A MEAN-VARIANCE THEORY OF OPTIMAL CAPITAL STRUCTURE AND CORPORATE DEBT CAPACITY

E. Han Kim*

I. INTRODUCTION

An issue of concern to the theory of business finance over the past two decades has been the effect of financial structure on the valuation of firms. The traditional presumption is that a firm's value is a concave function of its financial leverage, and that an optimal financial leverage exists where the slope of the function is zero.1 This argument is suspect to the extent that it attempts to value a firm's securities in isolation from the rest of the capital market. The pathbreaking works by Modigliani and Miller (MM) have provided the foundations for studying the effect of financial structure on the valuation of firms in equilibrium. MM (1958, 1969) establish that the total value of the firm, in the absence of taxes, remains constant across all degrees of financial leverage. Building on the foundations laid by MM, numerous authors2 have confirmed the MM no-tax thesis using a variety of equilibrium approaches. MM (1963) and some of these authors have shown further that a proportional corporate income tax provides sufficient economic incentive for firms to maximize their use of debt financing. However, in the five-year period from 1966 to 1970 the capital needs of nonfinancial corporations in the United States were financed approximately by two-thirds equity and one-third debt.3 Furthermore, the average corporate debt ratio (which reflects the valuation of equity at market value) is only approximately 20 percent.4 Even these highly aggregated figures suggest that an element of major importance to financial managers and the investing public is missing from the MM theory.

Robichek and Myers (1965, p. 20) and Hirshleifer (1970, p. 264) suggest that bankruptcy costs may represent the major missing element and that incorporating these costs within the foundations laid by MM may support the concept of an optimal capital structure. The importance of bankruptcy costs was particularly well demonstrated by Miller (1962) when he explicitly utilized bankruptcy costs to

* Visiting Associate Professor of Finance, Purdue University, on leave from the Ohio State University.

Acknowledgement. This paper is based on Chapter Five of my Ph.D. thesis [Kim (1974)] at State University of New York at Buffalo. I am grateful to my thesis committee, F. Jen (Chairman), A. Chen, and the late H. Samuelsson. I would also like to thank J. Boness, M. Gordon, R. Hagerman, R. Hamada, E. Kane, S. Kon, J. McConnell, and Steward C. Meyers, a referee of this Journal, for their helpful comments.

1. For example, see Gordon (1962) and Solomon (1963).
explain the phenomenon of credit-rationing on the basis of rational economic self-interest by lenders. Miller has stated:

The substantial costs and delays normally incurred in case of default and the fact that compensating increases in rates actually increase the probability that these costs will be incurred makes the loan contact a relatively inefficient instrument. (1962, pp. 487-488)

Explicit treatment of bankruptcy costs in the theory of capital structure is limited. Kraus and Litzenberger (1973) provide a state-preference model with wealth taxes and bankruptcy costs, and suggest a stochastic dynamic programming approach to search for an optimal capital structure. In a recent paper which appeared after this study was completed, Scott (1976) shows that, if investors are indifferent to risk, imperfect markets for physical assets (along with a constant liquidation value of the firm’s assets in bankruptcy) imply the existence of an optimal capital structure. Although these studies provide insight into the theory of optimal capital structure, their models are either too complex to implement (Kraus and Litzenberger) or ignore risk-aversion in the capital market (Scott).

More importantly, these studies fail to recognize that, if there are bankruptcy costs, debt capacity, defined as the maximum amount of borrowing allowed by the capital market, occurs well before the point of one-hundred-percent debt financing. Thus, the existence of bankruptcy costs presents another issue that was best characterized by Myers and Pogue (1974, p. 589): It is unclear whether “the lenders chicken out first” (i.e., debt capacity occurs first) or “the shareholders chicken out first” (i.e., optimal capital structure occurs first). For example, if “the lenders chicken out first,” the optimal amount of borrowing would not be obtainable and the question of an optimal capital structure would become irrelevant. Or, if the optimal debt level coincides with the firm’s debt capacity (i.e., the shareholders and the lenders chicken out together), the implication is the same as that of the MM tax model—the firm should simply borrow as much as possible. It is only when the optimal amount of debt is strictly less than the debt capacity that firms must search for the optimal trade-off between the tax advantage of debt and the costs of bankruptcy. Therefore, a logical progression requires analysis of the problem of debt capacity before consideration of the question of optimal capital structure.

This paper examines the issues of debt capacity and optimal capital structure when firms are subject to stochastic bankruptcy costs and corporate income taxes (accompanied by a parallel analysis with wealth-taxes in footnotes) in the context of the Sharpe (1964)-Lintner(1965)-Mossin (1966) Capital Asset Pricing Model (CAPM). After discussing the nature and the magnitude of bankruptcy costs in Section II, we analyze the effects of bankruptcy costs and corporate income taxation on firm valuation in Section III. Section IV shows that when firms are subject to bankruptcy costs their debt capacities will be reached prior to one-hundred-percent debt financing. Section V makes it clear that optimal capital structures involve less debt financing than the maximum amount of borrowing allowed by the capital

5. For example, Miller (1962) has shown that when bankruptcy costs exist, lenders will impose credit-rationing before bankruptcy becomes certain. Although the Miller model is restricted to the individual lender’s behavior, aggregation of such credit-rationing behavior across all investors (lenders) in the capital market may imply the existence of “credit-rationing” for corporate borrowing.
market, and, hence, shareholder-wealth-maximizing firms will search for optimal capital structures rather than simply maximize their borrowing.

Although less general, the CAPM is more amenable to implementation than the state-preference approach, and in Section VI we show that it provides a much simpler method for approximating optimal capital structures than Kraus and Litzenberger's approach. Since, unlike Scott's risk indifference approach, the CAPM incorporates risk-aversion, our method is more realistic but no more complex than the method suggested by Scott for determining optimal leverage. A numerical example, employing the method suggested in this paper, illustrates that not only is a firm's value a strictly concave function of its end-of-period debt obligations with a unique global maximum, but also that the maximum is reached prior to the firm's debt capacity.

This paper also provides some interesting results in other related areas. An explicit pricing model that values corporate debt directly with or without bankruptcy costs is derived in Section VI. This model provides additional insights into the operations of bond markets. In Sections III and VI, we compare and contrast different tax schemes and their effects on firm valuation. The results imply that the tax advantage of debt financing assumes much less importance in financial structure decisions than generally suggested.

II. Bankruptcy Costs

The cost of bankruptcy can be thought of as comprising three major components. First, depending on whether bankruptcy takes the form of liquidation or reorganization, there may be either the "short-fall" arising from the liquidation or the "indirect" cost of reorganization. Second, arising in the course of the bankruptcy proceedings, various administrative expenses must be paid to third parties. Third, firms lose tax credits which they would have received had they not gone bankrupt.

If bankruptcy takes the form of liquidation, the first type of cost is the "short-fall" arising from the liquidation of physical assets below their economic values at "distress" prices. This is mainly due to the imperfection of secondary markets for physical assets. In a recent paper, Van Horne (1975) argues:

In a distress sale, a finished good frequently brings only 30 to 70 per cent of the wholesale price. Depending on market conditions, a fixed asset may bring even less. While most bankruptcy auctions are honest, the very nature of the process coupled with some shady practices do not augur well for the seller. (p. 15-16)

If bankruptcy takes the form of reorganization, the first type of cost is what Baxter (1967) describes as the "indirect" costs of bankruptcy. The indirect costs include reduction in future sales due to customers' doubts of the reliability of the bankrupt firm as a supplier; difficulty in obtaining trade credit; higher production costs due to dislocations within the company and renegotiation of contracts for employees; and the time lost by executives in the reorganization procedure. In either form of bankruptcy, these costs are difficult to document although they may be the most important component of bankruptcy costs. To our knowledge, no one has attempted to measure the magnitude of these costs.

Embodied in the total "short-fall" of the liquidation or the total "indirect" costs
of the reorganization are serious delays in bringing about the liquidation and the reorganization. For example, Warner (1976) reports that bankruptcies for 11 railroad companies took on average 12.5 years to settle in the courts. These delays impose additional costs to debtholders with proven claims.

The second type of cost involves fees and other compensation to third parties (i.e., the lawyers, trustees, auctioneers, referees, accountants, appraisers, etc.) and represents the administrative expenses of bankruptcy. Payment for these costs receives the highest priority in a bankruptcy proceeding under the Bankruptcy Act. In a Brookings Institution study, Stanley and Girth (1971) examine these bankruptcy costs in a large sample of case analyses and interviews. They estimate that total administrative expenses in a business bankruptcy approximate 20 per cent of the estate and that about half of these expenses go to attorneys. This inference is supported by bankruptcy statistics from the Administrative Office of the U.S. Courts which show that in fiscal 1969 total administrative expenses were 23.4 per cent of the total realization from bankruptcies. However, Warner's (1976) analysis of 11 bankrupt railroad companies reports that the administrative expenses are on average 5.3 per cent of the market value of the firm at the time of bankruptcy. Warner attributes this discrepancy to the fact that he deals with entities of greater dollar size than Stanley and Girth's sample. Warner's sample shows that there are economies of scale with respect to bankruptcy costs. Collectively, these studies suggest that while some administrative expenses, such as the attorney fees, may not decline on a relative basis as the size of the estate increases, there are fixed costs associated with the bankruptcy process that do so decline.

The third type of bankruptcy costs is due to the tax court's refusal to grant tax credits for the tax losses of a bankrupt firm. Even if tax laws were lenient in this regard, bankruptcy usually is the result of several successive years of unprofitable operations, which lessens the possibility of carrying back these tax losses against previously paid taxes. Thus, in order for the bankrupt firm to receive even a fraction of tax credits for its losses, either the firm must merge with a profitable firm or it must carry-forward its tax losses after the bankruptcy. Section 269 of the U.S. Internal Revenue Code of 1954 prohibits firms from merging for the sole purpose of taking advantage of the tax law. Also, sections 269, 371, and 382 of the U.S. Internal Revenue Code of 1954 provide formidable barriers to a post-reorganization firm from carrying forward the losses incurred by a bankrupt firm, and past court cases have not allowed carry-overs of such tax losses. Hence, it is most likely that the creditors of a bankrupt firm will lose the tax credits to which the firm would have been entitled had it not gone bankrupt.

7. For the breakdown of specific expenses, see either Stanley and Girth (1971) or Van Horne (1975).
8. The breakdown of specific expenses is reported in Van Horne (1975).
9. Van Horne pointed out to the author that, as the size of the estate increases, the economic incentives to hire lawyers for legal suits increase, and hence attorney's fees tend to increase.
10. For an extensive discussion on this subject, see Holzman (1955) and Tobolowsky (1960).
III. FIRM VALUATION

A. Assumptions and Definitions

To distinguish the effect of financial from investment decisions, we assume that the firm already has selected its investments but has not yet decided on how to finance them. The firm's operating earnings will be \( \tilde{X} - A \), and the residual earnings after deducting interest payments will be subject to an income tax rate \( T \).

If the firm chooses to finance the entire investment by equity alone, the market value of the firm will be \( V_u = S_u \), where \( S_u \) is the market value of the unlevered firm's equity. Its taxable earnings at the end of the period will equal operating earnings, \( \tilde{X} - A \). Thus, one-plus-the-rate-of-return on a dollar invested in the firm's equity will be:

\[
\tilde{R}_u = \frac{\tilde{X} - T(\tilde{X} - A)}{S_u} = \frac{(1 - T)\tilde{X} + AT}{S_u}
\]  

If the firm chooses to finance part of the investment by borrowing \( D \) dollars, the market value of the firm will be \( V_f = S_f + D \), where \( S_f \) is the market value of the levered firm's equity. The levered firm will be bankrupt if it fails to meet its obligations to debtholders (the principal \( D \) plus promised interest payments) at the end of the period. The promised interest rate, \( r_P \), will depend on the amount of borrowing, \( D \).

If the levered firm does not go bankrupt, its taxable earnings will be \( \tilde{X} - A - (\tilde{r} - 1)D \), and one-plus-the-rate-of-return on a dollar invested in its equity will be:

\[
\tilde{R}_f = \frac{(1 - T)(\tilde{X} - \tilde{r}D) + T(A - D)}{S_f}
\]

One-plus-the-rate-of-return on a dollar invested in the firm's debt will be \( \tilde{r} \).

Bankruptcy occurs if the firm's terminal value is less than its total end-of-period debt obligation, i.e., \( \tilde{X} < \tilde{r}D \). In the event of an actual bankruptcy, stockholders will exercise their limited liability, and ownership of the bankrupt firm will be transferred to debtholders. Debtholders will have to pay the cost of bankruptcy out of \( \tilde{X} \). Therefore, one-plus-the-rate-of-return on a dollar invested in the securities of

12. This assumption implies that there is no debt outstanding prior to the financial structure decision. If the firm has already issued some positive amount of debt, it is necessary to assume an effective "me-first" rule. For an extensive discussion of "me-first" rule, see Fama and Miller (1972) and Kim, McConnell, and Greenwood (1977).

13. For the sake of brevity, we do not distinguish "bankruptcy" from "default". Whereas "default" merely describes the borrower's action, bankruptcy often involves lender initiatives. The question of whether or not bankruptcy proceedings should be initiated by creditors when the firm defaults on its debt obligation and the optimal timing of such proceedings are examined in detail by Van Horne (1976).
the levered firm may be defined as:

\[
\tilde{R}_t = \begin{cases} 
(1 - T)(\tilde{X} - \hat{r}D) + T(A - D) / S_t & \text{if } \tilde{X} > \hat{r}D \\
0 & \text{if } \tilde{X} < \hat{r}D,
\end{cases}
\]  

(2)

and

\[
\tilde{r} = \begin{cases} 
\hat{r} & \text{if } \tilde{X} > \hat{r}D \\
(\tilde{X} - \tilde{B}) / D & \text{if } \tilde{X} < \hat{r}D,
\end{cases}
\]

(3)

where \(\tilde{R}_t\) and \(\tilde{r}\) are one-plus-the-rate-of-return on the levered firm’s stocks and debt, respectively, and \(\tilde{B}\) equals bankruptcy costs.

Because tax credits are not included in the gross return to debtholders of the bankrupt firm, the third type of bankruptcy costs (i.e., the loss of tax credits) is already implicit in (3). Thus, \(\tilde{B}\) in (3) represents only the sum of the first two types of bankruptcy costs identified in Section II, and may be expressed as:

\[
\tilde{B} = \begin{cases} 
0 & \text{if } \tilde{X} > \hat{r}D \\
B(\tilde{X}) & \text{if } \tilde{X} < \hat{r}D,
\end{cases}
\]

(4)

where \(B(\tilde{X})\) is an implicit positive function of \(\tilde{X}\), but is no greater than \(\tilde{X}\), i.e., \(B(\tilde{X}) < \tilde{X}\). This upper limit on \(B(\tilde{X})\) provides debtholders with limited liability in a nominal asset case in which the proceeds from liquidation are entirely consumed in administrative expenses with no distribution to debtholders.

Finally, we assume that risky securities are priced according to the CAPM such that the equilibrium expected return on any risky security \(i\) is:

\[
E(R_i) = RF + \lambda \text{cov}(\tilde{R}_i, \tilde{R}_m),
\]

(5)

where \(RF\) is one-plus-the-rate-of-return on the riskfree asset, \(\tilde{R}_m\) is one plus the value-weighted rate of return on all risky securities in the market, and has an expected value of \(E(\tilde{R}_m)\) and a standard deviation of \(\sigma_m\): \(\lambda = [E(\tilde{R}_m) - RF] / \sigma_m^2\) is the market price of risk; and \(\text{cov}(\tilde{R}_i, \tilde{R}_m)\) is the covariance between \(\tilde{R}_i\) and \(\tilde{R}_m\) and represents the systematic risk of security \(i\).

B. Market Value of the Firm

If we define the bankruptcy operator as:

\[
\tilde{b} = \begin{cases} 
0 & \text{if } \tilde{X} > \hat{r}D \\
1 & \text{if } \tilde{X} < \hat{r}D,
\end{cases}
\]

(6)

14. Since we assume that firms are subject to bankruptcy costs, all returns on risky securities and portfolios, including \(\tilde{R}_m\), are defined as returns after bankruptcy cost. With this definition, Kim (1974, pp. 100–104) shows that the exact form of the CAPM holds when firms are subject to bankruptcy costs.

15. The assumption of the existence of a risk-free security is not substantive, provided zero-beta portfolios exist in which case the symbol \(RF\) in this paper may be replaced at every point by \(E(\tilde{R}_Z)\), where \(Z\) is the minimum variance portfolio with returns uncorrelated with the market return.
the market value of the firm may be expressed as: \(^{16}\)

\[ V_t = V_u + TD(R_F - 1)/R_F - T(A - D)V(\tilde{b}) - (1 - T)V(\tilde{B}), \]  
where \( V(\tilde{b}) = [E(\tilde{b}) - \lambda \text{cov}(\tilde{b}, R_m)]/R_F \) is the risk-adjusted present value of one dollar associated with the occurrence of bankruptcy, and 
\( V(\tilde{B}) = [E(\tilde{B}) - \lambda \text{cov}(\tilde{B}, R_m)]/R_F \) is the risk-adjusted present value of the bankruptcy costs.

In the absence of taxes and bankruptcy costs (i.e., \( T = 0 \) and \( \tilde{B} = 0 \)), but with a positive probability of bankruptcy, (7) reduces to \( V_t = V_u \), i.e., the market value of the firm is independent of its financial structure. With a positive income\(^{17}\) tax rate and positive bankruptcy costs, (7) illustrates that the market value of the levered

16. By combining (2), (3), (4), and (6), we can rewrite (2) as:

\[ R_t = [(1 - T)(\tilde{X} - \tilde{D}) + T(A - D)](1 - T)(1 - T)\tilde{B}]/S_t. \]  
(2a)

[To show that (2) and (2a) are identical, one only has to substitute (3), (4), and (6) back into (2a).] Having obtained (1) and (2a), derivation of (7) is straightforward. Substituting (1) into both sides of (5) gives:

\[ (1 - T)E(\tilde{X}) = S_tR_F - AT + (1 - T)\lambda \text{cov}(\tilde{X}, R_m) \]  
(a)

Substituting (2a) into both sides of (5), and because \( E(\tilde{r}) = R_F + \lambda \text{cov}(\tilde{r}, R_m) \), it follows that:

\[ (1 - T)E(\tilde{X}) = [S_t + D(1 - T)]R_F - T(A - D) + (1 - T)\lambda \text{cov}(\tilde{X}, R_m) \]
\[ + T(A - D)[E(\tilde{b}) - \lambda \text{cov}(\tilde{b}, R_m)] + (1 - T)[E(\tilde{B}) - \lambda \text{cov}(\tilde{B}, R_m)]. \]  
(b)

From (a) and (b), and from the definitions that:

\[ V_t = S_t + D \]  
and \( V_u = S_u, \)

we can obtain (7). It should be noted that, although the CAPM is necessary to provide a simple, but practical method to approximate the optimal capital structure, one does not need to assume the CAPM merely to derive (7). It can be shown that, given (1) and (2a), a similar result also holds for more general theoretical models such as the State Preference approach or Schall's (1972) Value Additivity Principle.

17. The tax structure considered here is an income-tax-system in which the interest payments are tax deductible but principal payments are not. The alternative tax structure that has been considered in finance literature is a net-terminal-wealth-tax system in which both the interest and principal payments of corporate debt are tax deductible. [For example, see Rubinstein (1973) and Kraus and Litzenberger (1973).] Clearly, an income-tax structure is more realistic than a wealth-tax structure in view of the United States tax code.

Nevertheless, if one assumes a net-terminal-wealth-tax-system and lets \( T \) be the wealth-tax rate, one-plus-the-rate-of-return on the unlevered and the levered firms' equities should be redefined as follows:

\[ \tilde{R}_u = (1 - T)\tilde{X}/S_u, \]  
(1a)

and

\[ \tilde{R}_t = \begin{cases} 
(1 - T)(\tilde{X} - \tilde{D})/S_t & \text{if } \tilde{X} > \tilde{D} \\
0 & \text{if } \tilde{X} < \tilde{D}. 
\end{cases} \]  
(2b)

By following the steps that led to (7), it can be shown easily that the market value of the levered firm under a wealth-tax system is:

\[ V_t^w = V_u^w + TD - (1 - T)V(\tilde{B}), \]  
(7a)

where \( V_u^w \) is the market value of the unlevered firm under the wealth-tax system.
firm is equal to the market value of the unlevered firm plus the present value of tax
deductibility of interest payments, \( TD(R_p - 1)/R_F \),\(^{18}\) minus the present value of the
loss of tax credits in the event of bankruptcy, \( T(A - D)V(\hat{b}) \), and the fraction
\((1 - T)\) of the present value of bankruptcy costs, \( V(\hat{B}) \).\(^{19}\)

With corporate income taxes and bankruptcy costs, the ex-ante market value of
\( \hat{X} \), which depends only on the firm’s investment decisions, is divided among four
parties: stockholders, debtholders, the government, and bankruptcy costs. Since
\( V(\hat{B}) \) increases as the probability of bankruptcy increases and the probability of
bankruptcy increases as financial leverage increases, \( V(\hat{B}) \) will also increase as
financial leverage increases. On the other hand, debtholders have claims to the
future earnings of the firm that are prior to the government’s claim, and hence
\( V(G) \), the value of the government’s claim to the future earnings of the firm,
decreases with increased financial leverage. Therefore, as the firm’s financial
leverage increases, the increase in \( V(\hat{B}) \) will be offset by a decrease in \( V(G) \), and
the sum of \( V(\hat{B}) \) and \( V(G) \) will either increase or decrease depending on the
particular degree of financial leverage.

The part of the market value of the firm that belongs to the suppliers of the
capital is the difference between the market value of \( \hat{X} \), \( V(\hat{X}) \), and the sum of
\( V(\hat{B}) \) and \( V(G) \), i.e., \( V_{f} = V(\hat{X}) - [V(\hat{B}) + V(G)] \). Since \( V(\hat{X}) = [E(\hat{X}) - \lambda \text{cov}(\hat{X}, \hat{R}_{m})]/\hat{R}_{F} \)
is independent of financial structure, the financial structure that
minimizes the sum of \( V(G) \) and \( V(\hat{B}) \) will maximize \( V_{f} \) in (7).

IV. CORPORATE DEBT CAPACITY

Since an optimal capital structure is a meaningful concept only if it can be shown
that the optimal debt is strictly less than debt capacity, in this section we present a
formal analysis of corporate debt capacity. Corporate debt capacity is defined as
the maximum amount that a firm with given investments can borrow in a perfect
capital market. Corporate debt capacity is denoted by \( \hat{D} \), and \( \hat{r}D \) represents the
amount that the firm must promise its debtholders to reach \( \hat{D} \).

\(^{18}\) Since the availability of an interest tax shield is contingent upon the actual payment of the
promised interest, with a positive probability of bankruptcy (and default of interest) the interest tax
shield becomes risky. \( TD(R_p - 1)/R_F \) is the certainty equivalent of the risky tax shield for interest payment
(or the certain tax shield if there is no probability of bankruptcy). Discounting this certainty equivalent
at the risk-free rate yields the present value of the tax shields, \( TD(R_p - 1)/R_F \). This quantity is
substantially smaller than the present value of tax savings in the original MM tax model, which states
that \( V_{f} = V_{a} + TD \). While the debt in our model matures at the end of a single period with no
presumption about the levered and the unlevered firm’s future financial structures beyond a single
period, the MM tax model assumes that debt has no maturity in an infinite horizon framework and that
the present financial structure of the levered and the unlevered firm will be maintained forever. If we
assert this perpetuity assumption, the present value of tax deductibility of interest payments becomes
\( \sum_{t=0}^{\infty} TD(R_p - 1)/(R_F)^t = TD \). This is identical to the present value of tax subsidy in the MM tax model.

Note that the present value of tax subsidy in (7a) is also \( TD \). [See footnote 17.] Hence, a wealth-tax in
a single-period framework yields the same present value of tax savings as an income-tax in MM’s
infinite horizon framework does.

\(^{19}\) The remaining T percent of \( V(\hat{B}) \) is borne by the government, as the premium added to the
promised interest rate due to bankruptcy cost reduces the government’s claim to the future earnings of
the firm by T percent.
In a perfect capital market, unless the firm already has reached its debt capacity, it can borrow more by promising to pay more to its potential lenders at the end of the period. Once it has reached its debt capacity, by definition, it can borrow no more regardless of how much more it promises to pay at the end of the period. Mathematically, it means that \( (dD/d\bar{D}) > 0 \) for \( \bar{D} < \bar{D} \), and \( (dD/d\bar{D}) = 0 \) for \( \bar{D} = \bar{D} \). Since in a perfect capital market, investors will lend exactly what the promise for future payment is worth according to the market pricing mechanism, \( dD/d\bar{D} \) will equal zero only when any further increase in \( \bar{D} \) does not create any additional value for lenders.

Such a debt capacity occurs before bankruptcy becomes certain if there are bankruptcy costs. Although Miller (1962) assumes that lenders display risk aversion when he utilizes bankruptcy costs to explain the phenomenon of credit-rationing, it is not risk aversion itself that causes occurrence of debt capacity before bankruptcy becomes certain. It occurs because: (1) the present value of bankruptcy costs, \( V(\bar{B}) \), increases as \( \bar{D} \) increases and (2) in the event of bankruptcy, the claims of bankruptcy costs to \( X \) must be satisfied prior to the claims of debtholders. These points can be demonstrated by assuming that lenders are risk-neutral. By assuming risk-neutrality we can distinguish the effect of bankruptcy costs from the effect of risk aversion. Then the market value of \( \bar{D} \) is simply the expected return to debtholders discounted at \( R_F \):\(^{22}\)

\[
D = E(\bar{D}) / R_F
\]  
(8)

Substituting (3) and (4) into (8) gives:

\[
D = \left\{ \bar{D} \left[ 1 - F(\bar{D}) \right] + \int_{-\infty}^{\bar{D}} \bar{X} f(\bar{X}) d\bar{X} - \int_{-\infty}^{\bar{D}} B(\bar{X}) f(\bar{X}) d\bar{X} \right\} / R_F
\]  
(9)

where \( f(\bar{X}) \) is the probability density of \( \bar{X} \) with an expected value of \( E(\bar{X}) \) and a standard deviation of \( \sigma \).

\( F(\bar{D}) = \int_{-\infty}^{\bar{D}} f(\bar{X}) d\bar{X} \) is the probability that the firm will be bankrupt at the end of the period.

Differentiating (9) with respect to \( \bar{D} \) yields:

\[
\frac{dD}{d\bar{D}} = \left[ 1 - F(\bar{D}) - B(\bar{D}) f(\bar{D}) \right] / R_F.
\]  
(10)

20. If bankruptcies are costless, corporate debt capacity is not an operative term in a perfect capital market, as the firm's maximum borrowing is reached only when bankruptcy becomes certain.

21. Miller addresses the issue of credit-rationing by individual lenders for personal borrowing rather than the phenomenon of credit-rationing by the capital market for corporate borrowing. However, his results have important implications for corporate borrowing, because it does not rely on the legal constraints of maximum contract rates of interest. Although credit-rationing for personal borrowing may be explained through the maximum ceiling on contractual rates of interest, credit-rationing for corporate borrowing may not. Due to the discounting of corporate debt securities in the capital market, there exists no practical ceiling on the maximum contractual rates of interest on corporate borrowing.

22. A more complete derivation of the market value of corporate debt within the context of the CAPM is provided in Section VI.
(10) illustrates clearly that any change in $\hat{r}D$ has both a positive and a negative effect on lender's expected return. On the one hand, an increase in $\hat{r}D$ means an increase in expected total return. This increment in expected return is represented by the term $1 - F(\hat{r}D)$, which is the probability that the firm will not be bankrupt. On the other hand, an increase in $\hat{r}D$ means an increase in expected bankruptcy costs, because a higher $\hat{r}D$ means a higher probability of bankruptcy. This increment in expected bankruptcy costs is represented by the term $B(\hat{r}D)\tilde{f}(\hat{r}D)$.

Since the probability of bankruptcy, $F(\hat{r}D)$, increases as $\hat{r}D$ increases, the first term in (10), $1 - F(\hat{r}D)$, decreases as $\hat{r}D$ increases. $B(\hat{r}D)$ varies directly with $\hat{r}D$ because $B(\tilde{X})$ is a positive function of $\tilde{X}$ [See (4)]. If $\tilde{X}$ is normally distributed, $1 - F(\hat{r}D)$ is equal to $B(\hat{r}D)\tilde{f}(\hat{r}D)$ before the probability of bankruptcy reaches one. That is, there exists an $\hat{r}\tilde{D}$ at which $(dD/d\hat{r}D) = 0$ and $F(\hat{r}D) < 1$. This result is shown in Appendix A. Appendix A also shows that the second-order condition is met at $\hat{r}\tilde{D}$. Therefore, the firm's borrowing reaches its maximum (the firm's debt capacity) while bankruptcy remains uncertain.

Appendix A also shows that $(dD/d\hat{r}D) < 0$ for $\hat{r}D > \hat{r}\tilde{D}$, which means that firms will face a decreasing D while $\hat{r}D$ is increasing once they reach their debt capacities. The intuitive explanation is as follows: Although debtholders bear the ex-post cost of bankruptcy [See (3)], in a perfect capital market lenders pass-on the entire ex-ante cost of bankruptcy, $V(\tilde{B})$, to stockholders and the government in the form of higher promised interest rates. However, when the firm increases its $\hat{r}D$ beyond $\hat{r}\tilde{D}$, the incremental increase in $V(\tilde{B})$ will dominate the incremental increase in lenders' expected returns. Consequently, lenders cannot pass-on the entire $V(\tilde{B})$ unless they reduce the total amount of funds lent, $D$. In a perfect capital market, perfect substitutes are always available. Hence, lenders will simply switch to other corporate debt securities which allow them to pass-on the entire $V(\tilde{B})$ to stockholders and the government and earn the market equilibrium return. These switches will reduce the demand for the firm's $\hat{r}D$, which in turn will force the market value of $\hat{r}D$ to decline. That is, $D$ will decrease as $\hat{r}D$ exceeds $\hat{r}\tilde{D}$.

If there is any chance of solvency at the end of the period, the market value of common equity at the beginning of the period will be positive and $D/V_i < 1$. Since debt capacity occurs while bankruptcy remains uncertain, the firm's financial structure at its debt capacity will be less than one-hundred percent debt financing.

V. OPTIMAL FINANCIAL STRUCTURE

Differentiating (7) with respect to $\hat{r}D$ gives

\[
\frac{dV_i}{d\hat{r}D} = \left[ T(R_F - 1)/R_F \right] \frac{dD}{d\hat{r}D} + TV(\tilde{b}) \frac{dD}{d\hat{r}D} - T(A - D) \frac{dV(\tilde{B})}{d\hat{r}D} - (1 - T) \frac{dV(\tilde{B})}{d\hat{r}D}.
\]

(11)

23. For example, $B(\hat{r}D) = C + c\hat{r}D$ if we assume that the costs of bankruptcy are the sum of a fixed charge of $C$ dollars and a variable charge equal to a fraction $c < 1$ of $X$ (i.e., $B(X) = C + cX$).
(11) illustrates that any change in \( r_D \) has both a positive and a negative effect on market value of the firm. On the one hand, an increase in \( r_D \) means an increase in the present value of tax savings (PVTS). On the other hand, it means an increase in the present value of bankruptcy costs (PVBC).

When \( r_D \) approaches the minimum possible value of \( \tilde{X} \), \( \frac{dD}{dr_D} \rightarrow 0 \), \( V(\tilde{b}) \rightarrow 0 \), \( \frac{dV(\tilde{b})}{dr_D} \rightarrow 0 \), and \( \frac{dV(\tilde{B})}{dr_D} \rightarrow 0 \); hence \( \frac{dV}{dr_D} \) would be strictly positive. That is, an increase in \( r_D \) means a greater increase in PVTS than in PVBC, and hence the market value of the firm increases.

When \( r_D \) is equal to \( r_D^* \), \( \frac{dD}{dr_D} = 0 \), and hence \( \frac{dV}{dr_D} = \frac{1}{RF} \left[ T(A-D) \frac{dV(\tilde{b})}{dr_D} + (1-T) \frac{dV(\tilde{B})}{dr_D} \right] \). If bankruptcy remains uncertain, an increase in \( \frac{dD}{dr_D} \) increases the probability of bankruptcy, and thus both \( V(\tilde{b}) \) and \( V(\tilde{B}) \) should increase as \( r_D \) increases, i.e., \( \frac{dV(\tilde{b})}{dr_D} > 0 \) and \( \frac{dV(\tilde{B})}{dr_D} > 0 \).24 Since bankruptcy remains uncertain at the firm's debt capacity, \( \frac{dV}{dr_D} \) should be strictly negative.25 That is, at debt capacity an incremental \( r_D \) means a positive increment in PVBC but a zero increment in PVTS (i.e., the first term in (11) is zero when \( \frac{dD}{dr_D} = 0 \)), and hence the market value of the firm will decline.

With \( V_\dagger \) rising when \( r_D \) is small, and \( V_\dagger \) falling when \( r_D \) is large, there must be an \( r_D \), say \( r_D^* \), at which \( \frac{dV}{dr_D} \) equals zero and \( V_\dagger \) attains a maximum, \( V^* \).

Furthermore, since \( \frac{dV}{dr_D} \) is strictly negative at \( r_D^* \), \( r_D^* \) should be strictly less than \( r_D^* \). \( r_D^* \) is the optimal end-of-period amount that the firm should promise to pay its debtholders in order to maximize \( V_\dagger \). Likewise, \( D^* \), the amount borrowed by promising \( r_D^* \), is the optimal amount to borrow. Since \( r_D^* \) is less than \( r_D^* \), \( D^* \) should also be less than \( \tilde{D} \). That is, the optimal capital structure involves less debt financing than the firm's debt capacity.

Therefore, a shareholder-wealth-maximizing firm will not maximize its borrowing. Instead, it will search for its optimal capital structure to attain \( V^* \). By setting (11) equal to zero, we obtain:26

\[
\left[ \frac{R_F - 1}{RF} + V(\tilde{b}) \right] \frac{dD}{dr_D} = \left( A - D \right) \frac{dV(\tilde{b})}{dr_D} + \frac{1-T}{T} \frac{dV(\tilde{B})}{dr_D}.
\] (12)

24. If bankruptcy is certain, \( V(\tilde{b}) = 1/RF \) and \( V(\tilde{B}) = \left( E[B(\tilde{X})] - \lambda \text{cov}[B(\tilde{X}), R_p] \right) / RF_p \), thus, any further increase in \( r_D \) will not affect \( V(\tilde{b}) \) and \( V(\tilde{B}) \), i.e., \( \frac{dV(\tilde{b})}{dr_D} = 0 \) and \( \frac{dV(\tilde{B})}{dr_D} = 0 \).

25. If a net-terminal-wealth-tax-structure is assumed and if \( T \) represents the tax rate, differentiation of (7a) [See footnote 17] gives:

\[
\frac{dV^*}{dr_D} = T \frac{dD}{dr_D} - (1-T) \frac{dV(\tilde{B})}{dr_D},
\] (11a)
which is also negative at the firm's debt capacity.

26. Under the wealth-tax-system, the optimal capital structure is the \( r_D \) which satisfies:

\[
\frac{dD}{dr_D} = \frac{1-T}{T} \frac{dV(\tilde{B})}{dr_D}
\] (12a)
without making both sides of (12a) be equal to zero.
The optimal capital structure is the \( \hat{r}D \) which satisfies (12) without making both sides of (12) be equal to zero.\(^{27}\)

VI. NORMAL DISTRIBUTION

We have shown that in a perfect capital market where firms are subject to corporate income taxes and costly bankruptcies, debt capacity occurs at less than one-hundred-percent debt financing and the optimal capital structure occurs before debt capacity. In this section we derive an explicit valuation model for risky debt with or without bankruptcy costs, and develop a simple method to approximate the optimal capital structure. We assume that (1) \( \tilde{X} \) and \( \tilde{R}_m \) are normally distributed; and (2) \( B(\tilde{X}) \) in (4) is the sum of a fixed charge of \( C \) dollars and a variable charge equal to a fraction \( c < 1 \) of \( \tilde{X} \).\(^{28}\)

\[
B(\tilde{X}) = C + c\tilde{X} 
\]

Substituting (6) into (3) gives:

\[
\hat{r}D = \hat{r}D(1 - \tilde{b}) + \tilde{b}\tilde{X} - \tilde{B}. 
\]

Substituting (14) into both sides of (5), and from \( E(\tilde{b}) = F(\hat{r}D) \) and \( E(\tilde{b}\tilde{X}) = \int_{-\infty}^{\hat{r}D} \tilde{X} f(\tilde{X}) d\tilde{X} \), it follows that:

\[
D = \left\{ \hat{r}D \left[ 1 - F(\hat{r}D) \right] + \int_{-\infty}^{\hat{r}D} \tilde{X} f(\tilde{X}) d\tilde{X} + \lambda \left[ \hat{r}D \text{cov}(\tilde{b}, \tilde{R}_m) 
- \text{cov}(\tilde{b}\tilde{X}, \tilde{R}_m) \right] \right\} / R_F - V(\tilde{B}). \hspace{1cm} (15)
\]

Substituting (B-1), (B-3), and (B-4) of Appendix B into (15) gives:

\[
D = \left[ E(\hat{r}D^0) - \lambda \text{cov}(\tilde{X}, \tilde{R}_m) F(\hat{r}D) \right] / R_F - V(\tilde{B}), \hspace{1cm} (16)
\]

where \( E(\hat{r}D^0) = \hat{r}D \left[ 1 - F(\hat{r}D) \right] + E(\tilde{X}) F(\hat{r}D) - \sigma^2 f(\hat{r}D) \) is the expected gross dollar return to debtholders in the absence of bankruptcy costs.

If there are no bankruptcy costs (i.e., \( \tilde{B} = 0 \)), (16) provides an explicit pricing model for corporate debt. It also states explicitly that the risk premium on corporate debt equals the borrowing firm’s operating risk premium, \( \lambda \text{cov}(\tilde{X}, \tilde{R}_m) \), multiplied by the

\(^{27}\) If bankruptcy is certain, both sides of (12) are equal to zero, but the firm’s total market value will be strictly less than \( V^* \). This is because \( dV_t / d\hat{r}D \) stays negative throughout all \( \hat{r}D > \hat{r}D^* \) until bankruptcy becomes certain: First, we know from the text that \( dV_t / d\hat{r}D < 0 \) for \( \hat{r}D^* < \hat{r}D < \hat{r}D^\circ \). Second, it also has been shown that, for \( \hat{r}D > \hat{r}D^\circ \), \( d\tilde{D}_t / d\hat{r}D < 0 \) and both \( d\tilde{V}(\hat{b}) / d\hat{r}D \) and \( d\tilde{V}(\hat{B}) / d\hat{r}D \) are positive, and hence, from (11), \( (d\tilde{V}_t / d\hat{r}D) < 0 \) until bankruptcy becomes certain.

\(^{28}\) To be more precise, the bankruptcy cost may be expressed as:

\[
B(\tilde{X}) = \begin{cases} 
C + c\tilde{X} & \text{if } C / (1 - c) < \tilde{X} < \hat{r}D \\
\tilde{X} & \text{if } \tilde{X} < C / (1 - c),
\end{cases}
\]

which provides the debtholders of a bankrupt firm with limited liability. However, for the sake of brevity we assume (13).
firm's probability of bankruptcy, \(F(\hat{D})\).\(^{29}\) The explanation for this measure of risk premium is straightforward. To the extent the firm is not bankrupt, debtholders receive a fixed amount \(\hat{D}\) which has no systematic relationship with the market. If the firm is bankrupt, debtholders receive \(\bar{X}\) which has a systematic risk of \(\text{cov}(\bar{X}, \bar{R}_m)\) that cannot be diversified away in debtholders' portfolios. Since this systematic risk is relevant to debtholders only to the extent that the firm is bankrupt, the relevant risk is the firm's total systematic risk multiplied by the probability of bankruptcy.

With positive bankruptcy costs, (16) shows that market value of risky debt is simply what it would have been in the absence of bankruptcy costs minus the present value of bankruptcy costs. In other words, lenders deduct the ex-ante costs of bankruptcy from the amount they would have lent had there been no bankruptcy costs. Substituting (6) and (13) into (4) gives \(\tilde{b} = \tilde{b} \tilde{b}\), which can be substituted into the definition of \(V(\tilde{b})\) in (7) to obtain:

\[
V(\tilde{b}) = \left\{ CF(\hat{D}) + c \left[ E(\bar{X}) F(\hat{D}) - \sigma^2 f(\hat{D}) \right] + \lambda \text{cov}(\bar{X}, \bar{R}_m) \left[ (C + c\hat{D})f(\hat{D}) - \sigma^2 f(\hat{D}) \right] \right\} / R_F. \tag{17}
\]

By substituting (17) into (16), we obtain the formal expression for the pricing of corporate debt with linear bankruptcy costs:

\[
D = \left\{ \hat{D} \left[ 1 - F(\hat{D}) \right] + (1 - c) \left[ E(\bar{X}) F(\hat{D}) - \sigma^2 f(\hat{D}) \right] - CF(\hat{D}) \right. \\
- \lambda \text{cov}(\bar{X}, \bar{R}_m) \left[ (C + c\hat{D})f(\hat{D}) + (1 - c) F(\hat{D}) \right] \right\} / R_F. \tag{18}
\]

Finally, to provide a method that can be used to search for the optimal capital structure, we solve for the remaining terms in equations (7) and (12). Substituting (B-3) of Appendix B into the definition of \(V(\tilde{b})\) in (7) yields:

\[
V(\tilde{b}) = \left\{ F(\hat{D}) + \lambda \text{cov}(\bar{X}, \bar{R}_m) f(\hat{D}) \right\} / R_F. \tag{19}
\]

Differentiating (17), (18), and (19) with respect to \(\hat{D}\) yields:

\[
\frac{dV(\tilde{b})}{d\hat{D}} = \left( C + c\hat{D} \right) \left[ 1 - \lambda \text{cov}(\bar{X}, \bar{R}_m) \frac{\hat{D} - E(\bar{X})}{\sigma^2} \right] f(\hat{D}) / R_F. \tag{20}
\]

\[
\frac{dD}{d\hat{D}} = \left\{ 1 - F(\hat{D}) - f(\hat{D}) \right\} \left( C + c\hat{D} \right) \left( 1 - \frac{\hat{D} - E(\bar{X})}{\sigma^2} \lambda \text{cov}(\bar{X}, \bar{R}_m) \right) \\
+ \lambda \text{cov}(\bar{X}, \bar{R}_m) \right\} / R_F, \quad \text{and} \tag{21}
\]

\[
\frac{dV(\tilde{b})}{d\hat{D}} = f(\hat{D}) \left[ 1 - \lambda \text{cov}(\bar{X}, \bar{R}_m) \frac{\hat{D} - E(\bar{X})}{\sigma^2} \right] / R_F. \tag{22}
\]

\(^{29}\) As the referee pointed out, this result may not depend on the assumption of normal distribution.
Note that all of the values in (17) through (22) are expressed in terms of \( f(\hat{r}D) \) and \( F(\hat{r}D) \). Thus, if we know the firm-specific and the market-wide parameters, all of the values in (17) through (22) for a given \( \hat{r}D \) can be found by using tables of ordinates and areas of the normal distribution. By trial and error, we can find the \( \hat{r}D^* \) which satisfies (12) without making both sides of (12) be equal to zero. The corresponding \( D^* \) and \( V^* \) can be obtained from (18) and (7).

For a hypothetical bi-variate distribution of \( X \) and \( R_m \), Table 1 contains the numerical values of \( V(\hat{b}) \), \( V(\hat{B}) \), \( D \), \( \hat{r} \), \( V_t \), \( D/V_t \), \( V_t^w \), \( D/V_t^w \) for various levels of \( \hat{r} \). Figure 1 depicts the behavior of \( D \), \( V_t \), and \( V_t^w \) as a function of \( \hat{r}D \).

From Table 1 and Figure 1 we see, first, that corporate debt capacity occurs prior to a one-hundred percent debt ratio, and any increase in \( \hat{r}D \) beyond \( \hat{r}D \) only decreases the firm’s borrowing. Hence, from the firm’s perspective, \( \hat{r}D > \hat{r}D \) is inferior to \( \hat{r}D < \hat{r}D \) and \( D > D \) is infeasible.

Second, regardless of whether one assumes income-taxes or wealth-taxes, the firm’s value is a strictly concave function of its end-of-period debt obligation with a unique global maximum, and the maximum is reached before debt capacity.

### TABLE 1

**Valuation Equations** (19), (17), (18), (7), and (7a) for a Hypothetical Bi-Variate Normal Distribution; \( E(X) = 1,300,000 \), \( \sigma_x = 300,000 \), \( E(R_m) = 1.15 \), \( \sigma_m = .20 \), \( p_{X,R_m} = .5 \), \( A = 1,113,636 \), \( C = 0 \), \( c = .4 \), \( T = .5 \), and \( R_F = 1.05 \).

<table>
<thead>
<tr>
<th>( z^a )</th>
<th>( \hat{r}D )</th>
<th>( F(z)^b )</th>
<th>( f(z)^c )</th>
<th>( V(\hat{b}) )</th>
<th>( V(\hat{B}) )</th>
<th>( D )</th>
<th>( \hat{r}D^d )</th>
<th>( V_e^e )</th>
<th>( D/V_t^e )</th>
<th>( V_t^{w^f} )</th>
<th>( D/V_t^{w^f} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>−4.33</td>
<td>0</td>
<td>.0000</td>
<td>.0000</td>
<td>0</td>
<td>0</td>
<td>1,113,636</td>
<td>.0000</td>
<td>583,333</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>−3.5</td>
<td>250,000</td>
<td>.0002</td>
<td>.0009</td>
<td>.0004</td>
<td>12</td>
<td>238,012</td>
<td>1.05</td>
<td>1,119,120</td>
<td>.2127</td>
<td>702,333</td>
<td>.3389</td>
</tr>
<tr>
<td>−3.0</td>
<td>400,000</td>
<td>.0013</td>
<td>.0044</td>
<td>.0023</td>
<td>271</td>
<td>380,445</td>
<td>1.05</td>
<td>1,121,712</td>
<td>.3392</td>
<td>773,420</td>
<td>.4919</td>
</tr>
<tr>
<td>−2.68</td>
<td>496,000</td>
<td>.0037</td>
<td>.0110</td>
<td>.0061</td>
<td>989</td>
<td>470,818</td>
<td>1.05</td>
<td>1,122,386</td>
<td>.4195</td>
<td>818,248</td>
<td>.5754</td>
</tr>
<tr>
<td>−2.5</td>
<td>550,000</td>
<td>.0062</td>
<td>.0175</td>
<td>.0091</td>
<td>1,810</td>
<td>520,985</td>
<td>1.06</td>
<td>1,122,138</td>
<td>.4643</td>
<td>842,921</td>
<td>.6181</td>
</tr>
<tr>
<td>−2.0</td>
<td>700,000</td>
<td>.0228</td>
<td>.0540</td>
<td>.0346</td>
<td>8,069</td>
<td>654,569</td>
<td>1.07</td>
<td>1,117,238</td>
<td>.5859</td>
<td>806,583</td>
<td>.7220</td>
</tr>
<tr>
<td>−1.5</td>
<td>850,000</td>
<td>.0668</td>
<td>.1295</td>
<td>.0945</td>
<td>26,857</td>
<td>769,524</td>
<td>1.10</td>
<td>1,102,263</td>
<td>.6981</td>
<td>954,667</td>
<td>.8061</td>
</tr>
<tr>
<td>−1.0</td>
<td>994,000</td>
<td>.1539</td>
<td>.2371</td>
<td>.2030</td>
<td>67,169</td>
<td>845,613</td>
<td>1.16</td>
<td>1,072,973</td>
<td>.7901</td>
<td>972,555</td>
<td>.8695</td>
</tr>
<tr>
<td>−1.0</td>
<td>1,000,000</td>
<td>.1587</td>
<td>.2420</td>
<td>.2088</td>
<td>69,450</td>
<td>847,795</td>
<td>1.18</td>
<td>1,071,335</td>
<td>.7913</td>
<td>972,506</td>
<td>.8718</td>
</tr>
<tr>
<td>−.5</td>
<td>1,150,000</td>
<td>.3085</td>
<td>.3521</td>
<td>.3776</td>
<td>142,290</td>
<td>874,384</td>
<td>1.32</td>
<td>1,018,131</td>
<td>.8588</td>
<td>949,380</td>
<td>.9210</td>
</tr>
<tr>
<td>0</td>
<td>1,300,000</td>
<td>.5000</td>
<td>.3989</td>
<td>.5712</td>
<td>237,132</td>
<td>851,277</td>
<td>1.53</td>
<td>940,401</td>
<td>.9052</td>
<td>890,406</td>
<td>.9561</td>
</tr>
<tr>
<td>.5</td>
<td>1,450,000</td>
<td>.6915</td>
<td>.7424</td>
<td>.331,083</td>
<td>801,091</td>
<td>1.81</td>
<td>851,144</td>
<td>.9412</td>
<td>818,337</td>
<td>.9789</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1,600,000</td>
<td>.8413</td>
<td>.2420</td>
<td>.8589</td>
<td>401,826</td>
<td>752,376</td>
<td>2.13</td>
<td>775,486</td>
<td>.9702</td>
<td>758,608</td>
<td>.9918</td>
</tr>
<tr>
<td>1.5</td>
<td>1,750,000</td>
<td>.9332</td>
<td>.3521</td>
<td>.7424</td>
<td>442,277</td>
<td>720,789</td>
<td>2.43</td>
<td>729,021</td>
<td>.9887</td>
<td>722,589</td>
<td>.9975</td>
</tr>
<tr>
<td>2.0</td>
<td>1,900,000</td>
<td>.9772</td>
<td>.5040</td>
<td>.9435</td>
<td>459,627</td>
<td>706,268</td>
<td>2.69</td>
<td>708,456</td>
<td>.9969</td>
<td>706,654</td>
<td>.9995</td>
</tr>
<tr>
<td>2.5</td>
<td>2,050,000</td>
<td>.9938</td>
<td>.0175</td>
<td>.9506</td>
<td>465,190</td>
<td>701,348</td>
<td>2.92</td>
<td>701,773</td>
<td>.9994</td>
<td>701,412</td>
<td>.9999</td>
</tr>
<tr>
<td>3.0</td>
<td>2,200,000</td>
<td>.9987</td>
<td>.0044</td>
<td>.9522</td>
<td>466,479</td>
<td>700,137</td>
<td>3.14</td>
<td>700,193</td>
<td>.9999</td>
<td>700,162</td>
<td>1.0000</td>
</tr>
<tr>
<td>3.5</td>
<td>2,350,000</td>
<td>.9998</td>
<td>.0009</td>
<td>.9524</td>
<td>466,672</td>
<td>699,952</td>
<td>3.36</td>
<td>699,968</td>
<td>1.0000</td>
<td>699,973</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

---

*a* \( z = [\hat{r}D - E(\bar{X})]/\sigma \) represents the standardized \( \hat{r}D \).

*b* \( F(z) = F(\hat{r}D) \) is the probability of bankruptcy.

*c* \( f(z) = \sigma f(\hat{r}D) \) represents the standardized normal density.

*d* \( \hat{r} = (\hat{r}D)/D \) is one plus the promised interest rate.

*e* Note that \( V_e = [(1 - T)E(\bar{X}) + A T - \lambda(1 - T) \text{cov}(\bar{X}, \bar{R}_m)]/R_F \).

*f* See footnotes 17, 25, and 26, and note that \( V_t^{w} = [(1 - T)E(\bar{X}) - \lambda(1 - T) \text{cov}(\bar{X}, \bar{R}_m)]/R_F \).

Third, while with wealth-taxes the difference between the market value of the firm at its optimal capital structure and the market value of the unlevered firm (i.e., \( V^w - V^u \)) is significantly large, the counterpart with income taxes (i.e., \( V^* - V^u \)) is rather insignificant. While both interest and principal payments of corporate debt are tax-deductible under wealth-taxes, only interest payments are tax-deductible under income-taxes. Hence, the tax-advantage of debt financing is much greater with wealth-taxes than with income-taxes, which in turn, implies a much steeper slope for \( V^w \) than \( V^* \) before they reach maximum values at the optimal capital structure. On the other hand, government's share of the firm is greater with wealth-taxes than with income-taxes, and thus the initial value for \( V^w \) (i.e., \( V^u \)) is smaller than the initial value for \( V^* \) (i.e., \( V^u \)).

These two alternative tax structures represent the two opposite extreme cases.

![Figure 1. \( D, V^*, V^w \), and \( V^w \) as a function of \( \delta_D \) for a hypothetical bi-variate normal distribution; \( E(X) = 1,300,000, \sigma_x = 300,000, E(R_m) = 1.15, \sigma_m = .20, \rho_{x,m} = .5, A = 1,113,636, C = 0, c = A, T = .5, \) and \( R_F = 1.05 \)](image)
With income-taxes the market value of the tax subsidy derives from the tax-
deductibility of only the one-period interest payment; but with wealth-taxes the 
market value of the tax subsidy is equivalent to the value associated with perpetual 
interest payments under an income-tax structure [See footnote 18]. Therefore, these 
two alternative tax cases in a single period framework provide the upper and lower 
bounds for the market value of the tax subsidy and hence for the optimal capital 
structure in a multi-period world with an income-tax.

VII. Conclusion

This paper has shown that, in a perfect capital market where firms are subject to 
income taxes and costly bankruptcies, debt capacity occurs at less than one-
hundred-percent debt financing and firms do have optimal capital structures which 
involve less debt financing than their debt capacities. The market value of the firm 
increases for low levels of debt and decreases as financial leverage becomes 
more extreme. With linear bankruptcy costs, a simple method to approximate the 
optimal capital structure was derived. A numerical example using this method 
shows that the market value of the firm is a strictly concave function of its total 
end-of-period debt obligations with a unique global maximum.

This is essentially the same as the traditionalist's position on the relationship 
between the value of the firm and corporate financial leverage. We have shown 
that the traditionalist's argument follows from the MM logic by allowing for the 
existence of corporate income taxes and bankruptcy cost. However, there are 
fundamental differences between the approach taken in this paper and the 
traditionalist's approach to the valuation of firms: While the traditionalist's 
approach is based on the notion that valuation of firms can be explained by 
considering securities in isolation from the rest of the capital market, our con-
clusions are derived within a theoretical framework based on capital market 
equilibrium.

Appendix A

Corporate Debt Capacity With Bankruptcy Costs and Normally Distributed \( \hat{X} \)

To prove that corporate debt capacity is reached before bankruptcy becomes 
certain, we must show that there exists a finite \( \hat{D} \) at which (10) equals zero (the 
first-order condition) and the market value of debt, \( D \), is at its maximum (the 
second-order condition).

Since at very low \( \hat{D} \), \( (dD/d\hat{D})\approx 1/R_F > 0 \), if \( (dD/d\hat{D})<0 \) for a large finite 
\( \hat{D} \), (10) must equal zero at a finite \( \hat{D} \). As \( \hat{D} \) increases, \( B(\hat{D}) \) increases and 
\( \sigma^2/(\hat{D} - E(\hat{X})) \) decreases. Hence, \( B(\hat{D}) > \sigma^2/(\hat{D} - E(\hat{X})) \) for a sufficiently large 
\( \hat{D} \). But for a normally distributed random variable, we know the following 
inequality [Feller (1968, p. 175)]:

\[
\frac{\sigma^2}{\hat{D} - E(\hat{X})} f(\hat{D}) > 1 - F(\hat{D}), \quad \text{for } \hat{D} > E(\hat{X}). \quad (A-1)
\]
Thus, there must exist a finite $\hat{r}_D$ at which:

$$B(\hat{r}_D)f(\hat{r}_D) > \frac{\sigma^2}{\hat{r}_D - E(\bar{X})} f(\hat{r}_D) > 1 - F(\hat{r}_D)$$

such that $\frac{dD}{d\hat{r}_D} < 0$.

To show that the second-order condition is satisfied, we differentiate (10) with respect to $\hat{r}_D$. Noting that $(df(\hat{r}_D)/d\hat{r}_D) = -((\hat{r}_D - E(\bar{X}))/\sigma^2) f(\hat{r}_D)$ for a normal distribution,

$$\frac{d^2D}{d\hat{r}_D^2} = - \left[ f(\hat{r}_D) - \frac{\hat{r}_D - E(\bar{X})}{\sigma^2} B(\hat{r}_D)f(\hat{r}_D) + \frac{dB(\hat{r}_D)}{d\hat{r}_D} f(\hat{r}_D) \right] / R_F \quad (A-2)$$

Since $(dB(\hat{r}_D)/d\hat{r}_D) > 0$ [See (4)], the second derivative is clearly negative if $\hat{r}_D < E(\bar{X})$. Substituting the first-order condition, $B(\hat{r}_D)f(\hat{r}_D) = 1 - F(\hat{r}_D)$, into (A-2) yields:

$$\left. \frac{d^2D}{d\hat{r}_D^2} \right|_{\hat{r}_D = \hat{r}_D} = - \left[ f(\hat{r}_D) - \frac{\hat{r}_D - E(\bar{X})}{\sigma^2} \left[ 1 - F(\hat{r}_D) \right] + \frac{dB(\hat{r}_D)}{d\hat{r}_D} f(\hat{r}_D) \right] \quad (A-3)$$

If $\hat{r}_D > E(\bar{X})$, (A-3) is also negative because of (A-1). Thus, $D$ is maximized at $\hat{r}_D$.

**APPENDIX B**

**Determination of Partial Means and Partial Covariances When $\bar{X}$ and $\tilde{R}_m$ are Normally Distributed**

From Winkler, Roodman, and Britney's equation (3.4) (1972, p. 294), we obtain the partial mean of $\bar{X}$,

$$\int_{-\infty}^{\hat{r}_D} \tilde{x} f(\tilde{x}) d\tilde{x} = E(\bar{X})F(\hat{r}_D) - \sigma^2 f(\hat{r}_D). \quad (B-1)$$

The covariance between $\tilde{b}$ and $\tilde{R}_m$,

$$\text{cov}(\tilde{b}, \tilde{R}_m) = E(\tilde{b}\tilde{R}_m) - E(\tilde{b})E(\tilde{R}_m)$$

$$= \int_{-\infty}^{\hat{r}_D} f(\tilde{x}) \left[ \int_{-\infty}^{\infty} \tilde{R}_m g(\tilde{R}_m | X) d\tilde{R}_m - E(\tilde{R}_m) \right] d\tilde{x}. \quad (B-3)$$

By theorem, the conditional mean of $\tilde{R}_m$ for a given value of $\tilde{X}$, $\int_{-\infty}^{\hat{r}_D} \tilde{R}_m g(\tilde{R}_m | X) d\tilde{R}_m = E(\tilde{R}_m) + \text{cov}(\tilde{X}, \tilde{R}_m)[\tilde{X} - E(\tilde{X})]/\text{var}(\tilde{X})$ [See Mood and Graybill (1963, p. 202)]. Therefore,

$$\text{cov}(\tilde{b}, \tilde{R}_m) = \text{cov}(\tilde{X}, \tilde{R}_m) \left[ \int_{-\infty}^{\hat{r}_D} \tilde{x} f(\tilde{x}) d\tilde{x} - E(\bar{X})F(\hat{r}_D) \right] / \sigma_X^2. \quad (B-2)$$
Substituting (B-1) into (B-2) gives:

\[ \text{cov}(\tilde{b}, \tilde{R}_m) = -\text{cov}(\tilde{X}, \tilde{R}_m) f(\hat{\tau}D). \]  
(B-3)

The covariance between \( \tilde{b}\tilde{X} \) and \( \tilde{R}_m \) can be written as:

\[ \text{cov}(\tilde{b}\tilde{X}, \tilde{R}_m) = E(\tilde{b}\tilde{X}\tilde{R}_m) - E(\tilde{b}\tilde{X})E(\tilde{R}_m) \]

\[ = \int_{-\infty}^{\hat{\tau}D} \tilde{X}f(\tilde{X}) \left[ \int_{-\infty}^{\tilde{R}_m} g(\tilde{R}_m|X)dR_m - E(\tilde{R}_m) \right] d\tilde{X}. \]

Using the theorem for conditional mean again, we obtain:

\[ \text{cov}(\tilde{b}\tilde{X}, \tilde{R}_m) = \text{cov}(\tilde{X}, \tilde{R}_m) \left[ \int_{-\infty}^{\hat{\tau}D} \tilde{X}^2 f(\tilde{X})d\tilde{X} - E(\tilde{X}) \right] \int_{-\infty}^{\hat{\tau}D} f(\tilde{X})d\tilde{X} \right] / \text{var}(\tilde{X}). \]

From Winkler, Roodman, and Britney's equation (3.4) (1972, p. 294), we also obtain

\[ \int_{-\infty}^{\hat{\tau}D} \tilde{X}^2 f(\tilde{X})d\tilde{X} - E(\tilde{X}) \int_{-\infty}^{\hat{\tau}D} f(\tilde{X})d\tilde{X} = \text{var}(\tilde{X})[F(\hat{\tau}D) - \hat{\tau}D f(\hat{\tau}D)]. \]

Therefore,

\[ \text{cov}(\tilde{b}\tilde{X}, \tilde{R}_m) = \text{cov}(\tilde{X}, \tilde{R}_m) \left[ F(\hat{\tau}D) - \hat{\tau}D f(\hat{\tau}D) \right]. \]  
(B-4)

REFERENCES