EC372 Economics of Bond and Derivatives Markets

Financial Intermediation II: Loans versus Bonds

Overview

This note explores why capital is sometimes raised directly from wealth holders (e.g., private investors) typically using the vehicle of marketable securities (such as bonds) and sometimes indirectly from wealth-holders via intermediaries (e.g., banks) typically as individually negotiated, non-marketable loans. Several possible explanations are available: here the focus is on the 'delegated monitoring' approach, in which some wealth-holders delegate to intermediaries the function of monitoring borrowers in order to determine the loan contracts offered to borrowers (some potential borrowers may be refused loans altogether).

Section 1 outlines three approaches to delegated monitoring, and describes the motivation for one of these in more detail.

Sections 2, 3 and 4 describe the models presented in Holmstrom, B. and Tirole, J. "Financial Intermediation, Loanable Funds, and the Real Sector", *Quarterly Journal of Economics*, vol. 112, August 1997, pp. 663–691, abbreviated as 'HT'. In particular, sections 2 and 3 present the core of this note – study these carefully.

1. Financial Intermediation as Delegated Monitoring

In the 'delegated monitoring' approach investors (but not necessarily *all* investors) delegate responsibility to an intermediary (e.g., a bank) to monitor the ultimate recipients (borrowers, e.g., firms) of investors' funds. Such delegation could result from several causes, including: (i) economies of scale (intermediaries gather funds from investors and lend to many borrowers); (ii) expertise (intermediaries have a comparative advantage in specialist monitoring skills); (iii) 'lumpy' (indivisible) investment projects (each firm's investment project may require the funds of many investors).

There is a separate, and different, sense in which 'monitoring' is used in the context of financial intermediation, namely monitoring *of* the intermediary by – or on behalf of – investors who provide the intermediaries funds (e.g., depositors). Such monitoring is typically undertaken in the form of government regulation (e.g., requirement of banks to hold specified proportions of their assets as equity capital). This aspect of monitoring is not considered further here: instead the focus is on monitoring *by* banks.

Freixas & Rochet[1] identify three aspects of 'delegated monitoring':

- *Screening* – the monitor's function is to assess the creditworthiness of the borrower in order to determine whether to grant a loan with pre-specified contractual terms; the monitor does not exert any other form of control over the borrower. The model considered in the notes for 'Financial Intermediation, I' is of this sort.

- *Prevention* – having decided to offer a loan, the monitor is able to influence – albeit imperfectly – the behaviour of the borrower, e.g., to encourage the borrower's effort (discourage 'shirking'), thereby raising the likelihood that the borrower will not default on the loan. The HT models adopt this approach.

[1] Freixas, X. & J-C. Rochet *Microeconomics of Banking*, Second edition, p. 30.

- *Auditing* – if the borrower fails to fulfil the terms of the loan contract in some respect (loan default is the most extreme), the monitor examines the borrower's behaviour, with the intention of imposing a penalty. (Thus, screening and prevention involve *ex ante* decisions on the part of the monitor, while auditing allows *ex post* action.)

Attractive though it might seem, a model of financial intermediation that incorporates all three elements, would be complex and probably too intractible to generate any insights. Hence it's no surprise that existing models address just one of the elements: this note considers the second only, using the two HT models.

1.a. The Principal-Agent Relationship: a Sketch

This sub-section digresses to show how principal-agent theory can be applied to delegated monitoring in the context of financial intermediation.

The so-called 'principal-agent relationship' is a device commonly used to analyse interactions between two individuals (or groups of individuals) where the decisions of one affect the welfare (broadly construed) of the other. The principal sets the terms of the relationship – a formal contract, a 'compensation scheme', or a loose, informal arrangement – so as to control the agent's actions, the consequences of which are evaluated according to the principal's preferences.

Central to the approach is that, for whatever reason, the principal cannot perfectly control the agent – otherwise there would be nothing to analyse; nothing, that is, using principal-agent models. The principal is assumed to pursue policies that guide – albeit indirectly – the agent to act in the interests of the principal. Under some circumstances the principal may be able to achieve a 'first best' outcome equivalent to complete control. More commonly, only limited influence is feasible.[2] Models of the principal-agent relationship thus seek to understand the interactions between the two decision-makers depending on their objectives and the environment in which they exist.

Viewed very generally, the relationship between principal and agent can be viewed as the outcome of a bargaining process between the two. The outcome could be a special case in which all the bargaining power is in the hands of the principal. Instead, the two may agree to split the benefits from the relationship, part of the agreement being that the principal has the right to determine some aspects of the contract made with the agent.

Theories of the principal-agent relationship are commonly divided into two classes, according to the reason why the principal has imperfect control over the agent:

(a) *Hidden information models.* In these models the agent has access to information which cannot be observed – or, if observed, cannot be verified objectively – by the principal.

In the insurance literature the 'hidden information' imperfection commonly leads to the problem of *adverse selection*. For example, a life insurance company ('the principal') may require someone applying to buy a policy ('the agent') to state whether he or she has undergone a test for Human Immunodeficiency Virus (HIV), the precursor of AIDS.

The company may be prohibited by law from offering insurance conditional on a negative test result (assuming, of course, that the company observes the outcome). Consequently, life insurance might be denied to anyone who has ever taken an HIV test, irrespective of whether

[2] In several variants of the theory, attention centres on the ways in which risk is shared between principal (typically assumed to be risk neutral) and agent (typically assumed to be risk averse). This is not so for the models considered here, where all decision-makers are assumed to be risk-neutral: the reason is that the purpose here is to study control mechanisms rather than risk-sharing.

the result was positive or negative. This being so, some individuals who would otherwise be offered insurance are 'adversely selected' and denied coverage because the company is not allowed to act on the result of the test (or might not trust the result even if it could observe it). If the insurance company had access to the 'hidden information' (and believed it), then insurance would be offered to applicants who take the test and are HIV-negative.

(b) *Hidden action models.* In these models the agent takes actions some aspects of which cannot be observed by the principal. They include models in which the actions cannot be verified objectively by an adjudicating third party, such as a court of law, even though the principal observes every relevant aspect of the agent's actions.

In the insurance literature, the 'hidden action' imperfection is referred to as a *moral hazard*. For example, a house-owner ('the agent') may be more careless about smoking cigarettes in bed if a policy covering fire damage has been purchased from an insurance company ('the principal') than if the house is uninsured. The insurance company might not be able to observe whether the individual smokes in bed. Even if the company could make this observation, it might not be able convince a court of law that, in the event of fire, carelessness was the cause.

While the distinction between the two approaches (hidden information and hidden action) tends to be less clear than the above comments may suggest, the classification is convenient in financial intermediation theory. Screening tends to follow the 'hidden information' approach (though banks may be allowed to deny loans according to the results of screening). Prevention and auditing analyses tend to be based on 'hidden action' principal-agent models.

The 'hidden information' approach also underlies the class of 'signalling models' in which firms seek ways to send information to potential lenders that the firm is creditworthy. Such models tend to be more suited to the study of corporate finance than financial intermediation and are not studied here.

The HT model described in the following section is motivated by 'moral hazard' on the part of firms, which need incentives in order to exert the effort necessary to secure a high probability of success in their investment projects – without such incentives, firms would tend to 'shirk' (the 'hidden action' in this application) thereby reducing the chance of success.[3] Moral hazard models are based on two crucial ingredients:

1. *Incomplete contracts* – an assumption that contracts cannot include provisions that specify, in advance, the agent's action to be taken in every conceivable circumstance. In particular, the 'principal' (lender) must leave at least some discretion to the 'agent' (borrower).

2. *Divergent objectives* – an assumption that the 'principal' and 'agent' have different preferences. For example, the 'principal' is likely to be interested only in the cash payoff to the principal, while the 'agent' would prefer to obtain some private benefits (perhaps just by making less effort than otherwise) that the principal cannot control.

In summary: if objectives of the two parties (borrower and lender) coincide, complete contracts are unnecessary (both parties would 'do the right thing' anyway); if contracts are complete, the principal would ensure that the agent does exactly what the principal desires (it does not matter that the agent's preferences differ, because the agent is left with no scope for manoeuvre to deceive the principal). Hence, both assumptions are needed: neither on its own is sufficient.

[3] In fact, in the HT model assumes two levels of shirking, both associated with a low probability of success – this assumption helps to clarify the distinction between 'direct' (bond) finance and 'indirect' (loan) finance.

2. The Holmstrom & Tirole Model

The two models of HT, 'fixed investment' and 'variable investment', are designed to serve different purposes:[4]

1. *fixed investment* model: each firm has access to an investment project of fixed size. Firms differ in their initial capital, A, available for investment: the value of A determines whether firms obtain additional funds from issuing bonds, borrowing from financial intermediaries, or both. (This model is suitable for studying why some firms issue bonds and/or obtain bank loans, while others do not; but it is awkward to extend to the aggregate credit market.)

2. *variable investment* model: again firms differ in the amount of initial capital, but can choose the scale of the investment project (which turns out to be proportional to the value of of each firm's initial capital). (This model is suitable for studying the aggregate credit market but not firms' access to the credit markets.)

The fixed investment model is studied first: extension to variable investment is then straightforward. The notation follows HT closely, with a few exceptions to avoid confusion when HT uses the same symbol to represent two different things. NOTE CAREFULLY: While the notation below is the same as in HT, 'starred' notation R^*, B^*, b^*, c^* does not appear in HT. The reason for including the asterisks here is that HT uses R, B, b, c to stand for two different sets of things (in the two different models they discuss).

There are three groups of decision makers – all of which are assumed to be risk-neutral (so that there is no difference between certainty-equivalents and expected values):

- **Firms** (entrepreneurs): described above. All firms face the same technology but can take actions that provide different levels of personal benefit to their owners – essentially 'benefit' (obtained by 'shirking') is reflected in reduced effort to devoted to the success of the firm: the greater is private benefit ('shirking'), the lower the probability of a successful investment project, resulting in lower expected profits and thus less attractive to lenders (investors or banks).

- **Investors**: can invest either in bonds issued by firms or in an alternative asset (e.g., a risk-free bond), but cannot control the effort devoted by firms whose bonds they hold – a bond's payoff depends on the firm's effort (because effort affects success) but the bond contract cannot directly stipulate how much effort is to be undertaken. HT sometimes calls investors *uninformed* because they cannot monitor firms.

  The alternative asset available to investors (as well as firms and financial intermediaries[5]) has a fixed payoff of γ per unit of investment. Thus the expected payoff of bonds must be at least equal to γ: competition among investors and risk neutrality mean that the expected payoff from bonds *equals* γ. Thus γ denotes the *opportunity cost of capital*. (A zero interest rate corresponds to $\gamma = 1$.)

- **Financial Intermediaries** (e.g., banks): lend to firms but, unlike investors, are able to influence firms' behaviour inasmuch as banks limit firms' private benefits ('shirking').

[4]The reason for using two models is pragmatic. Models intended to analyse both issues at the same time tend to be intractible, i.e. difficult to make sense of.

[5]HT's assumptions imply that, in equilibrium, firms and financial intermediaries will not hold the alternative asset because they can obtain higher payoffs than γ.

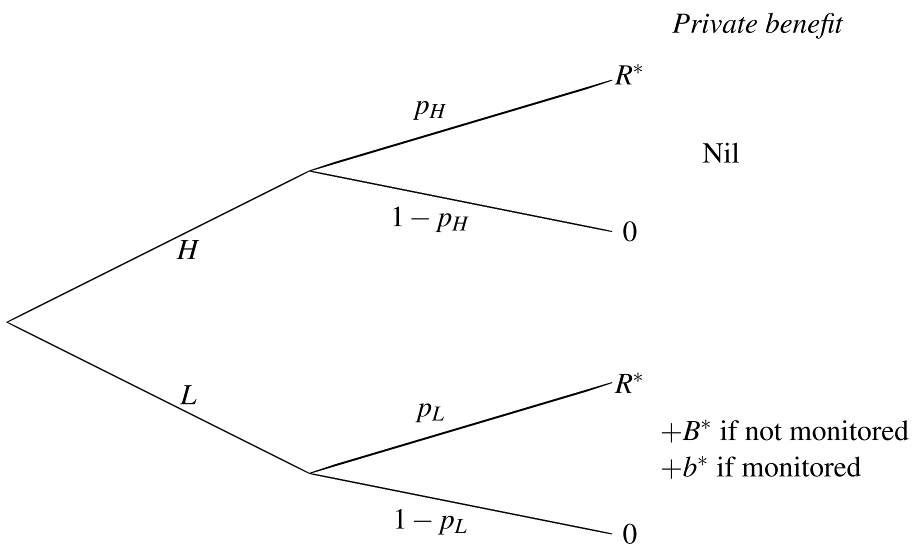

Figure 1: Firms' investment opportunities

Interpretation: Each firm chooses H or L (and B^* or b^*, if L is chosen); then 'Nature' chooses success or failure with the stated probabilities. Everyone is assumed to know the probabilities p_H and p_L that Nature will act in a particular way.

Firms' investment opportunities

Each firm chooses between 'high', H, and 'low', L effort. This choice determines the probability of success, p_H or p_L, with $p_H > p_L$. In the event of success, the payoff is R^*. In the event of failure, the payoff is zero. With high effort, the private benefit (to the entrepreneur) is zero. With low effort, there is a private benefit of B^* if the firm is not monitored (i.e., issues bonds) or $b^* < B^*$ if the firm is monitored (e.g., by a bank as a condition for making a loan). See Figure 1.

Given the terms on which intermediated loans are available, each firm's financing decisions depend on its initial capital, A. The value of A determines the sources of its investment funds.

Let I denote a firm's total investment (from its own capital, bond issuance or an intermediated loan). Basic assumptions:

- $p_H R^* - \gamma I > 0$: high effort investment is economically viable, in the sense that the expected payoff is sufficient to pay bond-holders γ and also leaves something for the firm;
- $p_L R^* - \gamma I + B^* < 0$: low effort is not economically viable in the sense that with L, investors would expect the payoff on bonds to be less than γ; hence they would not hold firms' bonds.

  Given that firms get even less benefit from intermediated funds $b^* < B^*$, it follows that $p_L R^* - \gamma I + b^* < 0$ too. (Also it will be shown that intermediated loans cost more than γ, reinforcing the unviability of L.)

3. Fixed Investment

HT distinguish between 'direct finance', where the firm obtains funds from bond issuance to uninformed investors – investors who cannot monitor the actions of firms. Each firm receiving 'indirect finance' will be subjected to monitoring by a financial intermediary.

In the 'fixed investment' model firms do not choose the size of the capital project: it is exogenously given. The 'variable investment' model in which firms choose the size of the project will be outlined later. (Although apparently more general, the variable investment model does not provide many more insights.)

3.a. Direct finance

The payoff, R^*, from a successful project is split between firms, which get R_f, and investors who get R_u: $R_f + R_u = R^*$. In the event of failure, the payoff both to investors and firms is zero – but firms still get B^* if they have shirked.

Investors know that firms will shirk (choose L) unless their expected payoff from not shirking, $p_H R_f$, is at least as great as from shirking, $p_L R_f + B^*$. That is, R_f must satisfy:

$$p_H R_f \geq p_L R_f + B^* \quad \text{or} \quad R_f \geq B^*/\Delta p \tag{1}$$

where $\Delta p \equiv p_H - p_L > 0$. HT call (1) the firm's 'incentive compatibility condition', which will be assumed to hold. It also defines the so-called 'pledgeable expected income', $p_H[R^* - B^*/\Delta p]$, which is the most that uninformed investors could be promised consistent with the firm choosing H.

Competition among investors will ensure that the expected payoff per unit of investment in the firm *equals* the payoff on the alternative investment, γ. Formally, the payoff to investors from an investment I_u in the firm is $\gamma I_u = \gamma[I - A]$. Thus in order for the firm to choose H and thus be able to deliver γ per unit of direct (bond) investment, the firm's initial capital, A, must be sufficiently large that:

$$\gamma I_u = \gamma[I - A] \leq p_H \left[R^* - \frac{B^*}{\Delta p}\right] \tag{2}$$

A firm will be able to borrow only if its A is large enough that it has incentive to 'behave' (choose H rather than L). If A is too low, the firm could not convince investors that it would not shirk – thus leaving the investors with nothing but the firm with its B^*. If A is large enough, the firm will need to borrow a small enough amount, $I - A$, that it can expect to receive at least as large a payoff, net of borrowing costs, to make it worthwhile to behave (choose H).

Formally, the minimum capital requirement is found by replacing (2) with an equality, and rearranging:

$$\overline{A}(\gamma) = I - \frac{p_H}{\gamma}\left[R^* - \frac{B^*}{\Delta p}\right] \tag{3}$$

Only firms for which $A \geq \overline{A}(\gamma)$ will be able to borrow from the bond market. Suppose that a firm has just enough initial capital to access the bond market, i.e., $A = \overline{A}(\gamma)$. Then it follows that $R_f = B^*/\Delta p$; in words, a firm with no 'excess' initial wealth, obtains a payoff on its investment project (if it is successful) exactly equal to the minimum required to induce the firm to behave – see equation (1).

Notice the role of γ: if the opportunity cost of funds increases, $\overline{A}(\gamma)$ decreases, so that fewer firms will be able to raise funds from bond issues.[6]

In what sense does initial wealth provide 'collateral' here? It is not in the sense that A would be available to repay investors in the event of failure of the investment project (whether or not the firm shirks). In the event of failure, the payoff is zero. Instead, initial capital provides collateral in the sense that if A is high enough, the firm has an incentive to generate enough (expected) income to

[6] HT provide a condition which guarantees that a firm must have positive initial capital in order to attract bond holders, i.e., $\overline{A}(\gamma) > 0$. This is equation (3) on p. 671, but notice the slip in the equation (the minus sign on the right-hand-side should not appear, or be replaced with a plus).

satisfy investors. This interpretation is one of the insights of the model.

3.b. Indirect finance

In the HT models 'indirect' finance corresponds to intermediary loans, most easily interpreted as bank loans (though other interpretations are possible). HT make assumptions such that the banks need to commit capital of their own, rather than merely 'certify' the behaviour of firms, for which the banks will require a rate of return, β, where $\beta > \gamma$ (because monitoring is costly).[7]

Now the payoff from a successful project is split three ways $R_f + R_u + R_m = R^*$, with firms receiving R_f, investors (bond holders) R_u, and banks R_m. With the presence of a bank to monitor the firm, the firm's private benefit from shirking becomes $b^* < B^*$. Thus, monitoring allows the bank to extract some of the surplus (payoff) that would otherwise have to be offered to firms in order to ensure them to behave – the 'bribe' is lower, or if you prefer, the amount of collateral needed to obtain finance is lower.

Using the same reasoning as for direct finance, the firm's 'incentive compatibility' condition becomes:

$$R_f \geq b^*/\Delta p \tag{4}$$

Because $b^* < B^*$, when monitoring is possible, firms need to be guaranteed a smaller payoff in order to behave (choose H).

Monitoring is not free. Assume that it costs c^* per firm (project) monitored. Then the bank must be guaranteed a minimum return in order to make monitoring worthwhile:

$$p_H R_m - c^* \geq p_L R_m \quad \text{or} \quad R_m \geq c^*/\Delta p \tag{5}$$

Suppose that the rate of return on bank loans (indirect finance) to the firm is given by β such that $\beta I_m = p_H R_m$ where I_m is the amount the intermediary offers to the firm in return for β and the right to limit the firm's private benefits to b^* (i.e., to monitor the firm).

If intermediaries will lend anything at all to a firm (they may choose not to do so), then competition among them will drive down the payoff to ensure $R_m = c^*/\Delta p$ so that:

$$I_m(\beta) = \frac{p_H c^*}{\beta \Delta p} \tag{6}$$

Under what conditions will an intermediary invest in a firm, i.e., $I_m > 0$? Only if the return, taking into account monitoring costs, c^*, is at least γ. As $c^* > 0$, this requires $\beta > \gamma$: 'indirect' finance is more expensive for the firm than 'direct' finance.

Firms with enough initial capital will be able to obtain direct finance (issue bonds). This source is preferred to indirect finance (intermediary funding such as bank loans) because indirect finance is more expensive (also it involves intrusive monitoring to make the firm behave). Hence, firms with lower initial capital find it necessary to obtain indirect finance. But they would accept only the smallest amount possible, preferring direct finance if they can obtain it.

It is possible that a firm will have so little initial capital that it cannot obtain even indirect finance.

[7] HT's assumption is essentially that all the projects financed by an intermediary either succeed or they all fail: the payoffs of different firms are perfectly correlated. Without this assumption, the financial intermediaries would be able to diversify the risks of firms' failure thus reducing the amount of capital needed, possibly to zero. To allow for such generality would require a much more complicated model.

The condition for this formally: $A \geq \underline{A}(\gamma, \beta)$. It can be shown that $\underline{A}(\gamma,\beta)$ is given by:

$$\underline{A}(\gamma,\beta) = I - I_m(\beta) - \frac{p_H}{\gamma}\left[R^* - \frac{b^* + c^*}{\Delta p}\right] \tag{7}$$

Firms with initial capital greater than $\underline{A}(\gamma,\beta)$ will be able to obtain indirect finance. They will then top this up with direct finance if $I > A + I_m(\beta)$.

If a firm's initial capital is high enough, then it will not need any indirect finance. The condition for this (see above) is $A \geq \overline{A}(\gamma)$. Can we be sure that the minimum initial capital needed to be eligible indirect finance is less than that for needed for direct finance, that is $\underline{A}(\gamma,\beta) < \overline{A}(\gamma)$? If not, then firms would never obtain indirect finance: they obtain bond finance or nothing – intermediary finance would be zero. Such a circumstance would occur if monitoring costs (measured by c^*) are too high relative to the 'benefit' of reducing firm's shirking (measured by the difference $B^* - b^*$).

Formally, indirect finance will be possible if c^* is sufficiently small:

$$c^*\Delta p < p_H[B^* - b^*] \quad \Longrightarrow \quad \underline{A}(\gamma,\beta) < \overline{A}(\gamma) \tag{8}$$

In words: if monitoring costs are low enough, there will be a range of initial capital levels that enable a firm to obtain loans from an intermediary.

3.c. Interpretation

The HT model shows how the availability of initial capital (serving as collateral) affects access of a borrower (firm) to credit markets. If initial capital is large enough, the firm will, without being monitored, be able to borrow 'directly' from uninformed investors. Nonetheless, borrowing costs will mean that even wealthy firms cannot afford to shirk. This is true in the HT model because the '*L*' (shirking) project is not economically viable, even with high private benefits (i.e., without monitoring).

At the other extreme are firms whose initial capital is so low that they cannot attract capital from either uninformed investors or intermediaries – even if monitored, they would still be expected to shirk. The amount of initial capital is not enough to convince even monitors otherwise (because intermediaries are not perfect: the do not have complete control of the firms they monitor).

Between the two extremes is a range of initial capital levels such that a firm in this range is eligible for indirect, but not direct, finance. It is as if the firm could obtain funds from an intermediary (at a high cost) to show its credibility to the uninformed investors. Note that because of the high cost of indirect finance compared with direct finance ($\beta > \gamma$), firms only ever borrow the minimum necessary from intermediaries: any amount needed in excess of the minimum will be borrowed directly from uninformed investors. In a sense, bond holders 'free ride' on the work of monitors – but they receive a lower rate of return. (As a result of competition, there is no difference in return when monitoring costs are taken into account.)

As HT explain, the formal model can be given more than one interpretation. The interpretation above is intended to illuminate a distinction between obtaining funds "from the market" (issuing bonds to uninformed investors) and obtaining funds "from intermediary loans" (borrowing from a bank, or some other entity such as a venture capitalist, that monitors those to whom it lends money).

An alternative interpretation is that uninformed investors lend to banks (representing the archetypal 'financial intermediaries'), which then lend to borrowers that they monitor – though the borrowers must be have enough initial capital (collateral) even to merit monitoring. Notice that in this interpretation banks do not simply channel funds from investors to borrowers: banks are obliged to lend their own capital too. HT's assumption that the payoffs from bank loans to different firms are perfectly correlated (mentioned above) means that they cannot 'diversify away' the risk of firms' failure – hence banks must commit their own capital. If the payoffs on bank loans were uncorrelated, then diversification could work, and the bank would invest no capital of its own and simply charge a fee for its monitoring services – but that would be a different model.

3.d. Credit Market Equilibrium

So far, the rates of return γ and β have been treated as parameters. In order to analyse market equilibrium it is necessary to make assumptions about the aggregate supply of the three sorts of capital:

- *Uninformed investors' aggregate capital*, K_u, is assumed to be determed according to an aggregate saving function, $K_u = S(\gamma)$, where $S'(\gamma) > 0$ (the higher the return, the higher the level of saving).
- *Informed investors' (banks') capital*, K_m, is assumed to be exogenous (fixed).
- *Firms' initial capital*: each firm's initial capital is assumed fixed. However, the *distribution* of this capital among firms may – and generally will – affect market outcomes. This complication is ignored here, though the resulting ambiguities will be discussed below. The aggregate of all firms' initial capital is denoted by K_f.

  In HT's 'variable investment' model (where firms choose the size of their investment projects), market equilibrium depends on only the *aggregate* of firms' initial capital – all firms become identical except for the scale of their investment projects, which are proportional to initial capital. (This aggregation property is essentially the only reason for constructing the variable investment model.)

Let the aggregate of firms' demand for uninformed and intermediary capital be denoted by:

$$\text{Firms' demand for uninformed capital:} \quad D_u(\underset{-}{\gamma}, \underset{-}{\beta}; \underset{+}{K_f}) \tag{9}$$

$$\text{Firms' demand for monitored loans:} \quad D_m(\underset{-}{\gamma}, \underset{-}{\beta}; \underset{+}{K_f}) \tag{10}$$

The symbols under the functions' arguments denote predicted directions of change. Thus, an increase in firms' capital, K_f, means that they have more collateral with which to justify higher greater borrowing from intermediaries and from uninformed investors. In a sense, bond issuance and intermediary loans (direct and indirect finance) are *complements*: when the cost of bank loans (β) increases, the aggregate demand by firms for investment funds goes down for both bank loans and from uninformed investors; similarly, when γ increases, the aggregate of firms demand for investment funds declines from both sources.

Take care to remember that the stated directions of firms' responses are *tentative*, especially with regard to the 'cross-effects' (response of D_u to β and D_m to γ) – i.e., the complementarity or substitutability of the two sources of finance. These are sensitive to the detailed assumptions (and depend particularly on the assumptions of the 'variable investment' model).

Equating demand and supply for funds gives the conditions for credit market equilibrium:

$$\text{Bond market (direct finance):} \quad D_u(\gamma, \beta\,; K_f) = S(\gamma) \tag{11}$$

$$\text{Intermediary loans (indirect finance):} \quad D_m(\gamma, \beta; K_f) = K_m \tag{12}$$

Equilibrium is illustrated in Figure 2 on page 10.

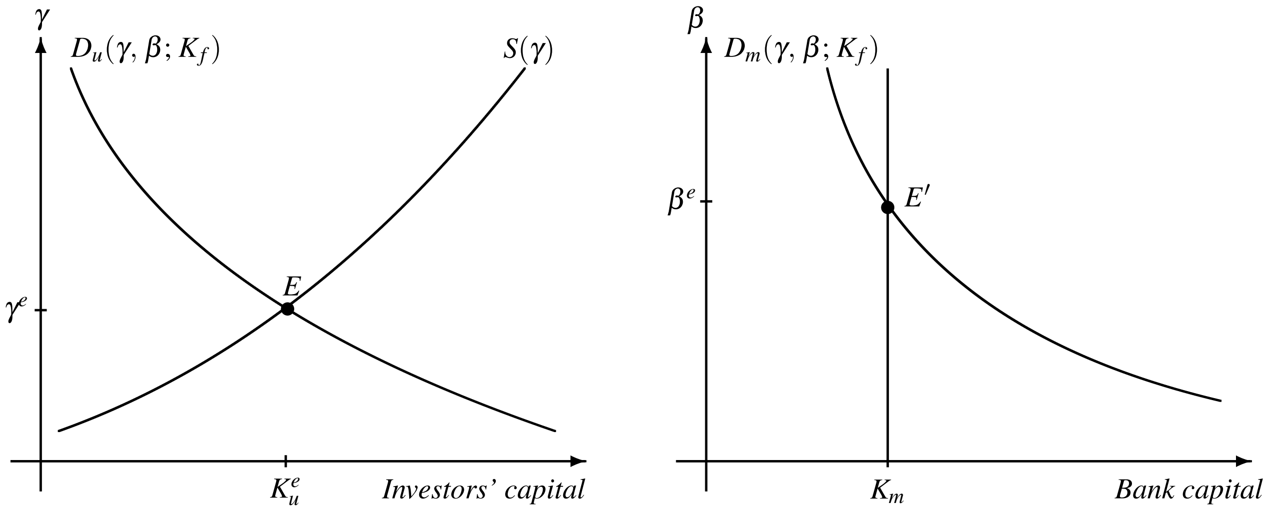

Figure 2: Credit Market Equilibrium

The left-hand panel illustrates equilibrium, E, in the market for uninformed capital (the bond market) in which firms' demand for funds equals uninformed investors supply of saving.

The right-hand panel shows the equilibrium, E', in the market for informed capital (bank lending) in which firms' demand for funds equals the exogenously given supply of informed capital. Note that the assumptions of the model guarantee that $\beta^e > \gamma^e$.

Now it is possible to study the effect of various 'shocks' including three that HT explore in detail:

A. *Credit crunch*: K_m falls – a reduction in intermediaries' capital.
B. *Collateral squeeze*: K_f falls – a reduction in firms' initial capital.
C. *Savings squeeze*: $S(\gamma)$ shifts to the left – saving is lower at each γ.

HT obtain the following results (rigorous derivation requires the variable investment model or additional assumptions in the fixed investment model):

A. *Credit crunch* $\Longrightarrow \beta \uparrow$: cost of bank loans increases; $\gamma \downarrow$: return on bonds decreases.
B. *Collateral squeeze* $\Longrightarrow \beta \downarrow$: cost of bank loans decreases; $\gamma \downarrow$: return on bonds decreases.
C. *Savings squeeze* $\Longrightarrow \gamma \uparrow$: return on bonds increases; $\beta \downarrow$: cost of bank loans decreases.

Each of these shocks can be illustrated and interpreted using the diagrams in Figure 2.

3.e. *Example*: a credit crunch.

In the HT model a credit crunch takes the form of a fall in banks' capital, say from K_m to $\widehat{K}_m$ in Figure 3.[8] See page 11. The fall in K_m forces up β, such that some firms (with low initial capital of

[8]Note that the aggregate level of banks' capital is exogenous in the HT model – the model is consistent with many possible reasons for a fall in banks' capital but is silent about their causes.

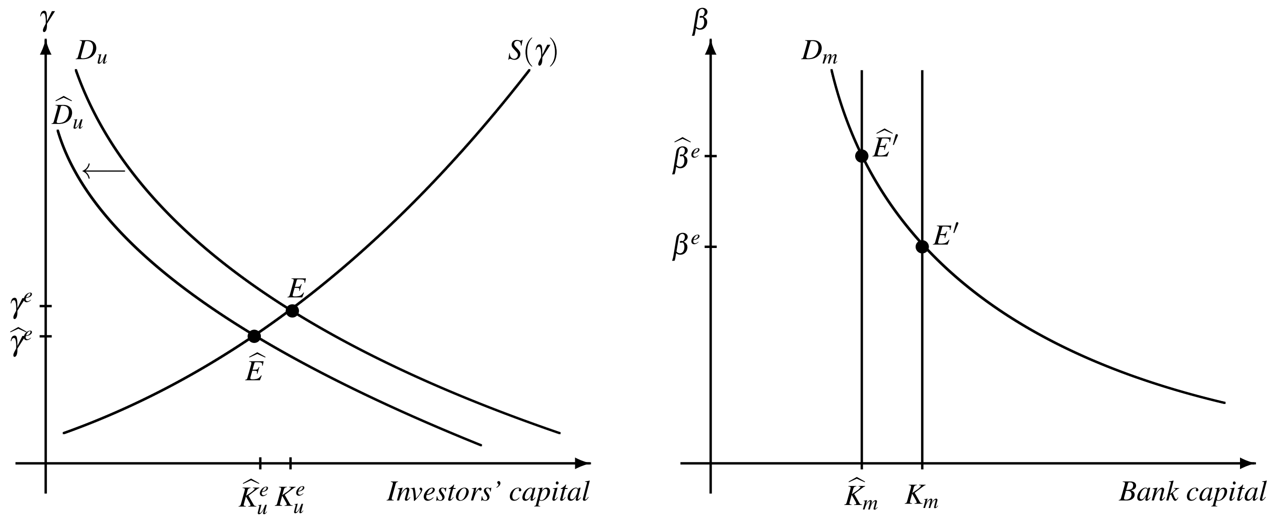

Figure 3: Credit Market Equilibrium: Credit Crunch

In a credit crunch the supply of intermediary capital falls from K_m to $\widehat{K}_m$, driving up firms' cost of indirect finance from β^e to $\widehat{\beta}^e$.

Without access to indirect finance, some firms (those with lower initial capital) are excluded also from the bond market, resulting in a leftwards shift in the aggregate demand for direct finance, $D_u \to \widehat{D}_u$, driving down the rate of return on bonds, $\gamma^e \to \widehat{\gamma}^e$.

their own) will no longer be able to obtain indirect finance.

As some firms are denied access to bank loans, they will no longer be able to top-up their borrowing by issuing bonds (remember that firms would have never obtained any more than the minimum amount of indirect capital, because it is more expensive than direct capital). Hence, given the complementarity assumption, that D_u varies inversely with β, the demand curve D_u shifts left. Consequently, γ falls – uninformed investors compete to purchase fewer bonds, forcing up the price of firms' bonds and reducing their yields: $\gamma\downarrow$.

BE CAREFUL: Figure 3 is incomplete in that it does not fully represent the interdependency of the two markets. The fall in γ will cause a 'second round' impact on D_m: because of the assumed complementarity, the demand curve D_m shifts to the right. This shift is not shown in Figure 3: its effect will be to reinforce the increase in β. There will also be a 'second round' impact on D_u, and thus yet further effects on each demand curve as γ falls and β rises.

In view of the interdependency of the two markets, the demand curves in Figure 3 should be interpreted as representing the final, overall effect of the reduction in K_m, with γ and β and their new equilibrium values, $\widehat{\gamma}^e$ and $\widehat{\beta}^e$.

4. Variable Investment

This section is optional.

Here is a brief overview of HT's model in which firms choose the size, I, of their investment projects. Perhaps it is 'more realistic' but provides few new insights compared with the fixed investment model. Its main role is to justify the market equilibrium analysis summarised earlier.

The values of R^*, B^*, b^* and c^* are no longer fixed but rather assumed proportional to the level of investment. Hence: $R^* = RI$, $B^* = BI$, $b^* = bI$, $c^* = cI$, where R, B, b and c are parameters.

HT now show that each firm's investment is proportional to its initial capital, A_0, so that all firms invest, albeit at different levels according to initial capital. Hence, the variable investment model is of no use in analysing what determines access to different capital markets (bonds or loans) but it simplifies analysis of the aggregate credit markets.

Much of the analysis of the previous section holds: just subsitute for the 'starred' values R^*, B^*, b^* and c^* in the expressions. The crucial conditions for each firm are as follows:

$A_0 + I_u + I_m \geq I$	Investment funds obtained from firms, investors & banks	(13)
$R_f + R_u + R_m \leq RI$	Total return divided among firms, investors & banks	(14)
$p_H R_u \geq \gamma I_u$	Minimum payoff for investors	(15)
$R_m \geq cI/\Delta p$	Minimum payoff for intermediaries	(16)
$R_f \geq bI/\Delta p$	Minimum payoff for firms	(17)
$p_H R_m \geq \beta I_m$	Minimum return on intermediary capital	(18)

Competition among firms, intermediaries and investors ensures that each of the above inequalities is replaced with *equality* in equilibrium. Routine substitutions then give:

$$\gamma I_u = p_H \cdot \left[R - \frac{b+c}{\Delta p} \right] \cdot I \qquad \text{Value of investors' capital} \tag{19}$$

$$\gamma = p_H \cdot \left[R - \frac{b+c}{\Delta p} \right] \cdot \frac{I}{I_m} \qquad \text{Return on investors' capital} \tag{20}$$

$$\beta = p_H \cdot \frac{c}{\Delta p} \cdot \frac{I}{I_u} \qquad \text{Return on intermediaries' capital} \tag{21}$$

Because firms' investment and financing are proportional to firms' initial capital, it is possible to replace A, I_u, I_m and I, with their corresponding market aggregates, K_f, K_u, K_m and K, respectively. Noting that $K = K_f + K_m + K_u$, and rewriting equations (19) – (21) market equilibrium is characterised by:

$$p_H K \left[R - \frac{b+c)}{\Delta p} \right] = \gamma(K_u) K_u \qquad \text{Determines level of saving, } K_u. \tag{22}$$

$$\gamma = \frac{p_H K}{K_u} \left[R - \frac{b+c)}{\Delta p} \right] \qquad \text{Determines cost of bond finance, } \gamma. \tag{23}$$

$$\beta = \frac{p_H c}{\Delta p} \cdot \frac{K}{K_m} \qquad \text{Determines cost of bank finance, } \beta \tag{24}$$

where $\gamma(K_u)$ denotes the inverse of the saving function – formally $K_u \equiv S(\gamma(K_u))$.

If equation (22) is solved for K, the resulting value can be substituted into (23) and (24) to provide solutions for the two crucial endogenous variables, γ and β (cost of bond and bank finance, respectively), as functions of parameters and exogenous variables, K_u and K_m (aggregate firms' initial capital and intermediary capital).

*****