No.176 (82-43)

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

by Fumio Hayashi

University of Tsukuba, Japan

December 1982

The author is grateful to the participants in workshops at the 1982 NBER Summer Institute, Osaka University and University of Tokyo for their commnets and discussios. Remaining errors are his own.

Real and Financial Decisions of a Firm

with Bankruptcy and Dafault: 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 the uncertain dividend stream. The main result of the paper is as follows: if the firm's after-tax profit is small relative to investment, the firm finances investment by retained earnings and debt; if it is large relative to investment, financing is done through new share issues and debt; in the intermediate case investment is financed entirely through debt. The optimality condition evolves around "Tobin's q" and it is claimed that the familiar optimality condition of equating the "cost of capital" and the "return to investment" is inappropriate as it holds only when the value of the firm is independent of its financial structure.

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

(0298) 53-5556

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 (6.12)

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 (3.11)

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 (3.6); 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 (6.11)

$\bar{\lambda}$: 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.

1. Introduction

The purpose of this paper is to provide a framework for analyzing the interaction between corporate investment and financial policies. Almost all of the huge literature on investment has been concerned with the optimal level of investment 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 entirely financed by retained earnings [Hayashi(1982)] or that a constant fraction of investment is financed by debt [Summers(1981) and Poterba and Summers(1982)]. The theory which posits that the optimal capital stock is determined at the equality of the "cost of capital" and the marginal product of capital makes similar assumptions in deriving the expression for the cost of capital [see, e.g., Chirinko and King(1982)]. On the other hand, much of the equally huge literature on corporate finance [with a possible exception of Gordon(1982)] has been concerned with the determination of 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 the level of investment is independent of how it is financed. This is 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 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.

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 result can be summarized as follows. If the firm's after-tax profit is low relative to the level of investment, the firm finances investment by retained earnings <u>and</u> debt. If it is abundant relative to investment, financing is done through new share issues <u>and</u> debt. In the intermediate case, investment is financed <u>entirely</u> through debt.

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 the homogeneous case where the production function exhibits constant returns to scale and where there are adjustment costs associated with investment. It is shown in section 5 that the value of the firm is proportional to its capital stock. The results derived in section 6 is sharp: the ratio of debt to the end-of-period capital stock is independent of investment or else the fraction of investment financed by debt is 100%. In section 7, we consider the case where there are no adjustment costs and where the production function shows decreasing returns to scale. There we will show 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 because it is a straightforward combination of the two cases considered in sections 5, 6 and 7. Neither will we consider the case without adjustment costs and with constant returns to scale because then the scale of the firm is indeterminate. Section 8 conctains concluding remarks.

2. Pricing Formula with Bankruptcy and Default

We consider the behavior of a competitive 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 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 $g_t Q_t$ 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 $d_t Q_t$ 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 rate c. All stockholders face the same tax rates, θ and c.[2]

If there are no arbitrage opportunities between p_t and p_t', we must have

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

i.e.,

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

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

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 current output and factor prices, but the dependence of Π_t on those variables is left implicit in $\Pi_t(.)$.

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.[3] 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.,

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

where

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

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

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

where

$$\gamma_{t,j} = \prod_{k=1}^{j} \beta_{t+k-1}\alpha_{t+k}, \quad \text{with } \gamma_{t,0} = 1, \tag{2.9}$$

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

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

We can also incorporate defaults. Bondholders can receive full amount $(1+i_t)B_t$ if the firm is not bankrupt at t+1. In the event of bankruptcy, 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:

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

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

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

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. We assume that the firm is competitive in the sense that it cannot influence the following "exogenous" variables: output and factor prices, the price of investment goods (v), tax rates (θ, θ', c, k), and the shock to the production function (e_t). We assume that these exogenous variables are part of a larger set of variables Z_t which follow a Markov process.[4] The vector Z_t will be referred to as the state of nature as it completely determines the stochastic properties of the exogenous variables that the firm will face in the future. The firm's current return $m_t X_t$ in (2.8) also depends on the firm's state variables: B_{t-1}, K_{t-1}. These are given to the firm at the beginning of period t. The firm's information set I_t consists of the state of nature Z_t and the firm's state variables B_{t-1}, K_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.[5] This function is called the firm's decision rule 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 the value of the firm, $V_t = p_t Q_t$. The functional form of the pricing operator L_t will depend on

Z_t. We assume that 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. In particular, the firm can figure out its highest possible value if it stays in business. That highest possible value of the firm will be a function of I_t, $V_t(I_t)$.

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:

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

If it stays in business, the firm's optimization problem is to find a sequance $(\mu_t^*, \mu_{t+1}^*, \ldots)$ of optimal decision rule that maximizes V_t. But there are a few constraints that the firm has to follow. 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 simply says dividends must be nonnegative.) In some countries (e.g., Great Britain) repurchase of equity is illegal. The second constraint, therefore, is that g be nonnegative.[6] The third constraint is the ever-present capital accumulation constraint. Thus the (currently operating) firm's problem is

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

where V_t is given by (2.8) and δ is the exponential rate of phisical depreciation. The value of V_t evaluated at the sequence of optimal decision rule $(\mu^*_t, \mu^*_{t+1}, \ldots)$ is, of course, $V_t(I_t)$.

In the firm's optimization problem just formulated, the current return $m_t X_t$ is not influenced by the firm's future action.[7] Furthermore, we can easily show from (2.6) and (2.8) that

$$(3.3) \qquad V_t = m_t X_t + \beta_t L^1_t(\alpha_{t+1} V_{t+1}),$$

and

$$(3.4) \qquad L^1_t(\alpha_{t+1} V_{t+1}) = p'_t Q_{t+1}.$$

Using this and the properties of L_t discussed in the previous section, it can be shown that the firm's current action (g_t, B_t, F_t) implied by the optimal decision rule μ^*_t solves the familiar DP (dynamic programming) algorithm:

$$(3.5) \quad V_t(I_t) = \max_{g,B,F} [m_t X_t + \beta_t L^1_t(\alpha_{t+1} V_{t+1}(I_{t+1}))]$$

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

The value of $L_t^1[\alpha_{t+1}V_{t+1}(I_{t+1})]$ is a function of B_t, K_{t+1}, and Z_t. We submerge the dependence in the functional form and write

$$(3.6) \qquad L_t^1[\alpha_{t+1}V_{t+1}(I_{t+1})] = W_t(B_t,K_{t+1}).$$

An increase in B_t affects W_t in four ways. The first is its direct effect on X_{t+1}. Second, it increases the probability of bankruptcy in t+1, which depresses the discounting factor $\gamma_{t+1,j}$ $(j \geq 1)$. Third, as a result of it the corporate bond rate i_t will go up. 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 direction.

From (2.2) and (3.4) we have

$$(3.7) \qquad d_tQ_t = X_t + g_tp_t'Q_t$$
$$= X_t + y_tW_t(B_t,K_{t+1}),$$

where

$$(3.8) \qquad y_t = g_t/(1+g_t).$$

With all these notation we can rewrite the optimization problem (3.5) as

$$(3.9)\quad \max_{y,B,F} \; m_t X_t(B,F) + (1-y+m_t y)W_t(B,(1-\delta)K_t+F)$$

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

$$y \geq 0,$$

where

$$(3.10)\quad X_t(B,F) = (1-u_t)\Pi_t(F,K_t) - (1-k_t)v_t F + B + R_t,$$

$$(3.11)\quad R_t = -[1+(1-u_t)i_{t-1}]B_{t-1}.$$

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 in 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 (3.9) that corporate equity policy is irrelevant in the taxless world as $m_t = 1$, so the constraint $X_t + yW_t \geq \bar{d}_t Q_t$ 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:

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

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

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

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

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

(4.2) reduces to (4.1) with $W_t^1(K_{t+1}) = L_t^1(Y_{t+1})$.[8] This completes the proof of the Modigliani-Miller theorem.

5. The Homogeneity Assumption

Properties of the solution to (3.9) 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 production function G(K,N,F,e) is linearly homogeneous in K,N,F for any given e.[9] The immediate implication of this is that the associated 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 $f_t = F_t/K_t$.[10] It then seems clear that the value function is also 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 <u>debt-capital</u> <u>ratio</u>. The reason for (5.2) is quite simple.[11] 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 (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 (often implicit but) popular in the corporate finance literature. For example Modigliani and Miller(1963) stated in footnote 15: "...we are referring in principle only to investments which increase the <u>scale</u> 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.

As we have seen, an increase in B_t (keeping K_{t+1} constant) has four effects on W_t. It decreases W_t directly through its appearance in X_{t+1}; it depresses the discounting factor $\gamma_{t+1,j}$ ($j \geq 1$) as bankruptcy becomes more likely, and as a result the corporate bond rate goes up. It also increases the probability that the firm will operate with the binding constraint $d = \bar{d}$. Therefore h_t must be a decreasing function of the debt-capital ratio λ. As the debt-capital ratio goes up, bankruptcy in period t+1 becomes almost certain, so h_t will reach zero at some high level of λ, say $\bar{\lambda}$.[12] When λ is equal to or less than $\bar{\lambda}$, the firm is bound to be bankrupt and 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., $\lambda < \bar{\lambda}$. If, on the other hand, the debt-capital ratio is small (or even negative), then the probability of bankruptcy will be zero. The corporate debt will then be default-free and the constraint $d \geq \bar{d}$ will not be binding. So the only effect of B_t on W_t when the debt-capital ratio is small is its direct effect on X_{t+1}. It follows that W_t has the following form if λ is small:

$$(5.3) \qquad W_t(B_t,K_{t+1}) = w_t^1 K_{t+1} - w_t^2 B_t,$$

where w^1 and w^2 are functions of Z_t but are independent of B_t and K_{t+1}. w_t^2 can be written as

$$(5.4) \qquad w_t^2 = L_t[m_{t+1}(1+(1-u_{t+1})i_t)].$$

In the taxless world with bankruptcy and default, this was equal to one. Since the corporate bond is default free when λ is small, its rate i_t satisfies (2.12) with $M_{t+1} = 1$:

$$(5.5) \qquad L_t[1+(1-\theta'_{t+1})i_t] = 1.$$

Therefore, if the tax rates u_{t+1}, m_{t+1}, and θ'_{t+1} are known in period t, the debt coefficient w_t^2 is equal to

$$(5.6) \qquad w_t^2 = m_{t+1}[1 + ((1-u_{t+1})/(1-\theta'_{t+1}))r_t]/(1+r_t),$$

where r_t is the nominal interest rate on a tax-free safe bond. The graph of $h_t(\lambda_t)$ is illustrated in Figure 1.

With these restrictions on W_t, we can now proceed to solve the DP algorithm (3.9). Without loss of generality we can set $Q_t = K_t$ and convert (3.9) into the following "per-capital" form:

$$(5.7)\quad \max_{b,y,f} \; mx(b,f) + (1-y+my)(1-\delta+f)h[b/(1-\delta+f)]$$

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

$$y \geq 0,$$

where the time subscripts are dropped for notational simplicity and

$$(5.8)\quad x(b,f) = (1-u_t)\pi_t(f) - (1-k_t)v_tf + b + R_t/K_t = X_t/K_t,$$

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

6. Financing Decisions and Investment Function

Probably the most instructive way to solve (5.7) is first fix f and investigate the optimal financing package b(f),y(f) as a function of f, although the optimization problem can alternatively be solved by choosing b,y,f simultaneously. 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 $y = 0$, (5.7) 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

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

so that the optimal debt issue given f is

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

where

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

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

$$(6.5) \qquad m[(1-u)\pi' - (1-k)v + \lambda_1] + h(\lambda_1) = 0,$$

i.e.,

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

where ω is the inverse function of $-\pi'(f)$. This is the investment function 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. Since π' is nonpositive, the argument in the investment function ω must be nonnegative. If taxes and the financial side of the firm are ignored, the investment function (6.6) becomes $f = \omega(h-v)$, so h cannot be less than v. This is probably why it has been often claimed that "Tobin's q," defined here to be $(h+\lambda)/v = (p_t'Q_{t+1} + B_t)/(v_tK_{t+1})$, is too low. The investment function (6.6), however, shows that if taxes and the financial side of the firm are taken into account, Tobin's q (as defined here) can be a lot lower than one.[13]

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 is independent of output and factor prices.

Case 2. If both constraints are binding, (5.7) reduces to

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

The optimal bond issue for given f is determined by the constraint $x = \bar{d}$. Thus

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

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

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

where λ is the debt-capital ratio $b_2(f)/(1-\delta+f)$. Comparing (6.6) and (6.9) we can see that $-h'$ in (6.9) plays the role of m in (6.6). This $-h'$ 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, (5.7) reduces to

$$(6.10) \qquad \max_{b,f} \quad x(b,f) + (1-\delta+f)h(b/(1-\delta+f)) + (m-1)\bar{d}.$$

The first order condition with respect to b is

$$(6.11) \qquad 1 + h' = 0,$$

so that the optimal debt issue given f is

$$(6.12) \qquad b_3(f) = (1-\delta+f)\lambda_3,$$

where

$$(6.13) \qquad h'(\lambda_3) = -1.$$

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

$$(6.14) \qquad f = \omega[(h(\lambda_3)+\lambda_3-(1-k)v)/(1-u)].$$

This is the investment function derived by Poterba and Summers(1982) for the "double tax view" which assumes 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 $x+y(1-\delta+f)h = \bar{d}$:

$$(6.15) \qquad y(f) = [\bar{d}-x(b_3(f),f)]/[(1-\delta+f)h(\lambda_3)].$$

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 $m=(1-\theta)/(1-c) < 1$.[14] 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 issue 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).[15] The graph of y(f) is drawn in panel (ii) of Figure 3.

The interpretation of the results 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 of corporate debt. If dividends cannot be cut any further, financing is done entirely by corporate debt issues. However, as soon as debt issues reach a critical level the rate of increase of debt issues will be cut back to λ_3. The graph of $b_2(f)$ represents the amount of debt issues that are 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 through new share issues $g_t p_t' Q_t$. 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 is the debt-capital ratio $\lambda_t = B_t/K_{t+1}$. 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 (i.e., even if the constraint $y \geq 0$ is absent).

7. Real and Financial Decisions without Adjustment Costs

We now go back to (3.9) 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}; if it is convex there will be no solution to the firm's optimization problem. Although not able to prove it, we conjecture that a sufficient condition for the concavity is that the production function shows decreasing returns to scale or that the firm is a monopolist in the product market with constant returns to scale technology.

It will be convenient and instructive to write the first order conditions in terms of Tobin's (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 $X + yW > \bar{d}Q$, (3.9) reduces to

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

The first order condition with respect to B is

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

where W_B is the partial derivative of $W_t(B_t,K_{t+1})$ with respect to $K_{t+1} = (1-\delta)K_t + F$. 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

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

i.e., the end-of-period capital stock K_{t+1} is optimal when Tobin's (marginal) q adjusted for investment tax credit is equal to m.

<u>Case 2</u>: If both constraints are binding, (3.9) reduces to

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

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

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

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

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

An interpretation of this marginal condition will be given later.

<u>Case 3</u>: If $X = \bar{d}Q$ and $y > 0$, (3.9) reduces to

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

The first order condition with respect to B is

$$(7.9) \qquad 1 + W_B = 0,$$

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

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

i.e., Tobin's q must be equal to one at the optimum. This is the condition derived by Gordon(1982) for the case where the firm can repurchase pre-existing shares.

As we have seen in section 3 and also in section 5, B_t affects $W_t(B_t, K_{t+1})$ in four ways and K_{t+1} affects W_t in exactly the opposite derections. 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.[16] 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 4. So the basic conclusion -- that investment is financed by cutting dividends <u>and</u> by debt if the firm's after-tax profit is large relative to investment, by new share issues <u>and</u> debt if it is 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 5 where Case 2 is followed by Case 1 as F passes F_2.

The above derivation evolves around Tobin's 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, but only under very special conditions. It is evident in the above derivation that (7.7) must hold in either case.[17] It turns out, not surprisingly, that $-W_B$ 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 B_t is small relative to K_{t+1}, then the probability of bankruptcy in period t+1 is almost zero and it it almost certain that Case 1 will 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 (3.6) 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)

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

and

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

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.[18] This result (7.15) and (7.16), which holds for any value of B_t and K_{t+1} (not just for small B_t), 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) $$-W_B = m_{t+1}[1+(1-u_{t+1})i_t]/(1+r_t),$$

and

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

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:

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

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

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

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:

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

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

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

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 never hold at the optimum. If B_t is 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 (7.17) and (7.18) [from which (7.24) is derived] invalid.[19] If B_t is not small relative to K_{t+1}, then W_K (and hence $q = W_K/v_t$) does not have a simple expression like (7.18) because a change in K_{t+1} (keeping B_t constant) alters the probability distribution of M_{t+1} and because it is not certain as of t which Case will occur in t+1. Furthermore, (7.17) captures only one effect of an increase in B_t on W_t, namely the direct effect of B_t on X_{t+1}. The other three effects -- the increase in the discounting factor $\gamma_{t+1,j}$ ($j \geq 0$) (which depresses the share price for any value of the corporate bond rate), the increase in the corporate bond rate, and the increase in the likelihood of $d_{t+1} = \bar{d}_{t+1}$ -- are also included in W_B but are not captured by the right hand side of (7.17).

9. Concluding Remarks

This paper has analyzed the interaction between corporate investment and the associated financing decisions in a model which explicitly allows bankruptcy and default. With a minimal set of assumptions on how the uncertain dividend stream is priced in the asset market, we were able to derive fairly sharp implications. One of the next steps will be to look at cross-section data on firms and test these implications.

In closing the paper, we should mention the things we did not do in this paper. Although it carries a rich menu of tax rates, the model does not consider depreciation allowances (for tax purposes) on investment expenditures. 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 result of the paper. 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 exactly how the pricing formula is affected by such a change. Analyzing it would require (like any other studies on the effect of taxes and inflation on corporate investment) a complete specification of preference, technology and expectations formation, which clearly is well beyond the scope of the paper. It is for this reason that we did not attempt comparative static exercises concerning tax rates and inflation.

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, $\{z_0, z_1, \ldots, z_t\}$, 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:

(A1.2) $L_t^j[L_{t+j}^k(x_{t+j+k})] = L_t^{t+k}(x_{t+j+k})$.

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) above, 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

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

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

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

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

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

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 linearly homogeneous 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 linearly homogeneous 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:

$$\text{the value of the firm with } (\lambda B^0_{t-1}, \lambda K^0_t) \text{ under } h^\lambda = \lambda V_t(B^0_{t-1}, K^0_t; Z_t).$$

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 $m_t X_t + W_t(B_t, K_{t+1})$ 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

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

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. If the dividend tax rate depends on the shareholders' income, θ represents the marginal tax rate for the shareholders. See Miller(1977).

3. 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 the "representative" consumer, δ = subjective rate of time preference, and C = consumption.

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

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

6. See the last paragraph of section 6 for what happens when this constraint is absent.

7. Otherwise the problem of time consistency (Lucas and Sargent(1981)) will arise.

8. To prove that $L_t^1(Y_{t+1})$ 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.

9. This does not necessarily mean that the shock is multiplicative.

10. This π' should not be confused with the marginal value product of capital.

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

12. If h reaches 0 only asymptotically, there will be no solution to the firm's optimization problem.

13. If depreciation allowances are explicitly taken into account as in Hayashi(1982), Tobin's q can be even lower.

14. 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 of 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 (5.4) or (5.6), is less than m_t (in absolute value).

15. 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.

16. The proof is essentially the same as footnote 14.

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

18. If Case 3 holds in period t+1, (7.15) and (7.16) hold with m_{t+1} replaced by one.

19. 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.

References

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

Benveniste, L.M. and J.A. Scheinkman (1979), On the differentiability of the value function in dynamic models of economics, Econometrica, pp. 727-732.

Brock, William A. (1978), Asset prices in a production economy, Social Science Working Paper 275, California Institute of Technology.

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.

Gordon, Roger H. (1982), 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, pp. 213-224.

Lucas, Robert E.. Jr. (1967), Adjustment costs and the theory of supply, Journal of Political Economy, pp. 321-334.

-------- and Edward C. Prescott (1971), Investment under uncertainty, Econometrica, 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, pp. 261-75.

Modigliani, Franco and Merton H. Miller (1958), The cost of capital, corporation finance, and the theory of investment, American Economic Review, pp. 261-97.

-------- and ------- (1963), Corporate income taxes and the cost of capital: a correction, American Economic Review, 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, No.3, pp. 453-475.

Stiglitz, Joseph E. (1969), A re-examination of the Modigliani-Miller theorem, American Economic Review, pp. 784-93.

Summers, Laurence H. (1981), Taxation and corporate investment: a q-theory approach, Brooking Papers on Economic Activity, pp. 67-127.

Tobin, James (1969), A general equilibrium approach to monetary theory, Journal of Money, Credit and Banking, pp. 15-29.

Figure 1

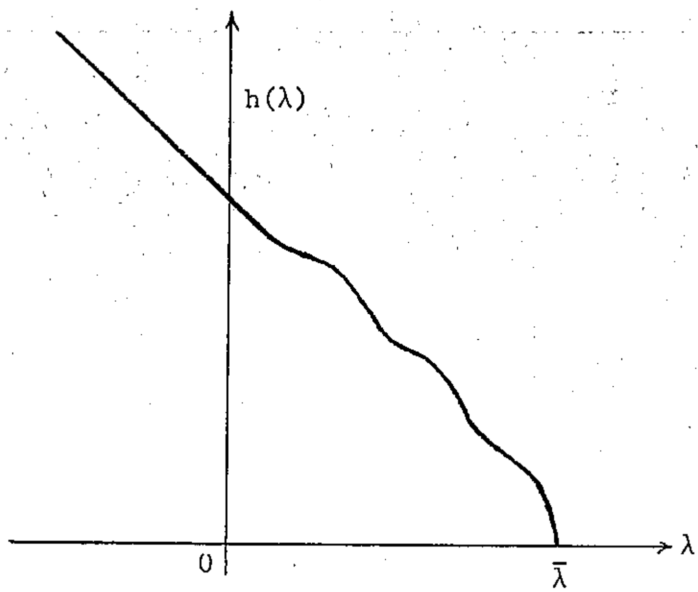
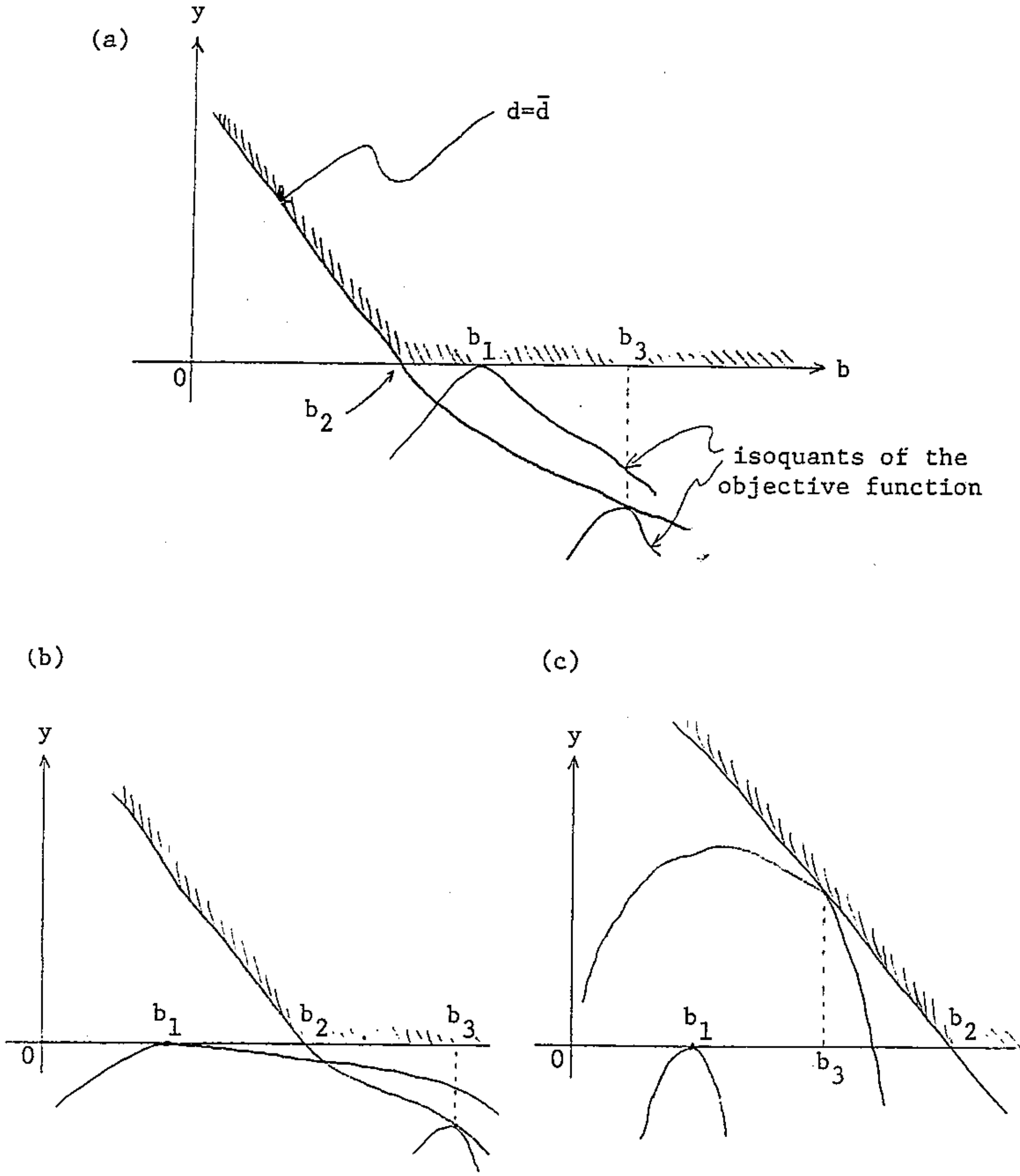


Figure 2



(a)
y
d=d̄
0
b
b_1
b_2
b_3
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

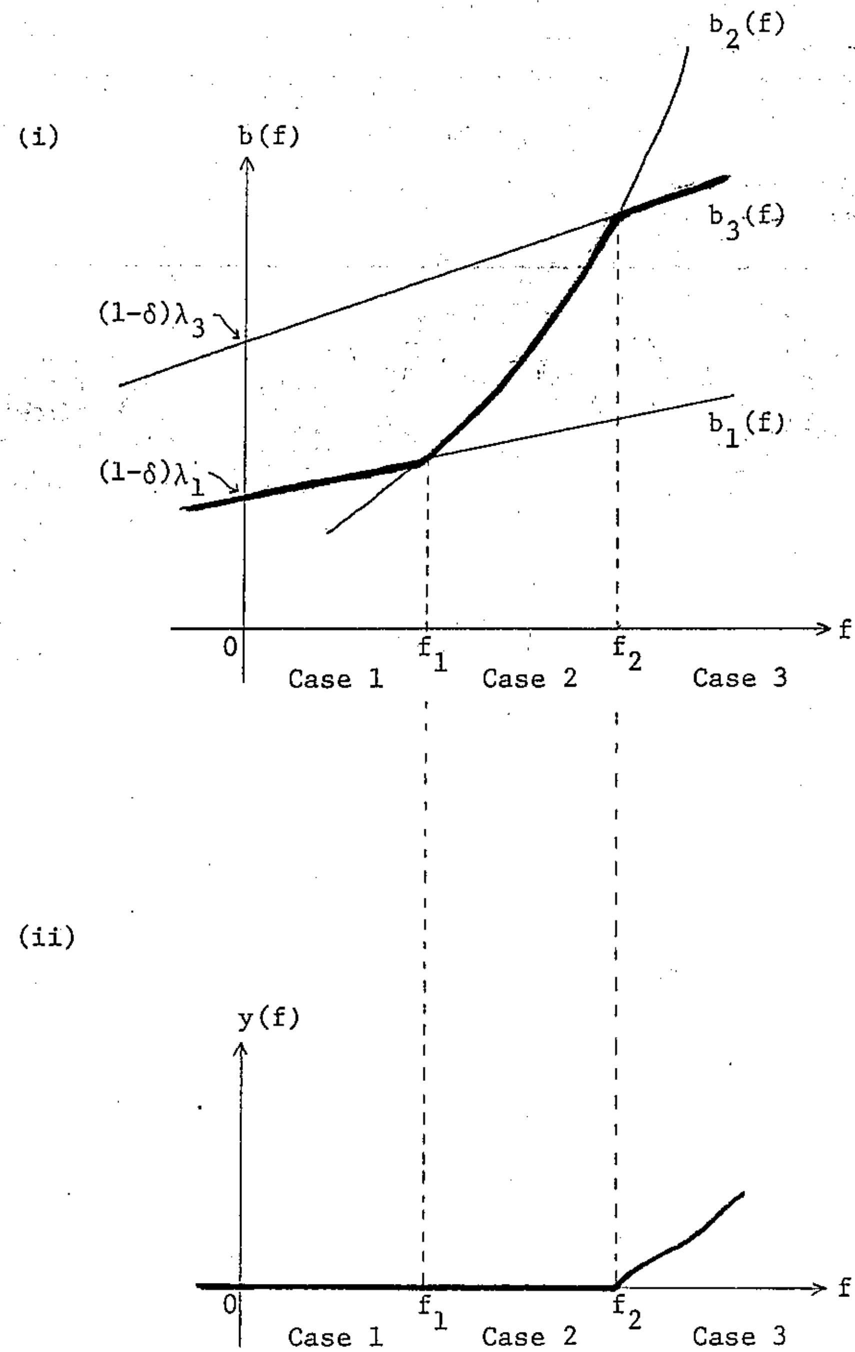
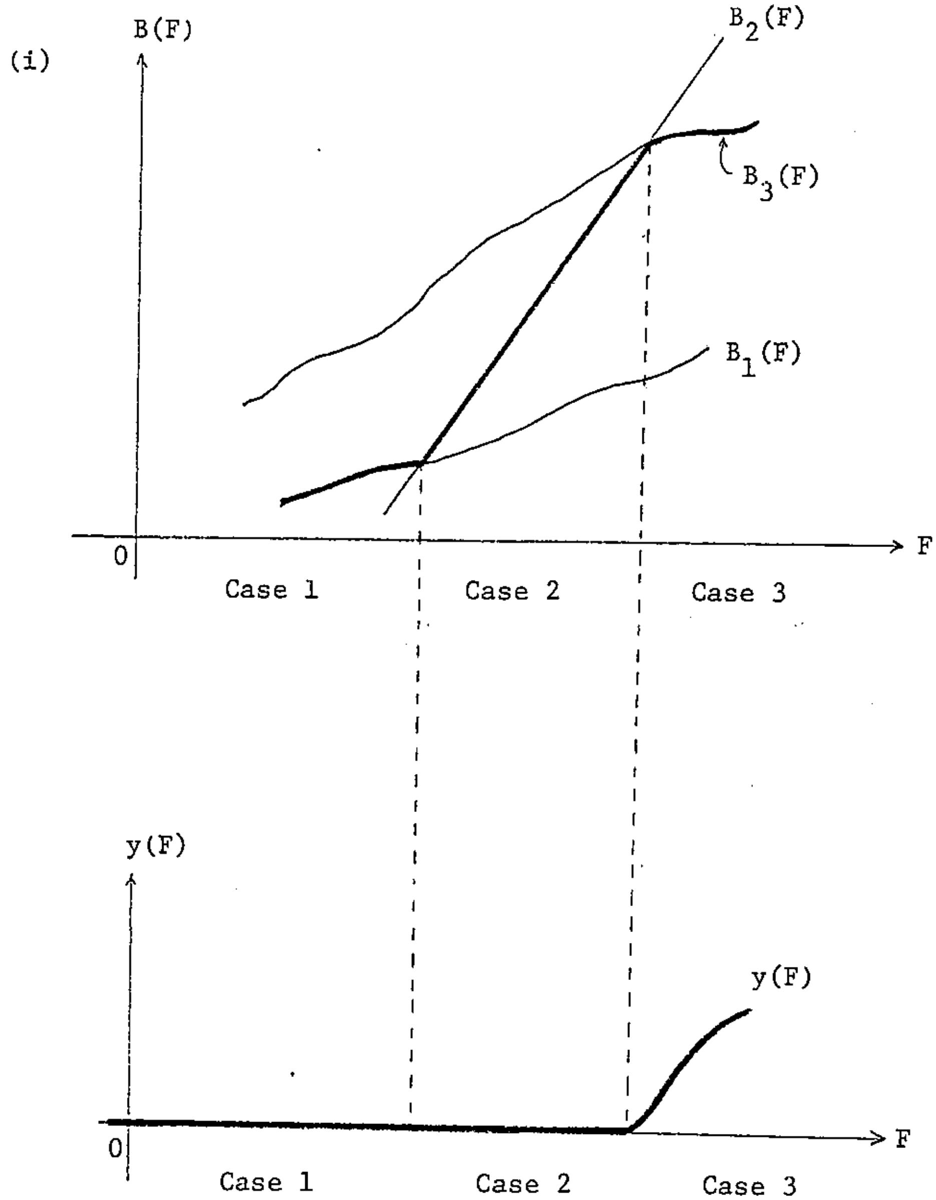

(i)
b(f)
b2(f)
b3(f)
(1-δ)λ3
b1(f)
(1-δ)λ1
0
f1
f2
f
Case 1
Case 2
Case 3
(ii)
y(f)
0
f1
f2
f
Case 1
Case 2
Case 3

Figure 4



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

Figure 5