{D}/A)$ becomes the same function as $\bar{C}_M(D/V)$ and, by extension, as $C(L)$. In this special case of no growth, therefore, all three company and industry-wide estimates will coincide.

This simple picture changes quite drastically, however, as soon as growth potential is introduced. The function $C(L)$ must now be expressed as

$$C(L) = \frac{\bar{X}^\tau - \tau\bar{R}}{V - \tau D - G}(1 - \tau L) = \frac{\bar{X}^\tau - \tau\bar{R}}{V - G} \cdot \frac{1 - \tau L}{1 - \tau\left(\dfrac{D}{V - G}\right)}, \quad (17)$$

where G is the market's current valuation of future growth potential. Hence, as can be seen by referring back to (15), there no longer exists any concept of L for which the function $C(L)$ will be the same as $\bar{C}_M(D/V)$. Note also that in the special case in which future growth potential constitutes the only major source of divergence between V and A, $(\overline{D/A}) \cong (\overline{D/V - G})$ so that our estimates of the average cost of capital $C(D/A)$ would be closely approximated by the ratio $(\overline{\bar{X}^\tau - \tau\bar{\bar{R}}/V - G})$. The actual sample mean values of that ratio (with G taken as the product of the growth coefficient in Table 1 and the mean value of our growth variable $\overline{\Delta A}/A)$ are shown in columns 13 and 14 of Table 4. As can be seen, the approximation to $C(\overline{D/A})$ is indeed quite close in 1956 and 1957; but it is less satisfactory in 1954 where the growth contribution is small, both in absolute terms and relative to the other sources of divergence between V and A.

Where the ratio $(\overline{\bar{X}^\tau - \bar{\bar{R}}/V - G})$ is a good approximation to $C(\overline{D/A})$, it will, of course, also follow that both measures will exceed $\bar{C}_M(\overline{D/V})$ which, as we saw above, is given approximately by $(\bar{\bar{X}}^\tau - \tau\bar{\bar{R}}/V)$. As for the relation between the popular $\bar{C}_B(\overline{D/A})$ and $C(\overline{D/A})$, note that we can express the ratio $(\overline{\bar{X}^\tau - \tau\bar{\bar{R}}/V - G})$ approximately as

$$\left(\overline{\frac{\bar{X}^\tau - \tau\bar{R}}{V - G}}\right) = \left(\overline{\frac{\bar{\pi}^\tau}{S - G} \cdot \frac{S - G}{V - G}}\right) + \left(\overline{\frac{Pdv}{P} \cdot \frac{P}{V - G}}\right)$$
$$+ \left(\overline{\frac{\bar{R}(1 - \tau)}{D} \cdot \frac{D}{V - G}}\right) \quad (18)$$
$$\cong \left(\overline{\frac{\bar{\pi}^\tau}{S - G}}\right)\left(\overline{\frac{\bar{B}}{A}}\right) + \left(\overline{\frac{Pdv}{P}}\right)\left(\overline{\frac{P}{A}}\right) + \left(\overline{\frac{\bar{R}(1 - \tau)}{D}}\right)\left(\overline{\frac{D}{A}}\right),$$

since the assumption $V - G \cong A$ implies $D/V - G \cong D/A$, $P/V - G \cong$

P/A and $S - G/V - G \cong B/A$. Comparison with (16) shows that the weights in the two expressions are essentially the same; but since $(\overline{\pi\tau}/S) < (\overline{\pi\tau}/S - G)$, $\overline{C}_B(\overline{D/A})$ too will necessarily fall short of $C(\overline{D/A})$ when growth is present, and the gap will be larger, the larger is the contribution of growth to the value of the shares.

Once again, then, we see that attempts to infer the cost of capital directly from market yields rather than by the more detailed, cross-sectional estimating procedures developed in this paper break down in the face of growth. Where growth is present, all the popular, short-cut approximations will underestimate the cost of capital; and where the market changes its evaluation of growth potential over time (as is inevitable in view of the nature of growth), the time path of the yield measures may give a quite misleading picture of the true changes in the cost of capital. In particular, in our sample it happens that the market's evaluation of growth increased substantially over the period, thereby causing the yield measures to understate seriously the rise in capital costs that appears to have been taking place at the same time. As noted earlier, it is somewhat paradoxical that these two changes should have occurred simultaneously, since an increase in the cost of capital should tend to reduce what the market is willing to pay for given investment opportunities. We can perhaps throw some light on this paradox by taking a closer look at our growth coefficients and their implicit components.

6. A FURTHER ANALYSIS OF THE VALUATION OF GROWTH

The growth term in our basic valuation equation (see section II 4) is of the form

$$k\overline{X}(1 - \tau)\left[\frac{\rho^* - C}{C(1 + C)}\right]T,$$

where k is the ratio of investment to tax-adjusted earnings, C is the average cost of capital, ρ^* is the tax-adjusted rate of return on new investment, and T is a measure of the length of time for which the opportunities to invest at the rate ρ^* are expected last. In the actual estimating equations, we have taken as our growth variable an estimate of $k\overline{X}(1 - \tau)$, the level of future investment opportunities. Hence, if one accepts the underlying model, the observed coefficients of the growth variable can be interpreted as an approximation to $[\rho^* - C/C(1 + C)]T$.

Now that we have estimates of C, the average cost of capital for a typical firm, we can attempt some further decomposition of these growth coefficients and, in the process, we hope, gain some additional insights into the market's appraisal of the growth potential in this industry. In particular, it should be possible, from what we know about past earnings and about the regulatory process governing earnings in the industry, to make at least a rough approximation of ρ^*.

An obvious first candidate as an approximation to ρ^* is, of course, the current, tax-adjusted rate of return on assets, $(\bar{X}^\tau - \tau\bar{R})/A$. Such a measure, however, is almost certainly an underestimate of ρ^* since we know that there are components of total assets—actually, of total liabilities—that regulatory commissions systematically exclude from the rate base.[14] Some idea of the extent of this underestimate is provided by our knowledge that during these years most of the state commissions were still setting the "reasonable return on the rate base" in the neighborhood of the classical 6 per cent. By contrast, the sample average values of $\bar{X}^\tau/A$—$\bar{X}^\tau$ rather than tax-adjusted earnings $\bar{X}(1-\tau)$ being the relevant earnings concept in rate setting—were only .054, .056, and .055 per cent in 1954, 1956, and 1957, respectively. One simple adjustment, therefore, would be to blow up each sample mean value of tax-adjusted earnings by the ratio of 6 per cent to the sample mean value of $\bar{X}^\tau/A$. The rates of return thus adjusted, as well as the original unadjusted rates of return, are presented in Table 5, along with the estimates of T they imply.

These results would seem to suggest the following as the resolution of the paradox described in the previous section. The observed rise in the market's valuation of the industry's growth potential, in the face of the sharp rise in the cost of capital during the period, cannot reasonably be attributed to any compensating increase in the expected rate of return on future investment. No sharp upward trend in earnings rates, adjusted or unadjusted, is visible in the data; nor would such a trend be expected in view of the regulatory controls over the level of earnings. What seems to have been happening rather is that early in the period investors came to recognize that the regulatory authorities were setting rates at levels

[14] A further word of caution is necessary because the so-called accounting rate of return (earnings after depreciation divided by net assets) is not the same as the ordinary internal rate of return when a firm is growing. This discrepancy does not seem likely to create any very serious problems insofar as the valuation equations or estimates of the cost of capital are concerned since in those equations assets appear only as a deflator and since an explicit growth variable is included. It may, however, raise difficulties for comparisons of the kind being attempted here. We say may, because the ρ^* in our formula is not the usual internal rate of return, but the so-called "perpetual rate of return" (see [9], p. 416), and the relations between that rate and the accounting rate of return have, to our knowledge, nowhere yet been explored. We are indebted to Sidney Davidson and Robert Williamson for some helpful discussions on this general point.

that would probably permit firms in the industry to earn somewhat more than the cost of capital on any new capital invested. The subsequent rise in the cost of capital narrowed the margin of gain somewhat; but its effects on valuation were more than offset by an increase in the length of time that favorable terms for new investment were expected to persist. The actual numerical estimates of the expected duration of growth opportunities, as presented in Table 5, are not, of course, to be taken seriously in view of the many approximations, theoretical and empirical, involved in their computation. But upward revaluation of growth prospects (or at least in increasing awareness of growth potential on the part of investors) is very definitely indicated.

IV. Some Concluding Observations

As emphasized at the outset, this paper should be regarded by economists as a first step toward developing historical estimates of the cost of capital relevant for investment decision-making by business firms. It will have adequately fulfilled its objectives if it has succeeded in convincing economists working with investment functions that there *is* a cost of capital problem, that some of the major econometric problems that have prevented progress in the area to date can be overcome, and that the averages and yield measures of the kind recommended in much of the traditional literature on corporation finance as measures of the cost of capital are likely to be treacherous and unreliable.

As for the direction to be taken by future research, clearly one urgent need is further testing of the basic, rational, behavior-perfect, capital market specification. Some confirmatory evidence for the model is provided in our unabridged paper and further tests on the same sample will be provided in sequel papers. But it would obviously be desirable to have independent tests by others and on samples which are a little fresher, both in age and extent of handling.

Even after a basic specification is agreed upon, there remain numerous perplexing problems of estimation, notably those connected with the crucial growth effect. If the market's expectation of future growth opportunities changed seldom, slowly, or only in response to movements of other more readily measurable variables such as sales, profits, or even dividends, then the task would be the difficult but still essentially straightforward one of extrapolation. In practice, however, the market's valuation of growth potential often changes abruptly, substantially, and with little readily apparent relation to changes in observable economic series.

TABLE 5

Analysis of the Growth Effect

Year	Growth Coefficient[a] (G)	Average Cost of Capital[b] $C(\overline{D/A})$	Average Tax-Adjusted Return on Assets (ρ_1^*)	Average Return Assuming 6 Per Cent Return After Taxes (ρ_2^*)	Implied Value of T[c] (in years) For ρ_1^*	Implied Value of T[c] (in years) For ρ_2^*
1957	1.36	.046	.047	.052	51	11
1956	.90	.045	.048	.052	12	6
1954	.30	.036	.046	.052	1	1

[a] From Table 1.

[b] From Table 4, column 1.

[c] Computed as $G\left[\frac{(C)(1+C)}{(\rho^*-C)}\right]$.

While we, and we hope others, will continue to experiment with the extrapolative approach, we suspect that a somewhat more indirect attack may in the long-run prove more fruitful. In particular, instead of attempting to correct for the elusive and changing growth component in the value of growth companies, effort might be directed to finding and identifying companies that the market seems to regard as essentially no-growth companies—and growth here does not mean expansion, but opportunities to invest large sums at rates of return above the cost of capital. The search for such no-growth firms might proceed either by further decomposing the growth term along the lines of the last part of the previous section, with a view to finding the firms for which $\rho^* \cong C$, or by constructing scatters similar to those in Figure 1 for a series of widely spaced years and observing which firms seem continually to cluster in the near neighborhood of the "pivots." Once a sample of no-growth firms has thus been obtained, estimates of ρ can be made relatively simply and quickly via the leverage-corrected yield route discussed in section III. This approach, or variants relying on a judicious interplay of time series and cross-sectional estimation, should enable us, within a reasonable span of time and with a reasonable amount of effort, to obtain a usable time series of the estimated cost of capital.

Needless to say, even these hoped-for time series estimates would have to be handled quite gingerly in investment studies. It is not to be expected, for example, that the desired capital stock will always adjust quickly to the current level of the cost of capital. Because of the substantial decision-making costs involved, the cut-off rate or required minimum yield on new investments is likely to be changed only infrequently by large firms (typically, but by no means exclusively, on the occasions when external financing is contemplated). By suitably smoothing or lagging the series, however, it should be possible to incorporate at least the major changes in the level of the cost of capital that have occurred as the economy has swung over the last forty years from boom to severe depression, to a postwar prosperity widely regarded as temporary, and finally to a long period of sustained prosperity interrupted by only minor recessions and with fears of major future depressions largely absent.

Finally, we should like to stress once more that the measure of the cost of capital described in this paper should prove primarily relevant for the investment behavior of large corporate enterprises. For smaller firms, other measures—including more conventional measures of interest rates and of those rather elusive factors that may be lumped under the heading of availability of funds—might be a good deal more to the point.

BIBLIOGRAPHY

1. Chow, G., *The Demand for Automobiles in the United States,* Amsterdam, 1957.
2. Durand, D., *Bank Stock Prices and the Bank Capital Problem,* Occasional Paper 54, New York, NBER, 1954.
3. Gordon, M., "Dividends, Earnings and Stock Prices," *Review of Economics and Statistics,* May 1959.
4. ———, *The Investment, Financing and Valuation of the Corporation,* Homewood, Ill., 1962.
5. Hirshleifer, J., "Investment Decision Under Uncertainty," *Quarterly Journal of Economics* (forthcoming).
6. ———, "On the Theory of Optimal Investment Decision," *Journal of Political Economy,* August 1958.
7. Lintner, J., "Dividends, Earnings, Leverage, Stock Prices and the Supply of Capital to Corporations," *Review of Economics and Statistics,* August 1962.
8. Madansky, A., "On the Efficiency of Three-Stage Least-Squares Estimation," *Econometrica,* January-April 1964.
9. Miller, M., and Modigliani, F., "Dividend Policy, Growth and the Valuation of Shares," *Journal of Business,* October 1961.
10. Modigliani, F., and Miller, M., "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review,* June 1958.
11. ———, "Reply," *ibid.,* September 1959.
12. ———, "Corporate Income Taxes and the Cost of Capital: A Correction," *ibid.,* June 1963.
13. Neilsen, S., *Market Value and Financial Structure in the Railroad Industry,* Hartford, 1961.
14. Theil, H., *Economic Forecasts and Policy,* 2d rev. ed., Amsterdam, 1961.