LOAN SIZE AND CREDIT RATIONING UNDER ASYMMETRIC INFORMATION

PIER GIORGIO ARDENI

MARCELLO MESSORI

No 195

Loan Size and Credit Rationing Under Asymmetric Information

Pier Giorgio Ardeni
Dipartimento di Scienze Economiche
University of Bologna

Marcello Messori
Dipartimento di Economia e Territorio
University of Cassino (FR)

Abstract

In this paper we analyze the effects of adverse selection due to asymmetric information on the optimizing behavior of risk neutral firms and banks in a competitive loan market. Realized firm returns have the *monotone likelihood ratio property* (MLRP) with respect to quality, in the sense of Milgrom (1981). This property encompasses the assumptions which characterize previous models such as Stiglitz and Weiss (1981) and De Meza and Webb (1987), allowing for a more general framework. Moreover, similarly to Milde and Riley (1988), the present model has loans of variable size, as opposed to the fixed loan size of Stiglitz and Weiss and De Meza and Webb. The references to a parameter of "organizational complexity" of the firm, defined by the prevailing type of economies of scale and costs, and to a Wilson (1977) construction of the contracting game, whereby uninformed lenders move first in a three-stage pure strategy game, differentiate the present model from Milde and Riley's with respect to two main elements: (i) while Milde and Riley's model is based on three different and unrelated cases, our model is able to subsume the various cases within a unified framework; (ii) while Milde and Riley's assumptions rule out the possibility of pooling equilibria, the Wilson construction in our model entails the possibility of both separating and pooling equilibria. In particular, pooling equilibria obtain whenever applicants of different quality have indifference curves so similar that banks are unable to screen out their projects. This result strengthens the possibility of credit rationing even in the presence of signaling through the size of the loan. Both separating and pooling equilibria entail the possibility of "type I" credit rationing of given quality types.

First Draft: January 26, 1994
Revised: April 8, 1994

1. Introduction.

In this paper we analyze the optimizing behavior of risk neutral firms and banks in a competitive credit market under conditions of imperfect (asymmetric) information and limited liability on the side of the firms. We assume that each firm aims at realizing a project of variable size and that, in order to finance the project chosen, it is constrained to borrow from a bank. Output uncertainty of the projects generates an incomplete information problem from both the firms and the banks viewpoint. Conversely, each firm knows the quality of its project which is unobservable to the banks. Hence, project quality generate an asymmetric information problem on the side of the lenders. We only consider the *ex ante* asymmetry of information, since we assume that banks observe firms' end-of-period returns. This asymmetry is sufficient to prevent banks from designing debt contracts contingent on the quality of projects. In addition, the assumption that firms have limited liability implies that debt contracts cannot compel firm owners to pay the bank more than the *ex post* end-of-period return of their investment. Asymmetric information and limited liability imply that lenders face an adverse selection and/or a moral hazard problem. This means that firms are not equal to each other and must be distinguished on the basis of their behavioral function.

Models of investment financing with asymmetric information in the credit markets have been proposed in the past, starting with Jaffee and Russell (1976) and followed by Stiglitz and Weiss (1981) (henceforth S-W) and Stiglitz and Weiss (1986b), De Meza and Webb (1987) (D-W), Milde and Riley (1988) (M-R). S-W considered the case of investment projects of fixed size with equal expected return but different riskiness in accordance with the *mean preserving spread* principle (Rothschild and Stiglitz (1970)). D-W considered both the case of fixed-size investment projects of different quality yielding the same actual return but with different probability of success and the case of fixed-size investment projects of different quality and different actual returns but equal expected return. Both S-W and D-W assumed that each bank offers a pooling standard debt contract to borrowers of unobservable different quality. Conversely, M-R assumed a variable project size, analyzing standard debt contracts contingent on the size of the loan and the interest rate. The variable size of the loan allows for signaling on the side of the borrowers, thus generating self-selection and the occurrence of separating equilibria for firms with projects of unobservable different quality.

As the size of the loan is allowed to vary, differences in returns between two investment projects can thus be due both to differences in quality and to differences in level. M-R characterize firms with a production function which depends on the level of input (loan) and on a firm-specific quality parameter unobservable to lenders, where projects of higher quality yield higher output levels. Depending on the relation between quality, investment level and output, M-R depict three different cases. In the first case, the levels of investment are increasing in quality. In the second, investment levels are invariant to quality, so that higher quality firms signal their better quality by accepting smaller rather than larger loans. In the third case, investment projects have different riskiness but yield equal expected output: thus, riskier projects which yield higher expected returns require larger rather than smaller loans.

This paper aims to take a further step in the direction followed by M-R. Both S-W and D-W pointed out the possibility of credit rationing in the face of asymmetric information, due to the differences in quality and/or riskiness of projects with equal investment level but different actual return. Yet, if projects differ in the size of the loan also, differences in actual returns cannot be unambiguously ascribed to differences in quality anymore, and a relation between quality and size of the loan must be defined. M-R have shown that, once we specify the relation between investment level, quality and output (a production function) and we allow for firm signaling, we will have either over-financing or under-financing of higher quality projects, under separating (second-best) equilibria. However, two conditions weaken M-R's results. In the first place, the technology adopted is known to both the firms and the banks, i.e. whether it is multiplicative or additive in the quality parameter. In the second place, such quality parameter is left unspecified, as it was also in S-W and in D-W.

In this paper we take a slightly different stance. As quality in not observable, we assume that realized firm profits have the *monotone likelihood ratio property* (MLRP) with respect to *quality*, in the sense of Milgrom (1981). The rationale for assuming a MLRP with respect to "quality" for the firm profit densities is that of exploiting the monotonicity property implied by the definition of quality. In the literature, two different *quality* definitions are used. With the first, higher quality implies a better profit distribution in the sense of first order stochastic dominance (e.g. Chan and Kanatas (1985), Besanko and Thakor (1987)). With the second, lower quality implies a mean preserving spread or higher risk in the sense of second order stochastic dominance (e.g. Stiglitz and Weiss (1981, 1986), Bester (1985)). In both cases *higher* quality means *better* outcomes, which is what matters for the monotonicity property of the definition of quality we need. But, as M-R have shown, these two definitions of quality imply different relations between quality and the level of inputs (loans), and so we have to distinguish them on an a priori observable basis. In other words, if we want to nest both definitions within a more general framework, we need to relate such unobservable quality features with some observable (technological or organizational) characteristic of the firm. As opposed to the approach adopted by M-R, in this paper we adopt a slightly different and more general approach.

From this point of view, the approach adopted by M-R, and the three models thereby elaborated, are not helpful, as they actually appear to be completely unrelated to each other and basically assume what actually ought to be explained. Here, we refer to a more general setting in which output depends not only on an unobservable *quality* parameter but also on an observable *technological* parameter. More specifically, we consider how different "technologies", meaning different ways of organizing production activities, can have different implications in terms of quality. An example may be clarifying. Let us refer to two ideal and opposite firms: firm A is vertically integrated and organized in various departments so that its R&D activity flows into one or more final goods by passing through several internal production phases; firm B is specialized in a particular activity which is not suitable to be organized in mass production. According to the neo-institutionalist parlance (e.g. Coase 1937, Grossman and Hart 1986, Milgrom and Roberts 1990), the organization of firm A is substituting entrepreneurial co-ordination and hierarchical relations for market exchanges, whereas the organization of firm B gives large room to market exchanges. The rationale for this different institutional setting is that firm A has *large internal economies of scale, high transaction* (or barganing) *costs*, and *small organizational* (or influence) *costs*, whereas firm B has *large external economies of scale, small transaction* (or bargaining) *costs* and *high organizational* (or influence) *costs*. This points out that the "quality" of an investment project of firm A will be increasing in its size (maybe only up to a certain point, and only in a decreasing fashion), and the "quality" of an investment project of firm B will not be increasing in its size.

Our previous example allows to state that if *internal* economies of scale and/or *transaction* or *bargaining* costs prevail a larger size of the project will signal higher firm quality, while if *external* economies of scale and/or *organizational* or *influence* costs prevail a larger size of the project will signal lower firm quality. It is important to stress that it should not be terribly controversial to assume that these characteristics of each firm that we define as technological or organizational be *observable* to outside lenders, in the sense that they may be gathered from the known "identity card" of the borrowers (industrial sector, market structure, internal organization, and so forth). In this paper we define an *organizational complexity* parameter H, which we first assume as taking on two values: either I, meaning that internal economies of scale and/or transaction costs are prevailing, or E, meaning that external economies of scale and/or organizational costs are prevailing.

The above institutional setting allows the strengthening of some credit rationing results. As is well known, S-W and D-W emphasized the possibility of credit rationing under the assumption of a common and fized loan size; on the other hand, M-R maintained that this possibility is not robust to the signaling of high quality borrowers through the size of their loan demand. This paper shows that credit rationing is still a possible outcome once the loan size is allowed to vary and hence signaling is permitted, but pooling equilibria are not ruled out by the equilibrium construction. Two basic ingredients underline this conclusion: the monotonicity of the lender payoff together with the MLRP, and the Wilson equilibrium resulting from a three-stage contracting game. The first ingredient implies that the optimal form of the financial contract between a bank and an entrepreneur is a standard debt contract which specifies both the size of the loan and the repayment schedule. The second ingredient implies that, as firm signaling is allowed, entrepreneurs of different but unobservable quality can sign either separating or pooling contracts[1]. More precisely, by allowing banks to use both the repayment schedule and the loan size, we allow for the possibility of banks to use the loan contract as a screening device. If this screening is effective, separating equilibria will obtain. However, as borrowers of different quality might have similar indifference curves, banks could still be unable to effectively screen loan applicants so that pooling equilibria obtain. Separating contracts are always Pareto efficient: hence, they flow into "first best" or "second best" equilibria. On the other hand, pooling equilibria can only be contrained efficient.

The results we obtain tend to confirm and extend the results in S-W, D-W and also M-R, at least at a quantitative level. In particular, we prove that, while some quality types

[1] A Wilson (1977) construction gives rise, in this case, to a wider spectrum of possible results than M-R. In particular, while M-R's adoption of a Riley (1979) construction rules out the possibility of obtaining pooling equilibria, our adoption of a Wilson game structure entails the possibility of both separating *and* pooling equilibria. On this point, see also De Meza and Webb (1989) and Innes (1991).

may be under-financed, some others will be over-financed so that some applicants will not get the amount of credit they would be willing to borrow. Furthermore, we prove that credit rationing is compatible with pooling as well as with separating equilibria. This means that our approach entails a robust possibility of "type I" credit rationing: all quality types get financed, but some quality types will obtain a smaller amount of credit than that desired at the contract equilibrium interest rate. Conversely, and differently from S-W, our reference to projects of varying loan size rules out the possibility of "type II" credit rationing: no quality type will ever be denied credit, even on a random basis.

2. The setting of the model.

A. The Firms.

We consider an arbitrarily large population of risk neutral entrepreneurs. Each firm i has the production function $y(I,q,H,\theta)$, where I is the firm production input in nominal terms, q is a *quality* parameter which may depend on the entrepreneur's managerial quality, H is a parameter representing the "organizational complexity" of the firm, and θ is a random shock which is i.i.d. across firms and has a positive support on $[0,\bar{\theta}]$. Without loss in generality, we assume that each entrepreneur has a null monetary wealth endowment[2]. Also, we assume that the only financing source comes from bank loans: no debt contracting with outside private investors is allowed in this model. Each entrepreneur knows the "quality" of his firm and/or his given project[3], which is unobservable by lenders, thus giving rise to adverse selection problems on the side of the lenders. We assume that H is a binary variable, i.e. $H \in [\mathrm{I},\mathrm{E}]$. That is, when the "organizational complexity" of the firm is such that *internal economies of scale*

[2] We could also refer to a population of entrepreneurs with positive and homogeneous wealth endowments, so that I would become the amount of firm's input financed through external funds. Following Leland and Pyle (1977), it can be shown that in such a case all entrepreneurs will invest all of their wealth in the firm before raising any funds from outside sources (see also De Meza and Webb (1987, p.289), Innes (1992, p. 1430)). Hence, signalling through wealth would still be impossible, and the assumption of positive and homogeneous wealth would not affect our conclusions. This implies that the assumption of a null monetary wealth endowment is not restrictive.

[3] In the following we use "firm quality", "project quality", and "entrepreneur quality" interchangeably.

tend to prevail, then H = I, whereas when *external economies of scale* tend to prevail, then H = E. Finally, we assume that entrepreneurs have a heterogeneous endowment of quality, either *low* or *high*, i.e. $q \in [l,h]$[4]. The relative frequency of different quality types is such that PROB(q = h) = p, and PROB(q = l) = 1-p, and it is common knowledge.

Each firm produces a stochastic end-of-period return $\pi = \pi(I,q,\theta)$, which is given by the product of the output price P and the firm's output $y(I,q,H,\theta)$. Given I, q and H, θ gives rise to probability density and distribution functions for π, $f(\pi | I,q,H)$ and $F(\pi | I,q,H)$, respectively. The function f(·) is assumed to be continuously differentiable on $[0,K(I,q,H)]$, where $K(\cdot)>0$ whenever $I>0$. In addition, we assume higher investment levels produce "better" net worth distributions in the sense of first-order stochastic dominance[5].

<u>Assumption A1</u>: $F_I(\pi) \leq 0$, for all π within $[0,K(I,q,H)]$

Defining the firm expected return, which is indexed by quality, as:

$$\bar{\pi}(I|q,H) \equiv \int_0^{K(I,q,H)} \pi f(\pi(I|q,H))\, d\pi, \tag{1}$$

then condition (A1) implies that $\bar{\pi}_I\ (I|q,H) > 0$. We will also assume that $\bar{\pi}_{II} < 0$, and $\bar{\pi}_I(I|q,H) = \infty$ as $I \to 0$ and $\bar{\pi}_I(I|q,H) = 0$ as $I \to \infty$. The latter constraints ensure positive and finite equilibrium investment levels for any quality type.

B. The Banks.

Banks are assumed to be competitive and risk neutral. On each loan, banks require an expected return of at least ρ, the return on a risk-free bond[6], that is, they only offer those contracts which are expected to earn them a non-negative mean profit. This return is obtained by ex-post firm payments that satisfy limited liability and which depend on the observables, P and y, and hence π, as specified in the financial contract defined below. Perfect competition

---

[4] The extension to the case $q \in [1,...,Q]$ is straightforward. In that case the relative frequency of quality types is g(q), and $\Sigma_q\, g(q) = 1$.

[5] Subscripts denote partial derivatives.

[6] We simply assume that banks obtain their funds from depositors at the safe rate of interest ρ.

in the loan market ensures that each bank satifies a zero-profit condition.

C. The Financial Contract and the Definition of Equilibrium.

The financial contract signed by the entrepreneur with the bank prescribes the size of the loan, I, and the amount of money which the bank is to be paid at the end of the period. Despite each bank would be willing to sign different contracts with different applicants, i.e. different quality types, it does not know on a *ex-ante* basis "who is who", as information about firm quality is asymmetric. Moreover, banks cannot observe *ex-post* neither the state of nature, θ, nor the quality of the project, q; they can only observe the firm's return, π, as well as H, at no cost of monitoring. Thus, independently of the optimality of its form, in our model any financial contract will set, beside I, the bank's payoff $B(\pi)$, i.e. the amount to be paid back by the firm[7]. This amount can be constrained by π since limited liability implies that $B(\pi)\leq\pi$, that is, firms cannot be compelled to pay more than the available return.

It has recently been shown (Innes (1993)) that, with a fixed and homogeneous loan size, a Nash equilibrium exists. A set of financial contracts is a Nash equilibrium if all contracts in the set earn banks an expected return of at least ρ, and there is no other set of contracts that, when offered in addition to the equilibrium set, earn banks an expected return greater than ρ. In that case, each quality-type class of borrowers subscribes a *pooling* contract which corresponds to a *pooling* single-contract equilibrium. Normally (e.g. Stiglitz and Weiss (1981)) it is exogenously assumed that the payoff function of this contract takes a standard debt form, namely, $B(\pi) = \min(\pi,z)$, with z representing the *fixed and predetermined loan payment*. The problem is that, when th size of the loan is allowed to vary and thereby to serve as a screening device, a Nash equilibrium will often fail to exist (as in Rothschild and Stiglitz (1976)). Thus, a number of studies have suggested alternative concepts of equilibrium, including Miyazaki (1977), Wilson (1977), and Riley (1979). At first, these concepts have been justified by assumptions on an uninformed agent's (i.e. a bank's) expectations about the plausible responses of their competitors (i.e. competing banks) to his actions. More recently, they have obtained game-theoretic foundations by Hellwig (1987) and Cho and Kreps (1987).

[7] For analytical purposes, some weak restrictions are placed on the forms which this function can take: $B(\pi)$ must be differentiable from the right, with positive first derivative for all $\pi>0$. Also, we assume $B(\pi)$ is monotonically nondecreasing (Mankiw (1986), DeMeza and Webb (1987)).

Following the latter approach (see also De Meza and Webb (1989) and Innes (1991)) we define a three-stage pure strategy game between banks and entrepreneurs, as this game structure has at least one Nash equilibrium overall (Kreps and Wilson (1982)). The extensive form of the game could be as follows:

Stage 1. The uninformed lenders (banks) move first and propose a menu of contracts. A contract specifies the repayment schedule and the size of the loan.

Stage 2. Each borrower chooses a contract within the offered menu.

Stage 3. After observing the application choices, lenders decide whether to accept or reject each application, and either sign or withdraw contracts.

Such a game structure can have many "sequential" Nash equilibria. A *variable loan size equilibrium* is defined by the set of contracts $\{B(\pi | q,H), I(q,H)\}$ that maximizes the utility of the higher quality entrepreneur subject to three constraints:

(E1) *Incentive compatibility*: lower-quality entrepreneurs weakly prefer their own contract to any higher-quality one, and viceversa.

(E2) *Low-quality rationality*: lower-quality entrepreneurs expected profit is at least as high as on a perfect-information contract, that is, for this quality-type of borrowers the perfect-information contract is the "benchmark" which earns lenders zero expected returns.

(E3) *Non-negativity of lenders profits*: lenders earn an expected return of at least ρ on each distinct contract.

The latter constraint characterizes a Wilson allocation[8]. Given the game structure posited above, the resulting equilibrium is thus a *Wilson equilibrium*. Obviously, this is not the only possible outcome. M-R refer, for instance, to a *reactive Riley equilibrium*. The main difference between these two alternative concepts of equilibrium is that while Riley's equilibrium elicits complete separation, whereby all agents reveal their type by their contract choice, Wilson's equilibrium admits pooling contracts that are shared by different quality-type borrowers, even when separating allocations are possible. Thus, a Wilson allocation can flow into either separating or pooling equilibria, and Wilson's and Riley's equilibria diverge only

[8] If (E3) was replaced with the assumption that banks earn an expected return of at least ρ on *all contracts taken together*, that would be a Miyazaki allocation, while if (E3) was replaced with the assumption that banks earn an expected return of at least ρ on *each entrepreneur*, that would be a Riley allocation. This means that a Riley allocation requires that at Stage 1 of the game the informed borrowers move first.

when the Wilson allocation is pooling. It is important to stress that, in the latter case, the Wilson allocation Pareto dominates the Riley equilibrium. This means that the reference to Riley's concept of equilibrium eliminates, by construction, the most favourable cases of credit rationing[9]. Hence, as we are mainly interested in analyzing the possibilities of credit rationing, the Wilson construction seems preferable[10].

As shown by Innes (1993), under asymmetric information any variable loan size equilibrium takes the standard debt form, provided the lender's payoff is monotonic. Thus, even for the Wilson equilibrium, the bank's payoff function must here be compatible with the standard debt form, namely $B(\pi) = \min(\pi, z(q,H))$, with $z(q,H)$ representing the *promised loan payment*, which depends on the unobservable quality parameter and on the observable firm organizational complexity H. In essence, given H, debt contracts maximize banks' payoffs with respect to low-quality profit distributions: since high-quality entrepreneurs, by signing a debt contract, can minimize the low-quality type incentive to "masquerade", a contract can be completely described by the pair (I,z).

D. Choices of the loan size.

Given H, the quality q entrepreneur's choice problem can be formally written as:

(2) $$\max_{(I)} V(I\,|q,H) = E\{\pi - B(\pi)\,|q,H\}$$

(2a) $$\text{s. t.} \quad E\{B(\pi)\,|\,q,H\} \geq (1+\rho)\,I$$

(2b) $$B(\pi) \leq \pi, \quad B_\pi(\pi) \geq 0 \quad \forall \ \pi \in [0, K(I,q,H)]\,.$$

Condition (a) gives the bank minimumm return requirement, while (b) gives the limited liability and monotonicity constraints on the contract form. In the absence of the limited

[9] Several authors (e.g. Bester (1987)) maintain that separating equilibria eliminate the possibility of credit rationing. However, Besanko and Thakor (1987) have built a model with separating equilibria and credit rationing.

[10] We maintain that the no-rationing results obtained by M-R largely depend on their arbitrary choice of a Riley allocation. Wilson's allocation has been criticized because it implies passive reactions by competing agents. The game-theoretic foundation of this allocation is, however, able to weaken this problem.

liability constraint, the solution to (2) would be the fixed-payment debt contract, $B(\cdot) = (1+\rho)I$, that yields "perfect information" choices of I, for given q (Shavell (1979)) and H. However, the entrepreneur's limited liability, together with the possibility of zero output levels, render any contract with payments that are invariant to output, and thus to profits (including the fixed-payment contract) unfeasible. Thus, in (2), $B(\pi) = \min(\pi, z(q,H))$.

The expected return, V, to the quality q entrepreneur of a firm H of a project of size I is thus given by:

$$V(I \mid q,H) = E\big(\max\left[\pi - z(q,H), 0\right]\big). \tag{3}$$

Equation (3) implies that, if the entrepreneur's project I is successful, then $V(\cdot)$ is equal to the positive difference between the project return and the promised loan payment, whereas, if the entrepreneur's project is unsuccessful, then $\pi < z(q,H)$, and $V(\cdot)$ is equal to 0. Given that the entrepreneur is risk-neutral, he maximizes $V(I \mid q,H)$ in (2). Given (1), this can be rewritten as:

$$\max_{(I)} V(I \mid q,H) = \int_{\pi^*}^{K(I,q,H)} [\pi - z(q,H)] f(\pi \mid I,q,H)\, d\pi, \tag{4}$$

where π^* satisfies:

$$\pi^* - z(q,H) = 0. \tag{4a}$$

That is, $V(\cdot)$ is the return of a successful investment times the probability of a "success"[11]. The probability of a "success" is given by the integral of the density function of π for $\pi \in [\pi^*, K(I,q,H)]$. But, as we have seen above, the density function of π, given I,q and H, depends on θ, which is i.i.d. across firms. Hence (4) is equivalent to:

$$\max_{(I)} V(I \mid q,H) = \int_{\theta^*}^{K(I,q,H)} [\pi(I,q,H,\theta) - z(q,H)] f(\pi(\theta) \mid I,q,H)\, d\theta, \tag{5}$$

where θ^* is such that:

[11] Thus $z(q)_{\pi \geq \pi^*} > (1+\rho)I > \pi(q)_{\pi \leq \pi^*}$. De Meza and Webb (1987) define R^s as the return of a successful investment project and R^f as the return of an unsuccessful project. They thus assume that $R^s > (1+\rho)I > R^f$.

(5a)
$$\pi(I,q,H,\theta^{*}) - z(q,H) = 0 .$$

In order to solve equation (5), it is necessary to specify condition (2a) above, i.e. banks expected profits. Banks expected profits are defined as:

(6)
$$R(I,z \mid q,H) = E\left[\min(\pi,z) \mid I,q,H\right] - (1+\rho)I .$$

Making use of (5) we can rewrite this expression as:

(7)
$$R(I,z \mid q,H) = z(q,H) - (1+\rho)I + \int_{0}^{z}(\pi - z)f(\pi \mid I,q,H)\,d\pi ,$$

where we use the fact that $z = \pi^{*}$. Substituting (7) for (2a) and (5) for (2), we can formally solve the quality-q entrepreneur's choice problem. The easiest way to handle this problem is to refer to entrepreneurs' and banks' *indifference* curves.

An entrepreneur's *indifference curve* is a set of (z,I) contracts yielding a common expected profit level[12]. Since entrepreneurs are always better off with a lower promised payment, z, given the project size, I, lower indifference curves correspond to higher expected profit levels for the firms. Formally, entrepreneurs' indifference curves are defined as:

(8)
$$V(I,z \mid q,H) = \int_{z}^{K(I,q,H)} [\pi(I,q,H) - z(q,H)]f(\pi \mid I,q,H)\,d\pi = \bar{V}(q,H) ,$$

a constant. Regardless of the entrepreneur's quality, indifference curves are *upward-sloping*, as:

(9a)
$$V_{I}(I,z \mid q,H) \equiv - \int_{z}^{K(I,q,H)} F_{I}(\pi \mid I,q,H)\,d\pi > 0 ,$$

(9b)
$$V_{z}(I,z \mid q,H) \equiv -(1 - F(z \mid I,q,H)) < 0 .$$

Intuitively, higher investment levels, given z, increase an entrepreneur's expected profit. Hence, to preserve the latter profit, an increase in I must be accompanied by an increase in the debt obligation, z. Because entrepreneurs "like" any increase in I, and "dislike" any

---

[12] Since we assumed that both banks and entrepreneurs are risk neutral, we can obviously refer to "indifference" of "iso-profit" curves. In the former and more general case, the set of contracts must yield a common expected utility level.

increase in payments to banks, z, indifference curves are upward-sloping (and lower curves correspond to higher firm profits).

Banks isoprofit curves are defined as the locus where banks expected profits, $R(I,z \mid q,H)$, are constant, that is:

$$(10) \qquad R(I,z \mid q,H) = E\left[\min(\pi,z) \mid I,q,H\right] - (1+\rho)I = \bar{R}(I,z \mid q,H).$$

Now, banks isoprofit curves are *upward sloping* and *convex*, with slope greater than $(1+\rho)$. The reason is that, given the increasing default risk of bank loans, when a bank commits an additional dollar of funds to loans, the increase in the promised loan payment z must be increasingly greater than $(1+\rho)$ dollars in order to satisfy the non negative condition of the bank expected profit. Due to the perfect competition in the loan market, we can limit our attention to banks iso-profit curves from loans to quality-q entrepreneurs, which imply:

$$(10a) \qquad \bar{R}(I,z \mid q,H) = 0.$$

These iso-profit curves will coincide with banks supply curves. Thus, if a bank is able to offer different contracts to borrowers of different quality, *separating* supply curve will prevail, while if a bank cannot discriminate between borrowers of different quality, a *pooling* supply curve will obtain.

Banks *separating* supply curves are defined as the locus where:

$$(11) \qquad E\left[\min(\pi,z) \mid I,q,H\right] = (1+\rho)I,$$

that is, making use of (7):

$$(12) \qquad z(q,H) - (1+\rho)I + \int_0^z (\pi - z) f(\pi \mid I,q,H)\, d\pi = 0.$$

Using equations (1) and (8) we can rewrite this expression as:

$$(13) \qquad \bar{\pi}(I \mid q,H) - V(I \mid q,H) - (1+\rho)I = 0.$$

Equation (13) implies that firm signaling (self-selection) will occur in this case.

Similarly, banks *pooling* supply curves are such that:

$$(14)\qquad \left[E\left[\min(\pi,z)\mid I,H,q=\mathrm{l}\right]-(1+\rho)I\right](1-p) + \left[E\left[\min(\pi,z)\mid I,H,q=\mathrm{h}\right]-(1+\rho)I\right]p = 0\,,$$

that is

$$(15)\qquad \left[z(q=\mathrm{l},H)-(1+\rho)I+\int_0^z(\pi-z)f(\pi\mid I,q,H)\,d\pi\right](1-p) + \left[z(q=\mathrm{h},H)-(1+\rho)I+\int_0^z(\pi-z)f(\pi\mid I,q,H)\,d\pi\right]p = 0\,.$$

This implies that the bank set of loan supplies will lie somewhere between the separating supply curves to the low-quality types and those to the high-quality ones. In other words, a bank pooling supply curve is the supply curve to the "average" quality-q entrepreneur[13], given H. Independently of the characteristics of the bank supply curve, the analysis of firms' and banks' iso-profit curves emphasizes that each entrepreneur can solve his choice problem by signing the debt contract which coincides with the tangency point between the bank supply curve and his lowest iso-profit curve.

3. Equilibrium Outcomes.

When the size of the loan is allowed to vary, the entrepreneur's choice of the contract becomes crucial (but recall that the setting of our model abstracts from moral hazard problems and refers to adverse selection problems only). When a bank offers an entrepreneur a menu of contracts, specifying both the loan size and the repayment schedule, it is for the purpose

---

[13] The extension to the case of more than two quality types is, again, straightforward. In that case, (17) and (18) can in fact be written, respectively, as:

$$(14')\qquad \sum_q \left[E\left[\min(\pi,z)\mid I,q,H\right]-(1+\rho)I\right]g(q) = 0\,,$$

and

$$(15')\qquad \sum_q \left[z(q,H)-(1+\rho)I+\int_0^z(\pi-z)f(\pi\mid I,q,H)\,d\pi\right]g(q) = 0\,,$$

where $g(q)$ is the proportion of entrepreneurs of quality q in the population.

of financing a *specific project* whose *specific return* will be due to a loan of specified size, to the random outcome of the production process, and to the quality q of the entrepreneur putting the project at work. It follows that, given the link between loan size and project expected return, the entrepreneur's contract choice can, in principle, signal his quality-type which is not directly observable by the lender because of the ex-ante asymmetry of information. The limits to this signaling derive from the fact that, as M-R have shown, the link between expected returns and loan size can basically flow into two different relations between quality and size of the loan: a positive one, and a negative one. In the first case, of two projects of (unobservable) different quality the "better" one yields a higher expected return for a given loan size: that is, the larger the size of the loan, the higher marginal returns; it follows that entrepreneurs with "better" projects can signal their higher quality by agreeing on larger increases in the loan interest rate, thus signing contracts specifying loans of larger size. On the other hand, in the second case, marginal returns are invariant to both the quality of the project and the size of the loan: however, since low-quality borrowers have higher default risk anyway, it follows that entrepreneurs with the "worse" project tend to agree on larger increases in the loan interest rate, thus signing contracts specifying loans of larger size (see also Innes (1991, p. 359)).

This means that the relation between quality and project marginal returns is specific to the entrepreneur. Hence, from the lender's perspective, ordering projects on the basis of the different expected returns from the loan is not sufficient to discriminate between borrowers of different quality and projects of equal expected return[14]. At this aim, we need to compare projects by means of some other *observable* characteristic: this is a key point, as the relation between quality and expected return from a project *has to be* independent of quality itself and observable. Similar relevant observable characteristics could be given by technological and organizational parameters. As discussed in Section 1 above, we could in principle distinguish between firms where internal economies of scale and/or transaction costs prevail and firms where external economies of scale and/or organizational costs prevail: it should not be terribly controversial to assume at this point that the prevailing parameters of each firm are observable

[14] This would also apply to models with projects with loans of fixed size. In the S-W model, the various projects are classified within groups of observationally equal riskiness, that is, by the equality of their expected returns. However, the subset of projects with greater men return is not necessarily the less (or the more) risky one. The same holds true for the D-W model: a higher R^s is not univocally related to a higher (or a lower) p.

to ouside lenders, which means that they are common knowledge. We can sum up these observable characteristics of a firm by defining an *organizational complexity* parameter, H: if H = I, then internal economies of scale and/or transaction costs are prevailing, whereas if H = E, then external economies of scale and/or organizational costs are prevailing[15]. When H = I, project returns and quality will be positively related, so that high-quality entrepreneurs will be willing to choose larger loans than low-quality entrpereneurs. When H = E, project returns and quality will be inversely related, so that high-quality entrepreneurs will be willing to choose smaller loans relatively to low-quality entrepreneurs[16]. The conclusion is that, *in the case of H =* I, *larger loans will signal higher firm quality*, while *in the case of H =* E, *larger loans will signal lower firm quality.*

Suppose two entrepreneurs apply for a loan offered by a given bank. To allow this bank to exploit borrowers' signaling through their choice of the loan size, we need to assume that for both applicants the relation between loan size and quality be the same, either positive or negative. Otherwise, the signals would be strongly disturbed and the comparative appreciation of the two entrepreneurs quality would be impossible. This basically amounts to say that it is convenient for the bank, when receiving an application for a loan offer, to screen first the borrowers according to their observable technological and organizational characteristics. Hence, we may assume that projects will be initially pooled by each bank according to the observable value of H, by which the relation between the size of the loan and the quality of applicants will be the same.

As we will see, we can thus exploit a suitable assumption of monotonicity of the loan size with respect to quality in the two cases of H = I and H = E by comparing signals in terms of relative favourableness, using the so-called *monotone likelihood ratio property*. Now, since the bank does know in advance which is actually the case, i.e. whether H = I or H = E, it will offer a menu of contracts which displays, say, four couplets of loan-repayment pairs (I, z) (among all the possible ones), lying on the three offer curves corresponding to the *low-quality*, *pooling*, and *high-quality* types (see Figure 1). For firms with H = I, a *larger* loan signals *higher* quality, and the choice will be among the couplets indicated by the points A,

[15] We initially assume that H can take on only two values, I or E. The extension to a continuous set of values is sketched below (see Section 3.C).

[16] It is worth noting that our setting leads to a parameter configuration which is slightly different from M-R's. As mentioned above, in the latter model marginal project returns are invariant to quality.

F or C, and E, while for firms with H = E, a *smaller* loan signals *higher* quality, and the choice will be among B, G or C, and D.

A. Internal economies of scale prevailing: H = I.

Consider first the case of prevailing internal economies of scale, i.e. H = I. In this case of positively related loan size and quality parameter, we have that marginal project returns are higher the higher the quality of firm entrepreneurs. In addition, we assume higher quality produces "better" net worth distributions in the sense of statistical "good news" (Milgrom (1981)).

Assumption A2: $\dfrac{\partial}{\partial \pi} \dfrac{f_q(\pi \mid I,q,H)}{f(\pi \mid I,q,H)} > 0$, for all π within $[0,K(I,q,H)]$.

The inequality in (A2) has been called the *monotone likelihood ratio property* (MLRP) and is a condition slightly stronger than first order stochastic dominance (FOSD). Except in the case of two point distributions, the MLRP is in fact not implied by FOSD; however, the MLRP characterizes a wide class of distributional specifications (see Milgrom (1981, p. 383)). The MLRP states that, if a signal x is *more favourable than* another signal y about a random parameter $\tilde{\phi}$, and if $\tilde{\phi}$ takes on two particular values ϕ_1 and ϕ_2, with $\phi_1 > \phi_2$, then

$$f(x \mid \phi_1) f(y \mid \phi_2) > f(x \mid \phi_2) f(y \mid \phi_1) .$$

As shown by Milgrom, this is a necessary and sufficient condition for x to be more favourable than y both in the sense of first-order and second-order stochastic dominance, for every increasing concave function $U(\theta)$ and nondegenerate prior distributions $G_1(\phi)$ and $G_2(\phi)$[17]. Hence, assumption (A2) above implies that the posterior distribution $G(\phi \mid \pi_1)$ dominates the posterior distribution $G(\phi \mid \pi_2)$ both in the sense of FOSD and in the sense of SOSD, and that any signal U_1 is *more favourable than* U_2 in either sense. In our case, this means that, for all

[17] In Milgrom's words, "a signal x is *more favourable than* another signal y if for every nondegenerate prior distribution G for ϕ, the posterior distribution $G(\cdot \mid x)$ dominates the posterior distribution $G(\cdot \mid y)$ in the sense of strict FOSD" (1981, p.382).

possible price realizations, higher "input" levels are assumed to yield superior output (return) distributions.

When internal economies of scale prevail (H = I), higher loan sizes signal higher quality (higher quality implies higher loan sizes), and the rationale for assuming a MLRP with respect to quality for the firm profit densities is that of exploiting the monotonicity property implied by the definition of investment. Given a stochastic environment which yields an uncertain level of output, higher loans will "produce" higher levels of output. Conversely, when external economies of scale prevail (H = E), as we will see below, larger loan sizes signal lower quality (better quality needs smaller loans), the opposite would hold true, and smaller loans would "produce" higher levels of output. But in that case, too, the rationale for a MLRP for the profit densities would exploit the monotonicity of investment with respect to quality. The conditions in (A1) and (A2) are thus amended by the further assumption that:

(A3) $$F_{Iq}(\pi) < 0 ,$$

where the negative cross-partial derivative ensures that higher quality is associated with "better" *marginal* project returns, that is, higher quality yields first order stochastically dominant returns from marginal loans. Then:

(16) $$\frac{\partial \pi_I(I \mid q, H = \mathrm{I})}{\partial q} > 0 .$$

This implies that the marginal rate of substitution between I and z is greater, the higher the firm entrepreneurs' quality, the steeper the indifference curve[18]. Formally, (A1), (A2), (A3), (9a), and (9b) imply:

(17) $$\frac{\partial}{\partial q}\frac{\mathrm{d}z}{\mathrm{d}I}\bigg|_{\bar{V}} = \frac{\partial}{\partial q}\left(-\frac{V_I(I,z \mid q,H=\mathrm{I})}{V_z(I,z \mid q,H=\mathrm{I})}\right) > 0 .$$

Also, the increase in the promised loan payment, z, needed to satisfy the bank's expected zero-profit condition will be higher the lower the quality of the applicant entrepreneur, because of the latter higher default risk. Hence, banks isoprofit curves are upward sloping, with slope decreasing with quality. Similarly to M-R, we can restate this by saying that the marginal

[18] That is, in each and every point in the space (I,z) where high-quality and low-quality indifference curves cross, the former are steeper than the latter curves.

increase in promised loan payment that a borrower is willing to accept in order to receive a larger loan is greater the higher investment quality. Since we are considering a competitive banking sector, the relevant banks' isoprofit curves are zero-isoprofit curves: the latter represent banks' supply curves.

In equilibrium, the quality-q entrepreneur's indifference curve and the quality-q bank's supply curve are tangent at:

$$V_I(I,z \mid q,H=\mathrm{I}) = \bar{\pi}'(I \mid q,H=\mathrm{I}) - (1+\rho). \tag{18}$$

For given levels of (z,I), this is the point where net expected profits of a quality-q entrepreneur are maximized, subject to the non-negativity of bank's expected profits from loans to that entrepreneur[19]. We denote this point with E^*_q. This equilibrium point will be a *first best* (perfect information) contract if it satisfies the feasibility constraints of a Wilson equilibrium. We have thus the following Proposition:

Proposition 1. In a world of asymmetric information, only a *separating* equilibrium can be a "first best" equilibrium (see Figure 2).

Suppose that, as each firm is either high (q = h) or low (q = l) quality, banks know that entrepreneurs can only be either high or low-quality. Consider two entrepreneurs, a high-quality and a low-quality one, who choose two optimal contracts (I^*_h,z^*_h) and (I^*_l,z^*_l), respectively. We will have a separating set of equilibria E^*_h and E^*_l (i.e. points E and A, Fig. 2) if (i) each lies on a separating supply curve, and (ii) the low-quality contract is weakly preferred by the low-quality entrepreneur. Condition (i) is necessary but not sufficient to have a separating equilibrium, since the high-quality contract (I^*_h,z^*_h) can lie below the low-quality indifference curve which yields the low-quality preferred contract (I^*_l,z^*_l). In this case, condition (ii) would not be satisfied. Thus, the high-quality optimal contract must lie above the low-quality "optimal" indifference curve. In this case, the entrepreneurs' optimal choice elicits a "first best" equilibrium (a separating equilibrium with firm signaling).

The converse of Proposition 1, however, is not true, in the sense that we can have separating equilibrium which are not "first best" (see Figure 3). The optimal equilibrium is

---

[19] At this point, high-quality entrepreneurs will never choose a low-quality contract.

not feasible whenever (i) the high-quality preferred contract lies below the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. The consequent "suboptimal" equilibrium will be separating if (ii) the high-quality indifference curve which is tangent with the bank pooling supply curve[20] crosses the bank supply curve to the high-quality types at a point higher than the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. The latter point of intersection will be the high-quality contract pair. In more formal terms, suppose that, given the two first best contracts (I_l^*,z_l^*) and (I_h^*,z_h^*), the latter contract lies below the low-quality preferred indifference curve, thus attracting low-quality entrepreneurs. To avoid the possibility of negative expected profits on loans to low-quality entrepreneurs, within the menu offered by the bank there is another contract (I_h,z_h) which lies at the intersection of the low-quality indifference curve and the bank supply curve to the high-quality types (point H). Thus, low-quality entrepreneurs will choose their "first best" contract (I_l^*,z_l^*), point A, whereas high-quality entrepreneurs will choose their "second best" contract (I_h,z_h), point H. The resulting equilibrium will be a (suboptimal) separating equilibrium. Notice that under this contract, high-quality firms get *more* than what they would have like to: thus, when a suboptimal separating equilibrium attains, we have *overinvestment* by high-quality firms.

However, if condition (i) above is satisfied while condition (ii) is not, no separating equilibrium obtains, either optimal or suboptimal. To see this, consider the contract pair (I_p,z_p) determined by the tangency point between the pooling supply curve and the high-quality indifference curve. The breaking down of condition (ii) above implies that the latter indifference curve crosses on the right the bank supply curve to the high-quality types at a point which is lower than the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. In this case, the (I_p,z_p) contract becomes a pooling equilibrium since the (I_h^*,z_h^*) contract (the high-quality "first best" contract, point E) is clearly unfeasible, as it would attract low-quality entrepreneurs, whereas the (I_h,z_h) contract (which lies on the low-quality indifference curve, at the tangency point H) would not be chosen by high-quality entrepreneurs. On the other hand, low-quality entrepreneurs prefer a contract like that given by point C to that in point A: hence, both quality-types entrepreneurs

[20] Recall that, since a borrower can be low-quality with probability (1-p) and high-quality with probability p, the pooling supply curve is determined by equation (15) above.

will choose the (I_p,z_p) contract, point C (see Figure 4).

Therefore, it is clear that in order to have a pooling equilibrium, indifference curves ought to be strictly convex functions of z, the promised loan payment. Hence, we will have a *pooling equilibrium* whenever, in addition to conditions (A1), (A2), and (A3), the following holds:

Assumption A4: The indifference curves of different quality types are convex in z.

Notice that the MLRP of firm returns with respect to quality (assumption A2) and the negative cross-partial derivative (assumption A3) together imply second-order stochastic dominance with respect to the loan size[21]. That is, for each quality type, indifference curves are concave in I. However, indifference curves need not be convex in z, as different quality types may have indifference curves that intersect somewhere. Assumption A4 implies that this will not occur[22], in which case quality signaling will be impossible[23].

Intuitively, a pooling equilibrium occurs when the separating offer curves are far apart and indifference curves of high and low quality types are similar. The first conditions implies large differences in default risks between quality types, whereas the second implies small differences in marginal project returns which offset differential default risks to produce similar tradeoffs between I and z. In these conditions, the bank is not able to exploit the differences in the shapes of entrepreneurs' indifference maps, and a pooling equilibrium emerges. Notice that with conditions (A1) and (A4) alone, we could have equilibria that are not suboptimal: contracts would not be "ranked" in a MLRP sense with respect to quality, thus violating the incentive compatibility constraint (E1) and the low-quality rationality constraint (E2). It is the joint imposition of conditions (A1), (A2), (A3), and (A4) that ensures the existence of pooling (sub-optimal) equilibria. Now, when a pooling equilibrium attains, we have that low-quality

[21] Provided returns are not only increasing but concave in quality.

[22] Condition (A4) is actually sufficient but not necessary. If convexity holds, signaling through quality would be impossible.

[23] In fact, it becomes impossible to distinguish a shift in the curve (an increase in firm returns due to larger loans) due to an increase in quality from a shift simply due to an increase in the size of the loan for a given quality. As a consequence of assumption A3, higher quality implies lower indifference curves, for given levels of I or z.

firms will obtain loan amounts *larger* than their "first best" levels, while high-quality firms will obtain loan amounts *smaller* than their "first best" level (see Figure 4). In other words, better investment projects subsidize poorer projects. Thus, we have type I *credit rationing* of high-quality firms (they get less than what they would like to).

Notice that while a Riley construction of the game structure would lead to the same results as the Wilson construction we have adopted here in the cases where separating equilibria occur, it would not do so when a Wilson construction leads to a pooling equilibrium. In the Riley construction, firms move first, and thus a pooling contract is not viable. Thus, if assumption (A3) holds and indifference curves have the same slope at the tangency points, the incentive compatibility constraint and the low-quality rationality constraint will assure that the two equilibrium points E_l^* and E_h will attain, i.e. points A and H, a separating equilibrium. In that case, while low-quality firms get their first-best loan size (which is less than the "pooling" level), high-quality firms get more than their first-best (and obviously more than the "pooling" level). Credit rationing in a Riley construction never occurs.

This stresses that the sequence of moves in this game structure is very important since, before an equilibrium is reached, all the feasibility constraints must be satisfied. Thus, when the bank first starts by offering the menu given by {A,F,E}, if the indifference curves are as in Figure 2, then a separating equilibrium obtains (low-quality applicants choose A, while high-quality applicants choose E), as the bank will earn non-negative expected profits on both contracts. Conversely, if both applicants choose E, then the bank withdraws the offer, and gives out a second offer, like {A,F,H}. If some applicants choose A while others choose H, then a separating equilibrium obtains (low-quality are better off, while high-quality applicants are over-financed but can signal their quality)[24]. If instead some applicants will prefer F (most probably the high-quality ones) to H, it will be the case that also low-quality will do that, and a pooling equilibrium will obtain. Firm signaling in such case will be impossible (Figure 4).

The extension of the above results to the case of *Q* quality types is straightforward. We can thus state the following Propositions.

---

[24] It could never happen that low-quality applicants choose A while high-quality choose F, since indifference curves are upward sloping, and hence H must lie on a higher high-quality indifference curve than F.

Proposition 2. Under conditions (A1), (A2) and (A3), the equilibrium set of separating contracts has the properties:

(i) the lowest quality investment that is financed is a "first best" contract;

(ii) each contract breaks even;

(iii) each contract (I_q, z_q) is increasing in q: if $(I_{q+1}, z_{q+1}) > (I^*_{q+1}, z^*_{q+1})$ (the "first best" contract for the quality-(q+1) entrepreneur), then we have a suboptimal separating contract, otherwise we have a "first best" separating contract;

(iv) each indifference curve is increasing in q.

Proposition 3. Under conditions (A1), (A2), (A3) and (A4) the equilibrium pooling contract has the properties:

(i) no single contract breaks even but the average contract does;

(ii) no project that is financed is a "first best" contract: "lower-than-average" quality projects are over-financed whereas "higher-than-average" quality projects are under-financed.

We can conclude that, as we claimed in Section 1 above, our model specification encompasses M-R's case in which the size of the project is increasing in firm quality, thus satisfying assumptions (A1), (A2), and (A3). M-R base this case on the arbitrary assumption that the firm production function has a multiplicative form; for, if the production function had an additive form, the other M-R's case, where the size of the project is invariant to quality, would hold. Our model, by referring to the "organizational complexity" parameter H, which is common knowledge, is, on the other hand, able to satisfy the same conditions without imposing constraints on the form of the production function. In particular, it could apply as well if the production function was additive, as assumption (A2) implies that, for given levels of z and increasing quality, the same level of output obtains from lower loan sizes. Moreover, M-R rule out the possibility of pooling equilibria. Our model, by adoptin a Wilson construction, leads instead to this type of equilibria whenever condition (A4) is satisfied. In summary, once conditions (A1), (A2), (A3), and (A4) are fulfilled, pooling equilibria obtain without imposing additional constraints on the form of the production function. What is needed here is simply a representation which attributes "better" distributions to "good news" and a Wilson game construction.

B. External economies of scale prevailing: H = E.

Consider now the case of prevailing external economies of scale. In this case, to preserve their expected returns, entrepreneurs of different qualities can give up the *same amount of expected profit* in exchange for one extra dollar of loans. However, since low-quality types have higher default risk, this homogeneous marginal transfer of expected returns is associated with a larger increase in the promised loan payment z for lower quality types. It follows that low-quality entrepreneurs will have steeper indifference curves than high-quality entrepreneurs. This means that, when H = E, we get the same results obtained in the M-R's case where the firm production function has an additive form (see M-R (1988, p. 111-113)): larger loan sizes signal lower entrepreneurs' quality. Assumption (A2) is then replaced by

<u>Assumption A5</u>: $\frac{\partial}{\partial \pi} \frac{f_q(\pi \mid I,q,H)}{f(\pi \mid I,q,H)} < 0$, for all π within $[0,K(I,q,H)]$.

Assumption (A5) implies that lower quality produces "better" net worth distributions in the sense of statistical "good news". Condition (A3) is in turn replaced by the assumption:

(A6) $$F_{Iq}(\pi) > 0$$

where the positive cross-partial derivative ensures that lower quality is associated with higher *marginal* project returns. Assumption (A6) implies that the marginal rate of substitution between I and z is lower, the higher the quality of a project[25]. Formally, (A1), (A5), (A6), (9a) and (9b) imply:

(19) $$\frac{\partial}{\partial q} \left. \frac{\mathrm{d}z}{\mathrm{d}I} \right|_{\bar{V}} = \frac{\partial}{\partial q} \left(- \frac{V_I(I,z \mid q,H=\mathrm{I})}{V_z(I,z \mid q,H=\mathrm{I})} \right) < 0 .$$

Similarly to M-R, once again, we can restate this by saying that the marginal decrease in promised loan payment that a loan applicant requires in order to accept a smaller loan is

[25] That is, as we just noted, in each and every point in the space (I,z) where low-quality and high-quality indifference curves cross, the former are steeper than the latter curves.

lower the higher the entreprenur's quality.

The equilibrium conditions in the case where H = E mirror the conditions derived above for H = I (see Figure 5). In particular, Proposition 1 still holds, in the sense that we will have a *separating* set of "first best" equilibria E_h^* and E_l^* (i.e. points D and B) if (i) each lies on a separating supply curve, and (ii) the low-quality contract is weakly preferred by the low-quality entrepreneur. Condition (i) is necessary but not sufficient to have a separating equilibrium, since the high-quality contract (I_h^*, z_h^*) can lie below the low-quality indifference curve which yields the low-quality preferred contract (I_l^*, z_l^*). In this case, condition (ii) would not be satisfied. Thus, the high-quality optimal contract must lie above the low-quality "optimal" indifference curve. In this case, the entrepreneurs' optimal choice elicits a "first best" equilibrium (a separating equilibrium with firm signaling).

Again, the converse of Proposition 1, is not true, in the sense that we can have separating equilibria which are not "first best" (Figure 6). The optimal equilibrium is not feasible whenever (i) the high-quality preferred contract lies above the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. The consequent "suboptimal" equilibrium will be separating if (ii) the high-quality indifference curve which is tangent to the bank pooling supply curve crosses on the right the bank supply curve to the high-quality types at a point lower than the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. The latter point of intersection will be the high-quality contract.

In more formal terms, consider the two first-best contracts (I_l^*, z_l^*) and (I_h^*, z_h^*), and suppose the latter contract lies below the low-quality preferred indifference curve, thus attracting low-quality entrepreneurs. To avoid negative expected profits on loans to low-quality entrepreneurs, the bank will then offer a new contract (I_h, z_h) at the intersection of the low-quality indifference curve and the high-quality offer curve. Thus, low-quality entrepreneur will choose their "first best" contract (I_l^*, z_l^*), whereas high-quality entrepreneurs will choose their "second best" contract (I_h, z_h). The resulting equilibrium will be a (suboptimal) separating equilibrium: under this contract, low-quality firm will get their first-best loan size, while high-quality firm will be under-financed.

Conversely, when condition (i) above is satisfied while condition (ii) is not, we do not get any separating equilibrium, neither optimal nor suboptimal. To see this, consider the contract (I_p, z_p) at the tangency point between the bank pooling supply curve and the high-quality indifference curve (see Figure 6). The breaking down of condition (i) above implies that the latter indifference curve crosses on the right bank supply curve to the high-quality types at a point which is higher than the intersection between the low-quality preferred indifference curve and the bank supply curve to the high-quality types. In this case, the (I_p,z_p) contract becomes a pooling equilibrium since the (I^*_h,z^*_h) contract (the high-quality "first best" contract) is clearly unfeasible, as it would attract low-quality entrepreneurs, whereas the (I_h,z_z) contract (which lies on the low-quality indifference curve tangent to **) would not be chosen by high-quality entrepreneurs. On the other hand, low-quality entrepreneurs prefer the contract at point ** than that at point **. Hence, both quality-types entrepreneurs will choose the (I_p,z_p) contract.

Therefore, in order to have a pooling equilibrium, indifference curves ought to be, as before, strictly convex functions of z. Hence, we will have a *pooling equilibrium* whenever, in addition to conditions (A1), (A5), and (A6), assumption (A4) also holds. In this case, lower quality is assumed to yield superior output distributions[26]. Notice that, again, the MLRP of returns with respect to quality (assumption A5) and the positive cross-partial derivative (assumption A6) together imply second-order stochastic dominance with respect to the project size (or loan size)[27]. That is, for each quality type, indifference curves are concave in I. However, indifference curves need not be convex in z, as different quality types may have indifference curves that intersect somewhere. Assumption A4 implies that this will not occur, in which case quality signaling will be impossible[28].

Now, if a pooling equilibrium attains, we have that low-quality firms will obtain loan amounts *smaller* than the "first best" levels, while high-quality firms will obtain loan amounts *higher* than the "first best" level (see Figure 7). In other words, low-quality projects subsidize high-quality projects. Hence, we have type I *credit rationing* of low-quality firms (they get less than what they would like to). Notice that, as in the case of H = I, this rationing result stemming from a pooling equilibrium differs from the result reached under a Riley game

[26] And, again, higher quality implies lower indifference curves, for given levels of I or z.

[27] Provided projects returns are not only increasing but concave in quality.

[28] Again, as in the previous case, it becomes impossible to distinguish a shift in the curve (an increase in returns due to a smaller project) due to an increase in quality from a shift simply due to an decrease in the size of the project for a given quality. As a consequence of assumption A5, when H = E, higher quality implies lower indifference curves, for given levels of I or z, like in case of H = I.

construction which necessarily leads to separating equilibria. However, differently from the case of H = I, when H = E, credit rationing occurs even with a Riley separating equilibrium. Under a Riley construction, while low-quality firms elicit their first-best loan size (which is obviously larger than the "pooling" level), high-quality firms obtain a loan amount smaller than their first-best level (and even smaller than the "pooling" level). Given that the Riley and the Wilson separating equilibria coincide, our model with H = E thus displays a rich set of possible credit-rationing equilibria: with a pooling equilibrium we obtain a type I credit rationing for low-quality borrowers, while with separating equilibria we obtain a type I credit rationing for high-quality borrowers.

The extension of the above results to the case of Q quality types is straightforward. We can thus state the following Propositions.

Proposition 4. Under condition (A1), (A5) and (A6), the equilibrium set of separating contracts has the properties:

(i) the project of lowest quality that is financed is a "first best" contract;

(ii) each contract breaks even;

(iii) each contract (I_q,z_q) is decreasing in q: if $(I_{q+1},z_{q+1}) < (I^*_{q+1},z^*_{q+1})$ (the "first best" contract for the quality-(q+1) entrepreneur), then we have a suboptimal separating contract, otherwise we have a "first best" separating contract;

(iv) each indifference curve is increasing in q.

Proposition 5. Under conditions (A1), (A4), (A5) and (A6), the equilibrium pooling contract has the properties:

(i) no single contract breaks even but the average contract does;

(ii) no project that is financed is a "first best" contract: "lower-than-average" quality projects are under-financed whereas "higher-than-average" quality projects are over-financed.

Again, we conclude by stating that, as we claimed in Section 1 above, our model specification encompasses M-R's case based on the arbitrary assumption of an additive firm production function, as it can satisfy assumptions (A1), (A5), and (A6) without imposing any other constraint. In particular, conditions (A5) and (A6) are compatible with a multiplicative production function and, for varying levels of z, with an additive production function as well. Moreover, while M-R rule out the possibility of pooling equilibria, our model leads to this type of equilibrium whenever assumption (A4) is satisfied and with no additional constraints on the form of the production function. Together with the conclusions drawn above for the case of H = I, we may conclude that our model is more general than M-R's. In this respect, it is worthwhile noticing that whereas M-R's model depicts different and unrelated "cases" and link them through the arbitrary assumption that each lender knows *ex-ante* the form of the borrowers' production function, our model offers a unified framework. For the differences between H = I and I = E reduce to some global and clearly observable characteristic. With repeated sampling by lenders, these characteristics are summarized by the signs of the MLRP or project returns with respect to quality and by the signs of the cross-partial derivatives of the return distributions with respect to I and q (both negative in the former and bothe negative in the latter case). Thus, it appears that, from the lender's perspective, observing the value taken on by the parameter H is (statistically) "sufficient" to know how to evaluate the derivatives of the posterior distribution functions, and hence to exploit borrowers' willingness of signaling. Yet, even though this is an extremely simplified representation of the information contraints, it appears to be less controversial than assuming that each lender knows *ex-ante* the form of each firm production function. As we will see below, it is quite difficult to go from a production function multiplicative with respect to quality to a production function that is additive in quality, as there is no way to devise anything between the two. These are opposite and alternative (and indeed specific) technologies: but can indeed everything be subsumed within these two forms? Conversely, by defining a technological or organizational parameter *which can vary* we can certainly be able to accomplish with this framing a more general and compelling way to settle a great deal of different possibilities.

D. "Complexity" as a continuous variable

As we have mentioned above, the assumption of our model, too, can be generalized. In particular, the assumption that H is a discrete binary variable is a rather strict one. Yet, letting H be a continuous variable makes the problem increasingly more complex to solve. In this sub-section we suggest that these technical complications do not alter our main results.

The treatment of H as a binary variable has basically allowed us to state that if internal economies of scale and/or transaction costs prevail, then the sign of the MLRP of project returns with respect to quality is positive, while if external economies of scale and/or organizational costs prevail, then the sign of this MLRP is negative. Consequently, the cross-partial derivative of profit densities with respect to loan size and quality is negative in the former case (H = I) and positive in the latter (H = E). We have also seen that, when H = I, the higher is project quality, the greater is the marginal rate of substitution between I and z (i.e. high-quality indifference curves are steeper), while, when H = E, the higher is project quality, the smaller is the marginal rate of substitution between I and z (i.e. low-quality indifference curves are steeper). Moreover, we have seen that a sufficient condition for a pooling equilibrium is that indifference curves be convex, which implies that the marginal rates of substitution between I and z are equal in correspondence of parallel tangency points.

These elements explain the specific features of both our pooling and separating equilibria when either H = I or H = E. As we have seen above, a separating equilibrium will emerge (i) if banks' separating supply curves are not so far apart, that is, if there are small inter-quality differences in default risks and (ii) if the indifference curves of low- and high-quality entrepreneurs are dissimilar, that is, if inter-quality differences in projects marginal returns do not offset different default risks to produce different tradeoffs between loan size and repayment. As banks' separating supply curves shift farther apart, which means that inter-quality differences in default risks get larger, and indifference curves of different types gets more similar, which means that inter-quality differences in projects marginal returns offset differential default risks to produce similar tradeoffs between loan size and repayment, a pooling equilibrium emerges.

It can thus be shown, at least on a intuitive ground, that if the "organizational complexity" variable H takes more than two values, or even a continuum of values in a definite range, the features of the different emerging equilibria are not substantially altered. So, it is apparent that, if continuous, H can take a value by which the *mlrp* of profits with respect to quality will be zero, so that the cross-partial derivative of profit densities with respect to investment and quality will be zero too. Then, *a fortiori*, this must hold true even if H is a discrete (binary) variable, in the sense that it becomes equal to I whenever it is lower than some arbitrarily given value ϕ, and equal to E whenever it is higher than that same value ϕ; for there will be a point $H = \phi$ at which $H \neq$I,E. In both these cases, larger loans *do not* signal neither higher nor lower quality: they would be simply invariant to quality. This happens when the two indifference maps of low-quality and high-quality borrowers are the same, or at least when these maps coincide in the relevant range of contracts offered by the bank. A pooling contract will be the only possibility for the bank in this case.

This result perfectly fits with the equilibrium features as stated by (i) and (ii) above. Let us compare Figure 2 with H = I with Figure 5 with H = E. As H changes, the two figures depict two opposite situations: the low-quality indifference map shifts from the left-bottom part north-westward along the bank supply curve to the low-quality types, while the high-quality indifference map shifts from the right-top part south-westward along the bank supply curve to the high-quality types. Hence, we may presume that as H changes from I to E the two indifference maps of the two quality types shift in opposite directions becoming more and more similar and the bank supply curves get farther apart as far as H reaches a given value. In the neighborhood of this value, H is neither I nor E, the bank supply curves are the farthest apart, and the indifference maps of the two quality types coincide. In this interval, a pooling equilibrium emerges. On the other hand, at the two extremes (H = I or H = E), we can think of no differences in inter-quality default risks and of large difference in projects marginal returns. In this case one sungle supply curve will be sufficient for the bank to elicit a separating equilibrium at two points lying on the same curve.

4. A comparison with previous model settings

In the previous Sections we have often claimed that our paper offers a more general and unifying framework for the definition of credit market equilibria. However, the comparison with other models has until now been limited to a rather quick critical analysis of the assumptions and of the equilibria characterizing the M-R model. For the sake of completeness, it seems necessary to extend this comparison to other models, and in particular to the S-W and D-W ones, as well as to other topics.

There are several key issues that make this comparison actually rather difficult. Although the main purpose of the S-W, D-W, and M-R models as well as ours is quite similar (to show the possibility of credit rationing in competitive credit markets), the setting of each of these models is rather different. In the first place, different models allow for projects either of fixed size or of variable size. Secondly, different models resort to different variables in order to gather projects of different "nature" within homogenous pools, which are thus determined and ordered according to different dominance criteria. The direct comparison between non-homogeneous ordering variables or dominance criteria is unfeasible in raw terms. At this aim, it seems necessary to define a common framework by which to relate the realized outcomes of different projects in the different models, as well as a principle of generalization. In this section we deepen our focus with respect to two key issues: the comparison between the *quality* and the *riskiness* of a given project, and the comparison between FOSD, SOSD, MPS and the MLRP.

A. Low-risk versus high-risk investment.

In all the cases described above, we had that larger loans implied higher firm expected returns, while higher quality implied higher firm expected returns if H = I and lower firm expected return if H = E. As we already mentioned, in the S-W model, borrowers (projects) are different in riskiness, which is unobservable to lenders. In particular, "the bank is able to distinguish projects with *different mean returns*" (1981, p. 395). It follows that all the projects in the same risk pool have the same expected return but some are riskier than others in the sense of mean-preserving spread (MPS). In the notation of the present paper we can say that $q_i<q_j$ indicates that projects of quality q_i are *riskier* than those of quality q_j, so that *a decrease in q represents an increase in riskiness.* As is well known, S-W considered projects of fixed loan size, while we apply their case to variable loan requirements, and we will show that S-W analysis gets to results similar to ours.

Consider a population of firms of different increasing riskiness, indexed by q. We say that if, for a nondegenerate prior distribution G for q, the posterior distribution $G(\cdot\,|\,\rho_2)$ dominates the posterior distribution $G(\cdot\,|\,\rho_1)$ in the sense of MPS, then ρ_1 is *riskier* than ρ_2. This is the definition of a MPS increasing risk which, as we know, implies but it is not implied by SOSD. Thus, if a signal ρ_1 is *less favorable (riskier)* than ρ_2, then it follows that the conditional densities $\{f(\cdot\,|\,q)\}$ have the MLRP for every $\rho_1 < \rho_2$ and $q_i < q_j$. In other words, we restrict our attention to the case of two entrepreneurs i and j of different riskiness whose expected profit is the same, but the first is riskier (it has "fatter right tail"). For the same given level of expected profits, an increase in risk (a decrease in quality) will induce an increase in the demand for loans. Hence, we are back to the case of H = E in our model above, and condition (A5) will keep holding. Also, for the profit distribution to be a MPS risk increasing it must be that assumption (A6) holds, as the cross-partial derivative of the cumulative profit distribution function with respect to I and q has to be positive (higher q means lower risk). We can thus parallel the conditions of the S-W model, as, for profit densities characterized by MPS differential riskiness, the MLRP implies SOSD (and thus MPS), while assumptions (A5) and (A6) will ensure that increase in risk and loan size be positively related.

Hence, similarly to the case of H = E depicted above, we have that applicants with riskier projects are willing to pay a larger z for a larger loan. Applicants with riskier projects can signal their quality type by accepting a larger rather than a smaller loan. This implies that the marginal rate of substitution between I and z is higher, the riskier the project. Formally, condition (19) still holds, where q now stands for "risk" rather than "quality". Similarly to M-R, once again, we can restate this by saying that the marginal decrease in promised loan payment that a loan applicant requires in order to accept a smaller loan is lower the riskier the project. Therefore, all the equilibrium conditions are thus similar to those depicted above for H = E, once we replace the concept of "lower quality" of a project with that of "higher risk".

The extension of the above results to the case of Q quality types is straightforward. We can thus state the following Propositions, paralleling Propositions 4 an 5 above.

Proposition 6. Under condition (A1), (A5), and (A6) the equilibrium set of separating contracts has the properties:

(i) the riskiest project that is financed is a "first best" contract;

(ii) each contract breaks even;

(iii) each contract (I_q, z_q) is increasing in riskiness: if $(I_{q+1}, z_{q+1}) < (I^*_{q+1}, z^*_{q+1})$ (the "first best" contract for the quality-$(q+1)$ entrepreneur), then we have a suboptimal separating contract, otherwise we have a "first best" separating contract;

(iv) each indifference curve is decreasing in risk.

Proposition 7. Under conditions (A1), (A4), (A5), and (A6), the equilibrium pooling contract has the properties:

(i) no single contract breaks even but the average contract does;

(ii) no project that is financed is a "first best" contract: projects of "lower-than-average" risk are over-financed whereas projects of "higher-than-average" risk are under-financed.

B. The characteristics of the MLRP.

To prove that also S-W model leads to results analogous to ours, it is of course not sufficient to claim that our model offers a more general an unifying framework to study credit rationing equilibria in the market for loans. The S-W model itself features key differences with respect to other models such as the D-W one. In this last respect, it suffices noticing that, even if both these models assume a fized loan size, they widely differ with respect to the variables explaining how projects are pooled and the related dominance criteria.

We just recalled that, in S-W, all projects in the same pool have the same expected return but different riskiness. On the other hand, D-W assume that all projects in the same pool "yield *the same return* R^S *if successful and* R^f *if not.* [...] What distinguishes projects is the probability of success" (1987, p. 282). This means that projects in the same pool have *different mean returns* in D-W while projects in the same pool have *different returns if successful* in S-W. As a consequence, we can single out along these lines the main elements which can be held as given in a simplified version of the two models. In S-W the bank screens out projects according to their *average* return, while in D-W the bank distinguishes projects according to their return *in the case of success.* In both cases, the probability of success is unknown to the lender. Hence, in S-W the bank pools together all projects with the same average return, R, *for a given* return in the case of a failure, R^f (often equalized to zero), whereas in D-W the bank pools together all projects with the same return in the case of success, R^s, *and* the same return in the case of a failure, R^f. Thus, the expected return of a project is *decreasing* in p, the probability of success, in the first case, and it is *increasing* in p in the second one[29]. Therefore, all projects in the same pool have the same mean and can be ranked according to their *dispersion* (the MPS criterion) in the S-W case, while they can

[29] Consider two projects of different quality/riskiness *in the same pool*. In the S-W setting, this implies that they yield the *same* mean return: hence, a *higher* probability of success means a *lower* return in the case of success (as the mean must be the same). In the D-W setting, this implies that they yield the *same* return in case of success: hence, a *higher* probability of success means a *higher* return in the case of success (the mean does not have to be the same) or, equivalently, the *same* return in the case of success means a *higher mean* return.

be ranked according to their *different mean* (the FOSD criterion) in the D-W case.

All this makes clear that, to compare the S-W and the D-W models, it would be necessary to subsume the two different projects orderings across pools resulting from the different criteria of project ranking under a unique sorting criterion. However, as is well known, the MPS and the FOSD ranking are not directly comparable, as the former only applies when the latter cannot be conclusive (two distributions with the same mean cannot be ranked according to the FOSD criterion).

Yet, we can elaborate on these differences. If, for instance, we drop the assumption that projects yield the same R^f in the S-W model, then the expected return of a project will now be *increasing* in p, if a *higher* p implies a *higher* R^s (a lower R^f). Conversely, if we drop the assumption that projects yield the same R^f in the D-W model, then the expected return of a project will now be *decreasing* in p, if a *higher* p is still associated with the same *mean return* R (and thus higher R^s and lower R^f). Hence, the expected return could be increasing in p, and rising as p rises across projects for given levels of R^s, in both models. When projects are pooled according to their return in case of success, R^s, the problem then becomes how to compare projects with *different* returns in case of success. As we can see, the problem has three dimensions: R^s, R^f, and p. Suppose two projects have the same R^f: then a higher p could either imply the same R^s, and the expected return will be rising (D-W), or a lower R^s, and the expected return will be the same (S-W). Thus, for given levels of R^f and R^s, a *higher* expected return implies either a higher p (D-W) or a lower p (S-W). Obviously, once both p and R^s are given, a lower R^f implies a higher expected return. Suppose now two projects have the same R^f but different R^s: then a higher R^s could either imply the same p, and the expected return will be rising (D-W), or a lower p, and the expected return will be the same (S-W): but if R^f is not the same this will no longer be true. As we have seen, we can always associate a higher R^s with an equal of a higher p, and, provided R^f is equal o lower, we will have an higher expected return or an equal expected return, respectively. The first case implies first-order stochastic dominance, while the latter implies a mean-preserving increase in dispersion. This means that, even if the two models remain characterized by their original dominance criteria (MPS in S-W and FOSD in D-W), they could issue comparable *signals*[30].

[30] Here the word *signal* refers to an observable *carrying news* and is not to be counfounded to the *signaling* by agents of their quality type.

The above conclusion is important, since the presence of comparable signals allows a comparison between different dominance criteria according to their degree of generality. As we have seen in Section 2, the MLRP is implied by FOSD in the case of two-point distributions, as is the case we are dealing with now, but also by MPS, provided the increase in skewness is monotonic. As Milgrom points out, this is not always true. The converse is true, however, as we will state in the following propositions. If we extend the analysis to more-than-two-point and to continuous distributions, the MLRP is not implied by FOSD (and, *a fortiori*, not even by MPS). The MRLP is suited for comparing two *signals* and is a necessary and sufficient condition for a signal x to be *more favourable than* a signal y both in the sense of FOSD and SOSD for every increasing concave function of a parameter τ, $U(\tau)$, and non-degenerate prior distributions $G_1(\tau)$ and $G_2(\tau)$. As in the present case, the MLRP is useful when we have a continuum of values for an observable variable (e.g. the return of a project) with associated probabilities of occurrance.

Let G be a prior distribution for the random variable τ that assigns probabilities $g(\tau_1)$ and $g(\tau_2)$ to two possible values τ_1 and τ_2 of τ.

Proposition 8. If the densities $\{f(\cdot\,|\,\tau)\}$ have the MLRP for every $x > y$ and $\tau_1 > \tau_2$, such that:

(20) $$f(x|\tau_1)f(y|\tau_2) > f(x|\tau_2)f(y|\tau_1)$$

then $G(\tau\,|\,x)$ dominates $G(\tau\,|\,y)$ in the sense of FOSD.

Proof. By Bayes' theorem, (20) implies that:

(21) $$\frac{g(\tau_1|x)}{g(\tau_2|x)} = \frac{g(\tau_1)}{g(\tau_2)}\frac{f(x|\tau_1)}{f(x|\tau_2)} > \frac{g(\tau_1)}{g(\tau_2)}\frac{f(y|\tau_1)}{f(y|\tau_2)} = \frac{g(\tau_1|y)}{g(\tau_2|y)}$$

that is

(22) $$\frac{g(\tau_2|x)}{g(\tau_1|x)} < \frac{g(\tau_2|y)}{g(\tau_1|y)}\,.$$

Recall that $G(\tau\,|\,x)$ dominates $G(\tau\,|\,y)$ is the sense of FOSD if and only if for every τ, $G(\tau\,|\,x) \leq G(\tau\,|\,y)$, with strict inequality for some value of τ. Now, as $\tau_1 > \tau_2$, $G(\tau\,|\,x)$ dominates

$G(\tau \mid y)$ is the sense of FOSD if $g(\tau_1 \mid x) < g(\tau_1 \mid y)$ or $g(\tau_2 \mid x) < g(\tau_2 \mid y)$ or both. This implies that:

$$\text{(23)} \qquad \int_{\tau \le \tau_1} f(x \mid \tau)\, dG(\tau) < \int_{\tau \le \tau_1} f(y \mid \tau)\, dG(\tau) \quad \Rightarrow \quad g(\tau_1 \mid x) < g(\tau_1 \mid y)$$

and

$$\text{(24)} \qquad \int_{\tau \le \tau_2} f(x \mid \tau)\, dG(\tau) < \int_{\tau \le \tau_2} f(y \mid \tau)\, dG(\tau) \quad \Rightarrow \quad g(\tau_2 \mid x) < g(\tau_2 \mid y)$$

Also, since $\tau_1 > \tau_2$ obviously implies that $g(\tau_1 \mid x) > g(\tau_2 \mid x)$ and $g(\tau_1 \mid y) < g(\tau_2 \mid y)$, the inequality in (19) is easily established. Hence, the MLRP implies FOSD. Q.E.D.

Proposition 1 implies that for every increasing function $U(\tau)$, if the family of densities $\{f(\cdot \mid \tau)\}$ has the MLRP, then

$$\int U(\tau) dG(\tau \mid x) > \int U(\tau) dG(\tau \mid y).$$

Proposition 9. If the densities $\{f(\cdot \mid \tau)\}$ have the MLRP for every $x > y$ and $\tau_1 > \tau_2$, such that:

$$\text{(25)} \qquad f(x \mid \tau_1) f(y \mid \tau_2) > f(x \mid \tau_2) f(y \mid \tau_1)$$

so that

$$\text{(26)} \qquad \frac{g(\tau_2 \mid x)}{g(\tau_1 \mid x)} < \frac{g(\tau_2 \mid y)}{g(\tau_1 \mid y)},$$

then $G(\tau \mid x)$ dominates $G(\tau \mid y)$ in the sense of SOSD.

Proof. Recall that $G(\tau \mid x)$ dominates $G(\tau \mid y)$ is the sense of FOSD if and only if, for every τ,

$$\int_{\tau \le \tau^*} [G(\tau \mid x) - G(\tau \mid y)]\, d\tau \equiv \Delta(\tau^*) \le 0,$$

for all τ^*, with strict inequality for some value of τ. Now, $G(\tau \mid x)$ dominates $G(\tau \mid y)$ in the sense of SOSD if:

(27) $$\int_{\tau\leq\tau_1} G(\tau|x)\,d\tau < \int_{\tau\leq\tau_1} G(\tau|y)\,d\tau$$

or

(28) $$\int_{\tau\leq\tau_2} G(\tau|x)\,d\tau < \int_{\tau\leq\tau_2} G(\tau|y)\,d\tau$$

or both. This implies, as before, that:

(29) $$\int_{\tau\leq\tau_1} f(x|\tau)\,dG(\tau) < \int_{\tau\leq\tau_1} f(y|\tau)\,dG(\tau) \quad\Rightarrow\quad g(\tau_1|x) < g(\tau_1|y)$$

and

(30) $$\int_{\tau\leq\tau_2} f(x|\tau)\,dG(\tau) < \int_{\tau\leq\tau_2} f(y|\tau)\,dG(\tau) \quad\Rightarrow\quad g(\tau_2|x) < g(\tau_2|y)$$

Also, since $\tau_1 > \tau_2$ obviously implies that $g(\tau_1|x) > g(\tau_2|x)$ and $g(\tau_1|y) < g(\tau_2|y)$, the inequality in (26), and hence in (25), is easily established. Hence, the MLRP implies SOSD[31]. Q.E.D.

Proposition 9 implies that for every increasing concave function $U(\tau)$, if the family of densities $\{f(\cdot|\tau)\}$ has the MLRP, then

$$\int U(\tau)\mathrm{d}G(\tau|x) > \int U(\tau)\mathrm{d}G(\tau|y).$$

Proposition 10. If the densities $\{f(\cdot|\tau)\}$ have the MLRP for every $x > y$ and $\tau_1 > \tau_2$, where τ_i is a risk parameter and a larger i denotes smaller risk, such that:

(31) $$f(x|\tau_1)f(y|\tau_2) > f(x|\tau_2)f(y|\tau_1)$$

so that

[31] As Milgrom points out, the inequality in (20) or (25) is necessary and sufficient to conclude that x is more favorable than y in the sense of both first-order and second-order stochastic dominance. As we know, in fact, FOSD implies SOSD (but not the other way around).

(32) $$\frac{g(\tau_2|x)}{g(\tau_1|x)} < \frac{g(\tau_2|y)}{g(\tau_1|y)},$$

and $f(x|\tau_1)$ is a MPS of $f(x|\tau_2)$ such that $E(\tau|x) = E(\tau|y)$, then $G(\tau|x)$ dominates $G(\tau|y)$ in the sense that $G(\tau|y)$ is a MPS of $G(\tau|x)$, that is, $G(\tau|x)$ and $G(\tau|y)$ have the same mean but the latter is riskier (has fatter tails) than the former.

Proof. Recall that $G(\tau|x)$ dominates $G(\tau|y)$ is the sense of MPS if and only if, for every τ, we have that:

(i) $$\int_\tau [G(\tau|x) - G(\tau|y)]\, d\tau = 0,$$

(ii) $$\int_{\tau\le\tau^*} [G(\tau|x) - G(\tau|y)]\, d\tau \equiv \Delta(\tau^*) \le 0,$$

for all τ^*, with strict inequality for some value of τ. Condition (i) is a mean-preserving condition. Condition (ii) implies that $G(\tau|x)$ dominates $G(\tau|y)$ in the sense of SOSD (as we know, in fact, MPS implies SOSD) and it has been proved above. Since condition (i) is a mean preservation property, it implies that:

(33) $$G(-\infty|x) = G(-\infty|y) \quad \text{and} \quad G(+\infty|x) = G(+\infty|y).$$

Now, the equality in mean implies that $E(\tau|y) = E[(\tau|(x+\varepsilon)]$, where ε has zero-mean density $g(\tau|\varepsilon)$ such that $E[\tau|\varepsilon] = 0$. Then, it must be that:

(34) $$\begin{aligned}&\int \tau\, g(\tau|y)\, d\tau = \\ &= \int \tau\, g((\tau|(x+\varepsilon))\, d\tau = \int \tau\, (g(\tau|x) + g(\tau|\varepsilon))\, d\tau = \\ &= \int \tau\, g(\tau|x)\, d\tau,\end{aligned}$$

that is

(35) $$\int \tau\, G(\tau|y)\, d\tau = \int \tau\, G(\tau|x)\, d\tau.$$

Integrating by parts we have:

(36) $$\tau\, G(\tau|y)\Big|_{-\infty}^{+\infty} - \int G(\tau|y)\, d\tau = \tau\, G(\tau|x)\Big|_{-\infty}^{+\infty} - \int G(\tau|x)\, d\tau$$

from which we get

$$(37)\qquad \int G(\tau|y)\,d\tau = \int G(\tau|x)\,d\tau$$

so that

$$(38)\qquad \int_\tau [G(\tau|x) - G(\tau|y)]\,d\tau = 0\,.$$

Q.E.D.

Therefore, $g(\tau_1 | y)$ is a MPS of $g(\tau_1 | x)$, and $g(\tau_2 | y)$ is a MPS of $g(\tau_2 | x)$. Moreover, since $\tau_1 > \tau_2$ implies increase in risk (greater dispersion or fatter tails), we have that $g(\tau_1 | x)$ is a MPS of $g(\tau_2 | x)$, and $g(\tau_1 | y)$ is a MPS of $g(\tau_2 | y)$. Now, the MLRP implies that the likelihood ratio $f(x | \tau_1)/f(x | \tau_2)$ is monotone in x and increasing if $\tau_1 > \tau_2$. Also, as the proof of condition (i) above has shown, the mean preservation property is independent of the inequality in (32), whereas (32) implies SOSD. Hence, the MPS in the conditional densities is maintaned in the posterior distributions, while, as the MLRP implies SOSD, it will do so no matter the equality of the posterior means (in other words, the MLRP implies SOSD in all cases, and if the MPS holds for the conditional densities it will hold as well). If the likelihood ratio $f(x | \tau_1)/f(x | \tau_2)$ is monotone in x, increasing in risk, and $f(x | \tau_1)$ is a MPS of $f(x | \tau_2)$, then $G(\tau | x)$ dominates $G(\tau | y)$ in the sense that $G(\tau | y)$ is a MPS of $G(\tau | x)$, that is, $G(\tau | x)$ and $G(\tau | y)$ have the same mean but the latter is riskier than the former. Q.E.D.

Proposition 10 implies that for every concave function $U(\tau)$, if the family of densities $\{f(\cdot | \tau)\}$ has the MLRP *and* the MPS property, then

$$\int U(\tau)\mathrm{d}G(\tau|x) > \int U(\tau)\mathrm{d}G(\tau|y)\,.$$

Furthermore, the converse of Proposition 10 is also true, at least in the case of two-point distributions. Milgrom already showed that both FOSD and SOSD imply MLRP, provided a signal x is more favorable than another signal y in either senses. Now, as a signal x can still be more favorable than another signal y in either senses even if the latter is a MPS of the former, then MPS also implies MLRP, provided that of two signals one is comparably more favorable than the other (but only for two-point distributions).

Thus, we have a tool to generalize the settings of the D-W and S-W models to continuous distributions, provided we have signals that are comparable. The MLRP is obviously a stronger condition, as it implies concavity and a family of densities which is increasing in the observables (the signals). Nevertheless, it is a property which nests both FOSD and MPS under suitable conditions. Thanks to the MLRP, lenders could be able to infer the quality type by looking at projects returns for equal levels of loan sizes.

C. The MLRP with loans of variable sizes.

The above conclusions concerning models with fixed loan sizes, like the S-W's and the D-W's ones, can be extended to the case of loans of variable sizes, all the more so. In this sense, the M-R model, too, offers a generalization of both the S-W and the D-W models. However, as M-R treat the loan size as a choice variable which always allows for complete separation of projects of different quality, we propose a different generalization allowing for both separation and pooling of different quality types. To compare M-R's generalization of S-W's and D-W's models with ours, which we discussed above in Section 4.1, let us consider once more the two-point distribution example of these models. Two different projects could have the same R^s and different p, whereby the higher is such p the "better" will be the project. Or, they could have different R^s and the same p, in which case they will be equally risky for the lender. Thus, what really matters is the probability of success: if two projects have different R^s and different p, the correpondent loci of indifference between interest rate and size of the loan will differ in such a way that a higher p will imply a different (steeper) indifference curve. This is so because a higher p implies, *coeteris paribus*, a higher expected return. There will be a point at which the relation between expected return and repayment schedule is "optimal", and this will differ according to the probability of success, or, in terms of the M-R model, to the threshold level of repayment. The key point, which gives rise to different "models" in M-R, is the relation between quality and loan size, that is, between quality and return of a project. If such a relation is positive (higher quality implies a higher p, which implies a higher return, and a larger loan size can signal such better quality) there will be one type of separating equilibrium, whereby high-quality agents signal their type by accepting larger loans. If such a relation is negative, high-quality agents signal their type by accepting smaller loans. Notice that all this keeps holding even if the size of the loan is fixed: in the former case, high-quality agents will accept a higher interest rate, whereas in the latter they will accept a smaller rate of interest[32]. Hence, letting the loan size vary adds one dimension to the problem, allowing different pairs of interest rate and loan size for different applicants.

5. Summary and Conclusions

In this paper we have analyzed the different equilibria that emerge in a competitive credit market in which optimizing risk neutral firms demanding for loans for financing production projects whose quality is not observable from the outside and banks facing and adverse selection problem due to the lack of information concerning the quality of firm projects sign lending contracts contingent on realized firm output and a given interest rate. Firm actual returns, and thus their ability to repay the loans, are uncertain, since *ex-ante* uncertainty in the realized output of each project makes actual returns stochastic. Moreover, as a result of costly monitoring, banks cannot enforce contracts which specify the quality of a project, so that contracts are made contingent on realized returns. In addition, firms are assumed to have limited liability in their contracts with banks, which thus take the form of standard debt contracts. Limited liability and unobservability of project quality yield an asymmetric information problem which induces adverse selection on the side of the lenders. However, as both the size of the loan and the loan interest rate are allowed to vary, self-selection on the side of the borrowers is made possible, whereby high-quality firms can elicit better credit conditions than low-quality ones, avoiding negative expected returns for the banks on each contract. Assuming a contracting game of a Wilson (1977) type, we have shown that both separating (i.e. where self-selection by borrowers of different quality is possible) and pooling equilibria can emerge, and these latter will entail the possibility of credit rationing for some quality types.

The effects of adverse selection due to asymmetric information in the credit market and the conditions for the rationing of credit have already been studied in a large strand of

[32] This can be seen in the M-R pictures (see Milde and Riley (1988, Figure II, p. 108, and Figure III, p. 112)) by fixing L on the horizontal axis and plotting the families of indifference curves for the two types of applicants. The resulting interest rates for the two types will correspond to the two points where the indifference curves have the same slope at the given level of L.

literature, starting with the work of Jaffee and Russell (1976). Models of credit rationing due to asymmetric information have later been developed particularly by Stiglitz and Weiss (1981) and De Meza and Webb (1987) for the case of loans of fized size, and have been extended to the case of loans of variable size by Milde and Riley (1988). In this paper we have taken a slightly different stance with respect to these models, which we take as representing the fundamental references among the various existing models of production financing through debt contracting. Like M-R, we consider the more general case of contracts designed for projects of different quality by allowing for loans of variable size and variable interest rates. Project quality and realized output are obviously related in a monotonic way, either positive or negative. Differently from M-R, we are able to obtain both credit rationing pooling equilibria and separating equilibria whereby firms of different quality are able to elicit different debt contracts. Also, as quality in not observable, we have assumed that realized firm returns have the *monotone likelihood ratio property* (MLRP) with respect to quality, in the sense of Milgrom (1981). The rationale for this assumption is that of exploiting the monotonicity property implied by the definition of "quality".

Three main results of this paper are worth summarizing. In the first place, our approach encompasses the main features of the previous models such as S-W and D-W, which can be considered as particular cases (loans have fixed size) of a more general model (loans can vary in size). In the second place, by referring to an observable parameter of *organizational complexity* characterizing the borrowing firm, our model allows for a more general framework of analysis than M-R's one. In particular, we can replace with a unifying framework M-R's three different and unrelated "cases". In the third place, by adopting a Wilson construction of the contracting game, we have proved that our model entails the possibility of both separating and pooling equilibria, as opposed to M-R which only entails separating equilibria. Hence, we have strengthened the possibility of obtaining equilibrium credit rationing.

The core of our model construction can be subsumed in the firm *organizational complexity* parameter and in the Wilson construction of the contracting game. The organizational complexity parameter has been defined on the basis of a comparison between two extreme types (internal or external) of economies of scale, and/or between two kinds of costs (transactional costs or organizational costs). We have stated that, when internal economies of scale and/or transaction costs prevail, a larger size of the loan will signal a higher project quality, while when external economies of scale and/or organizational costs prevail, a larger size of the loan will signal a lower project quality. These results are actually quite close to those reached by M-R in their various cases; however, differently from M-R, in our model they have not been based on some arbitrary assumption on the functional form of the firm production function. This means that, in our model, the relation between project quality and project return that the size of the loan actually signals depends ultimately on the organizational complexity parameter, which is observable, and is independent of the functional form of the production function.

On the other hand, the Wilson construction of the game contracting, characterized by uninformed lenders moving first in a three-stage pure strategy game which designs a standard debt contract between lenders and borrowers, has allowed us to prove that our model generates a pooling equilibrium whenever the indifference curves of different quality types are convex in the predetermined and agreed loan payment. Thus, having the possibility of getting both pooling and separating equilibria, we are able to offer a richer and more robust framework for type-I credit rationing than M-R's one. In this paper we have proved that, in the case of prevailing internal economies of scale and/or transaction costs: (i) the feasible separating equilibria are suboptimal, since the low-quality projects (entrepreneurs) get their first-best contract but the high-quality projects get an amount of loans greater than the size they would have liked to; (ii) the feasible pooling equilibrium too is suboptimal, since the low-quality projects get a larger amount of loans while the high-quality projects get a lower amount of loans than their respective first-best levels. We have also proved that, in the case of prevailing external economies of scale and/or organizational costs: (iii) the feasible separating equilibria are suboptimal, since the low-quality projects (entrepreneurs) get their first-best contract but the high-quality projects get an amount of loans smaller than the amount they would have liked to; (iv) the feasible pooling equilibrium too is suboptimal, since the low-quality projects get a smaller amount of loans while the high-quality projects get a higher amount of loans than their respective first-best levels.

Hence, both (i) and (iii) imply type-I credit rationing for high-quality firms, while (iv) implies type-I credit rationing for low-quality firms. Moreover, (ii) and (iv) emphasize the possibility of type-I credit rationing with pooling equilibria, while (i) points out the possibility of type-I credit rationing with separating equilibria. As mentioned above, the debt contracts which lead to this rich menu of possible credit rationing equilibria have the standard form.

Yet, differently from S-W and M-R which take this form of contract as given, our approach allows to use Innes's results, whereby the optimality of the form of the debt contract is endogenized. For Innes (1993) has proved that, under asymmetric information and variable loan sizes, the monotonicity of the lender's payoff function implies that the standard debt contract is the optimal contract. Hence, a further element strengthening our treatment of credit rationing is that this analysis is not based on the arbitrary assumption of an otpimal form of debt contract taken as given.

Despite these achievements, this paper is still constrained by several simplifications, the main being that firms demand loans to finance the working capital needed to activate the production process. No role is left for fixed capital accumulation (but this is a drawback of the whole literature on the topic, for that matter), as there is no fixed capital as such in our model. Moreover, no debt accumulation is allowed, given the strict one-period horizon of the borrowing problem faced by the firm. Further research is called for at this stage, to analyze the problems relating to capital accumulation through debt (and equity) financing under asymmetric information in the capital markets. This will certainly be the next step in our research agenda. By and large, the appraisals of fixed capital financing and of debt financing appear to be the necessary steps toward a more comprehensive analysis of the importance of credit with respect to other financial instruments in capitalist economies.

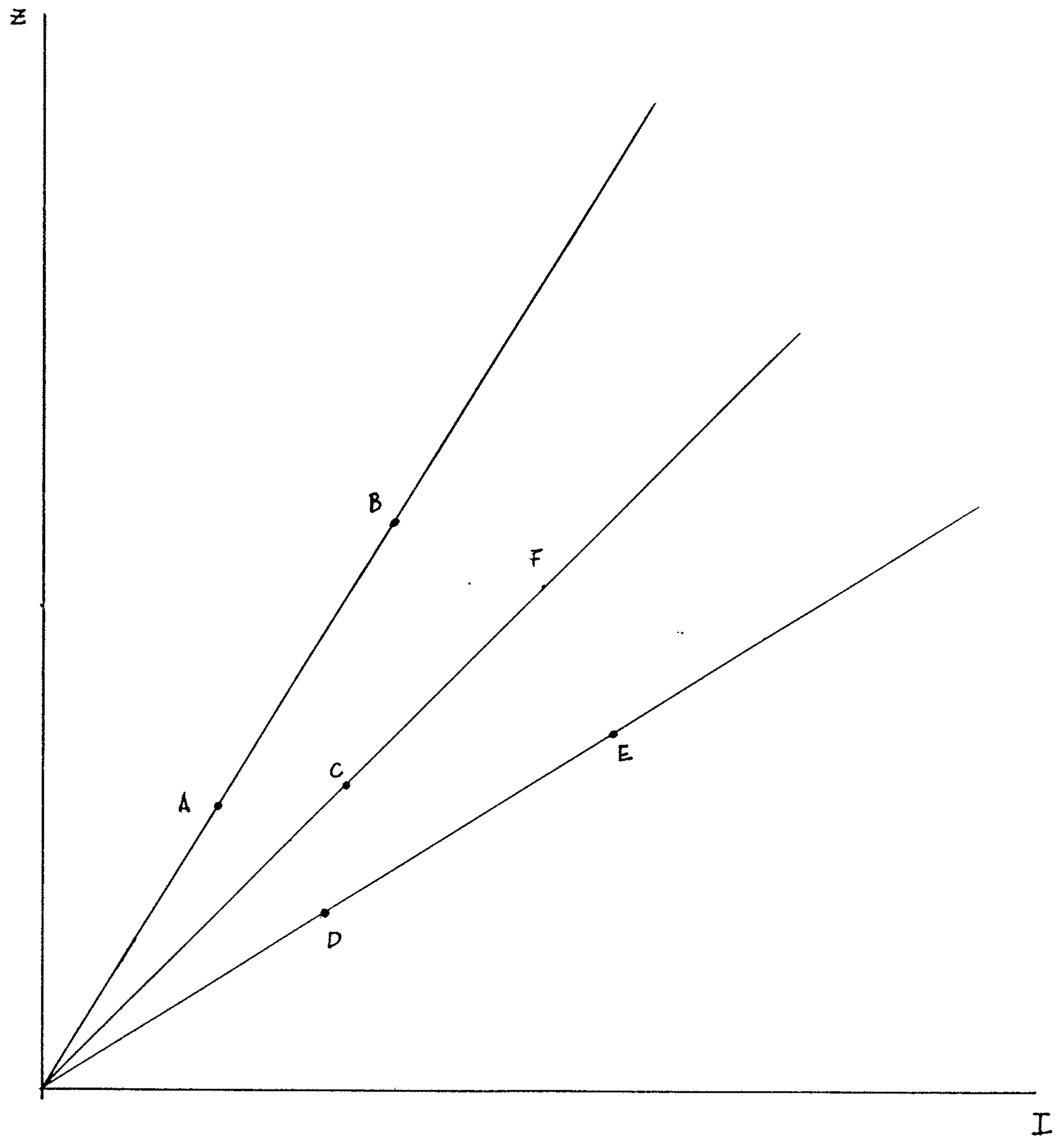

FIG. 4

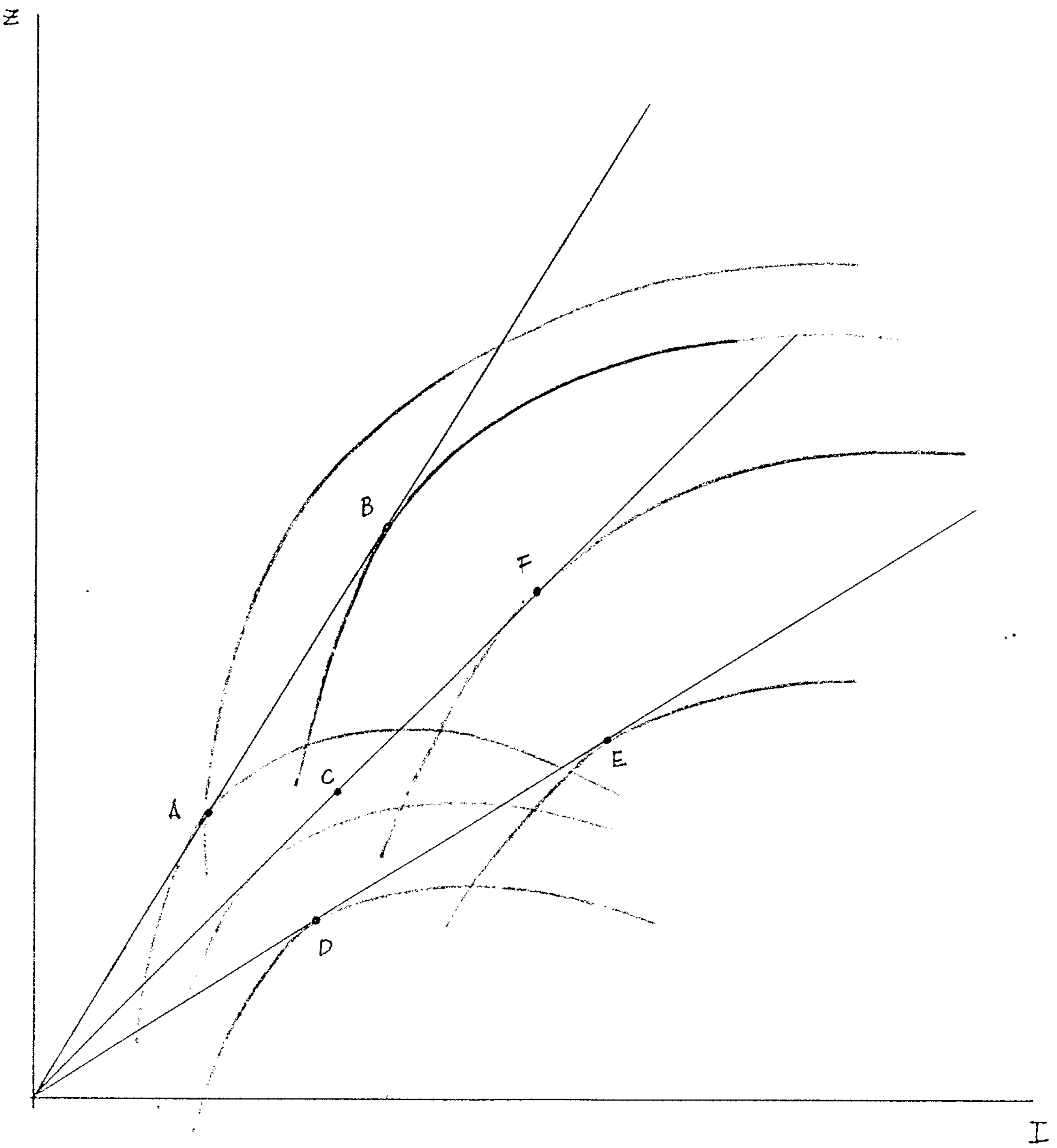

FIG. 2

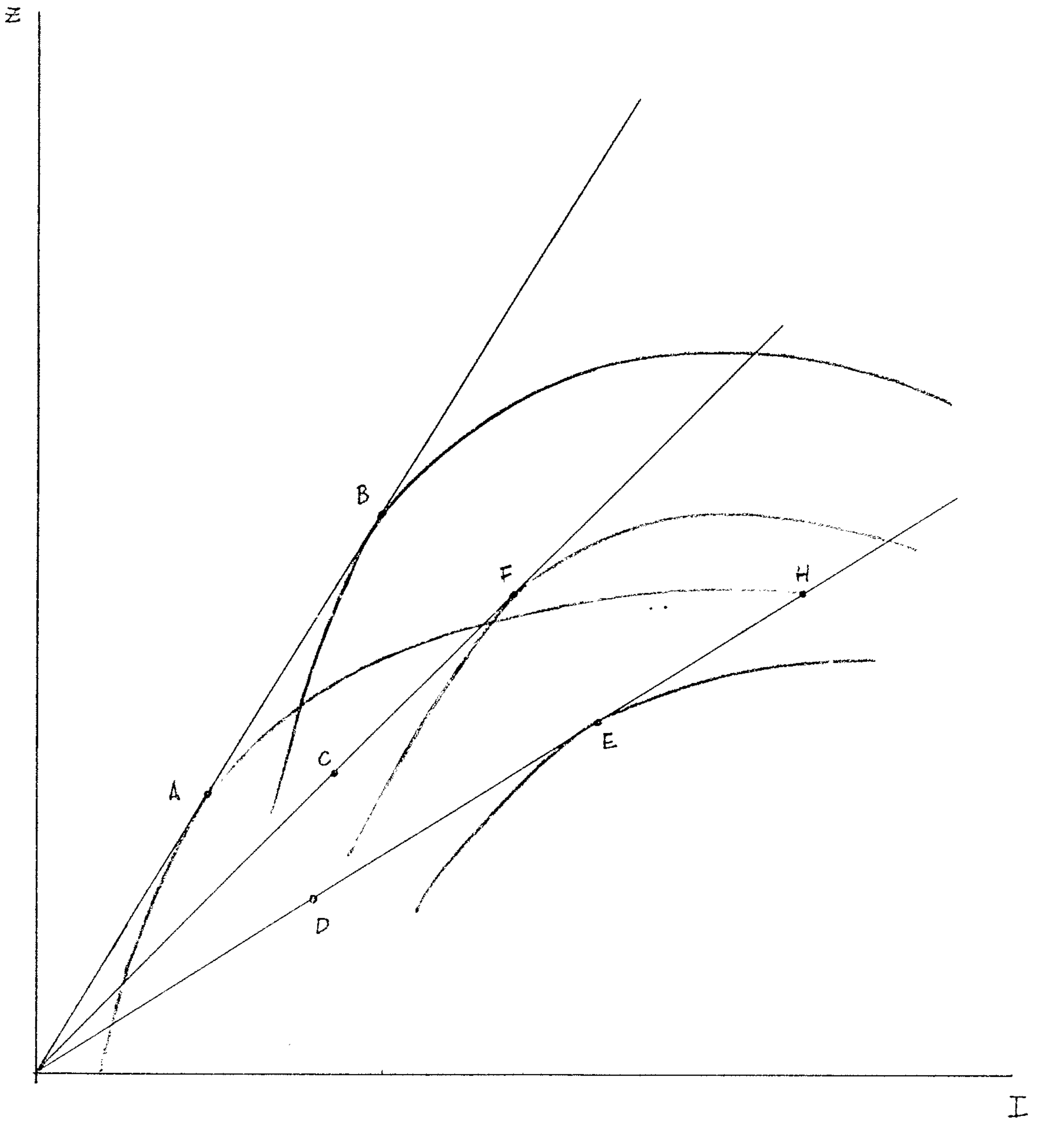

FIG. 3

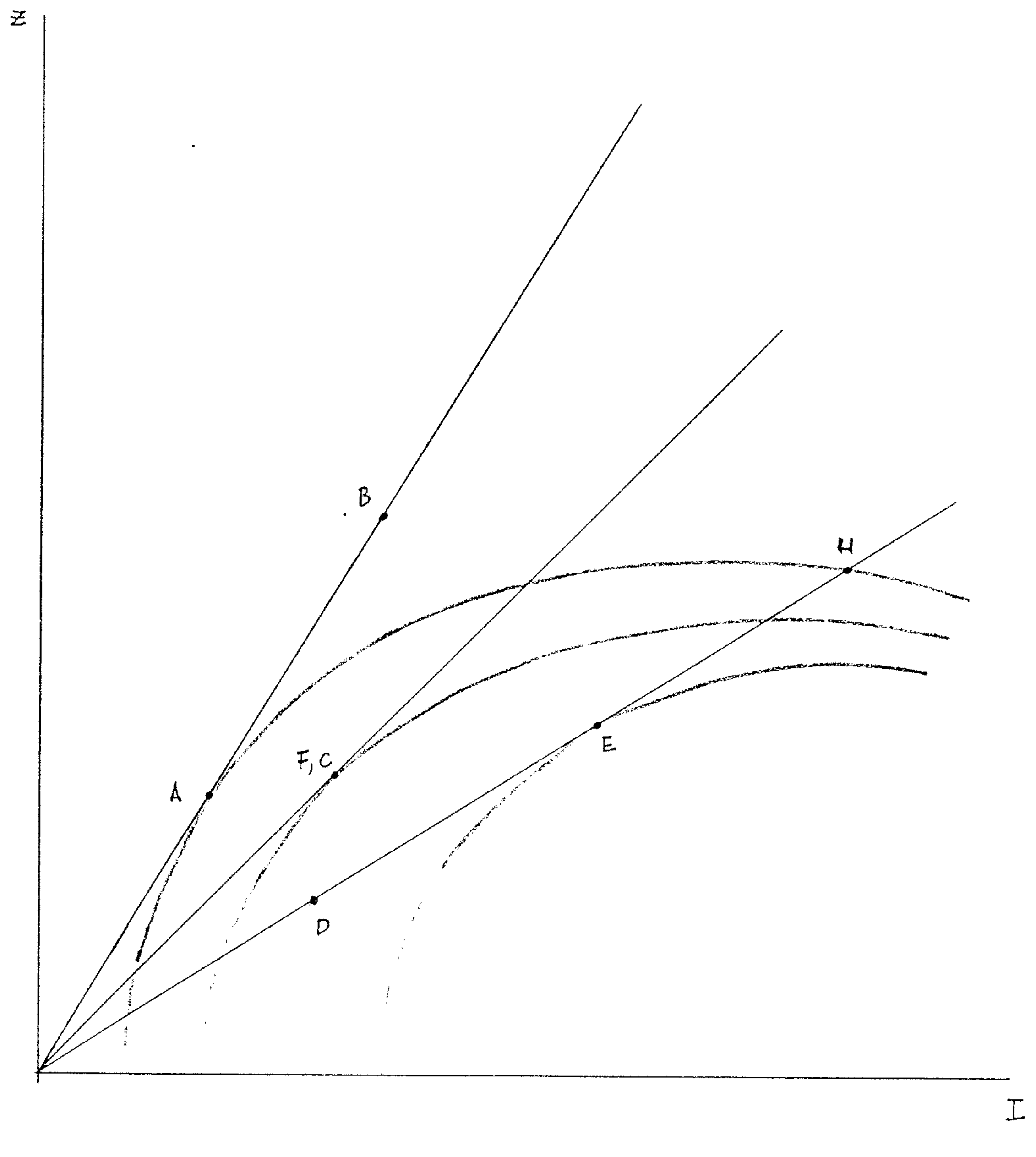

FIG. 4

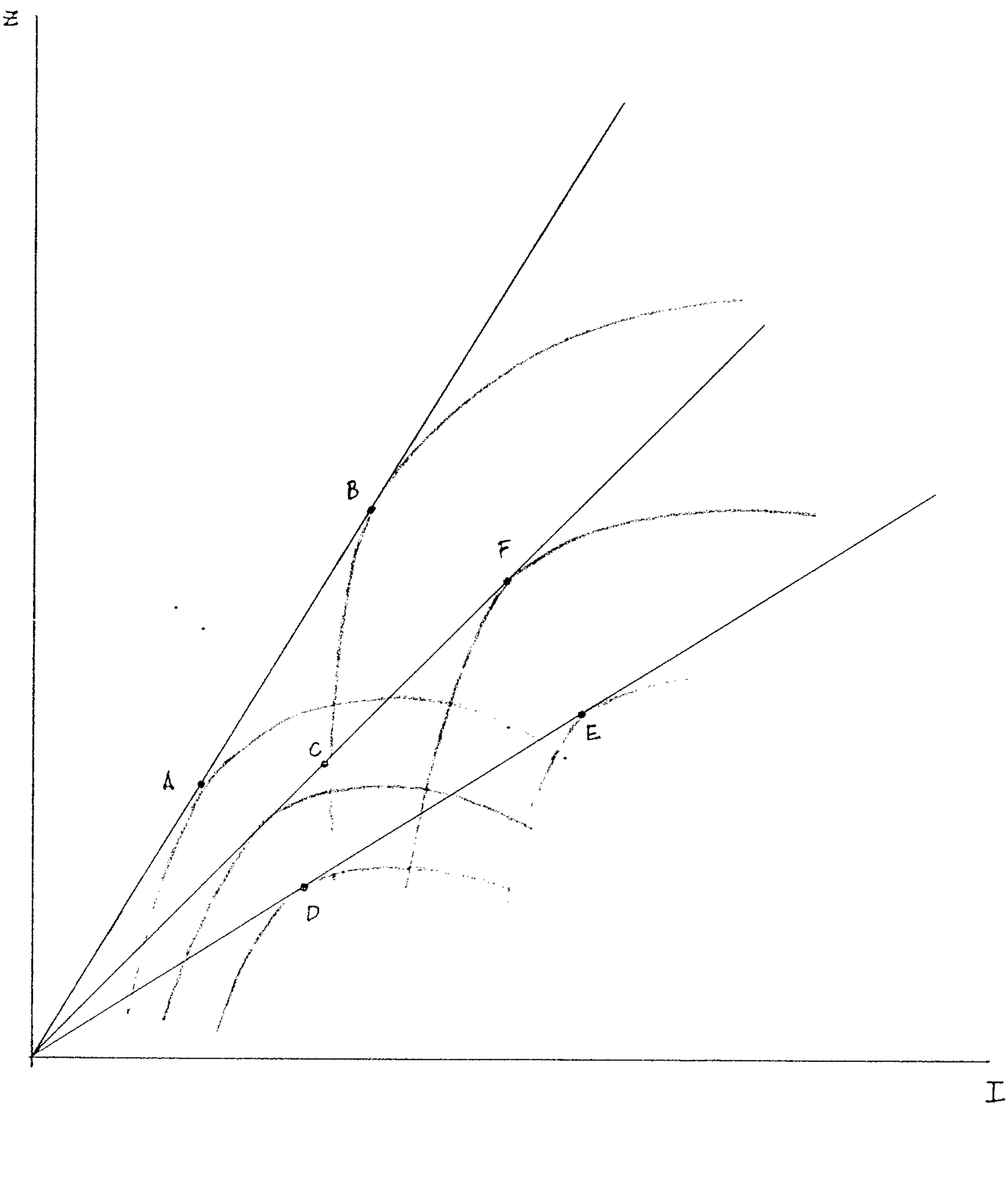

FIG. 5

References.

Besanko, D. and A. Thakor, 1987, "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets", *International Economic Review* 28: 671-79.

Bester, H., 1985, "Screening Versus Rationing in Credit Markets With Imperfect Information", *American Economic Review* 75: 850-5.

Chan, Y. and G. Kanatas, 1985, "Asymmetric Valuation and the Role of Collateral in Loan Agreements", *Journal of Money, Credit, and Banking* 17: 84-95.

Cho, I. and D. Kreps, 1987, "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics* 102: 179-222.

De Meza, D. and D. Webb, 1987, "Too Much Investment: A Problem of Asymmetric Information", *Quarterly Journal of Economics* 102: 281-92.

Hellwig, M., 1987, "Some Recent Developments in the Theory of Competition in Markets With Adverse Selection", *European Economic Review* 31: 319-325.

Innes, R., 1990, "Limited Liability and Incentive Contracting With Ex-Ante Action Choices", *Journal of Economic Theory* 29: 15-33.

Leland, H. and D. Pyle, 1977, "Informational Asymmetries, Financial Structure and Financial Intermediation", *Journal of Finance* 32: 371-87.

Mankiw, G., 1986, "The Allocation of Credit and Financial Collapse", *Quarterly Journal of Economics* 92: 71-93.

Milde, H. and J. Riley, 1988, "Signaling in Credit Markets", *Quarterly Journal of Economics* 103: 101-29.

Milgrom, P., 1981, "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics* 8: 394-416.

Miyazaki, H., 1977, "The Rat Race and Internal Labor Markets", *Bell Journal of Economics* 8: 394-416.

Riley, J., 1979, "Informationa Equilibrium", *Econometrica* 47: 331-59.

Shavell, S., "Risk Sharing and Incentives in the Principal and Agent Relationship", *Bell Journal of Economics* 10: 55-73.

Stiglitz, J. and A. Weiss, 1981, "Credit Rationing in Markets With Imperfect Information", *American Economic Review* 71: 393-410.

Stiglitz, J. and A. Weiss, 1986, "Credit Rationing and Collateral", in: J. Edwards, J. Franks, C. Mayer, and S. Schaefer, *Recent Developments in Corporate Finance*, Cambridge: Cambridge University Press.

Wilson, C., 1977, "A Model of Insurance Markets With Incomplete Information", *Journal of Economic Theory* 16: 167-87.