NBER WORKING PAPER SERIES

REAL AND FINANCIAL DECISIONS OF A FIRM WITH BANKRUPTCY AND DEFAULT: AN INTEGRATION

Fumio Hayashi

Working Paper No. 1097

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

March 1983

The author is grateful to the participants in workshops at the University of Minnesota, University of Wisconsin in Madison, the 1982 NBER Summer Institute, Osaka University and Tokyo University for their comments and discussions. Remaining errors are his own. The research reported here is part of the NBER's research program in Economic Fluctuations. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #1097
March 1983

Real and Financial Decisions of a Firm with Bankruptcy and Default: An Integration

ABSTRACT

This paper attempts to provide a framework for analyzing the interaction between real decisions (concerning investment and factor inputs) and financial decisions (concerning debt and new share issues) of a corporation. The model carries a rich menu of tax rates and explicitly incorporates bankruptcy and default. The firm's multi-period optimization problem is set up where real and financial decisions are simultaneously determined to maximize the value of the firm which is the market price of uncertain future dividends. The main results of the paper are as follows: if the firm's after-tax profits are small relative to investment, the firm finances new investment by retentions and debt; if they are large relative to investment, financing additional investment is done through new shares and debt; in the intermediate case, additional investment is financed entirely by debt.

Fumio Hayashi
Institute of Socio-Economic Planning
University of Tsukuba
Sakura, Ibaraki 305
JAPAN

(0298) 53-5556

1. Introduction

This paper is an attempt to provide an analytical framework for studying the interaction between corporate investment and financial policy. Almost all of the huge literature on investment has been concerned with the optimal level of investment or capital stock taking the firm's financial policy as given and independent of the level of investment. The theory which posits that investment is a function of Tobin's(1969) q assumes either that investment is financed entirely with retentions [Hayashi(1982)] or that a constant fraction of investment is financed with debt [Summers(1981) and Poterba and Summers(1982)]. The theory which posits that the optimal capital stock (as opposed to optimal investment) is determined at the equality of the "cost of capital" and the marginal product of capital makes similar assumptions to derive the expression for the cost of capital [see, e.g., Chirinko and King(1982)]. On the other hand, much of the equally large literature on corporate finance [with possible exceptions of Gordon(1981) and Auerbach(1982)] has been concerned with corporate financial structure taking the firm's investment and other real decisions as given.

Those who are accustomed to the tradition of Modigliani and Miller might think that investment is independent of how it is financed. This is indeed true if there are no taxes [Modigliani and Miller(1958) and Stiglitz(1969)] or if there are no bankruptcies and the corporate tax rate is equal to the individual marginal tax rate (with the dividend tax and the capital gains tax rates being zero) [Miller(1977)]. But if both taxes and bankruptcies are present, the value of the firm is not independent of its financial structure and the firm's investment and financial decisions are interrelated.

Using the Capital Asset Pricing Model with taxes, Gordon (1981) has analyzed the interaction between the optimal capital stock and the optimal debt-capital ratio. However, he appears to assume that profits are proportional the stock of capital without allowing for adjustment costs associated with investment, so the optimal capital stock is actually indeterminate in his model. He also appears to assume the fraction of new investment financed by debt is independent of the capital stock. Auerbach(1982) has considered a deterministic model of a value maximizing firm where profits are a additively separable function of the stream of past investments and where the corporate bond rate increases with the debt-capital ratio. He derived a few results on the interaction between the level of investment and the way it is financed for special cases.

The stochastic model of a corporation to be developed in this paper explicitly incorporates bankruptcy and default while allowing a rich menu of tax rates. The model assumes that the firm determines investment and financial policies so as to maximize its share price. The model is very general with respect to how the firm's uncertain future dividends are priced; no assumptions like certainty equivalence, constant risk premium, or the capital asset pricing model will be made. With a minimal set of assumptions on the pricing mechanism, we will derive fairly sharp results concerning how the firm finances investment and how the level of investment is affected by the way it is financed. The main results can be summarized as follows. If the firm's after-tax profits are small relative to the level of investment, the firm finances investment by retentions <u>and</u> debt. If they are large relative to the level of investment, addtional investment projects are financed with new shares <u>and</u> debt. In the intermediate case, additional finance comes <u>entirely</u> through debt. Somewhat surprizingly, these results are broadly consistent with Auerbach(1982)'s results which were derived for a model different from ours.

The organization of the paper is as follows. Section 2 presents a very general pricing formula for the shares of the firm under uncertainty with bankruptcy and default. The pricing formula is a considerable generalization of that in Auerbach(1979). Section 3 formulates the firm's problem of maximizing its share price and derives the associated dynamic programming algorithm. Section 4 makes a brief detour to the taxless world and verifies the Modigliani-Miller theorem.

Sections 5 and 6 consider what we call the homogeneous case where production and adjustment costs associated with investment exibit constant returns to scale and where the firm is competitive. It is shown in section 5 that the value of the firm is proportional to its debt and capital stock. The results in section 6 on the interaction between investment and financial policy are sharp: the ratio of debt to the end-of-period capital stock is independent of investment or else the fraction of additional investment financed by debt is 100%. In section 7, we derive similar but less sharp results for the case without adjustment costs. We will also show in section 7 that the notion of "cost of capital" loses its usefulness in a model such as ours where bankruptcy is explicitly incorporated. We will not consider the case with adjustment costs and without constant returns to scale, since it is a straightforward combination of the two cases analyzed in sections 5, 6 and 7. Section 8 lists qualifications.

2. Pricing Formula with Bankruptcy and Default

We consider the behavior of a firm in a discrete time, stochastic model. The firm is assumed to act so as to maximize its shareholders' wealth. At the beginning of the period, the firm decides whether or not to go bankrupt. If it decides not to go bankrupt (i.e., if the bankruptcy dummy $M_t = 1$), the (cum-dividend) share price at the beginning of the period is p_t. Let Q_t stand for the number of pre-existing shares. At the beginning of the period, the firm issues g_tQ_t units of new shares at the ex-dividend price p_t', issues new debt B_t (which, for the sake of simplicity we take to have a maturity of one period), pays interest and principal $(1+i_{t-1})B_{t-1}$ on corporate debt, and distributes total dividends d_tQ_t to pre-existing shareholders. After all this happens, new information hits the stock market and the ex-dividend price p_t' becomes $M_{t+1}p_{t+1}$ at the end of period t (i.e., the beginning of period t+1). If the firm decides to go bankrupt (i.e., if $M_t = 0$), shareholders receive nothing and bondholders take over the firm.[1] Dividends are taxed at rate θ at the personal level. Capital gains are taxed at a lower rate c. All stockholders face the same tax rates, θ and c.[2]

The equilibrium condition that links the cum-dividend price p_t and the ex-dividend price p_t' is the following:[3]

$$p_t = (1-\theta_t)d_t + p_t' - c_t(p_t' - p_t),$$

i.e.,

(2.1) $$p_t = m_t d_t + p_t', \quad \text{where} \quad m_t = (1-\theta_t)/(1-c_t).$$

The ratio m_t is less than one since $c_t < \theta_t$. Dividends per share, d_t, can be written as

(2.2) $$d_t = X_t/Q_t + g_t p_t',$$

where

(2.3) $$X_t = (1-u_t)\Pi_t - (1-k_t)v_t F_t + B_t - [1+(1-u_t)i_{t-1}]B_{t-1}$$

= cash flow + new debt issue
 - interest and principal on pre-existing debt[4]

Π = before-tax profits where variable factor inputs are already maximized out,

F = investment,

k = rate of investment tax credit,

v = price of investment goods,

u = corporate tax rate,

B = new issues of debt,

i = corporate bond rate.

Using (2.1) and (2.2) we can easily derive

(2.4) $$p_t = m_t X_t/Q_t + (1 + m_t g_t)p_t'.$$

Associated with before-tax profits Π_t in (2.3) is the production function $G_t(K_t,N_t,F_t,e_t)$, where N is a vector of variable factor inputs and e is the shock to technology. Here we follow Lucas(1967) and allow output to depend negatively on investment F: $\partial G/\partial F \leq 0$. This is how we introduce adjustment costs associated with investment. The firm's investment activity of volting down investment goods within the firm is a resource-using activity; as F increases, more and more fraction of K and N must be directed to the investment activity and as a result output goes down. We assume convex adjustment costs, i.e., $G_{FF} \leq 0$. Therefore the first and second partial derivatives of Π_t with respect to F_t is nonpositive:

$$(2.5) \quad \Pi_t = \Pi_t(F_t,K_t), \quad \Pi_F \leq 0, \text{ and } \quad \Pi_{FF} \leq 0.$$

The profit function should also involve the technology shock e_t and the parameters that characterize the demand and supply functions (for output and factor inpits) that the firm faces. If, in particular, the firm is competitive in the markets for output and factors of production, the profit function will involve output and factor prices as well as F and K. The dependence of profits on those variables is left implicit in the profit function.

Let $L_t^j(x_{t+j})$ be the price that would be given by the asset market as of t for an asset which pays (possibly stochastic) tax-free x_{t+j} dollars at t+j. Thus L_t^j is a mapping from the space of random variables to real numbers.[5] By definition, $L_t^0(x_t) = x_t$, and $L_t^j(1) = 1/(1+r_{jt})$ where r_{jt} is the nominal rate on a tax-free j-period default-free bond. Ross (1978) has proved that if there are no arbitrage opportunities left, the operator L_t satisfies

(1) linearity: $L_t^j(\mu_1 x + \mu_2 y) = \mu_1 L_t^j(x) + \mu_2 L_t^j(y)$ for any nonstochastic μ_1 and μ_2;

(2) iterative property: $L_t^j(L_{t+j}^k(x_{t+j+k})) = L_t^{j+k}(x_{t+j+k})$, $j,k \geq 0$.

Using this pricing operator, the relationship between p_t' and p_{t+1} is given by

$$L_t^1(M_{t+1}p_{t+1} - c_{t+1}(M_{t+1}p_{t+1} - p_t')) = p_t',$$

i.e.,

(2.6) $$p_t' = L_t^1(\alpha_{t+1}p_{t+1}),$$

where

(2.7) $$\alpha_{t+1} = (1-c_{t+1})M_{t+1}/(1-L_t^1(M_{t+1}c_{t+1})).$$

By doing recursion on (2.4) and (2.6) and using the above properties of L_t, we can derive (see Appendix 1):

(2.8) $$p_t Q_t = \sum_{j=0}^{\infty} L_t^j(\gamma_{t,j}\, m_{t+j}\, X_{t+j}),$$

where

$$(2.9)\qquad \gamma_{t,j} = \sum_{k=1}^{j} \beta_{t+k-1}\alpha_{t+k}, \quad \text{with } \gamma_{t,0} = 1,$$

$$(2.10)\qquad \beta_{t+j} = (1 + m_{t+j}g_{t+j})/(1 + g_{t+j}).$$

We note that $\gamma_{t,j} = 0$ if $M_{t+k} = 0$ for <u>some</u> k such that $j \geq k \geq 1$.

One simple way to allow for defaults is the following. Bondholders can receive full amount $(1+i_t)B_t$ if the firm is not bankrupt at t+1. In the event of bankruptcy in period t+1, bondholders take over the firm and attempt to sell the firm to the highest bidders. The market value of the firm at t+1 without obligation to pay interest and principal $(1+i_t)B_t$ is clearly equal to:

$$(2.11)\qquad p_{t+1}Q_{t+1} + [1+(1-u_{t+1})i_t]B_t.$$

This is what the bondholders can receive in period t+1 in the event of bankruptcy. Letting θ' stand for the tax rate on interest income, the corporate bond rate i_t must satisfy

$$(2.12)\qquad 1 = L_t^1[M_{t+1}(1+i'_{t+1}) + (1-M_{t+1})(1+i''_{t+1})],$$

where

$$i'_{t+1} = (1-\theta'_{t+1})i_t,$$

$$i''_{t+1} = (1-\theta'_{t+1})[p_{t+1}Q_{t+1}+(1-u_{t+1})i_tB_t]/B_t.$$

3. The Firm's Optimization Problem

With the pricing formula (2.8) at hand, we can now formalize the firm's optimization behavior. Our formulation will closely parallel Lucas and Sargent(1981). We begin with a few definitions. The exogenous variables are the variables that the firm cannot influence their values. They include: tax rates (u,θ,θ',c,k), the technology shock (e), and the parameters characterizing the demand and supply functions (for output, factor inputs and investment goods) that the firm faces.[6] We assume that these exogenous variables are part of a larger set of variables Z_t which follow a Markov process.[7] The vector Z_t will be referred to as the state of nature, because it determines not only the stochastic properties of the exogenous variables that the firm will face in the future but also the functional form of the pricing operator L_t^j $(j\geq 0)$.[8] The current return $m_t X_t$ in (2.8) depends on: the current value of the exogenous variables, B_t, F_t, i_{t-1}, B_{t-1}, and K_t. The last three variables are called the firm's state variables as they are historically given to the firm at the beginning of the period. The information set in period t, I_t, consists of the state of nature Z_t and the firm's state variables i_{t-1}, B_{t-1}, K_t. We assume that I_t is known to both the stock market and the firm at the beginning of period t. The firm's action is a vector (M_t, g_t, B_t, F_t) and the associated variable factor inputs. The firm determines its current action as a function of I_t.[9] This function is called the firm's <u>decision rule</u> and is denoted by μ_t. Since the number of pre-existing shares, Q_t, is historically given, maximizing p_t is equivalent to maximizing <u>the value of the firm</u>, $V_t = p_t Q_t$, whose expression is given by (2.8). Since the current information set I_t and a sequence of decision rules $(\mu_t, \mu_{t+1}, \ldots)$ completely determine the nature of the stochastic process for $\gamma_{t,j} m_{t+j} X_{t+j}$ $(j \geq 0)$, the value of the firm depends only on I_t and $(\mu_t, \mu_{t+1}, \ldots)$. We assume the firm knows the pricing operator L_t, so that it can correctly evaluate how the stock market would react to any hypothetical action contemplated by the firm.

Let $V_t(I_t)$ be the value of V_t that is maximized over $(\mu_t, \mu_{t+1}, \ldots)$ conditional on $M_t = 1$. Shareholders' wealth is $V_t(I_t)$ if the firm stays in business in period t, and zero if it goes broke. Thus the firm's bankruptcy decision is simply the following:

$$(3.1) \qquad M_t = \begin{cases} 1 & \text{if } V_t(I_t) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

This and (2.11) imply that the amount that bondholders can receive in the event of bankruptcy in period t is always less than the full amount $(1+i_{t-1})B_{t-1}$ that they can receive if the firm is in operation. Clearly, the probability of bankruptcy in period t increases with B_{t-1}. It is for this reason that the corporate bond rate i_t increases with B_t.

If the firm stays in business, the firm's optimization problem is to find a sequance $(\mu_t^*, \mu_{t+1}^*, \ldots)$ (with $M_t = 1$) of optimal decision rules that maximizes V_t. But there are a few constraints that the firm is subject to. The first is that (per share) dividends d_t must be greater than or equal to some lower bound $\bar{d}_t$. (If $\bar{d}_t = 0$, this constraint simply says dividends must be nonnegative.) In some countries (e.g., Japan and the United Kingdom) repurchase of equity is illegal. The second constraint, therefore, is that g be nonnegative.[10] The third constraint is, of course, the capital accumulation constraint. Thus the (currently operating) firm's problem is

$$(3.2) \quad \max_{\mu_{t+j},\ j\geq 0} V_t \quad \text{subject to } d_{t+j} \geq \bar{d}_{t+j}, \quad g_{t+j} \geq 0, \quad \text{and } K_{t+j+1} = (1-\delta)K_{t+j} + F_{t+j}, \quad (j=0,1,2,\ldots),$$

where V_t is given by (2.8), δ is the exponential rate of phisical depreciation of capital, and μ_t is constrained such that $M_t = 1$. The value of V_t evaluated at the sequence of optimal decision rules $(\mu_t^*, \mu_{t+1}^*, \ldots)$ (with $M_t = 1$) is, of course, $V_t(I_t)$.

Now, it immediately follows from (2.8) that

(3.3) $$V_t = m_t X_t + \beta_t L_t^1(\alpha_{t+1} V_{t+1}).$$

We note here that $L_t^1[\alpha_{t+1} V_{t+1}(I_{t+1})]$, which is the value of $L_t^1(\alpha_{t+1} V_{t+1})$ evaluated at the optimal future decision rules $(\mu_{t+1}^*, \mu_{t+2}^*, \ldots)$, is a function of B_t, K_{t+1} and Z_t.[11] We submerge the dependence of W_t on Z_t in the functional form and write

(3.4) $$W_t(B_t, K_{t+1}) = L_t^1[\alpha_{t+1} V_{t+1}(I_{t+1})].$$

It is clear from (2.6) and (2.8) that this W_t is equal to $p_t' Q_{t+1}$. Although it is a function of the future optimal decision rules, the value of W_t is independent of current decision rule μ_t as long as the values of B_t and K_{t+1} are given. Furthermore, the current return $m_t X_t$ is not influenced by the firm's future action.[12] Therefore, the firm's current action (g_t, B_t, F_t) implied by the optimal decision rule μ_t^* (with $M_t = 1$) solves the familiar DP (dynamic programming) algorithm:

(3.5) $$V_t(I_t) = \max_{g,B,F} [m_t X_t + \beta_t W_t(B_t, K_{t+1})]$$

$$\text{subject to } d_t \geq \bar{d}_t, \quad g \geq 0, \quad K_{t+1} = (1-\delta)K_t + F.$$

What restrictions can we place on W_t? An increase in B_t (with K_{t+1} held constant) affects W_t in four ways. The first is its direct effect on X_{t+1}: X_{t+1} decreases as interest and principal payments increases. Second, it increases the probability of bankruptcy in t+1, which lowers the discounting factor $\gamma_{t+1,j}$ $(j \geq 1)$ for any realization of Z_{t+j} $(j>0)$. Third, it follows from (2.12) that the increased likelihood of bankruptcy raises the corporate bond rate. Fourth, it becomes more likely that the constraint $d \geq \bar{d}$ is binding in period t+1. An increase in K_{t+1} affects W_t in exactly the opposite directions. Therefore, $W_t(B_t, K_{t+1})$ decreases with B_t and increases with K_{t+1}. If B_t is sufficiently small, the probability of bankruptcy will be zero. The corporate bond will then be default-free, so $(1-\theta_t')i_t$ must be equal to r_t, the nominal interest rate on a safe, tax-free one-period bond. Furthermore, the constraint $d \geq \bar{d}$ will not be binding. So the only effect of B_t on W_t is its direct effect on X_{t+1}. It then follows that $W_t(B_t, K_{t+1})$ has the following separable form (at least asymptotically) when B_t is sufficiently small:

(3.6) $$W_t(B_t, K_{t+1}) = W_t^1(K_{t+1}) - w_t^2 B_t,$$

(3.7) $$w_t^2 = L_t^1[\alpha_{t+1} m_{t+1}(1+(1-u_{t+1})i_t)].$$

Since $M_{t+1} = 1$ with probability one when B_t is sufficiently small, α_{t+1} is independent of B_t and K_{t+1}. Thus w_t^2 is exogenous to the firm, implying that $W_t(B_t,K_{t+1})$ is a linear function of B_t when B_t is sufficiently small. These information on the functional form of $W_t(B_t,K_{t+1})$ will turn out to be useful in later sections.

The DP algorithm (3.5) can be simplified somewhat if we notice

$$(3.8) \qquad \begin{aligned} d_tQ_t &= X_t + g_tp_t'Q_t \\ &= X_t + y_tW_t(B_t,K_{t+1}), \end{aligned}$$

where

$$(3.9) \qquad y_t = g_t/(1+g_t).$$

Thus (3.5) reduces to

$$(3.10) \qquad \max_{y,B,F} \; [m_tX_t(B,F) + (1-y+m_ty)W_t(B,(1-\delta)K_t+F)]$$

$$\text{subject to} \quad X_t(B,F) + yW_t(B,(1-\delta)K_t+F) \geq \bar{d}_tQ_t, \quad y \geq 0,$$

where

$$(3.11) \qquad X_t(B,F) = (1-u_t)\Pi_t(F,K_t) - (1-k_t)v_tF + B + R_t,$$

$$(3.12) \qquad R_t = -[1+(1-u_t)i_{t-1}]B_{t-1}.$$

4. The Taxless Case

It is useful at this stage to pause briefly and see what happens if there are no taxes while bankruptcy and default are possible. We prove here that the Modigliani-Miller(1958) theorem holds for this case by following the arguments in Stiglitz(1969) and Ross(1978). This exercise will put the results in the following sections in a proper perspective.

We first note from (3.10) that corporate equity policy is irrelevant in the taxless world as $m_t = 1$, so the constraint $X_t + \gamma W_t \geq \bar{d}_t Q_t$ is irrelevant. Proving the Modigliani-Miller proposition that the value of the firm is independent of leverage in the taxless world amounts to showing that $W_t(B_t,K_{t+1})$ can be written in a separable form (for any value of B_t and K_{t+1}, not just for small B_t):

$$(4.1) \qquad W_t(B_t,K_{t+1}) = W_t^1(K_{t+1}) - B_t,$$

for it follows from this that $V_t = X_t + W_t$ is independent of B_t. To prove (4.1), we note that $\alpha_{t+1} = M_{t+1}$ in the taxless world, so

$$(4.2) \qquad W_t(B_t,K_{t+1}) = L_t^1[M_{t+1}V_{t+1}(I_{t+1})]$$

$$= L_t^1(Y_{t+1}) - L_t^1[(1-M_{t+1})Y_{t+1}] - (1+i_t)B_tL_t^1(M_{t+1}),$$

where

$$(4.3) \qquad Y_{t+1} = V_{t+1}(I_{t+1}) + (1+i_t)B_t$$

$$= \Pi_{t+1} - (1-k_{t+1})v_{t+1}F_{t+1} + B_{t+1} + W_{t+1}(B_{t+1},K_{t+2}).$$

Since in the taxless world (2.12) becomes

$$(4.4) \qquad B_t = (1+i_t)B_tL_t^1(M_{t+1}) + L_t^1[(1-M_{t+1})Y_{t+1}],$$

(4.2) reduces to (4.1) with $W_t^1(K_{t+1}) = L_t^1(Y_{t+1})$.[13] This completes the proof of the Modigliani-Miller theorem.

5. The Homogeneity Assumption

Properties of the solution to (3.10) depend, of course, on the functional form of the value function $W_t(B_t,K_{t+1})$. In this section and the next, we focus on the case where the firm is a price-taker and the production function $G(K,N,F,e)$ is linearly homogeneous in K,N,F for any given e.[14] The immediate implication of this is that the assòciated profit function is linearly homogeneous in F,K and satisfies

$$(5.1) \quad \Pi_t(F_t,K_t) = \pi_t(f_t)K_t, \quad \pi_t' \leq 0, \quad \pi_t'' \leq 0,$$

where $f_t = F_t/K_t$.[15] It then seems clear that the value function, too, is linearly homogeneous in B_t,K_{t+1}:

$$(5.2) \quad W_t(B_t,K_{t+1}) = h_t(\lambda_t)K_{t+1},$$

where $\lambda_t = B_t/K_{t+1} = B_t/[(1-\delta+f_t)K_t]$ will be referred to as the debt-capital ratio. The intuition for (5.2) is quite simple.[16] When the initial condition is $(2B_t,2K_{t+1})$, it is feasible for the firm to double the future level of investment, employment and corporate debt that are optimal if the initial condition is (B_t,K_{t+1}). Since the probability of bankruptcy remains unchanged if the same corporate equity policy is followed, the value of the firm under this decision rule with $(2B_t,2K_{t+1})$ is at least twice $W_t(B_t,K_{t+1})$, i.e., it must be that $W_t(2B_t, 2K_{t+1}) \geq 2W_t(B_t, K_{t+1})$. Apply the same argument in the opposite direction to obtain $W_t(B_t/2, K_{t+1}/2) \geq W_t(B_t, K_{t+1})/2$. Thus we have $W_t(2B_t, 2K_{t+1}) = 2W_t(B_t, K_{t+1})$. This result, which is a generalization of Lucas and Prescott (1971) and Hayashi(1982), is proved in Appendix 2. We note from (2.6), (2.8) and (3.4) that h_t as defined by (5.2) is equal to the ex-dividend value of the firm divided by K_{t+1}, namely $p_t'Q_{t+1}/K_{t+1}$. So the result states that Tobin's(1969) "marginal q", $\partial W_t/(v_t \partial K_{t+1})$, is equal to the "average q", $W_t/(v_t K_{t+1})$.

This assumption of homogeneity has been popular but often implicit in the corporate finance literature. For example Modigliani and Miller(1963) stated in their footnote 15: "...we are referring in principle only to investments which increase the scale of the firm. That is, the new assets must be in the same 'class' as the old." (Italics original) Since the "return" of new assets is the same as that of the existing assets, an increase in the value of the firm due to investment, $\partial W_t/\partial K_{t+1}$, must be equal to the average value of the firm, W_t/K_{t+1}. In Gordon and Malkiel(1981), the marginal return on real investment is represented by s. But the same symbol represents the return from investing a dollar in equity. Their footnote 33 states: "Assume that the firm pays out as dividends p percent of its after-tax profits and reinvests the rest. Also assume that the investor with a marginal tax rate of m on interest payments has a marginal tax rate of n on dividends.... When investing a dollar in equity, the investor receives as dividends $ps(1-t)(1-n)$ after tax." (t is the corporate tax rate.) Unless all assets in the firm share the same return, this would not happen.

We have seen in section 3 that $W_t(B_t, K_{t+1})$ decreases with B_t and increases with K_{t+1} and that W_t is linear in B_t at least asymptotically when B_t is small. So $h_t(\lambda_t)$ must be a decreasing function of the debt-capital ratio λ_t and (at least asymptotically) linear when λ_t is small. We assume that $h_t(\lambda_t)$ comes down to zero as λ_t reaches $\bar{\lambda}$. If h_t does not come down to zero or if it does so only asymptotically, the firm's optimization problem will have no finite solution. When λ_t is greater than or equal to this $\bar{\lambda}$, the firm is bound to be bankrupt. Then nobody buys the corporate bond issued by the firm, because the rate of return on such bonds is dominated by that on a safe, tax-free bond. So the feasible debt-capital ratio must be less than $\bar{\lambda}$, i.e., $\lambda < \bar{\lambda}$. The graph of $h_t(\lambda_t)$ is illustrated in Figure 1.

With these restrictions on h_t, we are now ready to solve the DP algorithm (3.10). Without loss of generality we can set $Q_t = K_t$ and convert the DP algorithm into the following "per-capital" form:

$$(5.3) \quad \max_{b,y,f} \; mx(b,f) + (1-y+my)(1-\delta+f)h[b/(1-\delta+f)]$$

$$\text{subject to} \quad x(b,f)+y(1-\delta+f)h[b/(1-\delta+f)] \geq \bar{d},$$

$$y \geq 0,$$

where the time subscripts are dropped for notational simplicity and

$$(5.4) \quad x(b,f) = (1-u_t)\pi_t(f) - (1-k_t)v_t f + b + R_t/K_t = X_t/K_t,$$

$$\text{with} \quad b = B_t/K_t, \quad f = F_t/K_t.$$

6. Financing Decisions and the Investment Equation

Probably the most instructive way to solve the DP algorithm (5.3) is first calculate as functions of f the optimal financing package b(f),y(f) that maximizes the objective function for a given value of f and then change f to find the optimum of the objective function that is maximized out over b and y, although the optimization problem can alternatively be solved by choosing b,y,f simultaneously. For the most part, we will assume that $m = (1-\theta)/(1-c)$ is less than one. The following three cases can arise. Case 1: $y \geq 0$ is binding but $d \geq \bar{d}$ is not. Case 2: both constraints are binding. Case 3: $d \geq \bar{d}$ is binding but $y \geq 0$ is not. At least one constraint must be binding because otherwise the firm always gets better off by reducing new share issues.

Case 1: If $d > \bar{d}$ and $y = 0$, (5.3) reduces

$$(6.1) \qquad \max_{b,f} \; mx(b,f) + (1-\delta+f)h[b/(1-\delta+f)].$$

The first order condition with respect to b is

$$(6.2) \qquad m + h'(\lambda) = 0,$$

where λ is the debt-capital ratio B_t/K_{t+1}. So the optimal debt given f is

(6.3) $b_1(f) = (1-\delta+f)\lambda_1$,

where

(6.4) $h'(\lambda_1) = -m$.

The first order condition with respect to f with $b = b_1(f)$ is

(6.5) $h(\lambda_1)/m + \lambda_1 = (1-k)v - (1-u)\pi'(f)$,

i.e.,

(6.6) $f = \omega[(h(\lambda_1)/m+\lambda_1-(1-k)v)/(1-u)]$

$$= \omega[\frac{p_t' Q_{t+1}/m_t + B_t}{(1-u_t)K_{t+1}} - \frac{(1-k_t)v_t}{1-u_t}],$$

where ω is the inverse function of $-\pi'(f)$. This is the investment equation derived by Poterba and Summers(1982) for the "capitalization hypothesis" which assumes that the firm never issues new shares. In their derivation, the constant debt-capital ratio λ_1 is exogenously given; here, the ratio turns out to be constant as a result of optimization.

We note that ω also depends on current output and factor prices and the technology shock, since π' is a function of them as well as f. If the production function G takes the separable form $G(K,N,F,e) = G^1(K,N,e) - G^2(K,F,e)$, then ω does not involve output and factor prices.

Case 2. If both constraints are binding, (5.3) reduces to

$$(6.7)\qquad \max_{b,f}\ (1-\delta+f)h[b/(1-\delta+f)] \quad \text{subject to} \quad x(b,f) = \bar{d}.$$

The optimal debt given f is determined by the binding constraint $x = \bar{d}$. Thus

$$(6.8)\qquad b_2(f) = \bar{d} - R_t/K_t - (1-u)\pi(f) + (1-k)vf.$$

This is a convex function of f since $\pi'' \leq 0$. The first order condition with respect to f with $b = b_2(f)$ gives

$$(6.9)\qquad - h[\lambda(f)]/h'[\lambda(f)] + \lambda(f) = (1-k)v - (1-u)\pi'(f),$$

i.e.,

$$(6.10)\qquad f = \omega[(-h(\lambda)/h'(\lambda)+\lambda-(1-k)v)/(1-u)],$$

where $\lambda(f)$ is the debt-capital ratio $b_2(f)/(1-\delta+f)$. Comparing (6.6) and (6.10) we can see that $-h'(f)$ in (6.10) plays the role of m in (6.6). This $-h'(f)$ is essentially unobservable since in the present case there is no marginal condition involving $h'(\lambda)$.

Case 3. If $d \geq \bar{d}$ is binding but $y \geq 0$ is not, (5.3) reduces to

(6.11) $$\max_{b,f} \quad x(b,f) + (1-\delta+f)h[b/(1-\delta+f)] + (m-1)\bar{d}.$$

The first order condition with respect to b is

(6.12) $$1 + h'(\lambda) = 0,$$

so that the optimal debt given f is

(6.13) $$b_3(f) = (1-\delta+f)\lambda_3,$$

where

(6.14) $$h'(\lambda_3) = -1.$$

The first order condition with respect to f with $b = b_3(f)$ yields

(6.15) $$h(\lambda_3) + \lambda_3 = (1-k)v - (1-u)\pi'(f),$$

i.e.,

(6.16) $$f = \omega[(h(\lambda_3)+\lambda_3-(1-k)v)/(1-u)]$$

$$= \omega[\frac{p_t'Q_{t+1} + B_t}{(1-u_t)K_{t+1}} - \frac{(1-k_t)v_t}{1-u_t}].$$

This is the equation derived by Poterba and Summers(1982) for the "double tax view" which <u>assumes</u> that dividends are a constant fraction of profits and investment is financed with new share issues whenever necessary. In their derivation, a constant debt-capital ratio is assumed; in our derivation, it is a result of optimization. Given $b_3(f)$, the optimal value of y is determined by the binding constraint $x+y(1-\delta+f)h = \bar{d}$:

(6.17) $$y(f) = [\bar{d}-x(b_3(f),f)]/[(1-\delta+f)h(\lambda_3)].$$

This may or may not be an increasing function of f, but it is easy to show that the value of new shares issued, $g_t p_t' Q_t = y_t W_t$, increases with f.

It is easy to prove that $\lambda_3 > \lambda_1$ if $m=(1-\theta)/(1-c) < 1$.[17] Thus only three cases can happen concerning the ordering of b_1, b_2, b_3: (a) $b_2 \leq b_1 < b_3$, (b) $b_1 < b_2 \leq b_3$, and (c) $b_1 < b_3 < b_2$. These three cases are illustrated in panels (a)-(c) of Figure 2. It is clear that Cases 1,2,3 corresponds to cases (a),(b),(c), respectively. Therefore the graph of the optimal debt b(f) consists of pieces of $b_1(f)$, $b_2(f)$, $b_3(f)$, and will look like the solid line in Figure 3, panel (i).[18] The graph of y(f) is drawn in panel (ii) of Figure 3.

The interpretation of the results concerning the optimal financing package $b(f), y(f)$ we have just derived is quite clear. If dividends are greater than the lower bound $\bar{d}$, the firm finances additional investment projects by cutting dividends and issuing λ_1 dollars (per a dollar's worth of investment) of corporate bonds. If dividends cannot be cut any further, financing additional investment projects is done entirely by corporate debt. However, as soon as debt reachs a critical level, the rate of increase of bond issues will be cut back to λ_3 and marginal finance is done with new equity and debt. The graph of $b_2(f)$ represents the amount of debt that is necessary for the firm to deliver per share dividends $\bar{d}$ without resorting to new share issues. Therefore the vertical difference between $b_2(f)$ and the solid line $b(f)$ for $f > f_2$ is the amount of funds raised by new equity, $g_t p_t' Q_t$. Note that it is not the firm's optimal policy to finance investment entirely by cutting dividends even when it is feasible. This is because the increase in capital stock due to current investment makes bankruptcy less likely for any given level of debt; what determines the likelihood of bankruptcy in the homogeneous case is the debt-capital ratio $\lambda_t = B_t/K_{t+1}$.[19] This is also why debt finance coexists with equity finance in Case 3. We also note from panel (c) of Figure 2 that issuing new shares can be an optimal financial policy, even if repurchase of existing shares is legal.

It may be useful at this point to briefly look at the role of what we call the firm's <u>cash income</u> $(1-u)\pi(f) - R_t/K_t$ in financing decisions. If this increases by one dollar, the graph of $b_2(f)$ shifts downwards by exactly one dollar. If the increase in cash income does not change the functional form of $h(\lambda)$ -- which is what happens if, e.g., the increase is due to a purely transitory technology shock --, then the graphs of $b_1(f)$ and $b_3(f)$ do not shift, so that the critical levels of investment f_1 and f_3 that divide the three Cases in Figure 3 will both increase, making Case 3 less likely.

Turning now to the determination of investment given the financing package, we first note that the optimal investment must satisfy the marginal condition (6.5) for Case 1, (6.9) for Case 2, and (6.15) for Case 3. A natural way to interpret this marginal condition is to take the LHS (left hand side) as the marginal "benefit" of investment and the right hand side -- investment expenditure plus profits foregone due to investment -- as the marginal "cost" of investment.[20] Since $h(\lambda)$ decreases with λ and since $\lambda_1 \leq b_2(f)/(1-\delta+f) \leq \lambda_3$, the LHS of (6.5) is always greater than that of (6.15), and the LHS of (6.9) lies between the two in Case 2.[21] This is illustrated in Figure 4, panel (i). The upper horizontal line, labelled Line 1, is the graph of the LHS of (6.5) and Line 3 that lies below it is the graph of the LHS of (6.15).

The downward sloping line, lebelled Line 2, that cuts Line 1 and Line 3 from above is the graph of the LHS of (6.9). Thus the graph of the marginal benefit of investment consists of pieces of these three lines and is represented by the solid line in Figure 4. The graph of the marginal cost of investment is the upward sloping curve whose intercept is $(1-k)v - \pi'(0)$ and whose slope is $-(1-u)\pi''(f)>0$. The optimal investment f^* is determined at the intersection of the two curves. In panel (i), Case 2 is occuring. Now let's see what happens if there is an increase in the firm's cash income $(1-u)\pi(f) - R_t/K_t$. To make the story simple, suppose that the increase is caused by a (positive) technology shock of purely temporary nature that enters the current profit function π additively, so that the shock shifts neither Lines 1 and 3 through its effect on $h(\lambda)$ nor the marginal cost curve through its effect on $\pi'(f)$. Under these conditions Line 2 will slide horizontally to the right along Line 1 and Line 3, as the critical levels of investment f_1 and f_2 move to the right. This is illustrated in panel (ii) of Figure 4 where the new optimal investment is f^{**} which is greater than the old level f^*. We can thus conclude that a windfall gain which increases the firm's cash income but which is totally unrelated to the future exogenous variables can increase current investment. This may explain why corporate investment in the U.S. appears to be too sensitive to current profits.

We close this section by quickly looking at the case where $m_t = 1$. As was noted in footnote 3, this case arises if dealers in securities can deduct unlimited amounts of short-term capital losses against ordinary income and the stock prices are set by these dealers. It is clear from (5.3) that when $m = 1$ the objective function is independent of y and hence the optimal amount of new issues is indeterminate as long as the implied value of d is greater than or equal to $\bar{d}$. In particular, it is a rational behavior of the firm to simultaneously pay dividends and issue new shares. However, the optimal debt-capital ratio <u>is</u> determinate at λ_3 and independent of investment. Clearly, the investment equation is given by (6.16) and the graph of the marginal benefit of investmtnt collapses to Line 3 in Figure 4.

7. Real and Financial Decisions without Adjustment Costs

We now go back to (3.10) and investigate the firm's optimal decisions without adjustment costs. Thus the production function G and the associated profit function Π do not involve investment F. In order to make the optimization problem well-defined, we assume that the value function $W_t(B_t,K_{t+1})$ is concave in K_{t+1}; otherwise there will be no solution to the firm's optimization problem. Although unable to prove, we conjecture that a sufficient condition for the concavity is that the profit function is cancave in the capital stock.

It will be convenient and instructive to write the first order conditions in terms of marginal q, which we recall was defined to be

$$(7.1) \qquad q_t = \partial W_t(B_t,K_{t+1})/(v_t \partial K_{t+1}).$$

As in the homogeneous case, three cases arise. (In what follows the time subscript will be dropped whenever no confusions should arise.)

Case 1: If $y = 0$ and $X + yW > \bar{d}Q$, (3.10) reduces to

$$(7.2) \qquad \max_{B,F} \; mX(B,F) + W(B,(1-\delta)K_t+F).$$

The first order condition with respect to B is

$$(7.3)\qquad \bar{m} + W_B(B,(1-\delta)K_t+F) = 0,$$

where W_B is the partial derivative of $W_t(B_t,K_{t+1})$ with respect to B_t. This implicitly defines the optimal level of debt, $B_1(F)$, as a function of F. The first order condition with respect to F with $B = B_1(F)$ yields

$$(7.4)\qquad q_t/(1-k_t) = m_t,$$

i.e., the end-of-period capital stock K_{t+1} is optimal when marginal q adjusted for investment tax credits is equal to m.

<u>Case 2</u>: If both constraints are binding, (3.10) reduces to

$$(7.5)\qquad \max_{B,F}\; W(B,(1-\delta)K_t+F) \quad \text{subject to } X(B,F) = \bar{d}Q.$$

Since the level of debt must satisfy $X = \bar{d}Q$, we have

$$(7.6)\qquad B_2(F) = \bar{d}Q - (1-u)\,\Pi(K_t) + (1-k)vF - R_t,$$

which is a linear function of F. The first order condition with respect to F with $B = B_2(F)$ is

(7.7) $\quad q_t/(1-k_t) = -W_B(B,(1-\delta)K_t+F).$

An interpretation of this marginal condition will be given shortly.

Case 3: If $X = \bar{d}Q$ and $y > 0$, (3.10) reduces to

(7.8) $\quad \max_{B,F} \; X(B,F) + W(B,(1-\delta)+F) + (m-1)\bar{d}Q.$

The first order condition with respect to B is

(7.9) $\quad 1 + W_B(B,(1-\delta)K+F) = 0,$

which defines the optimal level of debt, $B_3(F)$, as a function of F. The first order condition with respect to F with $B = B_3(F)$ is

(7.10) $\quad q_t/(1-k_t) = 1,$

i.e., marginal q (adjusted for investment tax credits) must equal one at the optimum. This is the condition derived by Gordon(1981) for the case where the firm can repurchase pre-existing shares.

As we have seen in section 3, B_t affects $W_t(B_t, K_{t+1})$ in four ways and K_{t+1} affects $W_t(B_t, K_{t+1})$ in exactly the opposite directions. So it is reasonable to assume:

$$(7.11) \qquad \partial^2 W_t(B_t, K_{t+1})/(\partial B_t \, \partial K_{t+1}) < 0.$$

It then follows from (7.3) and (7.9) that $B_1(F)$ and $B_3(F)$ are increasing functions of F. Furthermore, it is easy to show that $B_1(F) \leq B_3(F)$ for any value of F.[22] If $B_1(F)$ and $B_3(F)$ are continuous functions of F, the optimal financing package B(F),y(F) will typically look like the solid lines in Figure 5. So the basic conclusion -- that marginal investment projects are financed by retentions and by debt if the firm's after-tax profits are large relative to investment, by new shares and debt if they are small relative to investment, and by debt alone in the intermediate case -- is the same as in the homogeneous case. If $B_1(F)$ and $B_3(F)$ are discontinuous functions of F, "case reversals" can occur. A typical example of case reversals is illustrated in Figure 6 where Case 2 is followed by Case 1 as F passes F_2.

If m is equal to one, the situation is basically the same as in the homogeneous case: from the viewpoint of the value maximizing firm, issuing new shares and cutting dividends are two perfectly indifferent ways of financing investment; the only difference from the homogeneous case is that the optimal debt-capital ratio $B_3(F)$ is not independent of investment.

The above discussion of the firm's real and financial decisions evolves around marginal q and we have not mentioned the "cost of capital" or the "return to investment" which are familiar concepts in the corporate finance literature. A natural question is whether or not they can be related to the above derivation of the necessary conditions for optimality. The answer is yes, with some important reservations. It is evident in the above derivation that (7.7) must hold in either case.[23] It turns out, not surprisingly, that $-W_B$ -- which is the decrese in the ex-dividend value of the firm $p_t^{\prime}Q_{t+1}$ when debt is increased by one dollar -- is closely connected to the "cost of capital" and q/(1-k) to the "return to investment." However, they cannot be expressed in terms of the corporate bond rate and the marginal value product of capital, unless B_t is small relative to K_{t+1}. If it is, the probability of bankruptcy in period t+1 is almost zero <u>and</u> Case 1 is bound to happen in t+1. What happens if it is known with certainty in period t that $M_{t+1} = 1$ and Case 1 occurs in period t+1? It follows from (3.4) that

$$(7.12) \qquad W_t(B_t, K_{t+1}) = L_t^1[(1-c_{t+1})V_{t+1}]/[1-L_t^1(c_{t+1})],$$

so that (assuming the order of taking derivatives and applying the pricing operator can be interchanged)

(7.13) $$W_B(B_t,K_{t+1}) = L_t^1[(1-c_{t+1})\partial V_{t+1}/\partial B_t]/[1-L_t^1(c_{t+1})],$$

and

(7.14) $$W_K(B_t,K_{t+1}) = L_t^1[(1-c_{t+1})\partial V_{t+1}/\partial K_{t+1}]/[1-L_t^1(c_{t+1})],$$

where $W_K(B_t,K_{t+1})$ is the partial derivative of $W_t(B_t,K_{t+1})$ with respect to K_{t+1}. But it is easy to show that

(7.15) $$\partial V_{t+1}/\partial B_t = -m_{t+1}[1+(1-u_{t+1})i_t],$$

and

(7.16) $$\partial V_{t+1}/\partial K_{t+1} = m_{t+1}[(1-u_{t+1})\partial \Pi_{t+1}/\partial K_{t+1} + (1-\delta)(1-k_{t+1})v_{t+1}],$$

if Case 1 holds in period t+1.[24] This result -- which holds for any value of B_t and K_{t+1} (not just for small B_t) for the case without adjustment costs -- is proved in Appendix 3. Combining (7.13) through (7.16) we can conclude that

(7.17) $$-W_B = L_t^1[(1-c_{t+1})m_{t+1}(1+(1-u_{t+1})i_t)]/[1-L_t^1(c_{t+1})],$$

and

(7.18) $$W_K = L_t^1[(1-c_{t+1})m_{t+1}[(1-u_{t+1})(\partial \Pi_{t+1}/\partial K_{t+1}) + (1-\delta)(1-k_{t+1})v_{t+1}]]/[1-L_t^1(c_{t+1})].$$

If we further assume that $c_{t+1} = 0$ and that m_{t+1}, u_{t+1}, k_{t+1} and v_{t+1} are known with certainty in period t, then (7.17) and (7.18) simplify to

$$(7.19) \qquad -W_B = \frac{m_{t+1}[1+(1-u_{t+1})i_t]}{1 + r_t},$$

and

$$(7.20) \qquad W_K = \frac{m_{t+1}[(1-u_{t+1})MVP_t/(1+\Delta_t) + (1-\delta)(1-k_{t+1})v_{t+1}]}{1 + r_t},$$

where r_t is the nominal interest rate on a safe, tax-free one period bond, MVP_t is the expected value of the marginal value product of capital:[25]

$$(7.21) \qquad MVP_t = E_t(\partial \Pi_{t+1}/\partial K_{t+1}),$$

and Δ_t is a sort of risk premium associated with the uncertain marginal value product of capital $\partial\Pi_{t+1}/\partial K_{t+1}$ as it is defined by

$$(7.22) \qquad L_t^1(\partial \Pi_{t+1}/\partial K_{t+1}) = MVP_t/[(1+\Delta_t)(1+r_t)].$$

Thus the optimality condition $q/(1-k) = -W_B$ reduces to

$$(7.23) \qquad \frac{(1-u_{t+1})MVP_t/(1+\Delta_t) + (1-\delta)(1-k_{t+1})v_{t+1}}{(1-k_t)v_t} = 1 + (1-u_{t+1})i_t.$$

If, on top of all this, we assume $k_{t+1} = k_t$ and $v_{t+1} = v_t$, then this simplifies to the familiar expression:

$$(7.24) \qquad \frac{MVP_t/(1+\Delta_t)}{(1-k_t)v_t} = i_t + \delta/(1-u_{t+1}).$$

This is (approximately) equivalent to the following expression which is even more familiar:

$$(7.25) \qquad \frac{E_t(\partial \Pi_{t+1}/\partial K_{t+1})}{(1-k_t)v_t} = i_t + \Delta_t + \delta/(1-u_{t+1}).$$

The left hand side is the "return to investment" and the right hand side is the "cost of capital."

However, apart from the assumptions (on tax rates and the price of investment goods) we have made, this familiar equality (7.25) will <u>never</u> hold at the optimum. If B_t is sufficiently small, the value function $W_t(B_t,K_{t+1})$ is linear in B_t [as (7.19) shows], so that the objective function $m_tX_t + W_t$ is linear in B_t. This implies that the optimal debt B_t is either infinitely negative or large enough to make (7.13) and (7.14) [from which (7.25) was derived] invalid.[26] If B_t is not small, neither W_B nor W_K (and hence $q = W_K/v_t$) has a simple expression like (7.13) or (7.14), because a change in B_t or in K_{t+1} alters the probability distribution (as of t) of M_{t+1} and because it is not certain as of t which Case will occur in t+1. Consequently, the cost of capital and the return to investment that appear in (7.25) capture only one effect of an increase in B_t and K_{t+1} on W_t, namely the direct effect on X_{t+1}. They do not capture the other three effects -- the resulting change in the discounting factor $\gamma_{t+1,j}$ ($j \geq 0$) (which affects the share price), in the corporate bond rate, and in the likelihood of $d_{t+1} = \bar{d}_{t+1}$. This seems to be a serious omission, particularly because the corporate bond rate in the real world <u>does</u> depend on leverage.

8. Qualifications

Although it carries a rich menu of tax rates, the model does not consider depreciation allowances for tax purposes on investment expenditure. We could incorporate them into the model along the line indicated in Hayashi(1982), but doing so would greatly complicate the analysis without altering the main results of the paper. It should however be noted that the sharp results we obtained in section 6 will not carry over to the case with depreciation allowances. The reason for this is as follows. Included in the firm's future cash flow is depreciation allowances for tax purposes yet to be claimed on past investments. If the market value of this stream of depreciation allowances were exogenous to the firm, we could subtract this market value from the firm's share price and carry out exactly the same analysis as we did in section 6. This will in fact be the case if there is a full loss offset and if shareholders can somehow secure that market value of depreciation allowances in the event of bankruptcy. Otherwise the firm's future action does influence the market value of depreciation allowances as it can alter the probability of bankruptcy. This is an element that was absent in the analysis in section 6.

The corporate bonds in our model have only one maturity. Relaxing this seems to be a rather straightforward task whose main part would be to formalize a rule to specify the shares of bondholders of various maturities in the event of bankruptcy.

Our analysis is a partial equilibrium one in the sense that the pricing mechnism in the asset market is taken as given. In order to analyze, for example, the effect of a change in tax rates or in the inflation rate on corporate behavior, we have to know how the pricing formula is affected by such changes. Analyzing it would require (like any other studies on the effect of taxes and inflation on corporate behavior) a complete specification of preference, technology and expectations formation, which clearly is well beyond the scope of the paper.

Footnotes

1. It is assumed for simplicity that the liquidation value of the firm is zero. In particular, as long as the liquidation value is proportional to the firm's capital stock, the analysis in sections 5 and 6 remains unchanged.

2. This is true in Japan where the dividend tax rate is .2 and the capital gains tax rate is zero for individuals. If the dividend tax rate depends on the shareholders' income, θ represents the marginal tax rate for the shareholders. See Miller(1977).

3. If $p_t > (1-\theta_t)d_t + p_t' - c_t(p_t'-p_t)$, individuals can make unlimited profits by selling the stock short cum-dividend and buy it back ex-dividend. If the reverse inequality holds, individuals can buy the stock cum-dividend and sell it ex-dividend. Thus $p_t = (1-\theta_t)d_t + p_t' - c_t(p_t'-p_t)$ if the stock price is set by these individual arbitragers. However, for corporations in most countries capital gains and losses are part of their corporate income. Furthermore, in the U.S. at least, short term capital gains for individuals are taxed as ordinary income. Thus a sizable fraction of the agents in the stock market are those for which $\theta_t = c_t$. If the stock price is set by such agents, then m_t in (2.1) should be one. See Kalay(1982) for more details. Most empirical studies [see Auerbach(1982) and Kalay(1982) for a survey] show that stock prices drop by significantly less than the value of the dividend on the ex-dividend day, implying that m_t is less

than one. What is crucial in the present paper is not that m_t is written as $(1-\theta_t)/(1-c_t)$, but that m_t is less than one. We will consider separately the case where $m_t = 1$.

4. We assume a full loss offset, so the firm can qualify for a rebate of $u_t i_{t-1} B_{t-1}$ dollars when profits for tax purposes, $(1-u_t)\Pi_t - (1+i_{t-1})B_{t-1}$, is negative. This assumption can easily be relaxed. Just define a dummy variable D_t for the sign of profits for tax purposes and replace u_t by $u_t D_t$; the formal analysis in this paper will hold without any modifications.

5. The capital asset pricing model is a special case of this. Another example is the so-called consumption based capital asset pricing model which implies $L_t^j(x_{t+j}) = E_t(y_{t,j} x_{t+j})$ where E_t is the conditional expectation operator and $y_{t,j} = \delta^j u'(C_{t+j})/u'(C_t)$ with $u(.)$ = utility function of a "representative" consumer, δ = subjective rate of time preference, and C = consumption.

6. If the firm is a price taker, output and factor prices and the price of investment goods are the parameters that characterize the demand and supply functions.

7. The Markovness assumption is not really crucial for the analysis that follows, but it clarifies it.

8. Thus Z_t for example includes r_t, the nominal interest rate on a safe, tax-free one-period bond. Macro variables such as money supply will be included in Z_t if they influence either the values of the exogenous variables or the function-

al form of the pricing operator.

9. So $K_{t+1} = (1-\delta)K_t+F_t$ effectively is in I_t.

10. See section 6 for what happens when this constraint is absent.

11. It is also a function of i_t, but, as (2.12) shows, i_t in turn is a function of B_t, K_{t+1}, Z_t, and $(\mu^*_{t+1}, \mu^*_{t+2}, \ldots)$.

12. Otherwise the problem of time inconsistency [Lucas and Sargent(1981)] will arise.

13. To prove that $L^1_t(Y_{t+1})$ is independent of B_t would require mathematical induction starting from the terminal period of the firm's horizon. The entire result of the paper will carry over to the case where the firm's planning horizon is finite rather than infinite.

14. This does not necessarily mean that the technology shock is multiplicative.

15. This π' should not be confused with the marginal value product of capital. Note that π also depends on the technology shock and output and factor prices.

16. The same line of proof was independently found by Andrew B. Abel.

17. Since λ_1 maximizes $m\lambda + h(\lambda)$ and λ_3 maximizes $m + h(\lambda)$, we have $m\lambda_1 + h(\lambda_1) > m\lambda_3 + h(\lambda_3)$ and $\lambda_3 + h(\lambda_3) > \lambda_1 + h(\lambda_1)$. This and $m < 1$ imply $\lambda_1 < \lambda_3$. We are assuming here that the maximizer of $m\lambda + h(\lambda)$ or $\lambda + h(\lambda)$ is unique, which is a reasonable assumption since the functional form of h depends only on Z_t; it will be only by accident that the

maximizer of $m\lambda + h(\lambda)$ or $\lambda + h(\lambda)$ is not unique. We also assume that λ_1 is positive. A sufficient condition for this is that the function $h(\lambda)$ is linear for nonpositive λ and its slope for nonpositive λ, which is equal to w_t^2 in (3.7), is less than m_t (in absolute value).

18. Since λ_1 and λ_3 are unique maximizer of $m\lambda + h(\lambda)$ and $\lambda + h(\lambda)$, respectively, "case reversals" cannot occur. For example, it cannot happen that Case 2 is followed by Case 1 as f keeps increasing.

19. So the corporate bond rate i_t depends on B_t and K_{t+1} only through their ratio B_t/K_{t+1}. This is not true in the non-homogeneous case.

20. Note that $\pi'(f)$ is negative. The cost of investment must not be confused with the "cost of capital" which we will define in the next section.

21. The proof is similar to the argument in footnote 17.

22. The proof is essentially the same as footnote 17.

23. For example for Case 1, (7.3) and (7.4) imply (7.7).

24. If Case 3 holds in period t+1, (7.15) and (7.16) hold with m_{t+1} replaced by one. (7.15) holds also for the case with adjustment costs.

25. If the firm is a price taker, $\partial \Pi / \partial K$ is equal to the marginal product of capital multiplied by the output price.

26. Of course, in the knife edge case where m_{t+1} is equal to the right hand side of (7.17), the optimal debt B_t is indeterminate.

APPENDIX 1

This appendix presents a formal definition of the pricing operator L_t and a proof of the pricing formula (2.8). The former closely parallels Ross(1978) and Hansen, Richard and Singleton(1982) and the latter Brock(1978).

Let M_{t+j} be a set of random nominal after-tax payoffs in period t+j. More formally, let $\{z_t\}$ (t=0,1,2...) be a sequence of random vectors defined on a probability space $(\Omega, \mathcal{F}, P)$. We call the sequence up to t, $\{z_0, z_1, .., z_t\}$, the information set at t and denote it by I_t. Let $\mathcal{F}_t$ be the sigma field generated by I_t. M_{t+j} is a set of functions from Ω to R that are $\mathcal{F}_{t+j}$-measurable. By definition, M_{t+j} is a linear space. Associated with M_{t+j} and I_t is a mapping L_t^j from M_{t+j} to R. We assume:

(A1.1) L_t^j is a linear operator so that

$$L_t^j(\mu_t x_{t+j} + \lambda_t y_{t+j}) = \mu_t L_t^j(x_{t+j}) + \lambda_t L_t^j(y_{t+j}),$$

for any μ_t and λ_t in I_t (i.e., any μ_t and λ_t that are $\mathcal{F}_t$-measurable.

Since $L_{t+j}^k(x_{t+j+k})$ is in M_{t+j}, it can be priced by L_t^j. We assume:

$$\text{(A1.2)} \qquad L_t^j(L_{t+j}^k(x_{t+j+k})) = L_t^{t+k}(x_{t+j+k}).$$

Ross(1978) has proved (i) that if M is a set of payoffs that can be spanned by available marketed assets and if L(x) is the market price of an asset whose payoff is x, then the absence of arbitrage opportunities implies that L satisfies (A1.1) and (A1.2), and (ii) that L can be extended to the space of payoffs that includes non-marketed assets as well as marketed assets (although the extension is not unique).

We now prove (2.8). By multiplying both sides of (2.4) by Q_t and using $Q_{t+1} = (1+g_t)Q_t$, we obtain

$$\text{(A1.3)} \qquad p_tQ_t = m_tX_t + \beta_tp_t'Q_{t+1}.$$

From (2.6) we get

$$\text{(A1.4)} \qquad p_t'Q_{t+1} = L_t^1(\alpha_{t+1}p_{t+1})Q_{t+1}$$

$$= L_t^1(\alpha_{t+1}p_{t+1}Q_{t+1}) \qquad \text{(since } Q_{t+1} \text{ is in } I_t).$$

Thus from (A1.3) and (A1.4) we obtain

$$\text{(A1.5)} \qquad p_tQ_t = m_tX_t + L_t^1(\beta_t\alpha_{t+1}p_{t+1}Q_{t+1}).$$

By shifting time forward by one period on (A1.3) and multiplying both sides by $\beta_t \alpha_{t+1}$ we obtain

$$\beta_t \alpha_{t+1} p_{t+1} Q_{t+1} = \beta_t \alpha_{t+1} m_{t+1} X_{t+1} + \beta_t \alpha_{t+1} \beta_{t+1} p'_{t+1} Q_{t+2}.$$

Apply L_t^1 on this to get

$$L_t^1(\beta_t \alpha_{t+1} p_{t+1} Q_{t+1}) = L_t^1(\beta_t \alpha_{t+1} m_{t+1} X_{t+1}) + L_t^1(\beta_t \alpha_{t+1} \beta_{t+1} p'_{t+1} Q_{t+2}).$$

This last term equals

$$L_t^1[\beta_t \alpha_{t+1} \beta_{t+1} L_{t+1}^1(\alpha_{t+2} p_{t+2} Q_{t+2})] \qquad \text{(by A1.4)}$$

$$= L_t^1[L_{t+1}^1(\beta_t \alpha_{t+1} \beta_{t+1} \alpha_{t+2} p_{t+2} Q_{t+2})] \qquad \text{(by A1.1)}$$

$$= L_t^2(\gamma_{t,2} p_{t+2} Q_{t+2}). \qquad \text{[by A1.2 and (2.9)]}$$

Thus we have

$$L_t^1(\gamma_{t,1} p_{t+1} Q_{t+1}) = L_t^1(\gamma_{t,1} m_{t+1} X_{t+1}) + L_t^2(\gamma_{t,2} p_{t+2} Q_{t+2})$$

By the same argument we can easily show that

$$\text{(A1.6)} \quad L_t^j(\gamma_{t,j} p_{t+j} Q_{t+j}) = L_t^j(\gamma_{t,j} m_{t+j} X_{t+j}) + L_t^{j+1}(\gamma_{t,j+1} p_{t+j+1} Q_{t+j+1})$$

By summing (A1.6) over j, we obtain

$$p_t Q_t = \sum_{j=0}^{N-1} L_t^j(\gamma_{t,j} m_{t+j} X_{t+j}) + L_t^N(\gamma_{t,N} p_{t+N} Q_{t+N}).$$

If we assume the transversality condition

$$\lim_{N \to \infty} L_t^N(\gamma_{t,N} p_{t+N} Q_{t+N}) = 0,$$

we obtain the desired result (2.8).

APPENDIX 2

This appendix proves the following theorem:

Theorem. If the production function $G_t(N_t, K_t, F_t, e_t)$ is homogeneous of degree one in (N_t, K_t, F_t) and if the firm is a price taker, then $V_t(B_{t-1}, K_t; Z_t)$ is homogeneous of degree one in (B_{t-1}, K_t).

Proof. Let $\{x^0_{t+j}\}$ $(j\geq 0)$ be the stochastic process generated by the optimal decision rule $(\mu^*_t, \mu^*_{t+1}, \ldots)$ with the initial condition $(B_{t-1}, K_t) = (B^0_{t-1}, K^0_t)$, where x_{t+j} stands for $(B_{t+j-1}, K_{t+j}, F_{t+j}, Y_{t+j}, M_{t+j}, X_{t+j})$. For the initial condition $(B_{t-1}, K_t) = (\lambda B^0_{t-1}, \lambda K^0_t)$ consider the following decision rule h^λ:

$$h^\lambda_{x,t+j}(B_{t+j-1}, K_{t+j}; Z_{t+j}) = \lambda h^*_{x,t+j}(B^0_{t+j-1}, K^0_{t+j}; Z_{t+j}) \qquad (x = B, F)$$

and

$$h^\lambda_{x,t+j}(B_{t+j-1}, K_{t+j}; Z_{t+j}) = h^*_{x,t+j}(B^0_{t+j-1}, K^0_{t+j}; Z_{t+j}) \qquad (x = y, M),$$

where h_x represents a decision rule for x. Let $\{x^\lambda_{t+j}\}$ be the stochastic process generated by the decision rule h^λ when the initial condition is $(\lambda B^0_{t-1}, \lambda K^0_t)$. Clearly $K^\lambda_{t+j} = \lambda K^0_{t+j}$ for any realization of $\{Z_{t+j}\}$ and for all $j\geq 0$. It then follows from the hypotheses in the Theorem that $X^\lambda_{t+j} = \lambda X^0_{t+j}$ for any realization of z and for all $j \geq 0$, so that the discounting factor $\{\gamma_{t,j}\}$ $(j \geq 0)$ takes the same value under the two decision rules h^* and h^λ for all $j \geq 0$. Therefore we can conclude that:

the value of the firm with $(\lambda B^0_{t-1}, \lambda K^0_t)$ under h^λ

$$= \lambda V_t(B^0_{t-1}, K^0_t; z_t).$$

But since the left hand side is less than or equal to the value of the firm with $(\lambda B^0_{t-1}, \lambda K^0_t)$ under the optimal decision rule h^*, we have

$$V_t(\lambda B^0_{t-1}, \lambda K^0_t; z_t) \geq \lambda V_t(B^0_{t-1}, K^0_t; z_t).$$

Exactly the same argument gives

$$V_t(B^0_{t-1}, K^0_t; z_t) \geq (1/\lambda) V_t(\lambda B^0_{t-1}, \lambda K^0_t; z_t).$$

These two inequalities imply the desired result.

Remark 1. To implement the decision rule h^λ at time $t+j$, the firm (with the initial condition $\lambda B^0_{t-1}, \lambda K^0_t$) has to know B^0_{t+j-1} and K^0_{t+j} which are functions of $(z_t, z_{t+1}, \ldots, z_{t+j-1})$. So if the firm knows just z_{t+j} but not its past realized values, the decision rule h^λ cannot be implemented. We can avoid this difficulty by redefining z_t to be $(z_0, z_1, \ldots, z_t)$.

Remark 2. Since $W_t(B_t, K_{t+1}) = L_t^1(\alpha_{t+1} V_{t+1}(B_t, K_{t+1}; Z_{t+1}))$ [see equation (3.6) in the text] and since α_{t+1} does not depend on the size of the firm [see (2.7) and (3.1)], our Theorem immediately implies that $W_t(B_t, K_{t+1})$ also is linearly homogeneous.

APPENDIX 3

This section proves the following theorem:

Theorem. Suppose the value function (for the case without adjustment costs) $V_t(B_{t-1},K_t;Z_t)$ is a concave function in (B_{t-1},K_t) in a neighborhood of (B^0_{t-1},K^0_t) and suppose the profit function $\Pi_t(K_t)$ is concave and differentiable in K_t. Then V_t is a differentiable function of B_{t-1},K_t at (B^0_{t-1},K^0_t) if either Case 1 or Case 2 occurs in the neighborhood of (B^0_{t-1},K^0_t). The derivatives are given by:

$$\partial V_t(B^0_{t-1},K^0_t;Z_t)/\partial B_{t-1} = -m_t[1+(1-u_t)i_{t-1}]$$

and

$$\partial V_t(B^0_{t-1},K^0_t;Z_t)/\partial K_t = m_t[(1-u_t)\partial\Pi_t/\partial K_t + (1-\delta)(1-k_t)v_t],$$

if Case 1 occurs in period t, and

$$\partial V_t(B^0_{t-1},K^0_t;Z_t)/\partial B_{t-1} = -[1+(1-u_t)i_{t-1}]$$

and

$$\partial V_t(B^0_{t-1},K^0_t;Z_t)/\partial K_t = (1-u_t)\partial\Pi_t/\partial K_t + (1-\delta)(1-k_t)v_t,$$

if Case 3 occurs in period t.

The proof of this theorem is essentially the same as the proof of Theorem 1 in Benveniste and Scheinkman(1979), so we do not repeat it here. We merely point out that V_t is m_tX_t +

$W_t(B_t, K_{t+1})$ if Case 1 occurs in period t and $X_t + W_t(B_t, K_{t+1}) + (m_t-1)\bar{d}_t Q_t$ if Case 3 occurs in period t, and that X_t can be written as

$$X_t = (1-u_t)\Pi_t(K_t) - (1-k_t)v_t[K_{t+1}-(1-\delta)K_t] + B_t - [1+(1-u_t)i_{t-1}]B_{t-1}.$$

References

Auerbach, Alan J. (1979), Wealth maximization and the cost of capital, Quarterly Journal of Economics, Vol.93, August, pp. 433-446.

--------- (1982), Taxes, firm financial policy and the cost of capital, NBER Working Paper No.955.

Benveniste, L.M. and J.A. Scheinkman (1979), On the differentiability of the value function in dynamic models of economics, Econometrica, Vol.47, No.3, pp. 727-732.

Brock, William (1978), Asset prices in a production economy, in McCall, John J. ed., The Economics of Information and Uncertainty, University of Chicago Press, pp. 1-46.

Chirinko, Robert S. and Stephen R. King (1982), Hidden stimuli to capital formation: debt and the non-adjustment of financial returns, Mimeo., Northwestern University, July.

Gordon, Roger H. and Burton G. Malkiel (1981), Corporation finance, in Pechman, Joseph A. and Henry J. Aaron, eds., How Taxes Affect Economic Behavior, Brooking Institution, pp. 131-192.

Gordon, Roger H. (1981), Inflation, taxation and corporate behavior, Mimeo., Bell Laboratories.

Hansen, Lars P., Scott F. Richard, and Kenneth J. Singleton (1982), Testable implications of the intertemporal capital asset pricing model, Mimeo., Carnegie-Mellon University, June.

Hayashi, Fumio (1982), Tobin's marginal q and average q: a neoclassical interpretation, Econometrica, Vol.50, No.1, pp. 213-224.

Kalay, Avner (1982), The ex-dividend day behavior of stock prices: a re-examination of the clientele effect, Journal of Finance, Vol.37, No.4, pp. 1059-70.

Lucas, Robert E.. Jr. (1967), Adjustment costs and the theory of supply, Journal of Political Economy, Vol.75, No.4, pp. 321-334.

-------- and Edward C. Prescott (1971), Investment under uncertainty, Econometrica, Vol.39, No.5, pp. 659-682.

------- and Thomas J. Sargent (1981), Introduction to Rational Expectations and Econometric Practice, University of Minnesota Press.

Miller, Merton H. (1977), Debt and taxes, Journal of Finance, Vol.32, No.2, pp. 261-75.

Modigliani, Franco and Merton H. Miller (1958), The cost of capital, corporation finance, and the theory of investment, American Economic Review, Vol.48, No.3, pp. 261-97.

------- and ------ (1963), Corporate income taxes and the cost of capital: a correction, American Economic Review, Vol.53, No.3, pp. 433-443.

Poterba, James M. and Laurence H. Summers (1982), Dividend taxes, corporate investment, and "Q", NBER Working Paper No.829.

Ross, Stephen A. (1978), A simple approach to the valuation of risky assets, Journal of Business, Vol.51, No.3, pp. 453-475.

Stiglitz, Joseph E. (1969), A re-examination of the Modigliani-Miller theorem, American Economic Review, Vol.59, No.5, pp. 784-93.

Summers, Laurence H. (1981), Taxation and corporate investment: a q-theory approach, Brooking Papers on Economic Activity, No.1, pp. 67-127.

Tobin, James (1969), A general equilibrium approach to monetary theory, Journal of Money, Credit and Banking, Vol.1, No.1, pp. 15-29.

Figure 1

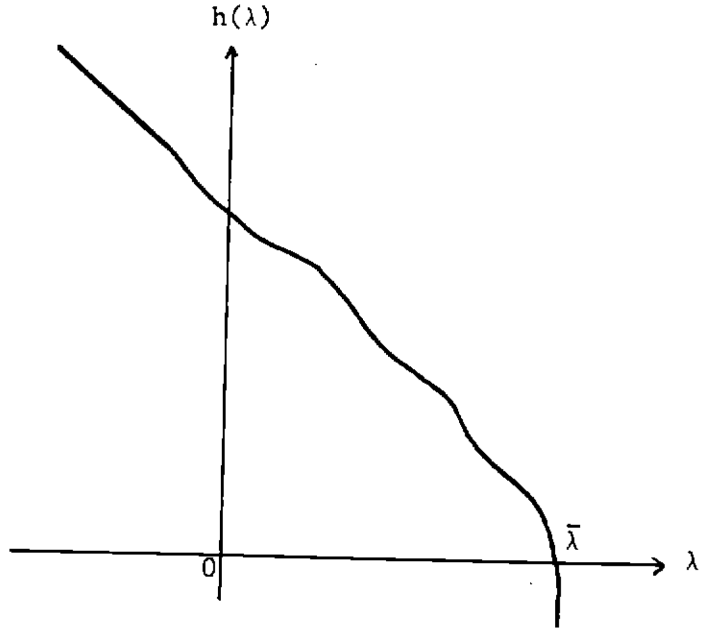


Figure 2

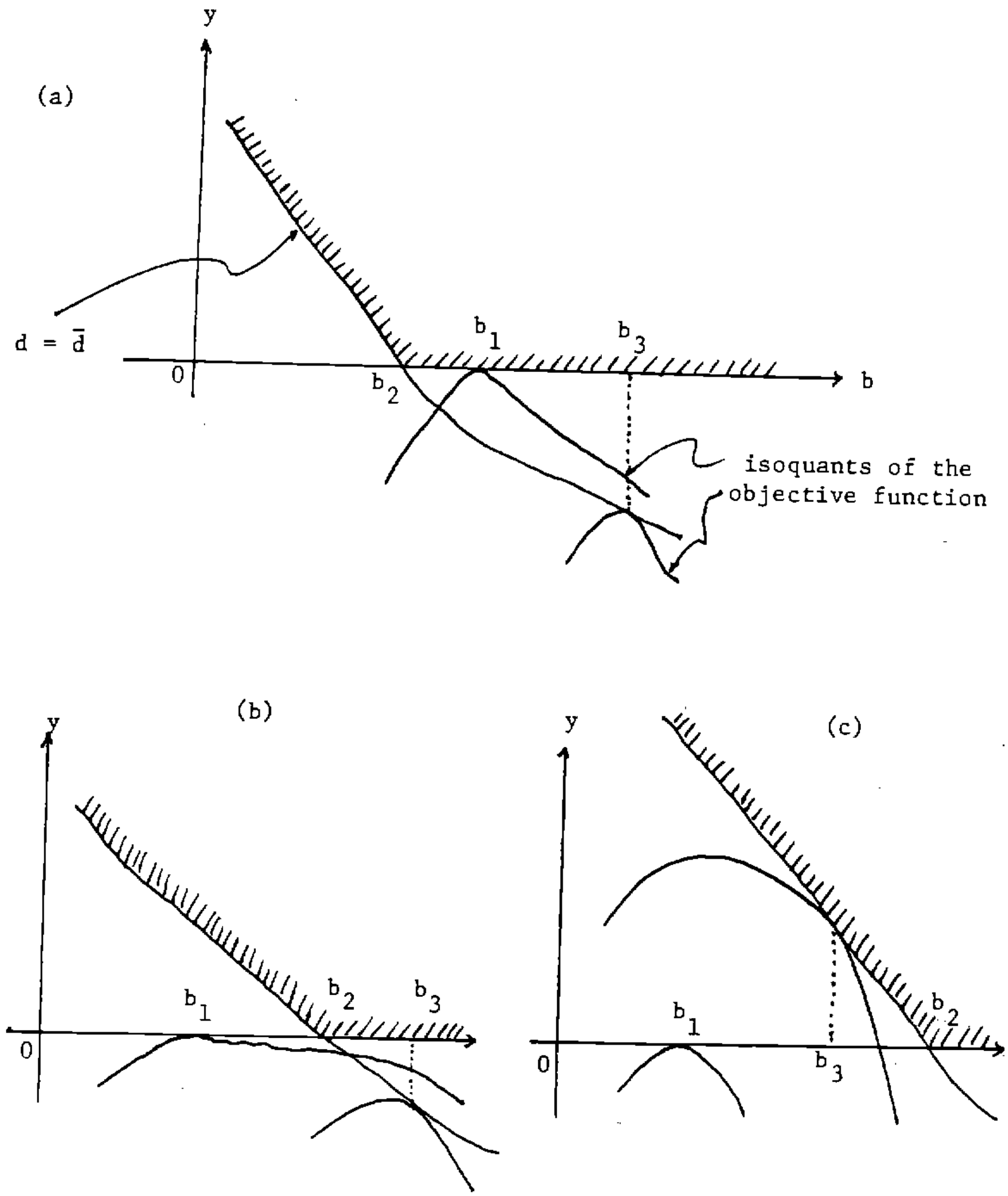

(a)
y
d = $\bar{d}$
0
b_1
b_2
b_3
b
isoquants of the
objective function
(b)
y
0
b_1
b_2
b_3
(c)
y
0
b_1
b_2
b_3

Figure 3

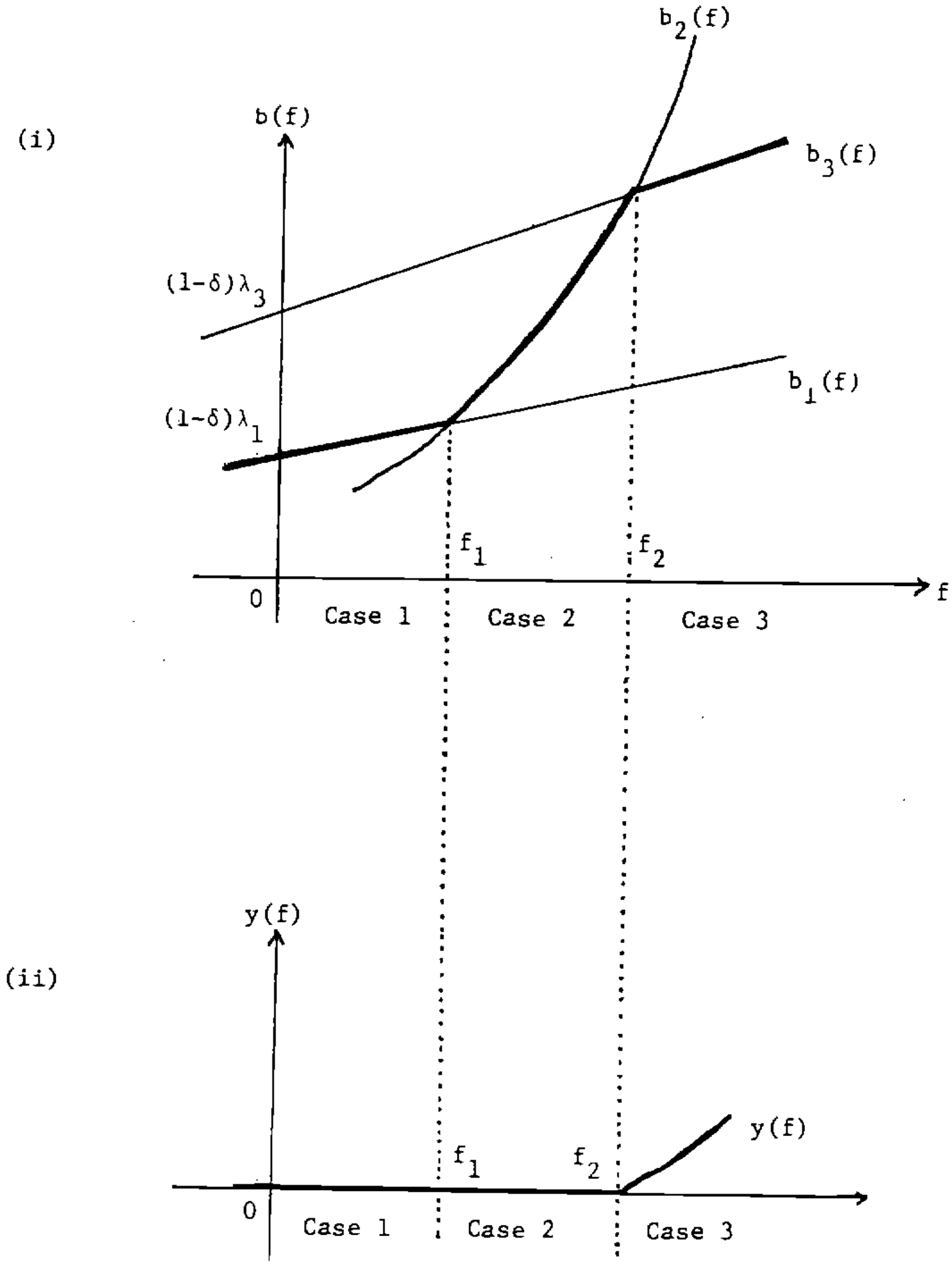

(i)
b(f)
$b_2(f)$
$b_3(f)$
$(1-\delta)\lambda_3$
$b_1(f)$
$(1-\delta)\lambda_1$
f_1
f_2
f
0
Case 1
Case 2
Case 3
(ii)
y(f)
y(f)
f_1
f_2
0
Case 1
Case 2
Case 3

Figure 4

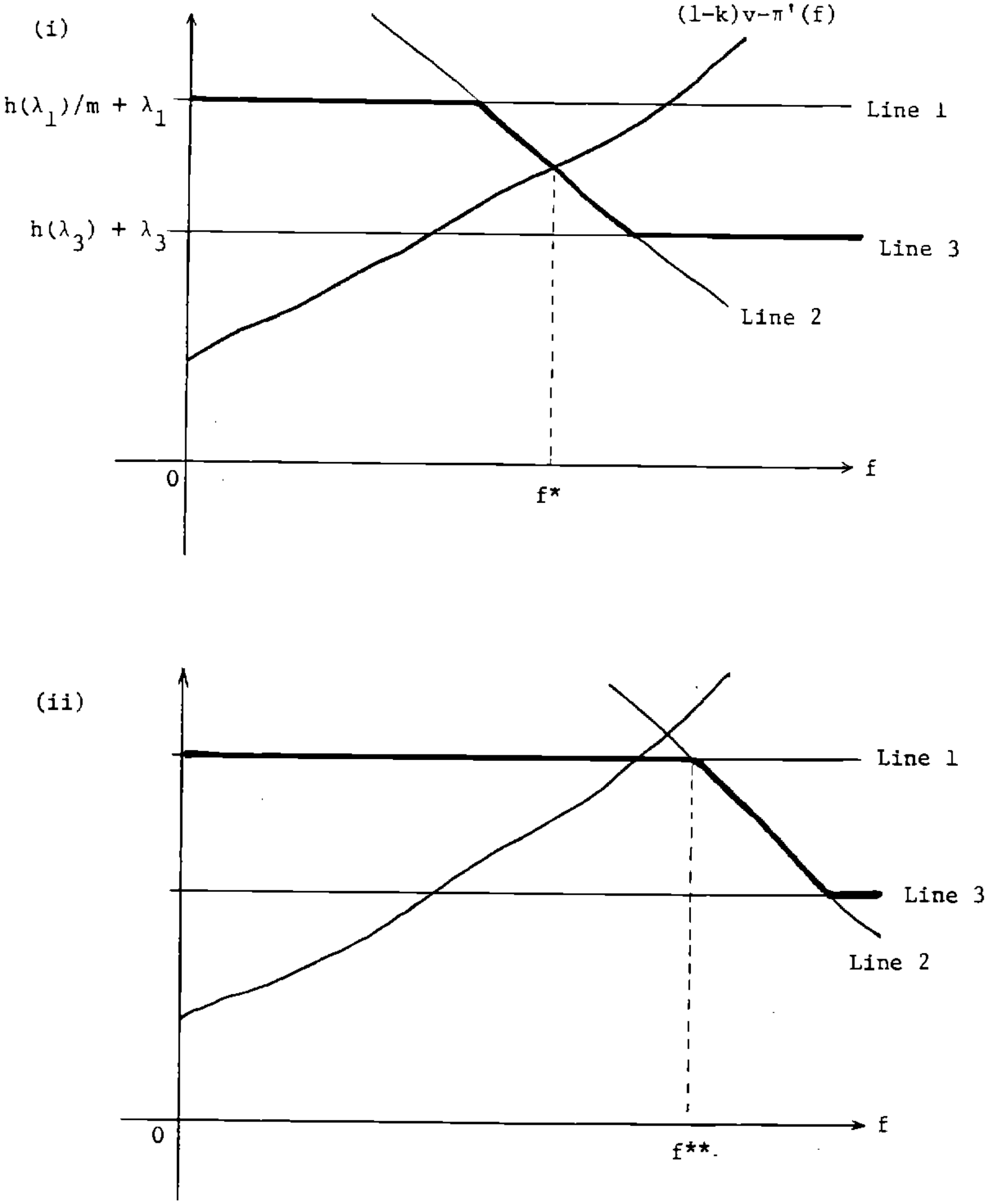

(i)
(1-k)v-π'(f)
h(λ1)/m + λ1
Line 1
h(λ3) + λ3
Line 3
Line 2
0
f*
f
(ii)
Line 1
Line 3
Line 2
0
f**
f

Figure 5

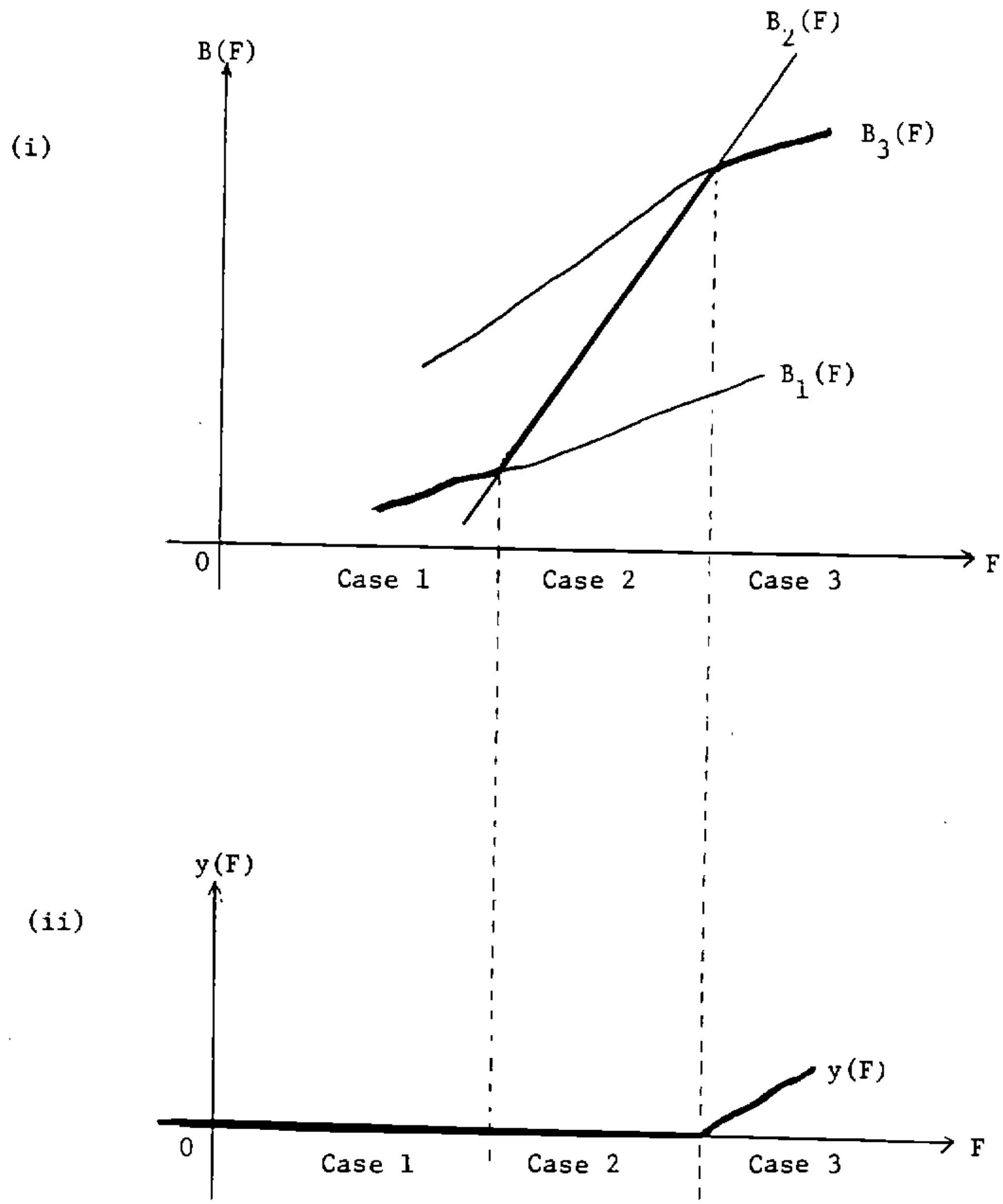
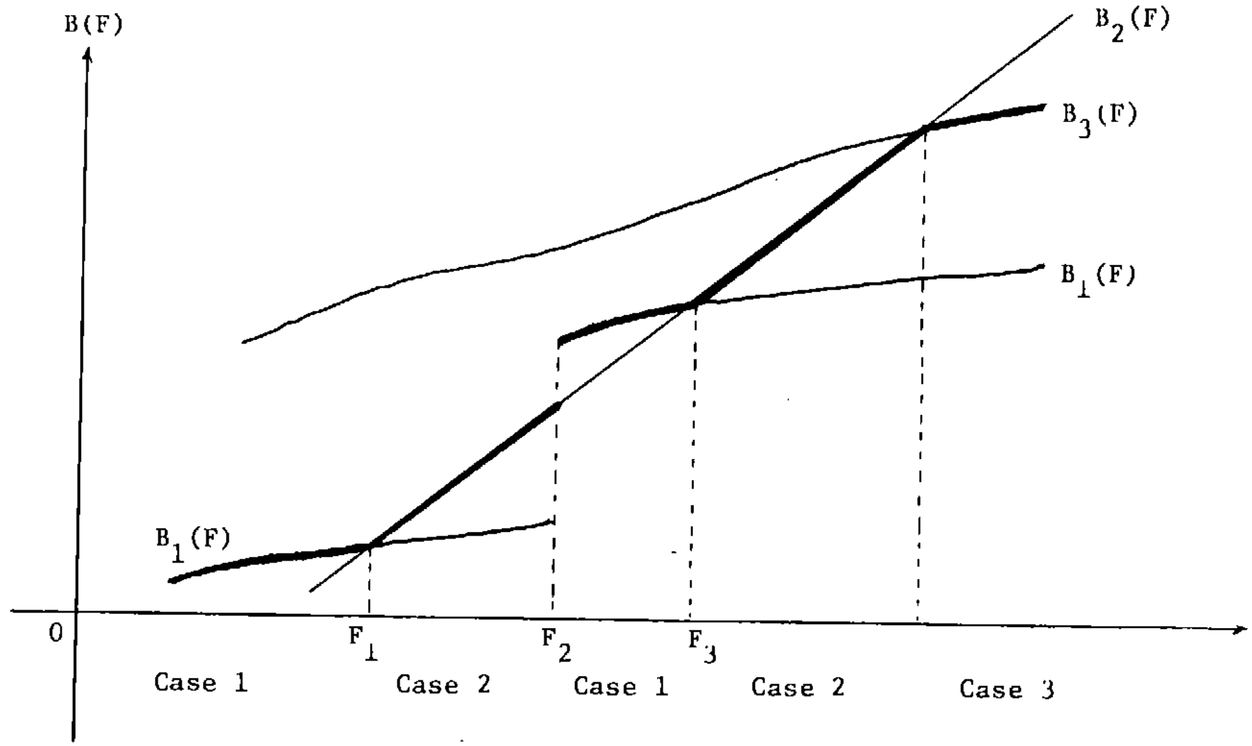

(i)
B(F)
$B_2(F)$
$B_3(F)$
$B_1(F)$
0
F
Case 1
Case 2
Case 3
(ii)
y(F)
y(F)
0
F
Case 1
Case 2
Case 3

Figure 6



LIST OF SYMBOLS

B_t: nominal value of corporate debt issued in period t

B_1: implicitly defined by (7.3)

B_2: defined by (7.6)

B_3: implicitly defined by (7.9)

b_t: B_t/K_t, ratio of B_t to the beginning-of-the-period capital stock

b_1: defined by (6.3)

b_2: defined by (6.8)

b_3: defined by (6.13)

c: capital gains tax rate

d: nominal amount of dividend per share

$\bar{d}$: lower bound for d

e: technology shock

F: investment

f_t: F_t/K_t, ratio of F to the beginning-of-the-period capital stock

G: production function

g_t: rate of growth of the number of shares, i.e.,

$g_t = (Q_{t+1}-Q_t)/Q_t$

h_t: defined by (5.2); equals $p_t'Q_{t+1}/K_{t+1}$

I_t: information set at the begenning of period t; consists of B_{t-1}, K_t, Z_t

i: corporate bond rate

K_t: capital stock at the beginning of period t

k: rate of investment tax credit

L: pricing operator

M: bankruptcy dummy; equals 0 if the firm goes broke and one otherwise

MVP: defined by (7.21); expected value of the marginal value product of capital

m: $(1-\theta)/(1-c)$

N: vector of variable factor inputs

p: share price of the currently operating firm

p': ex-dividend share price of the currently operating firm

Q_t: number of pre-existing shares at the beginning of the period

q: marginal q, defined by (7.1)

R: interest payment plus principal, defined by (3.12)

r_{jt}: nominal interest rate on a default-free, tax-free j-period bond

u: corporate tax rate

V: value of the firm; equals pQ

v: price of investment goods

W: defined by (3.4); equals $p_t'Q_{t+1}$

X: defined by (2.3); cash flow plus debt issue minus interest payments and principal

Y: defined by (4.3)

y_t: $g_t/(1+g_t)$

Z_t: state of the world for period t

α: defined by (2.7)

β: defined by (2.10)

γ: discounting factor, defined by (2.9)

δ: exponential rate of phisical depreciation

θ: tax rate on dividend income

θ': tax rate on interest income

λ_t: "debt-capital ratio," defined as B_t/K_{t+1}

λ_1: defined implicitly by (6.2)

λ_3: defined implicitly by (6.12)

$\bar{\lambda}$: upper bound for λ

μ: decision rule of the firm

Π: before-tax profits where variable factor inputs are already maximized out

π_t: Π_t/K_t

ω: inverse function of $-\pi(f)$

Δ: defined by (7.22); "risk premium" associated with the next period's before-tax profits.