OPTIMAL SECURITY DESIGN

by

Franklin Allen
and
Douglas Gale

(26-87)

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
The University of Pennsylvania
Philadelphia, PA 19104-6367

OPTIMAL SECURITY DESIGN

Franklin Allen*

University of Pennsylvania

and

Douglas Gale*

University of Pittsburgh
and
University of Pennsylvania

Address for all correspondence:

Finance Department,
The Wharton School,
University of Pennsylvania,
Philadelphia, PA 19104
U.S.A.

Telephone: (215) 898 3629

*Financial support from the NSF (Grant nos. SES-8420171 and SES-8520224 for the two authors respectively) is gratefully acknowledged.

Abstract

How should new securities be designed? Traditional theories have little to say on this: the literature on capital structure and general equilibrium theories with incomplete markets take the securities firms issue as exogenous. This paper explicitly incorporates the transaction costs of issuing securities and develops a model where the instruments that are traded are chosen optimally and the economy's market structure is endogenous. Among other things, it is shown that the firm's income stream should be split so that in every state all payoffs are allocated to the security held by the group that values them most.

1. Introduction

It is widely acknowledged there has been a significant increase in the amount of financial innovation in recent years (see e.g. Miller (1986)). A vast number of new securities with novel features have been introduced. These include not only corporate securities such as zero coupon bonds and floating rate bonds but also other types of security such as financial futures, options on indexes and money market funds to name but a few. An important question concerns how such securities should be optimally designed. In other words, how should the payoffs to a security be allocated across states of nature to maximize the amount the issuer receives? Unfortunately, with the exception of a few papers (see, e.g., Duffie and Jackson (1986) and Silber (1981) on the design of futures contracts and Williams (1986) on corporate securities as optimal contracts when there is asymmetric information) traditional theories have little to say on this issue.

The literature on firms' capital structure decisions assumes that the firm can only issue debt and equity. It does not consider the more basic question of whether debt and equity are the best securities the firm can issue. The result of Modigliani and Miller (1958) and subsequent authors such as Stiglitz (1969; 1974), Baron (1974; 1976) and Hellwig (1981) that capital structure is irrelevant when markets are complete (assuming no taxes or frictions) suggests that the form of securites issued in this case is also unimportant. Although the result that capital structure is important when markets are incomplete indicates the form of securities may also matter, it throws little light on their optimal design.

In order to develop a theory of optimal security design it is clearly necessary to develop a framework where markets are incomplete. Traditional general equilibrium theories with incomplete markets such as Diamond (1967), Radner (1972) and Hart (1975), have taken the markets which are open and those which are closed as exogenous. The usual justificaton for considering incomplete markets is the existence of transaction costs of some sort. However, the relationship between these transaction costs and the securities which it is optimal for agents to issue are not considered.

The purpose of this paper is to explicitly incorporate the transaction costs of issuing securities in order to consider their optimal design. We suppose that firms bear fixed costs of issuing securities and determine the allocation of payoffs to securities across states. Thus the securities that are traded are chosen optimally and the incomplete market structure is endogenous. The model is intended to be a very simple one so that it can act as a benchmark. It is an abstraction which is meant to give insight rather than a realistic description of what we observe.

Two important issues must be addressed when constructing this type of model. The first is who precisely can issue securities and what the costs of doing this are. We assume that firms can issue one security for some fixed cost and a second security for some additional amount. The payoffs to these securities are generated by the assets of the firm. Individual investors are unable to issue securities. In particular, we assume that individuals cannot short sell firms' securities. If individuals could do this without cost then no equilibrium could exist in which firms issue costly securities. This is because the short sellers are effectively able to expand the supply of a firm's security more cheaply than firms can. However, in order for a firm to be willing to issue securities it must be able to recoup the cost of doing so. But if it faces competition from short sellers who face no costs this will be impossible. The short selling of one firm's securities by another is also excluded as this leads to nonexistence in the same way.

The justification for these assumptions is that short selling is fundamentally different from issuing securities backed by real assets. In the latter case the buyer can directly observe the security's source of income whereas with short selling this is usually much more difficult. In reality, it appears that short selling is very costly and is only rarely done. For most of the analysis it is assumed that the costs are such that there is no short selling. However, at one point this assumption is relaxed to illustrate what happens when limited short selling is possible at a low cost.

The second issue concerns the knowledge that agents have about the prices of securities that in principle could be issued by firms but which in equilibrium are not. We assume that both firms and consumers are aware of these prices. The trading process that we have in mind is analogous to the Lloyd's of London insurance market where a price can be obtained for a policy insuring any risk whether or not such policies are actively traded. Our approach is therefore different from that taken by Hart (1979) who assumes that there is an asymmetry between firms and producers: firms know the prices that would be obtained for securities that are not issued; consumers do not.

We obtain five main results.

(i) The constrained efficiency (and existence) of equilibrium is demonstrated. This contrasts with the results of Hart (1980) and Makowski (1980) who in related work find that equilibrium is not constrained efficient. The difference is due to the fact that we assume both firms and consumers know market clearing prices for all securities that do exist and all those that might exist whereas they use Hart's (1979) approach of assuming firms know the prices of unissued securities but consumers do not.

(ii) In our model, the number of securities that are issued in equilibrium can exceed the number of states without there being complete markets in the sense that first-best risk sharing is possible. We show that an upper bound on the number of securities issued in equilibrium is $|J|.[|I|.(|I| - 1) + 1]$ where $|J|$ is the number of different types of firm and $|I|$ is the number of different types of consumer.

(iii) As a point of reference we consider examples where firms are constrained to use debt and equity securities but these are costly to issue. We show that Modigliani and Miller's irrelevance result does not hold: the value of firms has to depend on their financial structure to give them an incentive to issue the costly securities. In equilibrium both debt and equity are issued. This contrasts with the arguments by Jensen and Meckling (1976) and others that the existence of debt before the introduction of the corporate income tax suggests the importance of asymmetric information in determining firms' capital structures. In our model firms issue debt despite there being symmetric information and no taxes. In addition for some parameter values (ex ante) identical firms can have dissimilar capital structures, an empirical regularity which, as Myers (1984) stresses, is inconsistent with conventional theories.

(iv) If firms are not restricted to issuing debt and equity but can allocate their earnings in a particular state to their two securities in any way they choose then the optimal securities are not debt and equity. When the firm is split into two securities, each one is targetted at a particular clientele. In any state all the firm's output is allocated to the group that places the highest valuation on consumption in that state.

(v) The role of the assumptions concerning short sales is shown to be critical. As argued above, if costless short sales are possible then equilibrium fails to exist. Financial structure must affect firm value in order to provide an incentive for firms to issue securities. Without frictions of some kind or another there will be arbitrage opportunities which will cause equilibrium to fail to exist. We illustrate what happens when limited short selling is possible at low cost. In this case equilibrium may or may not exist depending on the cost of issuing securities.

The paper proceeds as follows. In Section 2 the model is outlined. Section 3 contains an analysis of the general model. The next two sections develop illustrative examples. For the purposes of comparison Section 4 shows what happens when there are costs of issuing securities but firms are constrained to issue debt and equity; Section 5 considers the case where firms are free to allocate their earnings in a particular state to the two securities in any way they desire. In Section 6 the role of short sale constraints is considered. Finally, Section 7 contains concluding remarks.

2. The model

A formal definition of the model is given in the technical appendix. An intuitive description of its main features is given in this section.

There are two dates $t = 0, 1$ and a finite set of states of nature $s \in S$ which occur with probability π_s. All agents have the same information structure: there is no information at the first date and the true state is revealed at the second. At each date there is a single consumption good. There is a finite set of types of producer $j \in J$ and a finite set of types of consumer $i \in I$. There is a continuum of agents of each type and for simplicity the measure of each type is assumed to be unity.

Producers

A type of producer is defined by a production plan, a set of financial structures and a cost function. A production plan of type $j \in J$ is represented by $Y_j = (Y_j(s))_{s \in S}$, where $Y_j(s)$ denotes the output of the good in state s at date 1 for every $s \in S$.

Any production plan can have at most two types of claims issued against it. These two claims are indexed by $k = 1, 2$. A specification of the claims against the production plan is called a financial structure. In the case where the claims are debt and equity the financial structure corresponds to the debt/equity ratio. The set of financial structures available to the producer is assumed (for ease of exposition in this section but not elsewhere) to be finite. The financial structures available to the firm are indexed by $e \in E_j$. The index set E_j is referred to as the set of available financial structures. For any financial structure $e \in E_j$ and $k = 1, 2$ let $r^k(e)$ denote the vector of dividends corresponding to the k-th claim in the financial structure e. Thus the properties of different financial structures and different claims are completely described by the dividend functions (r^1, r^2).

The dividend functions are assumed to satisfy certain natural conditions. Dividends are always non-negative: $r^1(e) \geq 0$ and $r^2(e) \geq 0$ for all structures $e \in E_j$. Inactivity is possible, that is, if the producer decides not to operate then $r^1(e) = r^2(e) = 0$. Except in the case of this trivial financial structure the dividends are assumed to exhaust the production plan, that is, $r^1(e) + r^2(e) = Y_j$. Also, it is possible to issue equity only: $r^1(e) = Y_j$ and $r^2(e) = 0$.

The cost function of producers of type $j \in J$ is denoted by C_j. For any financial structure $e \in E_j$, $C_j(e)$ is the cost of setting up a firm with a financial structure e. The cost of setting up the firm includes the cost of inputs to the production process (i.e., investment) as well as the costs of issuing securities. If a producer decides not to set up a firm and so issues no securities, no cost is incurred. The null structure corresponding to this action is denoted by $e = 0$ and so $C_j(0) = 0$.

Producers derive no utility from future consumption. Every producer maximizes his (current) profit taking as given the market value of firms with different financial structures. Let $v^k(e)$ denote the market value of the k-th claim on a firm with financial structure e for k = 1, 2 and $e \in E_j$. Then the market value of a firm with financial structure e is $v^1(e) + v^2(e) \equiv MV(e)$, say. A producer of type j chooses e to solve

$$\max_{e \in E_j} MV(e) - C_j(e) .$$

Since the producers' problem is non-convex there may not be a unique maximum. Let $\nu_j = (\nu_j(e))$ denote the distribution of financial structures chosen by producers of type j. That is, $\nu_j(e)$ is the measure of producers of type j who choose financial structure e for every $e \in E_j$.

It should be noted that there is not essential loss of generality assuming that producers maximize current profits. In any competitive market profit maximization in this sense is a necessary condition of equilibrium (see Hart (1979)).

Consumers

Let Ω denote the set of non-negative consumption bundles. For any consumption bundle $x \in \Omega$ it is useful to write $x = (x^0, x^1)$ where x^0 denotes consumption at date 0 and $x^1 = (x^1(s))_{s \in S}$ is the vector of consumption levels in different states at date 1. That is, $x^1(s)$ is the consumption level in state s at date 1 for every $s \in S$.

A consumer of type $i \in I$ is characterized by a von Neumann-Morgenstern utility function and an endowment of goods. The utility function of type i consumers is denoted by $\hat{U}_i$. For any consumption bundle $x \in \Omega$ the expected utility of x is denoted by $U_i(x)$ where

$$U_i(x) = \sum_{s\in S} \pi_s \hat{U}_i(x^0, x^1(s)) .$$

For some purposes it is more convenient to represent consumer preferences by U_i; for some purposes $\hat{U}_i$ is more convenient. The consumer's endowment is denoted by $w_i \in \Omega$. For simplicity, we assume that consumers' first period endowments are such that their consumption is strictly positive at $t = 0$.

Preferences are assumed in the technical appendix to be well behaved: U_i is strictly increasing and strictly quasi-concave.

A portfolio is represented by an ordered pair $\alpha = (\alpha^1, \alpha^2)$ of vectors. For any financial structure $e \in E \equiv \cup E_j$ and $k = 1, 2$, $\alpha^k(e)$ is the number of units of the k-th claim on a firm with financial structure e held in the portfolio. Since short sales are not allowed portfolios must be non-negative: $\alpha^1, \alpha^2 \geq 0$. Let A denote the set of non-negative portfolios.

Recall that $v^k(e)$ is the price in terms of consumption at date 0 of the k-th claim on a firm with financial structure e, for any $e \in E$ and $k = 1, 2$. For any security price function $v = (v^1, v^2)$, the value of a portfolio $\alpha \in A$ is denoted by $\alpha * v$ where

$$\alpha * v \equiv \sum_{k=1}^{2} \sum_{e\in E} \alpha^k(e)v^k(e) .$$

Similarly, the income at date 1 from a portfolio $\alpha \in A$ is denoted by $\alpha * \nu$ where

$$\alpha * \nu \equiv \sum_{k=1}^{2} \sum_{e\in E} \alpha^k(e)r^k(e) .$$

In this notation the budget constraints of a consumer of type i can be written as follows:

$$x^0 = w_i^0 - \alpha * v ,$$

$$x^1 = w_i^1 + \alpha * r .$$

(The budget constraints are written as equations because U_i is assumed to be strictly increasing.) For any security price function v and any choice of portfolio $\alpha \in A$ the budget constraints define a unique consumption bundle $x = \xi_i(\alpha, v)$. So the consumer's problem is to choose a portfolio α to maximize expected utility subject to the requirement that the consumption bundle $\xi_i(\alpha, v)$ implied by α and v is feasible. Formally,

$$\max_{\alpha \in A} U_i[\xi_i(\alpha, v)] \text{ S.T. } \xi_i(\alpha, v) \geq 0 .$$

<u>Equilibrium</u>

We define our equilibrium concept in two stages. First we outline a Walrasian type of equilibrium concept and then we add an extra condition. Our notion of equilibrium is similar to that suggested by Hart (1979) but with one important difference which is pointed out below.

In Walrasian equilibrium all agents maximize (expected) utility taking prices for both issued and unissued securities as given and all markets clear at the prevailing prices. Formally, a Walrasian equilibrium must specify a price function $v = (v^1, v^2)$, a portfolio $\alpha_i = (\alpha_i^1, \alpha_i^2)$ for each type of consumer $i \in I$, and a distribution ν_j of financial structures for each type of producer $j \in J$. The equilibrium conditions are:

(E1) for every type $i \in I$, α_i maximizes the consumers' utility subject to their budget constraints and the non-negativity constraints given the price function v;

(E2) for every type $j \in J$ and $e \in E_j$, if $\nu_j(e) > 0$ (i.e., if the structure e is chosen by some producers), then e maximizes the producers' profits given the price function v;

(E3) markets must clear: $\sum_{i \in I} \alpha_i = \sum_{j \in J} \nu_j$.

In a large economy the securities issued by any producer are negligible relative to the size of the economy as a whole. Without loss of generality we can assume that if a new security is issued it will be widely held and each consumer will hold a negligible amount of it. In that case a security is valued according to the marginal rates of substitution of those who hold it. The rational conjecture condition says that the price firms expect to receive if they issue a security is the maximum amount that any individual would be prepared to pay for a very small quantity of it. For any consumer type $i \in I$ with an equilibrium consumption bundle $x_i \in \Omega$ let $p_i(x_i)$ denote the vector of marginal rates of substitution defined by

$$p_i(x_i) = \frac{\partial U_i(x_i)}{\partial x^1} \Big/ \frac{\partial U_i(x_i)}{\partial x^0} .$$

If we write $p_i(x_i) = (p_{is}(x_i))$ then $p_{is}(x_i)$ is the amount of consumption a consumer of type i is willing to give up in exchange for an extra unit of consumption in state s at date 1, for any $s \in S$. Suppose a security offers a vector of dividends $r^k(e)$ at date 1. For some very small quantity ε of this security a consumer of type i ought to be willing to pay $\varepsilon p_i(x_i) \cdot r^k(e)$ units of consumption at date 0. In other words, the price he should pay is $p_i(x_i) \cdot r^k(e)$. The equilibrium price $v^k(e)$ should be the maximum such willingness to pay, that is

$$v^k(e) = \operatorname*{Max}_{i \in I} p_i(x_i) \cdot r^k(e)$$

for every claim k = 1, 2 and financial structure $e \in E$.

An equilibrium is defined to be a Walrasian equilibrium that satisfies the rational conjecture condition. This condition does not affect the equilibrium allocation in any substantive way. It is automatically satisfied in Walrasian equilibrium for every security that is actually issued. Because of the restrictions on short sales, this is not the case for securities that are not issued. However, if it is not satisfied for unissued securities in some Walrasian equilibrium then one can satisfy it simply by reducing the value of $v^k(e)$ appropriately without making any other change in the Walrasian equilibrium.

Hart (1979) distinguishes between markets that are open and markets that are closed. If a particular security is not issued by any firm he assumes that market to be closed. Also consumers are not allowed to trade in that security and they do not observe a market-clearing price for that security. On the other hand, producers know the price that would clear the market if a small quantity of a non-issued security were to be introduced.

An equilibrium as we have defined it differs from Hart's in one important respect. As far as producers are concerned the two concepts are the same. It really makes no difference whether producers see prices quoted for every possible security or have rational conjectures about what prices would be if a new security were issued. For consumers, on the other hand, it makes a difference whether they can trade every security at the quoted price, even if they choose not to trade some of them in equilibrium. In other words, there is a difference between a closed market and a market that is open but inactive because no trade takes place at the prevailing price. The importance of this difference is discussed in Section 3.

3. General properties of equilibrium

The first property we note is that the equilibria of our model are all constrained efficient. An equilibrium is said to be constrained efficient if a planner, who is subject to the same transaction costs as individual agents are, cannot make everyone better off. In this context the condition "subject to the same transaction costs" is interpreted to mean that the planner can only re-allocate securities and consumption at the first date. (The formal definition of constrained efficiency that we use is given in the Appendix.)

Under this interpretation the constrained efficiency of equilibrium is not surprising. Using the budget constraints to solve for second-period consumption we can express the utility of any agent in terms of first-period consumption and the portfolio of securities he holds at the end of the period. Then the economy is seen to be isomorphic to an Arrow-Debreu economy in which only first-period consumption and securities are traded. Pareto-efficiency in this Arrow-Debreu economy is equivalent to constrained efficiency in the original. So our result can be seen as a special case or application of the First Fundamental Theorem of welfare economics.
Thus we have the following result.

THEOREM 1: Every equilibrium is constrained efficient.

This result is not surprising when seen in the right perspective. Nonetheless it stands in sharp contrast to the claims of Hart (1980) and Makowski (1980). Both authors argue in the context of a product differentiation model formally similar to ours that "equilibrium" need not be efficient. The reason for this difference of opinion can be traced to their definition of equilibrium. It is important to understand the difference between their concept and ours because there are substantive modelling issues involved. Both Hart (1980) and Makowski (1980) use a concept of equilibrium introduced in Hart (1979). As outlined in Section 2, Hart's concept is as follows. Prices are quoted for securities that are actually issued in equilibrium. If a security is not issued there is no market for it. No price is quoted and consumers simply assume that it cannot be bought. On the other hand, firms must make conjectures about the price at which a security could be sold if it were to be introduced. Otherwise they could not make an optimal decision about which security to offer. These conjectures are assumed to be rational, that is, the conjectured price equals the maximum price at which the economy would absorb a small amount of the security. The rest of the definition is standard. Let us call Hart's concept a "conjectural equilibrium" to distinguish it from our "Walrasian equilibrium."

This sketch of conjectural equilibrium allows us to see the essential cause of market failure in the Hart-Makowski model and why it is absent in ours. There is a complementarity between products in their models that may lead to a pecuniary externality. This externality can be easily explained by an example. Suppose that nuts and bolts can only be produced by different firms. Nuts and bolts only have value if consumed together. If nuts are not produced the marginal value of a bolt is zero so bolts will not be produced either. The same argument shows that nuts will not be produced if bolts are not produced. Thus both types of firms are maximizing profits at zero output given the behavior of the other and their own (rational) conjectures. Yet this conjectural equilibrium is not efficient since a coordinated increase in the output of both products could make everyone better off.

The same kind of phenomenon could clearly occur in our model, if we analyzed it using the conjectural equilibrium concept. (Think of an economy with two types of firms, each of which has positive output in one state only.) The reason why it does not occur is that prices are quoted for every security whether it is actually issued or not. To revert to the nuts and bolts example, if the prices of nuts and bolts were both zero firms would be maximizing profit at zero output, but consumers would not be maximizing utility if they thought they could buy at the prevailing price and they still demanded none. So markets will not clear at a zero output level. In the Hart-Makowski model, on the other hand, the nuts and bolts markets are closed to consumers. We do not need to worry about clearing these markets.

At some level it probably seems natural to assume that a market does not exist if the corresponding good or security is not being produced. This kind of reasoning is encouraged by a tendency to equate markets with a gathering of people for the purchase and sale of commodities. But when we speak of a market existing in a Walrasian model, this is not what we have in mind. The existence of a market in this context is just another kind of equilibrium condition. A market "exists" if and only if a market clearing price prevails. That is, buyers and sellers can agree on a common price at which they all think they can trade any amount that is small relative to the economy as a whole. It is not immediately obvious why the existence of a market in this sense should depend on a non-zero amount of the corresponding commodity being traded. For example, at Lloyd's of London it is possible to obtain a price for insuring any risk even though no policy of that type is actively traded. Similarly, investment banks will quote prices for tailor-made securities.

There may be reasons why a market does not exist in this sense; but these reasons are not apparent in the Hart-Makowski model. There are no transaction costs. As soon as a firm decides to produce a positive amount of some commodity a market, that is to say, a market-clearing price, springs costlessly into existence. Nor is it clear that there are other obstacles to the free flow of information implied by a market-clearing price. After all, producers conjecture correctly the prices at which they can sell a nonexistent product even without the aid of a market. So it is not clear why, in equilibrium, there should not be a market-clearing price for every commodity. To say that some of these conditions will not be satisfied because the corresponding markets do not "exist" is tautologous. The existence of a market means that an equilibrium condition is satisfied, nothing more.

The next result concerns existence of equilibrium.

THEOREM 2: Under the maintained assumptions the equilibrium set is non-empty.

This theorem does not require comment except to note the role of the no-short-sales assumption in guaranteeing the existence of equilibrium. This assumption is needed to prevent the equalization of marginal rates of substitution everywhere. As we saw in the introduction, some divergence of marginal valuations between different consumer types is necessary to give firms an incentive to create costly securities.

The prohibition of short sales can also be used to avoid another well known non-existence problem, discussed in Hart (1975). This may suggest some similarity between the two problems but in fact they are rather different. Hart's problem arises from changes in the dimension of the linear space spanned by the securities. This problem cannot arise in a one-good economy. The present problem arises because short sales are only possible when there is a security to sell short. So there must be a discontinuity when the supply of the security hits zero if short sales are admitted.

The next question to be addressed is the number of securities that is needed in equilibrium. Without essential loss of generality it can be assumed that there is only one kind of firm. It will be clear that the argument generalizes easily to the case of $|J| > 1$ types. Let $e \in E$ be any particular split that is offered in equilibrium (i.e., supplied in a positive amount). Corresponding to this split are two securities $r^1(e)$ and $r^2(e)$. If the split is costly then no agent will want to hold both securities for he could hold the undivided firm more cheaply. In any case, we can assume that each type of consumer holds at most one of these securities. Then we can think of the set of firms offering the split e as being divided up among pairs of types of consumers. That is, for each firm offering the split e there is a pair of types $(i_1, i_2) \in I \times I$ such that the i_k types buy all the k-securities $r^k(e)$ and $i_1 \neq i_2$.

Now suppose there is a pair of consumer types (i_1, i_2) that invest in securities created by a number of different splits. By relabeling if necessary we can always assume that i_k buys the k-th security for $k = 1, 2$. Suppose this pair buys λ_1 units of a split e_1 and λ_2 units of a split e_2. If $r(\bar{E})$ is a convex set, where $\bar{E}$ denotes the set of non-trivial splits, there will exist a split $e_3 \in E$ such that $(\lambda_1 + \lambda_2)r(e_3) = \lambda_1 r(e_1) + \lambda_2 r(e_2)$. By buying $(\lambda_1 + \lambda_2)$ units of firms splitting according to e_3 agents of types i_1 and i_2 can attain the same consumption allocation. (Note the importance of assuming that consumers of type i_k hold all of the k-th security for any split e.) By extending this argument in the obvious way it is easy to see that only one split is needed for each pair of consumer types. In order for this reduction in the number of splits to leave firms no worse off, we must assume that costs are "convex" on $\bar{E}$ in the following sense: $(\lambda_1 + \lambda_2)C(e_3) \leq \lambda_1 C(e_1) + \lambda_2 C(e_2)$ for any $\lambda_1, \lambda_2 \geq 0$ and $e_1, e_2, e_3 \in \bar{E}$ such that $(\lambda_1 + \lambda_2)r(e_3) = \lambda_1 r(e_1) + \lambda_2 r(e_2)$. The number of split securities needed for an equilibrium allocation is no greater than the number of ordered pairs of consumer types. There are two split securities for every distinct pair so the number of split securities equals the number of ordered pairs. We may also need unsplit firms. Hence we have proved the following.

THEOREM 3: Suppose that $r(\bar{E}_j)$ is convex and C_j is convex on $\bar{E}_j$ for every $j \in J$. Then the total number of securities needed in equilibrium (or for any other constrained efficient allocation) is not greater than $|J| \cdot (|I| \cdot |I - 1| + 1)$.

This upper bound may be much too loose because many types of firms will not operate in equilibrium. In that case the bound is more usefully expressed as $(|I| \cdot |I - 1| + 1)$ securities per active type of firm.

We have chosen not to model firms' choice of production plan explicitly, preferring to simplify the analysis instead by treating each individual firm's output vector as a datum and allowing the firm to choose only whether to operate and its financial structure if it does operate. We can obtain an interesting extension of Theorem 3, however, if we place the model in a somewhat richer framework. Suppose that there is a continuum of firms all subject to a common technology. The technology is described by a compact, convex set y of output vectors and a cost function $C: E \to R_+$, where E denotes the set of feasible financial structures. Let $E = \{(\rho_1, \rho_2) \in R^S_+ \times R^S_+ \mid \rho_1 + \rho_2 = Y \in y\}$ and let $\bar{E}$ denote the set of non-trivial splits, that is, $\bar{E} = \{e \in E \mid r^k(e) \neq 0, k = 1, 2\}$. We can think of the model studied in this paper as an approximation of this richer setup. For example, the set of output vectors $\{Y_j: j \in J\}$ may be chosen arbitrarily close to y in the Hausdorff sense by introducing sufficiently many types of firm.

A production plan in this notation is a vector having the form $(- C(e), r^1(e) + r^2(e))$ for $e \in E$ and a firm chooses a production plan by choosing the financial structure e. One difference between the two approaches is that in this new setup the total measure of firms is given, rather than the measure of firms of a given type. If the measure of firms is large enough these constraints do not matter because not all firms operate. In that case there is effectively free entry and a zero profit condition holds.

With these assumptions and Theorem 3 behind us we can say something stronger about the number of securities required in equilibrium. Without loss of generality, consider a pair of consumer types i_1 and i_2 who jointly own firms with financial structures e_1 and e_2 in $\bar{E}$. We can assume, again without loss of generality, that consumers of type i_k own all of the k-th security in each case. Let λ_1 denote the measure of firms with structure e_1 owned by these consumers and let λ_2 denote the measure of firms with structure e_2. By convexity of y there exists a production plan (financial structure) e_3 such that

$$(\lambda_1 + \lambda_2)r^i(e_3) = \lambda_1 r^i(e_1) + \lambda_2 r^i(e_2)$$

for i = 1, 2. Thus we can replace the firms with financial structures e_1 and e_2 with an equal measure of firms with financial structure e_3 and leave consumers equally well off. To make sure firms are equally well off we must make the appropriate concavity assumption for C on $\bar{E}$. In this way we have shown that for each pair (i_1, i_2) of consumer types there need only be one production plan (financial structure) with two securities. We may also need firms with one security but only one production plan per consumer type. Then the total number of securities in equilibrium (or in any constrained efficient allocation) need not exceed $|I| \cdot |I - 1| + |I| = |I|^2$. Of course, this result depends on the existence of a single convex technology available to all potential firms. If there existed mutually exclusive technologies, each available only to a particular set of firms, then the upper bound given above would apply to each technology separately.

In any case we can now give Theorem 3 a more interesting interpretation. When choice of production plan together with convex technologies are assumed, the arguments that led to Theorem 3 imply a much tighter bound. The number of pairs of consumer types determines the securities needed irrespective of the number of types of active firms.

This version of Theorem 3 is much easier to compare to results obtained in standard stock-market economies (see Hart (1977)). When each firm issues only a single security (equity) the number of securities needed with a single convex technology is $|I|$, the number of consumer types. Each consumer type gets its own tailor made company (production plan) and that is clearly the best one can do.

4. Examples with debt and equity

This section develops illustrations of the case where firms are restricted to issue debt and equity. This situation can arise if the legal system is such that debt and equity are the only types of security that can be used. The defining characteristic of debt is that the par payment (i.e. the promised payment) is the same in all states. In states where the output of the firm is below the par payment the debtholders receive the entire output. If the firm's output is above the par payment in all states the debt is safe, otherwise it is risky.

The cost to a firm of issuing its first security is denoted c_1 and the marginal cost of issuing the second security is c_2 (i.e. the total cost of issuing two securities is $c_1 + c_2$). Example 1(a) is a specific illustration where c_1 is assumed to be zero and c_2 is strictly positive. Example 1(b) is a general version of this where c_1 and c_2 are both positive.

<u>Example 1(a)</u>

There are two states s = 1, 2 which are equally likely so $\pi_s = 0.5$. It is assumed

$$Y(1) = 1;\ Y(2) = 2.$$

There are two types of consumers i = a, n. For them

$$\hat{U}_a = x^0 + V(x^1) \qquad \text{with } V' > 0,\ V'' < 0\ ;$$

$$\hat{U}_n = x^0 + x^1\ .$$

When a specific functional form is used for V it is assumed

$$V(x^1) = 2\ \ln\ (1 + x^1)\ .$$

It is helpful to define the marginal utility of consumption of type i in state s

$$\mu_{is}(x^1) = p_{is}(x_i)\ /\ \pi_s\ .$$

It follows that $\mu_{as} > 1;\ \mu_{ns} = 1.$

The term μ_s is used to denote the marginal utility of consumption that is relevant for determining market values.

Since $c_1 = 0$ all firms set up and issue at least one security. First, consider the case where c_2 is sufficiently large that firms only issue one security which must be equity:

$$r^1 = Y(s) \quad \text{for all s.}$$

Its value is given by

where $\mu_s = \mu_{as} > 1$ for all s if endowments are such that only the risk averse group holds the security and $\mu_s = \mu_{ns} = 1$ for all s if they are such that the risk neutral group holds the security. Consider the case where both groups hold the security so that μ_s is given by the latter and $v^1 = 1.5$. The situation where the supply of the security is sufficiently small that only the risk averse group holds it can be similarly analyzed.

It can straightforwardly be shown that the risk averse group's demand for the firms' equity is:

$$\alpha^{1*} = 0.629$$

(where α^{1*} is used here and below to denote the demand of the risk averse group when $v^1 = 1.5$). The risk neutral group holds the remaining 0.371 of the security. The total value of a firm is $v^1 = 1.5$. The marginal utilities of consumption for the risk averse group are

$$\mu_{a1} = 1.228; \mu_{a2} = 0.886.$$

Suppose a firm issues debt so that the payoffs to its securities are:

Claim	Payoff		Price
	State 1	State 2	
r^1	1	1	1.057
r^2	0	1	0.500

Given all other firms are issuing equity which is priced with $\mu_s = \mu_{ns} = 1$ for s = 1, 2 it follows that the risk averse group will value this debt the most since $\Sigma\pi_s\mu_{as} = 1.057 > \Sigma\pi_s\mu_{ns} = 1.000$. Since they hold all of it, its price is determined by their marginal utilities of consumption μ_{as}. This firm's equity will be valued the most by the risk neutral group since $\mu_{n2} = 1.000 > \mu_{a2} = 0.886$ so they will hold all of it and its price will be determined by their marginal utilities of consumption. Thus the total value of such a firm would be $1.557 - c_2$.

It can be seen that the capital structure shown above is in fact the optimal one for a firm issuing debt and equity in this situation. The cost of issuing a second security is independent of its payoffs. Since the risk averse group holds the debt, its value would only increase by $\pi_2\mu_{a2} = 0.443$ for each unit of output the par payment on the debt is increased above 1. Since this is less than the $\pi_2\mu_{n2} = 0.5$ per unit of output the risk neutral group would be prepared to pay if the payoff were allocated to the equity they hold. Thus such a change would lower the value of the firm. Similarly a reduction in the par payment on the debt will also reduce the value of the firm since the risk averse group values these marginal payoffs more than the risk neutral group. A similar argument holds in all two-state examples: it is always optimal for a firm to issue the maximum amount of risk-free debt possible. It can easily be seen that this is a special feature of the two-state case: with three or more states firms can find it optimal to issue either risky or risk-free debt depending on the parameter values.

Thus for $c_2 > 0.057$ only equity will be issued by firms and for $c_2 < 0.057$ some proportion of the firms will issue both debt and equity so that three types of security will exist in total: the one-security firms' equity and the two-security firms' debt and equity. The risk neutral group of investors holds the equity of both the one- and two-security firms. The risk averse group holds the equity of the one-security firms and the debt of the two-security firms. The price of both types of equity is equal to its expected return and the price of the debt is such that firms are indifferent between issuing one or two securities. The debt is priced above its expected return in order to compensate firms for the cost that they bear in issuing it. For example, if $c_2 = 0.05$ then the price of debt is 1.05, the price of the equity of a two-security firm is 0.5 and of a one-security firm is 1.5. The proportion of firms that issues two securities is 0.111 and 0.889 issue one. The risk averse group holds 0.556 of the one-security firms' equity. As $c_2 \to 0$ more firms issue debt. When $c_2 = 0$ debt's price is equal to its expected return and the risk averse group just holds debt and the risk averse group holds the remainder of the securities. There is full risk sharing in the sense that marginal utilities of consumption at $t = 1$ are equated.

Example 1(b)

Consider next what happens when c_1 and c_2 are both positive. The example is otherwise the same as before except that now the only restriction on V is that $V'(1) = 1$ and $V'(0)$ is finite.

Figure 1 illustrates the relationship between the costs and the number of securities issued. Example 1(a) corresponds to the c_2 axis where $c_1 = 0$. For $c_2 > 0.5(1 - V'(2\alpha^{1*}))$ ($= 0.057$ in 1(a)) there are only one-security firms and so the total number of securities issued is one. For $c_2 < 0.5(1 - V'(2\alpha^{1*}))$ some firms issue debt and equity and the remainder only equity so that three securities are issued in total. The line marked "debt and equity" represents the remainder of the boundary between the regions where one and three securities are issued.

It is not worthwhile for any firms to set up and issue a security unless $c_1 < 1.5\, V'(0)$. For values of c_1 above this level no investment is undertaken and no securities are issued. As c_1 falls more firms are set up. Initially these are held entirely by the risk averse group since $V'(0) > 1$ and so μ_{a1} and μ_{a2} are greater than $\mu_{ns} = 1$ for all $s = 1,2$. Eventually there comes a value of c_1 where group a's demand is 0.5. At this point $\mu_{a2} = 1$. For $c_2 = 0$ it becomes worthwhile for some firms to issue debt and equity at this point since the risk averse group will be prepared to pay $\mu_{n2} = 1$ for the two-security firms' equity. The debt and one-security firms' equity is held by the risk averse group. As c_2 rises the boundary also rises because $\mu_{n2} - \mu_{a2}$ must be sufficiently large to allow two-security firms to recoup their additional costs. Along the boundary v^1 falls and α^1 increases. It is this that causes $\mu_{n2} - \mu_{a2}$ to rise.

When $c_1 = 1.5$ it is profitable for one-security firms to set up and be held by the risk neutral group. At this point there is <u>full investment</u> in the sense that all firms set up and issue at least one security. Since the risk neutral group hold at least some of the one-security firms' equity, it is priced at 1.5. There is no further scope for a fall in the price of the one-security firms' equity and so the boundary is horizontal as c_1 falls from 1.5 to 0.

The only region where two types of security are issued in total is along the c_1 axis where $c_2 = 0$ between 0 and 1.5. Here all firms set up and since the marginal cost of issuing a second security is zero they issue enough debt for the risk averse group to completely smooth their consumption across states. Thus in this region there is full investment and full risk sharing.

Another feature of this example is worth noting. Even though there are only two states there can nevertheless be three securities in existence without there being complete markets in the sense that full risk sharing is possible. Thus when there are costs of issuing securities the question of whether or not the number of securities is greater or less than the number of states is of less significance than when these costs are ignored.

The standard Modigliani-Miller theory, adapted to include corporate taxes and bankruptcy costs, suggests that firms choose their capital structure to trade off the tax advantage of debt against the costs of going bankrupt. Jensen and Meckling (1976), among others, have argued that a weakness of the theory is that it predicts that before the existence of the corporate income tax, debt would not be used at all since bankruptcy was costly. Of course, debt was widely used before the introduction of the tax. They argue that this demonstrates the importance of other factors, in particular asymmetric information, in determining firms' capital structures.

The above examples do not incorporate bankruptcy costs but it can readily be seen that including them does not change the nature of equilibrium significantly and debt will be used because of its role in improving the allocation of risk across consumers. Even though there is no tax advantage to debt and information is symmetric, the fact that firms bear a cost of issuing debt means it is priced so that in equilibrium they are indifferent between the optimal level of leverage and being unlevered. Thus when viewed in the context of a general equilibrium theory instead of the usual partial equilibrium one, the observation that debt was used by firms before the introduction of the corporate income tax does not necessarily imply that asymmetric information is an important determinant of firms' capital structures.

Myers (1984) has stressed that conventional theories of capital structure are inconsistent with the observation that similar firms in the same industry often have significantly different debt/equity ratios. However, it can be seen that the illustrative examples above have this feature. In equilibrium firms of the same type are indifferent between the optimal levered capital structure and being unlevered. In more complex examples with more types of consumers and firms it is possible for there to be many equally profitable optimal levels of leverage for firms of the same type.

5. Examples with optimal securities

In this section it is assumed that firms are not restricted to issuing debt and equity. Instead when they issue two securities they are free to allocate the firm's output in a particular state between the two securities in any way they wish. It is shown that the optimal securities are not debt and equity. Nevertheless, they do have a particularly simple form.

Example 1(a)

Consider first the initial example with c_2 such that both groups hold the equity of one-security firms. As before, group a's demand is 0.629, $\mu_{a1} = 1.228$ and $\mu_{a2} = 0.886$. If a firm were to issue debt as its second security it would obtain (0.5)(1.228) + (0.5)(0.886) = 1.057 for it and (0.5)(1) = 0.5 for its equity. However, this is not the best it can do. It can clearly increase its receipts by reducing the payment on its r^1 security in state 2 and increasing the payment on the other security. The r^1 security is held by the risk averse group who only value each unit of expected revenue in state 2 at 0.886 whereas the r^2 security is held by the risk neutral group who value each unit at 1.000. Similarly, the firm should not reduce the payment on the r^1 security in state 1 since it is held by the risk averse group who value each unit at 1.228 whereas the risk neutral group only value each unit at 1.000. This means the optimal pair of securities for a firm to issue have the following payoffs and prices.

Claim	Payoff		Price
	State 1	State 2	
r^1	1	0	0.614
r^2	0	2	1.000

The optimal securities thus have a particularly simple form. All the output in a particular state should be allocated to the security which is held by the group that values it most. This feature of optimal securities is quite general. It is due to the fact that firms' maximands are linear which arises from the competitive nature of the model. It does not depend on there only being two states and a risk averse and risk neutral group.

The total value of a firm which issues two securities when all other firms issue one security is $1.614 - c_2$. Hence firms will start issuing debt and equity at the critical cost $c_2^* = 0.114$. This contrasts with the case where firms issue debt and equity where the critical level $c_2 = 0.057$ is much lower. Thus optimal securities can permit a strictly better allocation of risk for a given (positive) issue cost than debt and equity.

<u>Example 1(b)</u>

Figure 1 illustrates what happens when c_1 and c_2 are both positive. The line marked "optimal securities" is the boundary in this case. The explanation of the form of the boundary is similar to before. The difference is that with optimal securities it is possible to allocate output more efficiently to those that value it most. With debt the problem is that firms cannot give less in state 2 than in state 1. As a result the boundary is shifted out. The only two points where they are the same is at the intersection with the c_1 axis at $1 + 0.5\, V'(0.5)$. The reason for this is that the critical factor here is when the amount of one-security firm equity held by the risk averse group is such that $\mu_{a1} = \mu_{n1} = 1$. This is unaffected by the form of the second security. It is only when $c_2 > 0$ and it is necessary that $\mu_{a1} = \mu_{n1} + c_2 > 1$ that this makes a difference.

Example 2

The only region in Figure 1 in which two securities in total are issued is when $c_2 = 0$ and c_1 is between 0 and 1.5. However, it is possible to construct examples where the total number of securities is two even when c_1 and c_2 are strictly positive. In order to illustrate this consider another example where both groups have the same (risk averse) utility function but differ in their endowment in the two states at $t = 1$. Group 1 receives 1 unit of output in state 1 and 0 in state 2. For group 2 the reverse is true. The output of the one type of firm that exists is

$$Y(1) = Y(2) = 1.$$

Hence the example is symmetric and it is this, as opposed to risk neutrality before, which simplifies it. There is one consumer of each type and one firm. The utility functions of consumers are as in Example 1(b).

Figure 2 illustrates the relationship between issuing costs and the total number of securities issued. When just one security is issued then it must be equity with a payoff of 1 in both states. Thus it will not be worth issuing just one security unless $c_1 < 0.5\, (V'(1) + V'(0))$. The best a two-security firm can do again involves splitting itself in such a way that all the output in a particular state is allocated to the group of securityholders that values it most. Thus the optimal securities have the following payoffs and prices:

Claim	Payoff		Price
	State 1	State 2	
r^1	1	0	V'(0)
r^2	0	1	V'(0)

It will be worth issuing at least two securities whenever $c_1 + c_2 < V'(0)$. It follows no securities at all will be issued to the right of the boundary with the kink at (0.5 (V'(1) + V'(0)), 0.5 (V'(0) - V'(1))).

For c_2 below the kink and to the left of this boundary firms will find it worthwhile to issue two securities as shown in Figure 2. In this region there is full risk sharing but not full investment: all the firms that set up allocate their output between their securityholders efficiently but not all firms that could set up do so. As c_1 falls more firms set up. When $c_2 = 0$ the point where all firms invest is reached when $c_1 = V'(1)$. To the left of this all firms still just issue two securities and there is both full risk sharing and full investment. For small values of c_2 those firms which issue two securities must do at least as well as those that issue one. For sufficiently small values of c_1 firms are indifferent between issuing one and two securities. But for large values it is not worthwhile issuing just one: firms can do strictly better by issuing two. The points at which all firms issue two securities but are indifferent between this and just issuing one gives the boundary between the regions where two and three securities are issued in total.

All firms find it optimal to just issue one security above and to the left of the kink at (0.5(V'(1) + V'(0)), 0.5(V'(0) - V'(1))). The boundary between this and the region where three securities are issued in total occurs at the points where all firms issue one security but are indifferent between this and issuing two. Thus when $c_1 = 0$ and all firms issue at least one security which is held equally by the two groups it occurs at the point where $c_2 = 0.5(V'(0.5) - V'(1.5))$. In the case shown in Figure 2 $V''' > 0$ and this is below $0.5(V'(0) - V'(1))$; otherwise it would not be. When c_1 is sufficiently small all firms set up and the boundary is horizontal. However, when $c_1 > 0.5(V'(0.5) + V'(1.5))$ this is not the case and the boundary slopes up as c_1 increases and the amount of equity held in the one-security region falls.

Figure 3 shows how the efficiency of investment and risk sharing are affected by issuing costs. It can be seen that provided the marginal cost of issuing the first security is sufficiently low there is full investment efficiency and provided the marginal cost of issuing the second security is sufficiently low there is full risk sharing efficiency. When both are high there is a compromise between the two. The only case where there is full investment and full risk sharing so the allocation is fully efficient is when $c_2 = 0$ and c_1 is sufficiently small.

In summary, this section has developed simple examples to illustrate the optimal design of securities. It has been shown that debt and equity are not the optimal securities even in cases where one group is risk neutral and one is risk averse. Instead it is optimal to split the firm in a particularly simple way. The firm should be marketed to the pair of groups of securityholders which leads to the highest valuation of the firm. The output in each state should be allocated to the group which values it most highly with the other getting zero in that state. Only in states where both groups value consumption equally is it optimal to have both securityholders having positive payoffs. Although this has only been demonstrated in the context of very simple examples it can be seen that it applies in much more complex environments with many states, types of firm and types of consumer.

Of course in reality firms do issue debt and equity. However, as pointed out in the introduction the model is meant to be an abstraction which gives insight: the result demonstrates that the basic principle of security design is that the firm should be split in such a way that in any state all the payoffs are allocated to the group that values it most. We have not taken into account all the relevant institutional details. By extending the model to allow for richer institutional and tax environments, more directly descriptive results should be obtainable. For example, we have not incorporated the tax deductibility of interest or possibilities for intermediation. What we might actually expect to obtain in a model incorporating these factors is firms issuing debt to get its tax advantages and intermediaries repackaging it to get any risk sharing advantages.

6. The role of short-sale constraints

With issuing costs, financial structure must matter in order to provide an incentive to firms to issue securities. This potentially creates arbitrage possibilities which must be ruled out by some sort of friction if existence of equilibrium is to be assured. In every case above it is assumed that short sales are so costly that they are not undertaken. In practice the costs of undertaking short selling are large and the actual amount that is done by non-members of exchanges is small (see, e.g., Pollack (1986)). The purpose of this section is to consider the case where there are costs but these are not sufficient to rule out all short sales.

to construct a perfectly hedged portfolio and earn a profit equal to c_2 for each firm that is short sold. By taking a sufficiently large position it will always be possible to more than cover c_2. Hence some marginal cost of short selling is necessary if an equilibrium is to exist with short sales. A constant marginal cost is not sufficient since this either rules out short sales completely if it is greater than c_2 per firm or fails to prevent the nonexistence problem if it is less than this. Hence the case of interest is where there is a nonlinear cost of short selling.

There are two obvious possibilities: either the costs depend on the number of shares short sold or the value of shares short sold. Given the arbitrariness of assuming any number of shares in a firm it is perhaps easier to adopt the latter approach here. In particular it is assumed that the costs of short selling, σ, are an increasing convex function of the total value of securities, Z, that is short sold. For the purposes of illustration it is convenient to assume:

$$\sigma = Z^2 .$$

In the equilibrium of Example 1(a) above the short sale constraints bind on both the risk neutral and the risk averse group. The risk neutral group would of course like to short sell the r^1 security with payoffs (1, 0) since its price is greater than its expected return. However, the risk averse group would also like to short sell the r^2 security with payoffs (0, 2). This is because by combining this with a long position in the one-security firm's equity they can effectively create the r^1 security at a price equal to its expected return. Thus what happens when there are nonlinear costs and only limited short selling is possible is that the n's expand the supply of the r^1 security and the a's expand the supply of the r^2 security.

Consider what happens at the critical value $c_2^* = 0.114$ where introducing a two-security firm at prices $v^1 = 1.000$ and $v^2 = 0.614$ becomes worthwhile. As soon as some firm issues r^1 and r^2 the two groups will short sell them thereby expanding the supplies of both. The price of r^1 in particular must then fall to equate demand and supply. But at a lower price it is not worthwhile any firm issuing two securities. Hence no equilibrium exists similarly to the case with unlimited short sales. In order for an equilibrium to exist, c_2 must be sufficiently small that at the price where demand equals supply for r^1 it is nevertheless worthwhile for a firm to issue two securities. For all values of c_2 below the critical level $c_2^{**} = 0.076$ where this becomes feasible, an equilibrium exists. At this point the n's short sell an amount 0.115 of r^1 and the a's short sell an amount 0.038 of r^2. However, between c_2^* and c_2^{**} no equilibrium exists as shown in Figure 4 which summarizes when an equilibrium exists and when it does not.

The possibility for short selling effectively introduces another way for securities to be issued. Since in the example the costs of shortselling are quadratic it is always worthwhile for individuals to issue a limited amount of securities because for small amounts they have a cost advantage over firms in expanding the supply of the security. At least for low values of c_2 this expanded low cost supply of the security leads to a better allocation of risk than would occur without short selling. The problem is that for intermediate values of c_2, the competition between firms and individuals can lead to the nonexistence of equilibrium.

7. Concluding remarks

This paper develops a framework for considering the question of optimal security design. Since this is only of interest when markets are incomplete, a model is used where the absence of full risk sharing possibilities is due to the existence of transaction costs for issuing securities. In constructing any model of this type two questions must be addressed. The first concerns the form of friction that must be introduced so that equilibrium can exist. The second involves the way in which the markets for unissued securities are represented. The approaches taken above of introducing costs of short sales and assuming that both firms and consumers know the prices of unissued securities are not the only alternatives for resolving these. Their main advantage is that they provide a simple and tractable benchmark. Within this framework many issues remain to be addressed. Among other things, an important question which we hope to pursue in future research concerns the optimal design of securities in multiperiod models. In addition to implications for capital structure, such analyses should also have implications for dividend policy. The role of the assumptions concerning short sales and unissued securities also points to the importance of research aimed at providing a more detailed understanding of the microstructure of markets.

Appendix

1. Formal definition of the model

There are two dates t = 0, 1 and a finite set of states of nature $s \in S$. All agents have the same information structure: there is no information at the first date and the true state is revealed at the second. At each date there is a single consumption good. Since commodities are distinguished by the date and state in which they are delivered the commodity space is $R^{|S|+1}$. There is a finite set of types of consumer $i \in I$ and a finite set of types of producer $j \in J$. To justify the assumption of perfect competition, we assume there is a continuum of agents of each type. For each $i \in I$ let $m_i > 0$ denote the measure of consumers of type i; for each $j \in J$ let $m_j > 0$ denote the measure of producers of type j.

Producers

A type of producer is defined by a production plan, a security set and a cost function. The production plan of type $j \in J$ is denoted by $Y_j : S \to R_+$, where $Y_j(s)$ denotes the output of the consumption good in state s at date 1 for every $s \in S$. The security set of type $j \in J$ is denoted by E_j. E_j is simply an index set. The properties of the different securities are described by two dividend functions (r^1, r^2) defined on the global security set $E = \cup_{j \in J} E_j$. Any production plan can have at most two types of claims issued against it. For any $e \in E$ these two claims are indexed by k = 1, 2. Thus $r^k(e)$ denotes the vector of dividends corresponding to the k-th claim of type e. Securities are assumed to satisfy the following conditions.

(A.1) (a) E_j is a compact metric space for every $j \in J$.

(b) For all $j \in J$ there exists $e \in E_j$ such that $r^1(e) = r^2(e) = 0$.

(c) For all $j \in J$ there exists $e \in E_j$ such that $r^1(e) = Y_j$ and $r^2(e) = 0$.

(d) For all $j \in J$ and all $e \in E_j$, $r^1(e) + r^2(e) = Y_j$ unless $r^1(e) = r^2(e) = 0$.

(A.2) $r^k : E \to R_+^{|S|}$ is a continuous function.

Since E is an index set, assumptions (A.1a) and (A.2) are technical conveniences. (A.1b) represents the possibility of issuing no securities, i.e. the producer decides not to set up a firm. (A.1c) says that it is possible to issue a single security (equity). (A.1d) says that except in the case where no securities are issued, the claims must exhaust the total product of the firm.

The cost function of producers of type $j \in J$ is denoted by $C_j : E_j \to R_+$. $C_j(e)$ is the cost in units of output at date 0, of issuing securities of type $e \in E_j$. Without essential loss of generality we may assume that these costs include the cost of setting up the firm in the first place. It is natural to assume that

(A.3) (a) $C_j(e) = 0$ if $e \in E_j$ and $r^1(e) = r^2(e) = 0$;

(b) $C_j : E_j \to R_+$ is continuous.

Every producer is assumed to maximize his profit, taking as given the market value of the different kinds of firms he can set up. Let V(e) denote the value of a firm of type $e \in E$ measured in units of consumption goods at date 0. Then a producer of type $j \in J$ chooses $e \in E_j$ to maximize $V(e) - C_j(e)$. Not all producers of the same type will choose the same set of securities.

The equilibrium choices of the j-th type are represented by a measure ν_j defined on E_j. A measure like ν_j is called a distribution; for any measurable set $H \subset E_j$, $\nu_j(H)$ is the measure of producers of type j who choose securities of type $e \in H$. A distribution ν_j is admissible iff $\nu_j(E_j) = n_j$, i.e., if all producers of type j are accounted for. N_j denotes the set of admissible distributions of producers of type $j \in J$. Then for each $j \in J$ we can define the optimal security correspondence ψ_j by putting

$$\psi_j(V) = \arg\max_{\nu \in N_j} \int_{E_j} (V - C_j)\, d\nu$$

for any value function V.

Consumers

Every consumer is characterized by a consumption set, a portfolio set, a utility function and an endowment. All consumers have the same consumption set $\Omega = R_+^{|S|+1}$ with generic element $x \in \Omega$. It is convenient to partition the consumption set $\Omega = \Omega^0 \times \Omega^1$ and to partition consumption bundles $x = (x^0, x^1) \in \Omega^0 \times \Omega^1$. ($x^0 \in R_+$ denotes consumption at date 0 and $x^1 \in R_+^{|S|}$ denotes consumption at date 1.) The utility function of consumers of type $i \in I$ is denoted by $U_i : \Omega \to R$; the endowment is denoted by $w_i \in \Omega$. It is assumed initial endowments are such that $x^0 > 0$.

A portfolio is a signed, vector-valued measure $\alpha = (\alpha^1, \alpha^2)$ defined on E. For any measurable set $H \subset E$ and any $k = 1, 2$ $\alpha^k(H)$ is the number of units of the k-th claim of type $e \in E$ held in the portfolio. No short sales are allowed so α has values in R_+^2. All consumers have the same set of admissible portfolios denoted by A.

A consumer of type $i \in I$ is characterized by the array (Ω, A, U_i, w_i). We assume each type satisfies the following properties.

(A.4) (a) $U_i : \Omega \to R$ is strictly quasi-concave and strictly increasing for every $i \in I$;

(b) U_i is C^1 on the interior of Ω and continuous at the boundary, for every $i \in I$.

Assumption 4 allows us to define a function $p_i : \text{int}\ \Omega \to R^{|S|}$ for each $i \in I$ by putting

$$p_i(x) = \frac{\partial U_i/\partial x^1}{\partial U_i/\partial x^0}$$

for every $x \in \text{int}\ \Omega$. Since U_i is strictly increasing $p_i(x) \gg 0$ for every $x \in \text{int}\ \Omega$.

A security price function is a function $v \equiv (v^1, v^2)$ defined on E to R^2. For any $e \in E$ and $k = 1, 2$, $v^k(e)$ is the price measured in units of consumption at date 0, of the k-th claim of type e. We restrict attention to continuous price functions. For any admissible portfolio $\alpha \in A$ the value of the portfolio is defined by $\alpha * v$ where

$$\alpha * v = \sum_{k=1}^{2} \int_E v^k \, d\alpha^k .$$

Similarly, the total income at date 1 from a portfolio $\alpha \in A$ is denoted by $\alpha * r$ where $r = (r^1, r^2)$ and

$$\alpha * r = \sum_{k=1}^{2} \int_E r^k \, d\alpha^k .$$

In this notation we can write the budget constraints for a consumer of type $i \in I$ as follows:

$$x^0 = w_i^0 - \alpha * v$$

$$x^1 = w_i^1 + \alpha * r .$$

(The budget constraints are equations because U_i is strictly increasing.) Using these equations we can write the consumption bundle x as a function of α and v. So we write $x = \xi_i(\alpha, v)$ for every $i \in I$. Then define the budget set for type $i \in I$ to be

$$\beta_i(v) = \{ \alpha \in A \mid \xi_i(\alpha, v) \in \Omega\}$$

for every price function v. The optimal portfolio correspondence is defined by putting

$$\phi_i(v) = \underset{\alpha \in \beta_i(v)}{\arg\max}\ U_i(\xi_i(\alpha, v))$$

for any price function v.

Equilibrium

We define our concept of equilibrium in two stages. First we outline a Walrasian type of concept; then we add an extra condition suggested by Hart (1979). In Walrasian equilibrium all agents maximize utility taking prices as given and all markets clear at the prevailing prices. An equilibrium must specify a portfolio α_i for each type of consumer $i \in I$, a distribution ν_j for each type of producer $j \in J$ and a continuous price function $v : E \to R_+^2$. We define a Walrasian equilibrium to be an array $\langle(\alpha_i)_{i \in I}, (\nu_j)_{j \in J}, v\rangle$ satisfying the following properties:

(E.1) $$\alpha_i \in \phi_i(v) \text{ for every } i \in I;$$

(E.2) $$\nu_j \in \psi_j(v^1 + v^2) \text{ for every } j \in J;$$

(E.3) $$\sum_{i \in I} m_i \alpha_i = \sum_{j \in J} \nu_j \text{ .}$$

(E.1) and (E.2) require utility maximisation on the part of consumers and producers respectively; (E.3) requires market clearing.

In a large economy the securities issued by any producer are negligible relative to the size of the economy as a whole. Without loss of generality we can assume that if a new security is issued it will be widely held and each consumer will hold a negligible amount of it. In that case a security is valued according to the marginal rates of substitution of those who hold it The rational conjecture condition can be written as follows:

(E.4) $v^k(e) = \max \{p_i(\xi_i(\alpha_i, v)) \cdot r^k(e)\}$ for every $e \in E$ and $j = 1, 2$.

An equilibrium is defined to be a Walrasian equilibrium that satisfies (E.4).

Condition (E.4) does not affect the equilibrium allocation in any substantive way. (E.4) is automatically satisfied in Walrasian equilibrium for every security that is actually issued. If (E.4) is not satisfied in some Walrasian equilibrium one can satisfy it simply by reducing the value of $v^k(e)$ appropriately without making any other change in the Walrasian equilibrium.

Efficiency

The appropriate concept of efficiency is constrained efficiency. We ask whether a central planner can make everyone better off using only the markets and technologies available to private individuals. (Under our assumptions there is no difference between making everyone better off and making some people better off and no one worse off.) The allocation at date 1 is determined by trades in securities at date 0. Therefore an equilibrium is constrained efficient iff it is impossible to make everyone better off by means of transfers of goods and securities at date 0. Formally, an equilibrium $\langle(\alpha_i)_{i\in I}, (\nu_j)_{j\in J}, v\rangle$ is constrained efficient iff there does not exist an obtainable allocation $\langle(\hat{\alpha}_i)_{i\in I}, (\hat{\nu}_j)_{j\in J}\rangle$ and transfers $(\tau_h)_{h\in I\cup J}$ satisfying these conditions:

(a) $\hat{\alpha}_i \in A$, $\hat{\nu}_j \in N_j$, $\tau_h \in R$ for every $i \in I$, $j \in J$ and $h \in I \cup J$.

(b) $\sum_{i\in I} m_i\hat{\alpha}_i = \sum_{j\in J} \hat{\nu}_j$ and $\sum_{i\in I} m_i\tau_i + \sum_{j\in J} n_j\tau_j = 0$;

(c) $\int_{E_j}(v^1 + v^2 - C_j)d\nu_j < \int_{E_j}(v^1 + v^2 - C_j)d\hat{\nu}_j + \tau_j$ for every $j \in J$;

(d) $U_i(\xi_i(\alpha_i, v)) < U_i[\xi_i(\hat{\alpha}_i, v) + (\tau_i, 0)]$ for every $i \in I$.

[The fact that we use the equilibrium price function v in the definition of constrained efficiency is immaterial. It simply allows us to use the notation introduced earlier.]

2. Proof of Theorems 1 and 2

THEOREM 1: Every equilibrium is constrained efficient.

Proof: Suppose not. Then in the previous notation

$$\xi_i^0(\hat{\alpha}_i, v) + \hat{\alpha}_i * v + \tau_i > w_i^0 \quad (i \in I)$$

and

(*) $$\int_{E_j}[v^1 + v^2 - C_j + \tau_j]d\hat{\nu}_j > \int_{E_j}[v^1 + v^2 - C_j]d\nu_j \quad (j \in J) .$$

Adding up we get

$$\sum_{i\in I} m_i\{\xi_i^0(\hat{\alpha}_i, v) + \hat{\alpha}_i * v + \tau_i\} > \sum_{i\in I} w_i^0 m_i$$

so attainability implies $\sum \int_E d\nu_j < 0$. From the equilibrium conditions we know that

$$\int_{E_j}[v^1 + v^2 - C_j]d\hat{\nu}_j \leq \int_{E_j}[v^1 + v^2 - C_j]d\nu_j$$

so for some $j \in J$, (*) is violated. ‖

THEOREM 2: Under the maintained assumptions the equilibrium set is non-empty.

Proof: For each q = 1, 2, ... let E_j^q be a finite subset of E_j satisfying the same properties as E_j and such that $E_j^q \to E_j$ as $q \to \infty$, where convergence of E_j^q is with respect to the Hausdorff metric. We prove existence for the economy with E_j replaced by E_j^q for all $j \in J$ and then the theorem follows by taking limits.

Let $E^q = \cup_{j \in J} E_j^q$ and let (A^q) denote the set of non-negative measures on E^q with the property that

$$\alpha^k(\{e\}) \leq \frac{(n_j + 1)}{\min\ \{m_i\}}$$

for any $e \in E_j$, k = 1, 2, $j \in J$, $\alpha \in A^q$. Thus any portfolio α observed in equilibrium must certainly belong to A^q. For each $i \in I$ let $\beta_i^q(v) = \beta_i(v) \cap (A^q)$ and define $\phi_i^q(v)$ by

$$\phi_i^q(v) = \arg\max_{\alpha \in \beta_i^q(v)} U_i(\xi_i(\alpha, v))$$

for any $e \in E^q$, k = 1, 2 and any consumption bundle x_i.

By a standard fixed point argument there exists $(z^q, v^q) \in R_+^{Eq} \times R_+^{Eq} \times V^q$ such that $z^q \in \zeta^q(v^q)$ and $v^q * z^q \geq V^q * z^q$. If $z^{qk}(e) > 0$ then $v^{qk}(e) = k$ but in that case $\alpha^k(e) = 0$ for any $\alpha \in \phi_i^q(v^q)$ and any $i \in I$, a contradiction. Conversely, $z^{qk}(e) < 0$ implies $v^{qk}(e) = 0$ in which case $\alpha^k(e) = (n_j + 1)/\min\ \{m_i\}$ for some j where $e \in E_j$ and $\alpha \in \phi_i^q(v^q)$. (Unless

$r^k(e) = 0$; in that case put $z^{qk}(e) = 0$ w.l.o.g.) Again we have a contradiction. So $z^q = 0$.

It is straightforward to check that conditions (a)-(c) of the definition of equilibrium are satisfied (i.e., none of the artificial constraints are binding). To ensure condition (d) simply replace $v^q(e)$ by

$$\min\{v^q(e), \max_{i \in I} p_i(\xi_i(\alpha_i, v^q))\}$$

where $\alpha_i \in \phi_i^q(v^q)$. (Note that strict quasi-concavity of U_i implies that $\xi_i(\alpha_i, v^q)$ is unique.) It is easy to check that this does not disturb the other equilibrium conditions for any $v\colon E^q \to R_+^2$. Let

$$V^q = \{v^1 E^q \to R_+^2 | v^1(e) + v^2(e) \le k\}$$

where k is very large. Then

$$\phi_i^q\colon \; V^q \to (A^q)$$

is u.h.c., non-empty and convex-valued. Similarly, define ψ_j^q by putting N_j^q equal to the set of non-negative real valued measures on E_j^q such that $\nu_j(E_j^q) = n_j$ for all $\nu_j \in N_j^q$ and then

$$\psi_j^q(v) = \arg\max_{\nu \in N_j^q} \sum_{k=1}^{2} \int_{E_j^q} (V - C_j) d\nu^k$$

for any $v \in V^q$. Then the usual arguments show

$$\psi_j^q\colon \; V^q \to M(E_j^q)$$

is u.h.c., non-empty and convex-valued. Define a correspondence

$$\zeta^q\colon \; V^q \to R^{Eq} \times R^{Eq}$$

by

$$\zeta^q(v) = \sum_{i \in I} m_i \phi_i^q(v) - \sum_{j \in J} \hat{\psi}_j^q(v)$$

for every $v \in V^q$ where $\hat{\psi}_j^q(v) = \{(\nu, \nu) | \nu \in \psi_j^q(v)\}$. Then ζ^q inherits the properties of (ψ_j^q) and (ϕ_i^q). Choose k in the definition of V^q so that in any attainable allocation, $p_i(x_i) \cdot r^k(e) < k$.

For each E^q we have an equilibrium $\{(\alpha_i^q), (\nu_j^q), v^q\}$ for the corresponding economy. The measures (α_i^q), (ν_j^q) can be extended to all of E in an obvious way; so can v^q and condition (d) tells us that this extension is continuous. Since (α_i^q) and (ν_i^q) are bounded there exist weakly convergent subsequences which we can take to be the original sequence. Along this sequence $\xi_i(\alpha_i^q, v^q)$ is unique for each i and q so we can choose a further subsequence along which $\xi_i(\alpha^q, v^q)$ converges. (Obviously $\{\xi_i(\alpha^q, v^q)\}$ is bounded.) Along this sequence v^q converges to a continuous function $v\colon E \to R_+^2$. We claim that the limit point $\{(\alpha_i), (\nu_j), v\}$ is the required equilibrium. By continuity conditions (c) and (d) are satisfied. Suppose, contrary to what we want to prove that (a) is not satisfied. Then there exists, for some $i \in I$, $\hat{\alpha}_i \in \beta_i(v)$ which is preferred to α_i. This is impossible unless $\hat{\alpha}_i * v > 0$ so without loss of generality we can assume that $w_i^0 - \hat{\alpha}_i * v > 0$ by continuity. Since $\alpha * r \geq 0$ for any $\alpha \in A^q$, for q sufficiently large we can find $\hat{\alpha}^q \in \beta_i^q(v^q)$ arbitrarily close to $\hat{\alpha}_i$ in the weak sense. (This requires the facts that E is compact and $v^q \to v$ uniformly on E.) But for q large enough and $\hat{\alpha}^q$ close enough to $\hat{\alpha}_i$, $\hat{\alpha}^q$ must be preferred to α_i^q, a contradiction. This proves (a) and the proof of (b) is similar. ‖

References

Baron, David P., 1974, Default Risk, Home-Made Leverage and the Modigliani-Miller Theorem, American Economic Review 62, 176-182.

Baron, David P., 1976, Default Risk and the Modigliani-Miller Theorem: A Synthesis, American Economic Review 64, 204-212.

Diamond, Peter A., 1967, The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty, American Economic Review 57, 759-766.

Duffie, Darrell, and Matthew O. Jackson, 1986, Optimal Innovation of Futures Contracts, Research Paper 917, Stanford Graduate School of Business.

Hart, Oliver D., 1975, On the Optimality of Equilibrium when Market Structure is Incomplete, Journal of Economic Theory 11, 418-443.

Hart, Oliver D., 1977, Take-Over Bids and Stock Market Equilibrium, Journal of Economic Theory 16, 53-83.

Hart, Oliver D., 1979, On Shareholder Unanimity in Large Stock Market Economies, Econometrica 47, 1057-1084.

Hart, Oliver D., 1980, Perfect Competition and Optimal Product Differentiation, Journal of Economic Theory 22, 279-312.

Hellwig, Martin F., 1981, Bankruptcy, Limited Liability and the Modigliani-Miller Theorem, American Economic Review 71, 155-170.

Jensen, Michael C., and William H. Meckling, 1976, Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, Journal of Financial Economics 3, 305-360.

Makowski, Louis, 1980, Perfect Competition, the Profit Criterion and the Organization of Economic Activity, Journal of Economic Theory 22, 222-242.

Miller, Merton H., 1986, Financial Innovation: The Last Twenty Years and the Next, Journal of Financial and Quantitative Analysis 21, 459-471.

Modigliani, Franco and Merton H. Miller, 1958, The Cost of Capital, Corporation Finance, and the Theory of Finance, American Economic Review 48, 162-197.

Myers, Stewart C., 1984, The Capital Structure Puzzle, Journal of Finance 39, 575-592.

Pollack, Irving M., 1986, Short-Sale Regulation of NASDAQ Securities, NASDAQ pamphlet.

Radner, Roy, 1972, Existence of Equilibrium Plans, Prices and Price Expectations in a Sequence of Markets, Econometrica 40, 289-303.

Silber, William L., 1981, Innovation, Competition and New Contract Design in Futures Markets, Journal of Futures Markets 1, 123-155.

Stiglitz, Joseph E., 1969, A Re-Examination of the Modigliani-Miller Theorem, American Economic Review 59, 784-793.

Stiglitz, Joseph E., 1974, On the Irrelevance of Corporate Financial Policy, American Economic Review 64, 851-866.

Williams, Joseph T., 1986, Bonds, Stocks and Options as Optimal Contracts between Corporate Claimants, mimeo, Graduate School of Business, New York University.

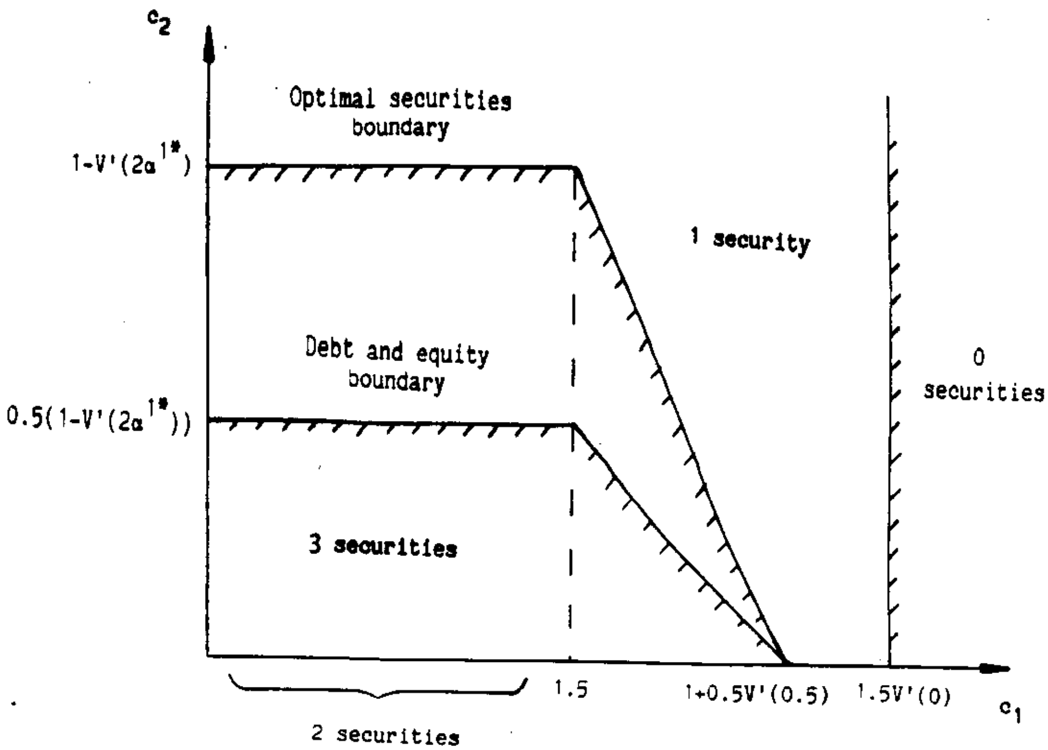

Figure 1

The relationship between issuing costs and the number of securities when utility functions differ

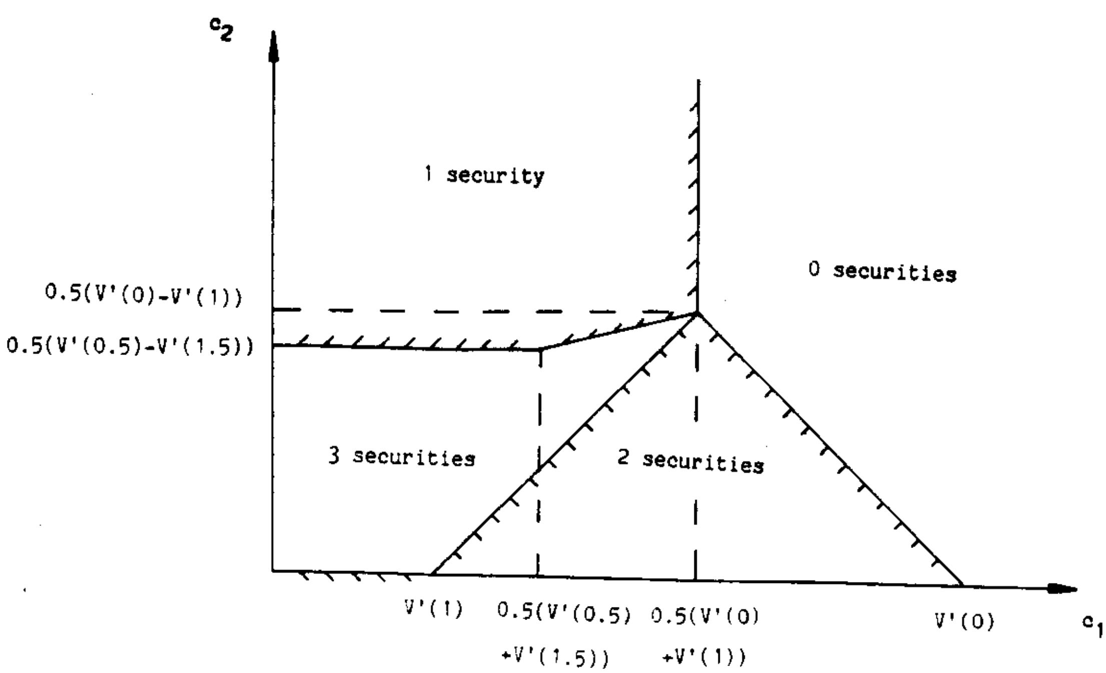

Figure 2

The relationship between issuing costs and the number of securities when endowments differ

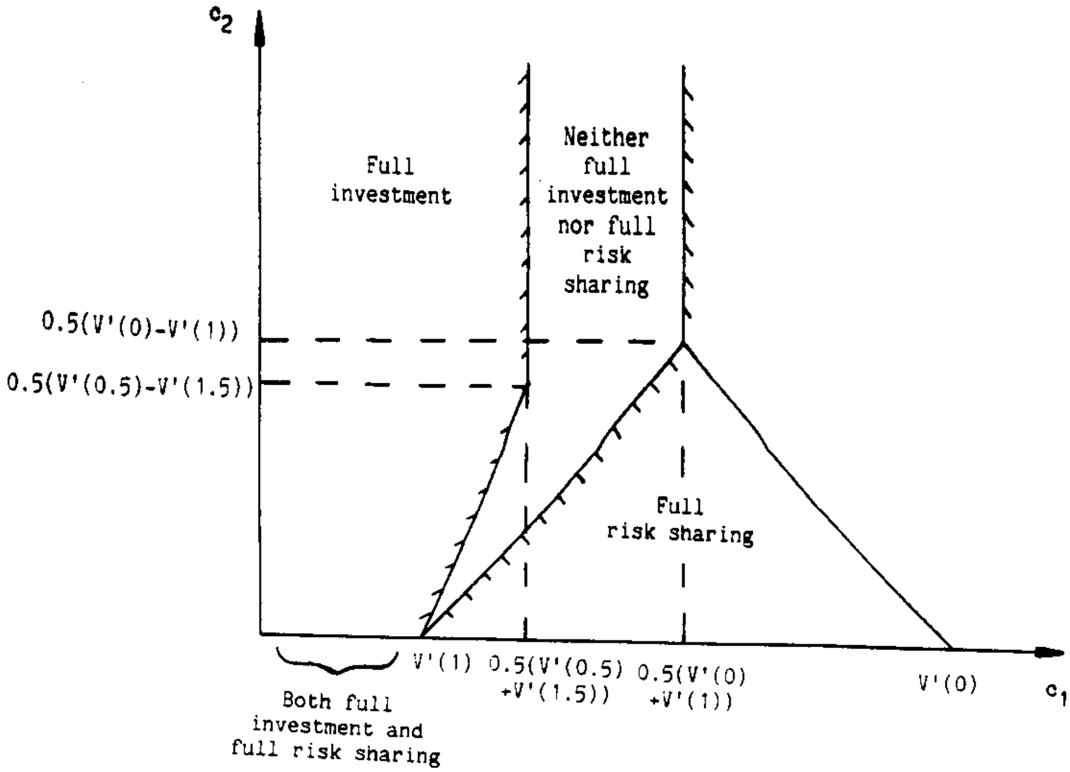

Figure 3

The relationship between issuing costs and investment and risk sharing

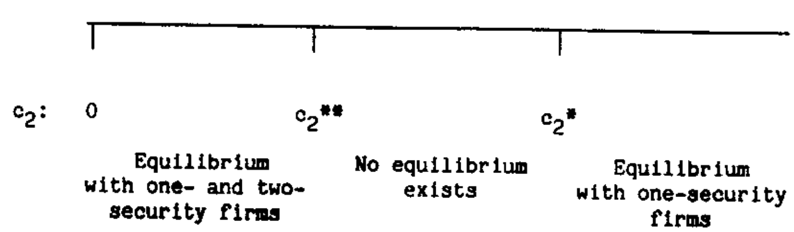

Figure 4

The relationship between c_2 and the existence of equilibrium when there is limited short selling