

Loan Size and Credit Market Equilibria Under Asymmetric Information*

Pier Giorgio Ardeni

Dipartimento di Scienze Economiche
University of Bologna

Marcello Messori

Dipartimento di Economia e Istituzioni
University of Roma "Tor Vergata"

Abstract

This paper analyzes the conditions under which optimal or sub-optimal equilibria obtain in a competitive credit market of risk-neutral optimizing agents where firms borrow from banks and information is asymmetric. Each firm has a heterogeneous endowment of *quality* and of a production *technology*, whose input can be financed by means of a bank loan only. The firm's output, which is increasing in the amount of the loan and is indexed by firm quality and technology, depends on an individual random component, thus varying across firms. As quality is unobservable by banks, an *ex ante* asymmetric information problem arises on the side of the lenders, which prevents banks from designing debt contracts contingent on firms' quality, and raises adverse selection effects. We assume that firm's technology can either be observable or not by lenders. In the former case, the adverse selection effect is due to the unobservability of quality only, while in the latter it is due to the combined interaction of both unobservable quality and technology. Credit market equilibria with loans of variable size and firm signaling can emerge, with either separating or pooling contracts, and are characterized by a rich set of over- and under-financing equilibria, whereby under-financing often implies credit rationing equilibria.

Mailing address:

Professor Pier Giorgio Ardeni
Dipartimento di Scienze Economiche
Università degli Studi di Bologna
Strada Maggiore, 45
40125 Bologna - Italy
Tel. +39-51-640 2649
Fax +39-51-640 2664
e-mail: "ardeni@boph01.cineca.it"

* This is a revised version of a paper that has circulated under the same title and that has been delivered at the IV International Financial Conference on *Asymmetric Information, Risk Management, Financial and Banking Innovation* held at the University of Rome "Tor Vergata" on Nov. 29 - Dec. 2, 1995. The paper draws on early works by the two authors, particularly: Ardeni, P.G. (1992), "Investment and Debt Contracting Under Asymmetric Information. I: The Static Case", Working Paper, Istituto di Scienze Economiche, Università di Urbino, *Quaderni di Economia, Matematica e Statistica*, no. 15, February; Ardeni, P.G., Messori, M. (1994), "Loan Size and Credit Rationing Under Asymmetric Information", Working Paper, Dipartimento di Scienze Economiche, Università di Bologna, *Collana Rapporti Scientifici*, no. 195, April. We thank all the people who have provided us with their comments and suggestions, as well as all the participants to conferences and seminars where our works have been presented. Financial support from CNR and MURST is acknowledged.

1. Introduction.

The aim of this paper is to analyze the conditions under which optimal or sub-optimal equilibria obtain in a competitive credit market of risk-neutral optimizing agents where firms borrow from banks and information is asymmetric. Each firm has a heterogeneous endowment of *quality* and of a production *technology*, whose input can be financed by means of a bank loan only. The firm's output, which is increasing in the amount of the loan and is indexed by firm quality and technology, depends on an individual random component, thus varying across firms. As quality is unobservable by banks, an *ex ante* asymmetric information problem arises on the side of the lenders, which prevents banks from designing debt contracts contingent on firms' quality, and raises adverse selection effects. We assume that firm's technology can either be observable or not by lenders. In the former case, the adverse selection effect is due to the unobservability of quality only, while in the latter it is due to the combined interaction of both unobservable quality and technology. We do not consider either moral hazard or *ex post* asymmetric information problems: each firm cannot change its quality or technological type after having signed a debt contract, and banks observe firms' end-of-period returns at no cost. Credit market equilibria with loans of variable size and firm signaling can emerge, with either separating or pooling contracts, and are characterized by a rich set of over- and under-financing equilibria, whereby under-financing often implies credit rationing equilibria.

Several models of credit market equilibria with asymmetric information 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 (1992), De Meza and Webb (1987) (D-W), Milde and Riley (1988) (M-R), Innes (1991). S-W analyzed banks' financing of loans of fixed size with equal expected firm returns but different riskiness across firms in accordance with the *mean preserving spread* principle (Rothschild and Stiglitz (1970)), and obtained pooling equilibria with type-II credit rationing. D-W too referred to the case of fixed size loans for firms of different quality. They considered both the case of loans to firms whose expected returns are equal and actual returns differ, which replicated the results reached by S-W, and the case of loans to firms whose returns are equal and the probability of success are different, which leads to a pooling equilibrium with over-financing of higher quality firms. Conversely, M-R and Innes (1991) considered loans of variable size and standard debt contracts contingent on realized firm returns. In their models, the variable size of the loan allows for borrowers' signaling, thus generating self-selection and separating equilibria for firms of different but unobservable quality. While Innes (1991) also obtains pooling equilibria with over- and under-financing, both signaling and separating equilibria weaken the occurrence of type-I credit rationing.

As the latter models highlight, when the size of the loan is allowed to vary, differences in firms' returns can be due both to differences in firms' quality and to differences in the loan size. Moreover, the variable loan size can imply the non-existence of Nash equilibria and the consequent reference to other concepts of equilibrium (e.g. Riley or Wilson equilibria). In M-R, for instance, each firm is endowed with a production function which depends on the amount of the loan and on a firm-specific quality parameter unobservable by lenders, and firms' returns are increasing in quality. The borrowers' signaling stems from

the relation between quality, loan size and output. M-R depict three different cases based either on different production technologies, which are known to both the firms and the banks, or on different orderings of the random returns. In the first case, the production function is multiplicative in the quality parameter so that the loan size is increasing in quality, and hence higher quality firms can signal their type by demanding larger loans; in the second case, the production function is additive in the quality parameter so that the loan size is invariant to quality, and hence higher quality firms can signal their type by accepting smaller loans¹; in the third case, firms projects have different riskiness but the same expected return, for any given amount of loan, so that lower quality firms with higher returns in case of success require larger loans. Under a Riley equilibrium construction, given asymmetric information, the latter two cases imply under-financing whereas the former case implies over-financing of firms of higher quality, and separating second-best equilibria.

In this paper we take two analytical steps forward with respect to M-R and Innes. First, we treat M-R's three separate cases within a general and unified setting in which firms' returns depend not only on an unobservable quality parameter but also on a *technological* one, and in which the Wilson equilibrium construction replaces Riley's one. More specifically, we argue that different ways of organizing production activities can have different implications in terms of firms' quality and can lead, in a three-stage Wilson contracting game, to a set of equilibria which is richer than those deriving from M-R's cases. Secondly, we do not resort to the strong informational assumption that each firm's technological parameter is always observable by banks, thus obtaining a new set of possible equilibria, as opposed to the set of equilibria that arise when technology is observable (like in M-R and Innes). In both cases we reach under-financing equilibria and type-I credit rationing which are stronger than those obtained by M-R.

For this purpose, as quality is not observable, we assume that firm realized returns have the *monotone likelihood ratio property* (MLRP) with respect to *quality*, in the sense of Milgrom (1981). The rationale for assuming this MLRP for the firm profit densities is that of exploiting the monotonicity property implied by the definition of quality. In the literature on credit rationing, two different quality definitions are used. With the first definition, 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. However, these two definitions of quality imply different relations between quality and the amount of loans, so that, as M-R have shown in their models, we have to specify these relations on an a-priori basis. In this paper we argue that this basis can be given by the way we look at how firms organize their production activities. Hence, we say that if *internal* economies of scale and/or *transaction* costs prevail, a larger size of the loan demanded will signal a higher firm quality, while if *external* economies of scale and/or *organizational* costs prevail a larger size of the loan demanded will signal a lower firm quality.

¹ Innes (1991), too, refers to these first two cases which, however, are stated as assumption and not derived from specific production functions (he only models the first case, anyway). Moreover, differently from M-R, Innes (1991) does not refer to a Riley but to a Wilson equilibrium, as we will see below. Innes (1990, 1993) contribution is important with respect to the determination of the optimal form of the debt contract. Even if this is not at issue in our paper, we will be able to replicate Innes' results by using the monotone likelihood ratio property (see below).

The MLRP tool has a crucial role to play also in the building up of the model when technology is assumed to be non-observable by banks. In this case, we do not simply assume that firm realized profits have the MLRP with respect to quality but we assume that the loan size has the MLRP with respect to quality *and* technology together. This assumption implies that banks are able to order all loan applicants, which have unobservable quality and productive organization, on the basis of their *marginal rate of substitution* (MRS) between the loan size and the interest rate factor. In particular, banks know that the high-quality firms with prevailing internal economies of scale have the greatest MRS, followed by the low-quality firms with prevailing external economies of scale, the low-quality firms with prevailing internal economies of scale and, finally, by the high-quality firms with prevailing external economies of scale.

The paper is organized as follows. In Sections 2A and 2B, we characterize the main features of borrowers and lenders in order to specify the debt contract. This last specification requires the definition of a three-stage pure strategy game between banks and entrepreneurs which leads to Wilson equilibria (Section 2C). Together with the entrepreneurs' indifference curves and the banks iso-profit curves, we can thus specify the set of contracts designed by banks and chosen by entrepreneurs (Section 2D). The introduction of the assumption that firm realized profits have the MLRP with respect to quality completes the setting of the model by defining the possible relations between the technological parameter, the quality of the firm and the size of its demand for loans (Section 2E). This setting is sufficient to determine the set of equilibria which can obtain when the technological parameter is observable by banks and when either internal or external economies of scale prevail (Section 3A). When the technological parameter is firms' private knowledge, the assumption of the MLRP with respect to quality and technology together is required. The previous setting of the model and this new MLRP determine the set of equilibria which can obtain when banks do not observe the technological parameter and make it possible to compare these equilibria with the previous ones (Section 3B). In the final Section we sum up the main results of the paper and we suggest how further research avenues could address a number of problems still unsolved.

2. The setting of the model.

A. *The Firms.*

We consider a competitive market for a single homogeneous good and an arbitrarily large population of risk neutral entrepreneurs with a null individual wealth endowment, so that the production of each firm must be financed by outside sources, i.e. bank *loans*. Each firm produces the level of output π , expressed in nominal terms, depending on the amount of "working capital" invested, L (the amount of loan received), which is expressed in nominal terms also². For every given L , the output π (the firm return)

² We could refer to a population of entrepreneurs with positive wealth endowments, so that L would become the amount of firm's input financed through external funds. Since no debt contracting with outside private investors is allowed in this model, 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, signaling through wealth (as well as through any other instrument, like e.g. collaterals) would still be impossible, and the assumption of positive wealth would not affect our conclusions.

depends on the *quality* of the firm. Each entrepreneur knows the “quality” of his firm³, which is unobservable by lenders, thus giving rise to an adverse-selection problem on the side of the lenders. Here we assume that entrepreneurs have a heterogeneous endowment of quality, either *low* or *high*, so that the *quality* parameter $q \in [q_l, q_h]$. The relative frequency of the two different quality types is such that $\text{PROB}(q = q_h) = p$, and $\text{PROB}(q = q_l) = 1 - p$, and it is common knowledge.

As M-R have shown, when the size of the loan is allowed to vary, differences in the level of output can be due both to differences in quality and to differences in the amount of loan received. Depending on the relation between quality, loan size and output, a larger amount of loan demanded does not necessarily indicate better quality: the amount of “working capital” used can be either increasing or decreasing in quality, so that firms of better quality will be willing to accept either larger or smaller changes in loans than those accepted by firms of lower quality, for given changes in the interest rate⁴. In this paper we aim to generalize M-R’s approach by couching their two different (and unrelated) cases based on different production functions within a unitary analytical framework. This can be achieved by considering that the relation between loan size and quality ultimately depends on the “technological” or “organizational” setting of the firm. In particular, we can say that, if *internal* economies of scale and/or *transaction* or *bargaining* costs prevail, then the larger the size of the loan a firm is willing to accept the higher the firm quality, while if *external* economies of scale and/or *organizational* or *influence* costs prevail, then the larger the size of the loan a firm is willing to accept the lower the firm quality. Accordingly, the relation between quality and output from a loan is either positive or negative⁵.

Obviously, there can be many different firm technological or organizational “settings”, each with a varying degree of *internalization* or *externalization* of economies of scale. Here we assume, for the sake of simplicity, that all the possible different settings are represented by a finite number of categories, each belonging to one of the two disjoint sets $[I, E]$. Hence, we define a parameter H , such that when *internal economies of scale and/or transaction costs* tend to prevail, then $H=I$, whereas when *external*

³ In the following we use “firm quality” and “entrepreneur quality” interchangeably: each firm is owned by only one entrepreneur.

⁴ It must be recalled that M-R presented two different models, in which the firm production function has either a multiplicative form or an additive one (with respect to the quality parameter). In both cases output is increasing in loan size. However, the former model implies that a given loan size will give rise to a higher level of output for firms of better quality, so that these firms will be willing to accept loans larger than those accepted by firms of lower quality. Conversely, the latter model implies that an increase in the loan size has a marginal effect on output which is invariant to quality, so that firms of lower quality will be willing to accept loans larger than those accepted by firms of better quality due to their higher default probability.

⁵ What we refer to here as *technology* are simply the different ways of organizing production activities, which naturally have different implications in terms of quality, and on the ways quality affects the final output. An example may be clarifying. Consider two ideal and opposite firms: firm A&A is vertically integrated and organized in various departments so that its intermediate output flows into one or more final goods by passing through several internal production phases; firm B&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 (1988)), the organization of firm A&A is substituting entrepreneurial co-ordination and hierarchical relations for market exchanges, whereas the organization of firm B&B gives large room to market exchanges. The rationale for these different institutional settings is that – for every given loan size – firm A&A has *large internal economies of scale*, *high transaction* (or bargaining) *costs*, and *small organizational* (or influence) *costs*, whereas firm B&B has *large external economies of scale*, *small transaction* (or bargaining) *costs* and *high organizational* (or influence) *costs*. This points out that the size of the loan for firm A&A will be increasing in its “quality”, while the size of the loan for firm B&B will not.

economies of scale and/or organizational costs tend to prevail, then $H=E$. The implications of this apparently “heroic” assumption can be weakened by considering that every shift from E to I increases the firm’s degree of *scale economy internalization*. In any case, by parameterizing the relations between loan size and firm quality, we can unambiguously state that, given the banks’ supply function, when $H=I$, output and quality will be positively related, so that high-quality entrepreneurs will be willing to choose larger loans than low-quality entrepreneurs; whereas, when $H=E$, output and quality will be unrelated, so that high-quality entrepreneurs will be willing to choose smaller loans relatively to low-quality entrepreneurs due to their different default probabilities. Thus, *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*. This implies that the non-random component of the relation between the amount of “working capital” L (the loan of size L) and the firm output, π , is not unique, depending on two parameters: the *technological* parameter, H , and the *quality* parameter, q . In what follows, we will first assume that while only the entrepreneur knows the quality of his firm, the firm’s technological and organizational setting is publicly observable⁶. When observable, in fact, banks face two separate groups of applicants, and are able to elicit either separating or pooling equilibria between low and high-quality firms within each of these two groups; while when not observable, all applicants are treated as identical, and separating or pooling equilibria among all the four “types” will prevail.

The end-of-period return for each firm, π , i.e. nominal output “sold in the market”, has a stochastic component due to a random shock θ (e.g. a demand shock) which is i.i.d. across firms and has a positive support on $[0, \bar{\theta}]$. Hence, for each firm, $\pi = \pi(L, H, q, \theta)$. Given L , q and H , θ gives rise to probability density and distribution functions for π , $f(\pi | L, q, H)$ and $F(\pi | L, q, H)$, respectively. The function $f(\cdot)$ is assumed to be continuously differentiable on $[0, K(L, q, H)]$, where $K(\cdot) > 0$ whenever $L > 0$. In addition, we assume that larger loan sizes produce “better” net worth distributions in the sense of first-order stochastic dominance:

ASSUMPTION⁷. For all $\pi \in [0, K(L, q, H)]$:

$$(AA) \quad F_L(\pi) \leq 0.$$

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

⁶ It should not be terribly controversial to assume that the characteristics of each firm that we define as “technological” or “organizational” can be *observable* by outside lenders, in the sense that they may be gathered from the known “identity card” of the borrowers (industrial sector, market structure, firm size, number of employees, internal organization, and so forth) which is obviously common knowledge. In fact, we may think of the *technological setting* as, for instance, the degree of vertical integration, the size of the firm (in terms of plants and machinery), the number of people employed, and so on: something which is, in any case, publicly observable. Conversely, we may think of *quality* as, for instance, the managerial capacity of entrepreneurs, which is not observable from outside. Obviously, such two features of the production process strictly interact, in the Coasian spirit (the production process is something more than the production function). In the second part of the paper, however, we will relax the assumption that the “technological setting” of the firm be observable by banks.

⁷ Subscripts denote partial derivatives.

$$(A) \quad \bar{\pi}(L|q, H) \equiv \int_0^{K(L,q,H)} \pi f(\pi(L|q, H)) d\pi,$$

then condition (aa) implies that $\bar{\pi}_L(L|q, H) > 0$. We will also assume that $\bar{\pi}_{LL} < 0$, $\bar{\pi}_L(L|q, H) = \infty$ as $L \rightarrow 0$ and $\bar{\pi}_L(L|q, H) = 0$ as $L \rightarrow \infty$. The latter constraints ensure positive and finite equilibrium loan sizes for any quality and technology type.

B. The Banks.

Banks are assumed to be competitive and risk neutral. On each loan, banks require an expected unit return of at least $(1+\rho)$, where ρ is the return on a risk-free bond⁸, that is, banks 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 observable π , as specified in the financial contract defined below. Perfect competition in the loan market ensures that each bank satisfies a zero-profit condition.

C. The Debt Contract and the Definition of Equilibrium.

The financial contract signed by the entrepreneur with the bank prescribes the size of the loan, L , 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 and technology types, it does not know on an *ex-ante* basis “who is who”, as information about firm quality, and possibly about technology, is asymmetric. Banks cannot observe *ex-post* neither the “state of nature”, θ , nor the firm’s quality q and possibly the firm’s technology H ; they can only observe the firm’s return, π , (possibly H as well), at no cost of monitoring. Thus, in our model, the financial contract will set, beside L and independently of the optimality of the form of the contract, the bank’s payoff $B(\pi)$, i.e. the amount to be paid back by the firm; in other words, it will be a *debt contract*⁹. The bank’s payoff $B(\pi)$ is constrained by π since limited liability implies that $B(\pi) \leq \pi$, that is, firms cannot be compelled to pay more than the available return¹⁰.

It has already been shown (Innes (1993)) that, with a fixed and homogeneous loan size across firms, a Nash equilibrium exists. A set of debt contracts is a Nash equilibrium if all contracts in the set earn banks an expected unit return of at least $(1+\rho)$, and there is no other set of contracts that, when offered in addition to the equilibrium set, earn banks an expected unit return greater than $(1+\rho)$. In that case, each quality-type class of borrowers subscribes a single *pooling* contract which leads to a *pooling* equilibrium. Normally (e.g. S-W (1981)) it is exogenously assumed that this contract takes a *standard* debt form,

⁸ We simply assume that banks pay depositors the “safe” rate of interest ρ .

⁹ Hence, such a contract will be contingent on the firm’s technology, besides output, when technology is observable by banks.

¹⁰ 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, like DeMeza and Webb (1987) (see below, equation (bb)).

namely, $B(\pi) = \min(\pi, z)$, with z representing the *fixed and predetermined loan payment*¹¹. The problem is that, when the 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 the competitors (i.e. competing banks) to his actions. More recently, they have obtained game-theoretic foundations by Hellwig (1987) and Cho and Kreps (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 $\{B(\pi | q, H), L(q, H)\}$ 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 of the contracts designed for higher-quality.

(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 profits.

(E3) *Non-negativity of lenders profits*: lenders earn an expected unit return of at least $(1+\rho)$ on each distinct contract.

The latter constraint characterizes a Wilson allocation¹². Given the game structure posited above, the resulting equilibrium is thus a *Wilson equilibrium*. 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: Riley's equilibrium elicits complete separation, whereby all agents reveal their type by their contract choice, even when pooling equilibria would dominate separating equilibria; 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 when the Wilson allocation is pooling. Most importantly, in the latter case, the Wilson allocation Pareto dominates the Riley equilibrium. This means that

¹¹ It must be noted that Innes (1990) specifies under which conditions the optimal form of any fixed loan size contract is the standard debt form.

¹² If (E3) was replaced with the assumption that banks earn an expected unit return of at least $(1+\rho)$ 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 $(1+\rho)$ on *each entrepreneur*, that would be a Riley allocation (see also below). This means that a Riley allocation requires the informed borrowers move first, i.e. at Stage 1 of the game.

the reference to Riley's concept of equilibrium eliminates, by construction, the most favourable cases of credit rationing¹³. Hence, as we are mainly interested in analyzing the possibilities of credit rationing, the Wilson construction seems more promising¹⁴.

As shown by Innes (1993), under asymmetric information the optimal form of any variable loan size contract is the standard debt form, provided the lender's payoff is monotonic. Thus, even for the Wilson equilibrium, the bank's payoff function must be here compatible with the *standard* debt form, namely $B(\pi) = \min(\pi, z)$, with $z=z(q,H)$ representing the *promised loan payment*, which depends on the firm quality and technology parameters q and H , respectively. 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 (z, L) .

D. Choices of the loan size.

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

$$\begin{aligned} \text{(B)} \quad & \max_{(L)} V(L|q, H) = E\{\pi - B(\pi)|q, H\} \\ \text{(ba)} \quad & \text{s. t. } E\{B(\pi)|q, H\} \geq (1+\rho)L \\ \text{(bb)} \quad & B(\pi) \leq \pi, \quad B_{\pi}(\pi) \geq 0 \quad \forall \pi \in [0, K(L, q, H)]. \end{aligned}$$

where V is the expected profit of the quality- q entrepreneur of a firm H for a loan of size L . Condition (ba) gives the minimum bank return requirement, while (bb) gives the limited liability and monotonicity constraints on the contract form. The entrepreneur's limited liability, together with the possibility of zero-output levels, make any contract with payments that are invariant to firm's returns (including the fixed-payment contract), unfeasible¹⁵. Thus, in (b), $B(\pi) = \min(\pi, z)$.

The expected profit V of the quality- q entrepreneur of a firm H for a loan of size L is thus given by:

$$\text{(C)} \quad V(L|q, H) = E\left[\max\left[\pi(L, q, H) - z(q, H), 0\right]\right].$$

Equation (c) implies that, if the firm is "successful", then $V(\cdot)$ is equal to the positive difference between the firm return π and the promised loan payment z , whereas, if the firm is "unsuccessful", and then $\pi < z$, $V(\cdot)$ is

¹³ Several authors (e.g. Bester (1987)) maintain that separating equilibria eliminate the possibility of credit rationing. However, Besanko and Thakor (1987) have been the first to build a model with separating equilibria and credit rationing.

¹⁴ We maintain that the weak rationing results obtained by M-R largely depend on their choice of a Riley allocation. Wilson's allocation has been criticized because it implies passive reactions by competing agents. However, the game-theoretic foundation of this allocation is able to weaken this problem.

¹⁵ In the absence of the limited liability constraint, the solution to (b) would be the fixed-payment debt contract, $B(\cdot)=(1+\rho)L$, that yields "perfect information" choices of L (Shavell (1979)), for given q and H .

0. Since the entrepreneur is risk-neutral, he maximizes $V(L|q,H)$ in (b). Hence, given (a), (b) can be rewritten as:

$$(D) \quad \max_{(L)} V(L|q,H) = \int_{\pi_0}^{K(L,q,H)} [\pi(L,q,H) - z(q,H)] f(\pi|L,q,H) d\pi,$$

where π_0 satisfies:

$$(E) \quad \pi_0 - z(q,H) = 0.$$

Basically, $V(\cdot)$ is the profit of a firm when successful times the probability of a “success”¹⁶. The probability of a “success” is given by the integral of the density function of π for $\pi \in [\pi_0, K(L,q,H)]$ ¹⁷. In order to solve equation (d), it is necessary to specify condition (ba) above, i.e. banks expected profits. Banks expected profits are defined as:

$$(F) \quad R(L,z|q,H) = E[\min(\pi,z)|L,q,H] - (1+\rho)L,$$

that is, as:

$$(G) \quad R(L,z|q,H) = z(q,H) - (1+r)L + \int_0^{p^*} [p(L,q,H) - z(q,H)] f(p|L,q,H) dp.$$

Substituting (g) for (ba) and (d) for (b), we can formally solve the choice problem of the entrepreneur of quality q and technology H . The easiest way to handle this problem is to refer to entrepreneurs’ *indifference* curves and banks’ *supply* curves.

An entrepreneur’s *indifference curve* is a set of (z, L) contracts yielding a common expected profit level¹⁸. Since entrepreneurs are always better off with a lower promised payment, z , given the loan size, L , lower indifference curves correspond to higher expected profit levels for firms. Formally, entrepreneurs’ indifference curves are defined as:

$$(H) \quad V(L,z|q,H) = \int_z^{K(L,q,H)} [\pi(L,q,H) - z(q,H)] f(\pi|L,q,H) d\pi = \bar{V}(q,H),$$

a constant. Regardless of firm quality and technology, indifference curves are *upward-sloping*, as:

¹⁶ Thus $z(q|\pi \geq \pi_0) > (1+\rho)L > \pi(q|\pi \leq \pi_0)$. De Meza and Webb (1987) define R^s as the return of a successful project and R^f as the return of an unsuccessful project. They thus assume that $R^s > (1+\rho)L > R^f$.

¹⁷ As we have seen above, the density function of π , given L, q and H , depends on θ , which is i.i.d. across firms. Hence (d) could also be written as:

$$\max_{(L)} V(L|q,H) = \int_{\theta_0}^{K(L,q,H)} [\pi(L,q,H,\theta) - z(q,H)] f(\pi(\theta)|L,q,H) d\theta,$$

where θ_0 is such that:

$$\pi(L, q, H, \theta_0) = z(q, H).$$

¹⁸ Since we assumed that both banks and entrepreneurs are risk neutral, we can obviously refer to “indifference” or “iso-profit” curves. Indifference curves (a more general concept) imply that the set of contracts must yield a common expected utility level.

$$(Ia) \quad \frac{\partial V(L, z|q, H)}{\partial L} \equiv - \int_z^{K(L, q, H)} F_L(\pi|L, q, H) d\pi > 0,$$

$$(ib) \quad \frac{\partial V(L, z|q, H)}{\partial z} \equiv -(1 - F(z|L, q, H)) < 0.$$

Intuitively, larger loan sizes, given z , increase an entrepreneur's expected profit. Hence, to preserve the latter profit, an increase in L must be accompanied by an increase in the debt obligation, z . Because entrepreneurs "like" any increase in L , and "dislike" any increase in payments to banks, z , indifference curves are upward-sloping (and lower curves correspond to higher firm profits)¹⁹.

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

$$(J) \quad R(L, z|q, H) = E[\min(\pi, z)|L, q, H] - (1 + \rho)L = \bar{R}(L, z|q, H).$$

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 profit²⁰. Thanks to perfect competition in the loan market, we can limit our attention to banks iso-profit curves from loans to quality- q entrepreneurs, which imply:

$$(K) \quad \bar{R}(L, z|q, H) = 0.$$

The latter iso-profit curves will coincide with banks' *supply* curves.

Whenever a bank is able to offer different contracts to borrowers of different quality, *separating* supply curves prevail, whereas if a bank cannot discriminate between borrowers of different quality, a *pooling* supply curve obtains.

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

$$(L) \quad E[\min(\pi, z)|L, q, H] = (1 + \rho)L,$$

that is, making use of (g):

$$(M) \quad z(q, H) - (1 + \rho)L + \int_0^z (\pi - z)f(\pi|L, q, H)d\pi = 0.$$

Using equations (a) and (h) we can rewrite this expression as:

$$(N) \quad \bar{\pi}(L|q, H) - V(L|q, H) - (1 + \rho)L = 0,$$

or:

¹⁹ Notice also that the second derivative with respect to L is negative, implying that indifference curves are concave.

²⁰ 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' iso-profit curves are not only upward sloping and convex, but their slope is decreasing with quality.

$$(Oa) \quad \bar{\pi}(L|q_l, H) - V(L|q_l, H) = (1 + \rho)L,$$

$$(ob) \quad \bar{\pi}(L|q_h, H) - V(L|q_h, H) = (1 + \rho)L.$$

Only by having a different supply curve for each quality type we can have firm signaling (self-selection).

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

$$(P) \quad \{E[\min(\pi, z)|L, H, q_l] - (1 + \rho)L\}(1 - p) + \{E[\min(\pi, z)|L, H, q_h] - (1 + \rho)L\}p = 0,$$

that is

$$(Q) \quad \left[z(q_l, H) - (1 + \rho)L + \int_0^z (\pi - z) f(\pi|L, q_l, H) d\pi \right] (1 - p) \\ + \left[z(q_h, H) - (1 + \rho)L + \int_0^z (\pi - z) f(\pi|L, q_h, H) d\pi \right] p = 0.$$

Using equations (oa,b) we can rewrite this as:

$$(R) \quad [\bar{\pi}(L|q_l, H) - V(L|q_l, H)](1 - p) + [\bar{\pi}(L|q_h, H) - V(L|q_h, H)]p = (1 + \rho)L.$$

This implies that the bank set of contracts offered will lie somewhere between the separating supply curve to the low-quality types and that to the high-quality ones. In other words, a bank pooling supply curve is the supply curve to the “average” quality- q entrepreneur, given H .

Each entrepreneur will solve his optimal choice problem by signing the debt contract which coincides with the tangency point between the bank supply curve and his lowest iso-profit curve.

E. The relation between quality, loan size and firm profits.

Given the “technological” parameter H , we can exploit a suitable assumption of monotonicity of the loan size with respect to quality by comparing signals in terms of relative favourableness, using the so-called *monotone likelihood ratio property* (MLRP) (Milgrom (1981)). The MLRP states that, if a signal x is *more favourable than* another signal y about a random parameter Φ , and if Φ takes on two particular values ϕ_1 and ϕ_2 , with $\phi_1 > \phi_2$, then

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

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 $U(\Phi)$ and nondegenerate prior distributions $G_1(\Phi)$ and $G_2(\Phi)$ ²¹.

In the case of prevailing internal economies of scale, i.e. $H=I$, as loan size and quality are positively related, the “better” of two firms of (unobservable) different quality yields a higher expected return for a given loan size: that is, the larger the size of the loan, the higher marginal returns so that firm

²¹ 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 $G(\cdot|x)$ dominates the posterior distribution $G(\cdot|y)$ in the sense of strict FOSD” (1981, p.382).

marginal returns are higher the higher the quality of the firm. In this case, we assume that higher quality produces “better” net worth distributions in the sense of statistical “good news” (Milgrom (1981)). Conversely, in the case of prevailing external economies of scale, i.e. $H=E$, entrepreneurs of different quality, to preserve their expected profits, can give up the *same amount of expected return* in exchange for one extra dollar of loan, and marginal returns are invariant to both the quality of the firm and the size of the loan. However, this given marginal transfer of expected returns must be associated with a larger increase in the promised loan payment z for lower quality types due to their higher default risk. This means that larger loan sizes signal lower entrepreneurs’ quality and that lower quality produces “better” net worth distributions in the sense of statistical “good news”. Hence we have the following assumption.

ASSUMPTION. For all $\pi \in [0, K(L, q, H)]$, if $H = I$, then:

$$(AB) \quad \frac{\partial \left(\frac{f_q(\pi|L, q, H)}{f(\pi|L, q, H)} \right)}{\partial \pi} > 0$$

while if $H = E$, then:

$$(AC) \quad \frac{\partial \left(\frac{f_q(\pi|L, q, H)}{f(\pi|L, q, H)} \right)}{\partial \pi} < 0.S$$

Each of the inequalities in (ab) and (ac) corresponds to the *monotone likelihood ratio property* (MLRP). Either assumptions (ab) and (ac) imply that the posterior distribution $G(\Phi | \pi_1)$ dominates the posterior distribution $G(\Phi | \pi_2)$ 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, when $H=I$ (assumption (ab)), larger loan sizes signal better quality and yield higher marginal returns, while when $H=E$ (assumption (ac)), larger loan sizes signal lower quality and smaller loans give rise to superior return distributions²². Given a stochastic environment which yields uncertain returns, the rationale for assuming a MLRP with respect to quality for firms profit densities is that of exploiting the monotonicity property implied by assumption (aa), so that higher loans will give rise to higher returns.

Consider first the case of prevailing internal economies of scale ($H=I$). Given (aa), condition (ab) implies that:

$$(T) \quad F_{Lq}(\pi) < 0,$$

where the negative cross-partial derivative shows that higher quality yields first order stochastically dominant returns from marginal loans, that is, higher quality is associated with “better” *marginal* returns.

Then:

²² This obviously implies that better quality needs smaller loans. Notice also that, when $H=I$, our results partially coincide with those obtained by M-R in the case where the firm production function has a multiplicative form (with respect to the quality parameter), whereas, when $H=E$, our results partially coincide with those obtained by M-R in the case where the firm production function has an additive form (see M-R (1988, pp. 111-113))

$$(U) \quad \frac{\partial}{\partial q} \left(\frac{\partial \pi(L|q, H)}{\partial L} \right) > 0.$$

This implies that the marginal rate of substitution between L and z is greater, and the steeper the indifference curve, the higher the entrepreneurs' quality. That is, in each and every point in the space (z, L) where the high- and the low-quality indifference curves cross, the former are steeper than the latter ones, i.e. $MRS(z, L|q_h) > MRS(z, L|q_l)$. Formally, (t), (ia), and (ib) together imply:

$$(V) \quad \frac{\partial dz}{\partial q dL} \Big|_{\bar{v}} = \frac{\partial}{\partial q} \left(\frac{\frac{\partial V(L, z|q, H)}{\partial L}}{\frac{\partial V(L, z|q, H)}{\partial z}} \right) > 0.$$

Similarly to M-R, we can restate this by saying that the marginal increase in promised loan payment that a borrower is willing to accept in order to receive a marginal increase in the loan size is greater the higher firm quality.

Consider now the case of prevailing external economies of scale ($H=E$). Given (aa), condition (ac) implies:

$$(W) \quad F_{Lq}(\pi) > 0,$$

where the positive cross-partial derivative now ensures that lower quality is associated with higher *marginal* project returns. Then:

$$(X) \quad \frac{\partial}{\partial q} \left(\frac{\partial \pi(L|q, H)}{\partial L} \right) < 0.$$

This implies that the marginal rate of substitution between L and z is greater, and the steeper the indifference curve, the lower the entrepreneurs' quality. That is, in each and every point in the space (z, L) where low-quality and high-quality indifference curves cross, the former are steeper than the latter curves, i.e. $MRS(z, L|q_h) < MRS(z, L|q_l)$ ²³. Formally, (w), (ia), and (ib) together imply:

$$(Y) \quad \frac{\partial dz}{\partial q dL} \Big|_{\bar{v}} = \frac{\partial}{\partial q} \left(\frac{\frac{\partial V(L, z|q, H)}{\partial L}}{\frac{\partial V(L, z|q, H)}{\partial z}} \right) < 0.$$

Similarly to M-R, we can restate this by saying that the marginal increase in promised loan payment that a borrower is willing to accept in order to receive a marginal increase in the loan size is greater the lower firm quality.

Now, from (ab) and (ac) and from the above definitions, it derives that, as the degree of internalization increases (decreases) -i.e. H "shifts" from E to I (from I to E), the derivative in (ab) (in (ac))

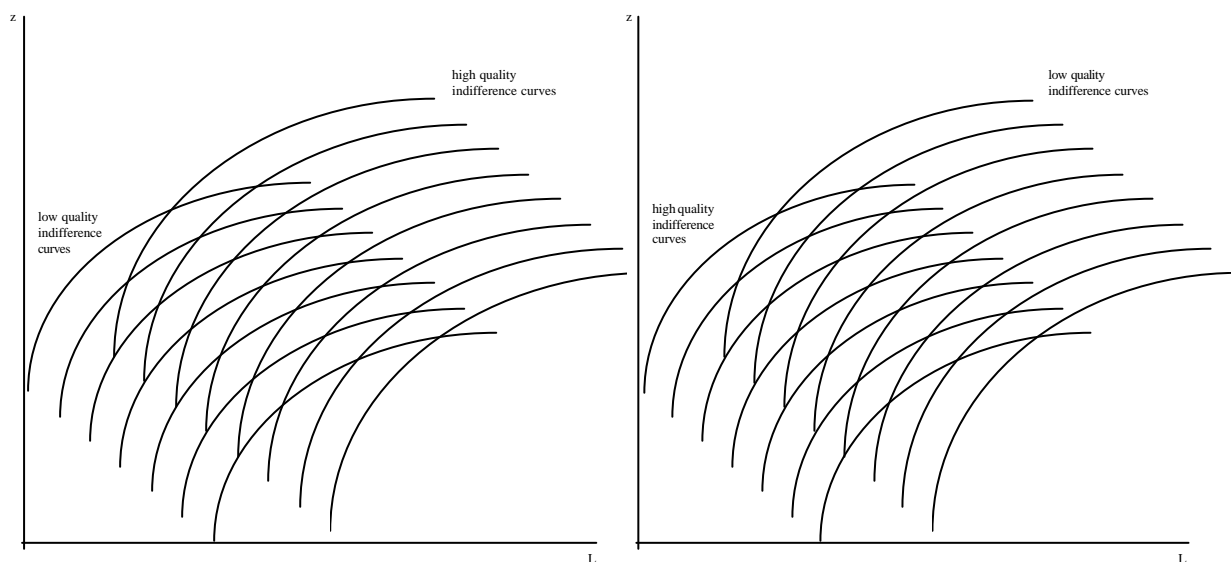
²³ Notice that either the MLRP of firm returns with respect to quality (ab) and the negative cross-partial derivative (equation (t)) or the MLRP of firm returns with respect to quality (ac) and the negative cross-partial derivative (equation (w)) together imply SOSD with respect to the loan size, provided returns are not only increasing but also concave in quality.

increases (decreases). Thus, as the degree of internalization increases (decreases), the slope of the high-quality indifference curves becomes steeper and steeper (flatter and flatter), while the slope of the low-quality ones becomes flatter and flatter (steeper and steeper). In other words an increase (decrease) in the degree of internalization implies that the high-quality indifference curves “rotate” in a counter-clockwise (clockwise) fashion, whereas the low-quality indifference curves “rotate” in a clockwise (counter-clockwise) fashion (see Figures 1a, 1b).

3. Equilibrium Outcomes.

As M-R have shown, when the size of the loan is allowed to vary, the entrepreneur’s choice of the contract becomes crucial²⁴. A bank offers an entrepreneur a menu of contracts, each specifying both the loan size and the repayment schedule, with the aim of financing a *specific firm* whose *specific return* will be due to the combination of the size of the loan actually obtained by this firm, its stochastic return, its “technological setting” and its quality-type. Since in general there is a relation between the loan size obtained and the return expected by the firm, the entrepreneur’s contract choice could signal his quality as well as his technology type, which are 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 shown above, the relation between loan size and firm marginal returns is specific to the firm, in the sense that it can be either increasing



(a) $H=I$ (b) $H=E$ **Figure A. Firms' indifference curves**

or non-increasing in the quality parameter. This implies that the relation between firm’s quality and its loan size can be either a positive or a negative one. If firm marginal returns are increasing in quality, quality and size of the loan will be *positively related*; and, if the bank knows this relation (by observing that $H=I$),

²⁴ Even abstracting from moral-hazard problems and referring to adverse-selection problems only.

“better” firms can signal their higher quality by agreeing on a higher loan interest rate, thus signing contracts based on loans of larger size. On the other hand, if firm marginal returns are non-increasing in quality, quality and size of the loan will be *negatively related*; and, if the bank knows this relation (by observing that $H=E$), “worse” firms agree on a higher loan interest rate, thus signing contracts based on loans of larger size (see also Innes (1991, p. 359)). Hence, from the lender’s perspective, ordering loan applicants on the basis of their expected marginal returns, without knowing H , would not be sufficient to discriminate between borrowers of different quality²⁵.

In what follows we will first assume that the “technological” parameter of each firm is observable by outside lenders (that is, it is common knowledge). This assumption allows the bank to exploit the borrowers’ signaling through their choice of the loan size. Therefore, when receiving applications for the menu of loans offered, the bank finds it convenient to screen first the borrowers off according to their observable “technology”. The bank will initially group all applicant firms according to their observable value of H (either $H=I$ or $H=E$). Once firms have been split into these two groups, we will see that either separating or pooling equilibria (between low and high quality firms within each group) will obtain, depending on the relative slope of the indifference curves of the firms. It must be noticed that, if an optimal separating equilibrium cannot obtain, a sub-optimal separating equilibrium can still obtain, where the high quality firms elicit a “second-best” contract which implies a loan size either higher (when $H=I$) or lower (when $H=E$) than the optimal one. Conversely, a pooling equilibrium attains.

When the “technology” of the firm is not observable by banks, firms cannot be grouped according to whether their parameter $H=I$ or $H=E$. However, as we will see, the bank is still able to exploit the borrowers’ signaling through their choice of the loan size, but only in a limited number of cases. Optimal separating equilibria can obtain, provided the slopes of the indifference curves are different enough for firms of different quality and technology. If that is not the case, different types of sub-optimal separating equilibria can obtain, in which high-quality firms elicit contracts that imply loan sizes higher (when $H=I$) or lower (when $H=E$) than the optimal ones and/or low-quality E -firms get more or less than the amount of credit they would have like to. Conversely, pooling equilibria attains.

A. *Observable technology (H is common knowledge).*

An *optimal* equilibrium contract obtains whenever the indifference curve of a given quality- q firm and the bank’s supply curve to that quality- q type are tangent at:

$$(Z) \quad \frac{\partial V(L, z|q, H)}{\partial L} = \frac{\partial \bar{\pi}(L|q, H)}{\partial L} - (1 + \rho),$$

given H .

²⁵ This would also apply to models with loans of fixed size across firms. 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 mean return is not necessarily the less (or the more) risky one. The same holds true for the D-W model: a higher “successful” return is not univocally related to a higher (or a lower) p .

For given values of (ζ, L) , these are the points where expected profits of a quality- q entrepreneur are maximized, subject to the non-negativity of bank's expected profits from loans to that entrepreneur, given H . Such equilibrium contracts will be *first-best* (perfect information) contracts if they satisfy the feasibility constraints of a Wilson equilibrium, as defined above (see (E1), (E2) and (E3), p. 7). We have thus the following proposition:

PROPOSITION A. In a world of asymmetric information, only a *separating* equilibrium can be a *first best equilibrium*.

Consider first the case of $H=I$. Suppose that two entrepreneurs, a low-quality and a high-quality one, choose the two different contracts $\{z_l^*, L_l^*\}$ and $\{z_h^*, L_h^*\}$, respectively (see Figure 2a, points A and C). The conditions for the two contracts to be a *separating* equilibrium are that:

- (i) each lies on a bank's separating supply curve ($S^q(I); q=l, h$);
- (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. For instance, the contract $\{z_h^*, L_h^*\}$ chosen by the high-quality entrepreneur (C in Figures 2a, 2b, and 2c) can lie below the low-quality indifference curve which yields the low-quality preferred contract $\{z_l^*, L_l^*\}$ (A in Figures 2). If this was the case, as in fact it is in Figure 2b, condition (ii) would not be satisfied for this contract. Thus, the entrepreneurs' optimal choice elicits a *first-best* equilibrium (a *separating equilibrium* with firm signaling) only if the high-quality *optimal* contracts lie *on* the high-quality preferred indifference curve, $V^*(q_h, H=I)$ (the one that is tangent to the supply curve to the high-quality types) and *above* the low-quality preferred indifference curve, $V^*(q_l, H=I)$. In Figure 2a we see that a separating optimal equilibrium (contracts A and C) obtains since the indifference curves are such that the high-quality firm chooses contract C from the offered menu of contracts without having the low-quality firm choose the same contract²⁶.

The converse of Proposition 1 is not true, however, in the sense that we can have separating equilibria which are not "first-best" (see Figure 2b). Suppose that 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, thus attracting low-quality entrepreneurs, so that an optimal equilibrium is not feasible. The consequent equilibrium will be separating if the high-quality indifference curve $V^p(q_h, I)$ that is tangent to the bank pooling supply curve²⁷, crosses the bank supply curve to the high-quality types, at a point (D) higher than the intersection between the low-quality preferred indifference curve, $V^*(q_l, I)$, and the bank supply curve to the high-quality type (point B). Since the latter point of intersection (point B) will be the contract chosen by the high-quality types, it allows for their signaling and avoid the possibility of bank's negative expected profits on loans to low-quality entrepreneurs. Thus, low-quality entrepreneurs will choose their "first-best" contract $\{z_l^*, L_l^*\}$, point A, whereas high-quality entrepreneurs will choose their

²⁶ This happens exactly as the tangency point between the high-quality preferred indifference curve $V^*(q_h, I)$ and the supply curve to the high-quality type, point C, lies above the intersection point between the low-quality preferred indifference curve $V^*(q_l, I)$ and the supply curve to the high-quality type (point B).

²⁷ 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 (q) above.

“second-best” contract (z_B, L_B) , point B. The resulting equilibrium will be a *suboptimal separating equilibrium*. Notice that under this contract, high-quality firms get a loan *larger* than what they would have like to: thus, when a suboptimal separating equilibrium attains, we have *over-financing* of high-quality firms.

Our conclusion shows that we can have separating equilibria that are *suboptimal* (i.e. “second-best”) since they allow costly signaling of the high-quality firms. With a given contract menu offer, whether we get an optimal or a suboptimal separating equilibrium depends on the slope of the indifference curves of the different types borrowers. Given that the low-quality firms choose their first-best contract (points A in Figures 2), all depends on the shape and the curvature of the indifference curves of the high-quality types as opposed to the low-quality preferred indifference curve. Notice that the $V(q_h, I)$ curves are less steep in Figure 2b than in Figure 2a: the slope of $V^*(q_h, I)$ in C is smaller in Figure 2b than in Figure 2a²⁸. In other words, since with $H=I$, for any given z or L , $MRS(z, L|q_h) > MRS(z, L|q_l)$, the closer $MRS(z, L|q_h)$ is to $MRS(z, L|q_l)$ the more unlikely it becomes for an optimal separating equilibrium to obtain.

The closeness of $MRS(z, L|q_h)$ to $MRS(z, L|q_l)$ can imply that, while condition (i) above is satisfied, condition (ii) is not. In this case no separating equilibrium obtains, either optimal or suboptimal. To see this, consider the contract pair $\{z_E, L_E\}$ determined by the tangency point between bank’s pooling supply curve, and the high-quality indifference curve, $V^p(q_h, I)$ (see Figure 2c, point E). The breaking down of condition (ii) above implies that this latter indifference curve would cross on the right the bank supply curve to the high-quality types, at a point (D) lower than the intersection between the low-quality preferred indifference curve, $V^*(q_l, I)$, and bank supply curve to the high-quality type (point B). In this case, the high-quality first-best contract $\{z_h^*, L_h^*\}$ (point C) would be clearly unfeasible, as it would attract low-quality entrepreneurs, whereas the $\{z_B, L_B\}$ contract (point B) would not be chosen by high-quality entrepreneurs with respect to $\{z_E, L_E\}$ or $\{z_D, L_D\}$. On the other hand, low-quality entrepreneurs would prefer a contract like point E to that in point A (it lies on a lower indifference curve). Hence, since high-quality entrepreneurs would be indifferent between point D and point E and low-quality entrepreneurs would prefer point E to point A, both quality-types entrepreneurs would choose the $\{z_E, L_E\}$ *pooling* contract, point E (see Figure 2c).

Intuitively, a pooling equilibrium occurs when not only indifference curves of high and low quality types are similar but also bank’s separating supply curves are far apart. Bank’s separating supply curves are far apart whenever there are large differences in default risks between different quality-type firms, whereas indifference curves of different types are similar whenever the differences in marginal firm returns are small and offset differential default risks to produce similar tradeoffs between L and z . In these conditions, while the bank is not able to exploit the differences in the shapes of entrepreneurs’ indifference maps, the high-quality firms find it too costly to signal their type, so that a pooling equilibrium emerges. Now, when a pooling equilibrium attains, we have that low-quality 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

²⁸ If the $V(q_h, I)$ curves are such as those in Figure 2a, then the high-quality types can elicit their first best choice, C (it lies above $V^*(q_h, I)$). Conversely, if the $V(q_h, I)$ ’s are such as those in Figure 2b, then the high-quality types cannot elicit their first best choice, C (it lies below $V^*(q_h, I)$ and it would attract the low-quality, too).

(see Figure 2c). In other words, the “better” firms subsidize the “worse” ones: high-quality firms are under-financed, in the sense that they get less than what they would like to.

Can we say that this under-financing is equivalent to *type I credit rationing* of high-quality firms? Independently of the specific case analyzed, it is possible to distinguish between credit rationing *strictu sensu* and under-financing. We say that under-financing and type I credit rationing coincide whenever an entrepreneur obtains an amount of loan smaller than the amount desired, even if he is willing to sign a debt contract with the bank prescribing a larger loan amount; and we say that under-financing does not imply type I credit rationing whenever an entrepreneur obtains an amount of loan smaller than the amount desired, but he is not willing to sign alternative debt contracts prescribing a larger loan amount²⁹. It follows from these two definitions that, in the case analyzed above, we have under-financing *but not* credit rationing of high-quality firms, for the latter choose to sign an under-financing contract (Figure 2C, point E) in order to avoid the cost of signaling their quality-type to the lending bank (Figure 2C, point B).

The analysis of the possible links between under-financing and credit rationing equilibria lead to a more general question: what are the relations among the different equilibria? In the case of $H=I$, looking at equation (v), which shows that the MRS between z and L is increasing in quality, we see that as the left-hand side become smaller and smaller, we move from an optimal to a suboptimal separating equilibrium and then to a pooling one³⁰. As far as conditions (A1), (A2) and the equations thereby derived, (t) to (v), are fulfilled, we can get *suboptimal* pooling equilibria. Notice that with condition (A1) 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 (aa) and (A2) (and hence (v)) that ensures the existence of pooling (sub-optimal) equilibria.

It must be noticed 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 so that a pooling contract is not viable, as there is no pooling supply curve offered by banks. While assumptions (aa) and (A2) keep holding, the incentive compatibility constraint and the low-quality rationality constraint ensure that in Figures 2 either the two contracts A and C will attain a separating optimal equilibrium, or the two contracts A and B will attain a separating sub-optimal equilibrium. In a Riley construction, in fact, no indifference curve crossing $S^h(I)$ at D can be preferred by the high-quality entrepreneur, it not being tangent to any supply curve (as there is no pooling supply curve in this case). In

²⁹ It can be given several justifications to this behavior: maybe our entrepreneur considers the alternative debt contracts either too costly or Pareto dominated by the under-financing equilibrium. In any case, as it will become clearer below, under-financing, as opposed to type-I credit rationing, is characterized either by the firm’s willingness to “signal” its quality-type or by the firm’s willingness to avoid the signaling cost. On the other hand, under-financing coincides with credit rationing in all the situations where the rationed firm has no alternative choices to overcome its quantity constraint.

³⁰ In the extreme case, we can have a pooling equilibrium where indifference curves are strictly convex functions of z as q changes ($MRS(z, L|q_i) = MRS(z, L|q_j)$). In such case, both the cross-partial derivative in equation (t) and the derivative in equation (u) would be zero. However, indifference curves need not be strictly convex in z (different quality types may have indifference curves that intersect somewhere). Moreover both (t) and (u) can keep holding and yet we can have a pooling equilibrium.

other words, contract D would never be proposed in the first place, and hence would never be chosen, even if indifference curves were such as those in Figure 2c. In this latter case, as in the case of Figure 2b, low-quality firms get their first-best loan size (which is less than the “pooling level”) and high-quality firms get more than their first-best (and obviously more than the “pooling level”).

This stresses that the sequence of moves in the structure of the contracting game is very important since, before an equilibrium is reached, all the feasibility constraints must be satisfied. Thus, when the bank first starts by offering a contract menu, if some applicants choose A while others choose C, implying that indifference curves are like those in Figure 2a, then an optimal separating equilibrium obtains (low-quality applicants choose their first-best contract A, while high-quality applicants choose their first-best contract C), as the bank will earn non-negative expected profits on both contracts. If some applicants choose A while others choose B, implying that indifference curves are like those in Figure 2b, then a suboptimal separating equilibrium obtains (low-quality are better off, while high-quality applicants can signal their quality by being over-financed). Finally, suppose some applicants prefer C (most probably the high-quality ones) and this attracts all other applicants. Then the high-quality ones would shift to D, which is preferred to B. If that is the case, it means that indifference curves are as in Figure 2c. Then the low-quality applicants would shift to E, and so would the high-quality who are indifferent between D and E. Hence, a pooling equilibrium would obtain. Firm signaling in such case would be inconvenient for high-quality firms.

To conclude, we can state the following propositions.

PROPOSITION B. When $H=I$ and is common knowledge, under conditions (aa) and (A2) the equilibrium set of separating contracts has the properties:

- (i) the lowest-quality firm that is financed gets a *first-best* contract;
- (ii) each contract breaks even for the bank;
- (iii) the loan size in each contract $\{z_q, L_q\}$ is increasing in q : if $\{z_{q+1}, L_{q+1}\} > \{z_{q+1}^*, L_{q+1}^*\}$ (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.

PROPOSITION C. When $H=I$ and is common knowledge, under conditions (aa) and (A2), the equilibrium pooling contract has the properties:

- (i) no single contract breaks even for the bank but the average contract does;
- (ii) no firm that is financed gets a first-best contract: “lower-than-average” quality firms are over-financed whereas “higher-than-average” quality firms are under-financed.

Consider now the case of prevailing external economies of scale ($H=E$). When $H=E$, larger loans signal lower quality (better quality needs smaller loans). Assumption (ac) implies that lower quality produces “better” net worth distributions in the sense of statistical “good news”. The equilibrium conditions in the case where $H=E$ mirror the conditions derived above for $H=I$. In particular, Proposition 1 still holds, in the sense that we will have a *separating* set of first-best equilibria if:

- (i) each lies on a separating supply curve ($S^q(E)$; $q=l, h$);
- (ii) the low-quality contract is weakly preferred by the low-quality entrepreneur.

Condition (i) is necessary but not sufficient to have an optimal separating equilibrium. It is also necessary that the high-quality optimal contract (point A, Figure 3a) lies above the low-quality optimal indifference curve (point C, Figure 3a). In this case, the entrepreneurs' optimal choice elicits a "first-best" equilibrium. However, the high-quality contract $\{z_h^*, L_h^*\}$, point A, can lie below the low-quality indifference curve which yields the low-quality preferred contract $\{z_l^*, L_l^*\}$, point C (see Figure 3b). In this case, condition (ii) would not be satisfied for this contract.

The latter possibility points out that, again, the converse of Proposition 1 is not true, in the sense that we can have separating equilibria which are not "first best". Suppose that the high-quality preferred contract (point A, Figure 3b) attracts the low-quality firms, lying on a low-quality indifference curve which is below the low-quality preferred indifference curve, so that an optimal equilibrium would not be feasible. The consequent "suboptimal" equilibrium will be separating if the high-quality indifference curve which is tangent to the bank pooling supply curve, lies above the high-quality indifference curve passing through the point of intersection between the low-quality preferred indifference curve and bank supply curve to the high-quality types (point B, Figure 3b). In the case of Figure 3b, the latter condition is satisfied and the point of intersection B is the high-quality contract. This contract allows for the signaling of high-quality firms, and avoids the possibility of bank's negative expected profits on loans to low-quality firms. Thus, low-quality entrepreneur will choose their "first best" contract $\{z_l^*, L_l^*\}$, point C, whereas high-quality entrepreneurs will choose their "second best" contract $\{z_h, L_h\}$, point B. The resulting equilibrium will be a *suboptimal separating equilibrium*. Notice that under this contract, high-quality firms will be under-financed. However, according to our previous definitions, this under-financing *does not imply* type I credit rationing: high-quality firms choose to sign an underfinancing contract in order to signal their quality-type to the lending bank.

As with $H=I$, we can still have that condition (i) above is satisfied while condition (ii) is not. In this case no separating equilibrium obtains, either optimal or suboptimal (cf. Figure 3c). To see this, consider the $\{z_E, L_E\}$ contract determined by the tangency point between the bank pooling supply curve and the high-quality indifference curve (point E). The breaking down of condition (ii) above is implied by the fact that this latter indifference curve lies below the low-quality preferred indifference curve and the high-quality indifference curve which pass through point B. In this case, the high-quality "first best" contract $\{z_h^*, L_h^*\}$, point A, is clearly unfeasible, as it would attract low-quality entrepreneurs; moreover the high-quality entrepreneurs are indifferent between D and E, which are preferred to B. On the other hand, low-quality entrepreneurs prefer contract E to contract C. Hence, both quality-types entrepreneurs will choose the *pooling contract* $\{z_E, L_E\}$ (see point E, Figure 3c). When a pooling equilibrium attains and $H=E$, we have that low-quality firms will obtain loan amounts *smaller* than their first-best levels, while high-quality firms will obtain loan amounts *larger* than their first-best level (see Figure 3c). In other words, "worse" firms get less than what they would like to since they subsidize "better" firms. According to our previous definitions, this under-financing of low-quality firms coincides with type I credit rationing: low-quality firms would be ready to sign a debt contract for a larger loan amount, but they have no alternative choice to overcome their quantity constraint.

This link between under-financing and credit rationing leads, again, to a more general question: what are the relations among the different equilibria in the case of $H=E$? Looking at equation (y), which shows that the MRS between z and L is decreasing in quality, we see that as the left-hand sides become smaller and smaller, we move from an optimal to a suboptimal separating equilibrium and then to a pooling one³¹. As far as conditions (aa) and (ac), and the equations (w) to (y) are fulfilled, we can get *suboptimal* pooling equilibria. Again, notice that with (aa) 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 (aa) and (ac) (and hence (y)) that ensures the existence of pooling (sub-optimal) equilibria.

Notice also that, like in the case of $H=I$, with $H=E$ 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, but it would not do so when a Wilson construction leads to a pooling equilibrium. Even if indifference curves were such as those in Figure 3c, in Riley construction contract D would not even appear, and the choice would again be the separating couplets of contract C and B. In that case, while low-quality firms get their first-best loan size (which is more than the “pooling level”), high-quality firms get less than their first-best (and obviously even less than the “pooling level”). Hence, under-financing does occur in a Riley construction for the high-quality types. This stresses once more that the sequence of moves in the structure of the contracting game is very important since, before an equilibrium is reached, all the feasibility constraints must be satisfied. Thus, when the bank first starts by offering a contract menu, if some applicants choose A while others choose C, implying that indifference curves are like those in Figure 3a, then an optimal separating equilibrium obtains (low-quality applicants choose their first-best contract C, while high-quality applicants choose their first-best contract A), as the bank will earn non-negative expected profits on both contracts. If some applicants choose A while others choose B, implying that indifference curves are like those in Figure 3b, then a suboptimal separating equilibrium obtains (low-quality are better off, while high-quality applicants can signal their quality by being under-financed). Finally, suppose some applicants prefer A (most probably the high-quality ones) and this attracts all other applicants. Then the high-quality ones would shift to D, which is preferred to B. If that is the case, this means that indifference curves are as in Figure 3c. Then the low-quality applicants would shift to E, and so would the high-quality who are indifferent between D and E. Hence, a pooling equilibrium would obtain. Firm signaling in such a case would not be convenient for the high-quality firms.

To conclude, we can state the following propositions.

PROPOSITION D. When $H=E$ and is common knowledge, under conditions (aa) and (ac), the equilibrium set of separating contracts has the properties:

- (i) the firm of lowest quality that is financed gets a “first best” contract;

³¹ As in the case of $H=I$, with $H=E$ we can have a pooling equilibrium where indifference curves are strictly convex functions of z as q changes ($MRS(z, L|q_h) = MRS(z, L|q_l)$). In such case, both the cross-partial derivative in equation (w) and the derivative in equation (x) would be zero. However, indifference curves need not be strictly convex in z (different quality types may have indifference curves that intersect somewhere). Moreover both (w) and (x) can keep holding and yet we can have a pooling equilibrium.

- (ii) each contract breaks even for the bank;
- (iii) the loan size in each contract (z_q, I_q) is decreasing in q : if $(z_{q+1}, I_{q+1}) < (z_{q+1}^*, I_{q+1}^*)$ (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.

PROPOSITION E. When $H=E$ and is common knowledge, under conditions (aa) and (ac), the equilibrium pooling contract has the properties:

- (i) no single contract breaks even for the bank but the average contract does;
- (ii) no firm that is financed gets a “first best” contract: “lower-than-average” quality firms undergo an under-financement, which implies type I credit rationing, whereas “higher-than-average” quality firms are over-financed.

B. Non-observable technology (H is private knowledge).

When H is not observable by outside lenders, the borrowing firms cannot be grouped by banks according to their technological typology³². Banks know the distribution of the firms’ population in terms of their degree of internalization, but they do not know the specific degree of internalization of each firm³³. If H belongs to two disjoint sets $[I, E]$, each bank will know the respective share of I -firms and E -firms on firms’ population but it *will not know* “who is who”. Thus, since banks cannot discriminate between firms with prevailing internal economies of scale (I -firms) and firms with prevailing external economies of scale (E -firms), all firms are grouped together and treated as *ex ante* identical (i.e. independently of their technological type). This implies that lending banks are no longer able to take an increasing amount of loan demanded as unambiguously representing either a higher or a lower quality, respectively. A non-observable technology means, in fact, that an increase in the loan size can signal either that quality is better or that quality is worse. The result is that each bank must simultaneously offer a set of four different menus of contracts, instead of only two, designed for both high-quality firms (the one with $H=I$ and the one with $H=E$) and for both low-quality firms (the one with $H=I$ and the one with $H=E$).

Nevertheless, the bank is still willing to exploit its knowledge of the borrowers’ distribution between I -firms and E -firms, as well as the monotonicity assumptions implied by conditions (ab) and (ac) above. To implement the latter possibility, we introduce the following assumption:

ASSUMPTION. For all $\pi \in [0, K(L, q, H)]$:

$$(AD) \quad \frac{\partial}{\partial \pi} \left(\frac{f_{qH}(\pi|L, q, H)}{f(\pi|L, q, H)} \right) > 0.$$

³² In other words, in this case the asymmetry of information concerns technology too.

³³ It is worth recalling that the degree of internalization of a given firm relates to the “location” of its parameter H between E and I . In particular, firm’s degree of internalization increases when there is a shift from E to I . Obviously, if H belongs to two disjoint sets $[I, E]$, the related firms’ distribution is limited to the two values $H=I$ and $H=E$. Thus, saying that banks know the distribution of firms according to their technology, means that banks know the *proportion* of firms with $H=I$ and that of firms with $H=E$.

Given (ab) and (ac), assumption (ad) is a suitable mild assumption of monotonicity of the loan size with respect to quality *and* technology together. Given (aa), it implies that larger loan sizes yield superior return distributions when they signal better quality ($H=I$) while they yield inferior return distributions when they signal lower quality ($H=E$). Such monotonicity can be exploited by the bank when ordering the signals in terms of their favourableness. Let us refer to the four different contract menus $(z, L|q, H)$ (with $q=q_l, q_h; H=I, E$) offered by a given bank to the four different types of firms. Since firms would possibly choose four different (z, L) couplets, six possible relations among the contract couplets are thus established. The monotonicity condition, specified by assumptions (aa)-(ad), implies that these relations must be characterized by the following ordering:

$$(AAa) \quad \text{MRS}(z, L|q_h, I) > \text{MRS}(z, L|q_l, I),$$

$$(aab) \quad \text{MRS}(z, L|q_h, E) < \text{MRS}(z, L|q_l, E),$$

$$(aac) \quad \text{MRS}(z, L|q_h, I) > \text{MRS}(z, L|q_l, E),$$

$$(aad) \quad \text{MRS}(z, L|q_h, E) < \text{MRS}(z, L|q_l, I),$$

$$(aae) \quad \text{MRS}(z, L|q_l, I) < \text{MRS}(z, L|q_l, E),$$

$$(aaf) \quad \text{MRS}(z, L|q_h, I) > \text{MRS}(z, L|q_h, E).$$

Conditions (aaa) and (aab) derive from assumptions (ab) and (ac), respectively, while conditions (aae) and (aaf) derive from assumption (ad). Also, since the derivatives in both (ab) and (ac) are increasing in the firms' degree of internalization, conditions (aac) and (aad) are implied.

As in the previous cases, for any given H , an *optimal* equilibrium contract obtains whenever the preferred indifference curve of each quality- q entrepreneur and the bank's supply curve to the corresponding quality- q firm are tangent at:

$$(BB) \quad \frac{\partial V(L, z|q, H)}{\partial L} = \frac{\partial \bar{\pi}(L|q, H)}{\partial L} - (1 + \rho).$$

For given values of (z, L) , this is the point where expected profits of a firm of quality- q and technology H are maximized, subject to the non-negativity of bank's expected profits from loans to that firm. Such *first best* (perfect information) equilibrium contract can be realized if the above defined feasibility constraints of a Wilson equilibrium are satisfied.

PROPOSITION F. In a world of asymmetric information, where both quality and technology are non-observable by lenders, only a *separating* equilibrium can be a *first best equilibrium*. Such equilibrium exists and is feasible.

Consider four entrepreneurs, two of which have scale-economy internalizing firms ($H=I$) of high and low-quality, respectively, whereas the other two have scale-economy externalizing firms ($H=E$), of high and low-quality too, respectively. Firms' quality and technology are not observable by lenders, but the

quality and technology distributions across the firms' population are common knowledge. Suppose that the four entrepreneurs choose the four different optimal contracts $\{z_h^*(I), L_h^*(I)\}$ and $\{z_l^*(I), L_l^*(I)\}$, and $\{z_h^*(E), L_h^*(E)\}$ and $\{z_l^*(E), L_l^*(E)\}$, respectively (see Figure 4, points D, B, A, C).

The conditions for the four contracts to be a *separating* equilibrium are that:

- (i) each lies on a separating supply curve;
- (ii) the low-quality contracts are weakly preferred by the low-quality entrepreneurs.
- (iii) the contract designed for the low-quality *I*-firm is weakly preferred by the entrepreneur of this technology and quality type with respect to the contracts designed for the *E*-firms, and the latter contracts are weakly preferred by scale-economy externalizing firms with respect to the contract designed for the high-quality *I*-firm.

In order for each firm's preferred indifference curve to be tangent to the corresponding separating supply curve of the bank, both the two bank supply curves to the low-quality firms and the two bank supply curves to the high-quality ones must not be too "close". Now, while the supply curves to the low-quality types will always lie *above* the supply curves to the high-quality ones (see section 2D), assumptions (aa) to (ad) and the related ordering stated by equations (27) imply that the bank supply curve to the low-quality *E*-firm lies *below* the bank supply curve to the low-quality *I*-firm, and that the bank supply curve to the high-quality *I*-firm lies *below* the bank supply curve to the high-quality *E*-firm (see Figure 4)³⁴.

Thus, an optimal separating equilibrium will exist and will be feasible if and only if: (1) the high-quality *optimal* contracts $\{z_h^*(I), L_h^*(I)\}$ and $\{z_h^*(E), L_h^*(E)\}$ lie *on* the corresponding high-quality preferred indifference curves, $V^*(q_h, I)$ and $V^*(q_h, E)$, respectively; (2) both these contracts (Figure 4, points D and A, respectively) lie *above* the low-quality preferred indifference curves, $V^*(q_l, I)$ and $V^*(q_l, E)$; (3) the contract D lies *above* the high-quality preferred indifference curve $V^*(q_h, E)$; (4) the low-quality *optimal* contract $\{z_l^*(E), L_l^*(E)\}$ (point C, Figure 4) lies *above* the low-quality preferred indifference curves, $V^*(q_l, I)$. By satisfying (1)-(4), the entrepreneurs' optimal choice elicits a *first-best* equilibrium. This optimal separating equilibrium (contracts A, B, C and D, Figure 4) implies that: the high-quality *I*-firms chooses contract D within bank's offered menu of contracts without having the high-quality *E*-firm and the two low-quality firms choose the same contract; the high-quality *E*-firm chooses contract A within bank's offered menu of contracts without having the two low-quality firms choose the same contract; the low-quality *E*-firm chooses contracts C within bank's offered menu of contracts without having the low-quality *I*-firm chooses the same contract.

The supply curves and the indifference curves depicted in Figure 4 show that there does exist a possibility to get an optimal separating equilibrium, even though technology is non-observable by lenders. Again, the converse of PROPOSITION F is not true, in the sense that we can have separating equilibria which are not "first-best" and, yet, imply self-selection of firms. Actually, we can have several different types of

³⁴ We claim that assumption (ad) is not a strong assumption but a mild one, indeed. Consider that bank supply curves tend to "diverge" whenever we get close to a situation leading to a pooling equilibrium, since in that case the slopes of firms' indifference curves tend also to be similar. The case depicted in Figure 4 (and implied by our assumptions) is thus one that would favor separating equilibria for both the $H=I$ and the $H=E$ firms when considered separately. We would have a supply curve to the low-quality types lying very close (almost overlapping) to a supply curve to the high-quality ones if for either the $H=I$ or the $H=E$ firms, taken separately, we had a possibly pooling equilibrium. But then, this would not be a candidate for a separating equilibrium anyway, even under a simultaneous four-supply curve representation.

suboptimal separating equilibria. In particular, some of the first-best equilibria that could obtain under perfect information about technology, are no longer feasible and become “second-best” separating equilibria when information about technology is asymmetric; on the other hand, some of the sub-optimal equilibria obtained when technology was observable do remain viable when technology becomes unobservable by lenders³⁵.

Consider the first-best equilibria under perfect information about H which are no longer viable under asymmetric information about H . We can have three main cases. First, an optimal equilibrium is not feasible whenever the preferred contract of the high-quality E -firm (point A) lies below (to the right along the bank’s corresponding supply curve) the intersection between the preferred indifference curve of low-quality I -firm and the bank supply curve to the high-quality E -type (point F, Figure 5). Second, an optimal equilibrium is not feasible whenever the preferred contract of high-quality I -firm lies below (to the left along the bank’s corresponding supply curve) the intersection between the preferred indifference curve of low-quality E -firm and the bank supply curve to the high-quality I -type (point G, Figure 6). Third, an optimal equilibrium is not feasible whenever the preferred contract of low-quality E -firm lies below (to the left along the bank’s corresponding supply curve) the intersection between the preferred indifference curve of low-quality I -firm and the bank supply curve to the low-quality E -type (point M, Figure 7).

In addition, there are two additional situations in which a previously optimal equilibrium becomes not viable under non-observable H . One of these situations obtains whenever at least two of the three main cases occur simultaneously. For instance, we can refer to the first two cases (see Figures 5 and 6, respectively): the preferred contract of the high-quality E -firm lies below the intersection between the preferred indifference curve of low-quality I -firm and the bank supply curve to the high-quality E -type and, at the same time, the preferred contract of high-quality I -firm lies below the intersection between the preferred indifference curve of low-quality E -firm and the bank supply curve to the high-quality I -type (see Figure 8). The other additional situation happens whenever the preferred contract of high-quality I -firm lies below the intersection between the preferred indifference curve of high-quality E -firm and the bank supply curve to the high-quality I -types. However, it is unlikely that this situation, which is based on quite peculiar assumptions³⁶, can occur independently of the three main cases analyzed above.

In principle, when H is non-observable, six different pooling supply curves are conceivable. Hence the five cases of non-feasibility of the optimal separating equilibria, analyzed above, can lead to suboptimal separating equilibria only if the latter dominate the set of possible pooling equilibria; and this dominance condition will be satisfied only if (1) the high-quality (whether of a I -firm or of a E -firm) indifference curve that is tangent to a bank pooling supply curve crosses the bank supply curve to the high-quality types at a point higher than the intersection between the low-quality (whether of a I -firm or of a E -firm) preferred indifference curve and the same supply curve, allowing for signaling of the high-quality types, (2) the high-quality I -firm indifference curve that is tangent to a bank pooling supply curve crosses

³⁵ It is worth noting that other separating equilibria that were suboptimal under perfect information about technology are no longer feasible, and become pooling equilibria (see below).

³⁶ In particular, the shapes of the indifference curves of the firms of different types must be very close to each other and/or the slope of the bank supply curve to the high-quality I -firms must be very flat with respect to the slopes of the three other separating supply curves.

the bank supply curve to the high-quality I -firm at a point above the intersection between the high-quality E -firm preferred indifference curve and the same supply curve, allowing for signaling of the high-quality I -type; (3) the low-quality E -firm indifference curve that is tangent to a bank pooling supply curve crosses the bank supply curve to the low-quality E -firm at a point higher than the intersection between the low-quality I -firm preferred indifference curve and the same supply curve, allowing for signaling of the low-quality E -type.

Points (1)-(3) show that the equilibria, which were optimal under perfect information about technology but are no longer feasible when H is non-observable, lead to *suboptimal separating equilibria* (as illustrated in Figures 4, 5, 6, and 7) provided the low-quality entrepreneurs, or *at least* the low-quality I -entrepreneur, choose their (his) “first-best” contracts and provided the high-quality entrepreneurs, or the latter *and* the low-quality E -entrepreneur, choose “second-best” contracts so that all these contracts dominate the pooling ones. It is worth stressing that under these suboptimal separating contracts either the low-quality E -firm gets *more* (Figure 7), or the high-quality I -firm gets *more* (Figure 6), or the high-quality E -firm get *less* (Figure 5), or else the high-quality I -firm gets *more* and the high-quality E -firm get *less* (Figure 8), than what they would have like to. Thus, when a suboptimal separating equilibrium attains, we have either *over-financing* or *under-financing* or *both* of the high-quality firms, or else *over-financing* of the low-quality E -firm, *that would have not occurred were technology observable by lenders*; however, according to our previous definitions of under-financing and type I credit rationing (see above, section 3A), these cases of under-financing cannot be identified with credit rationing. It is also worth noting that it can happen that two of the suboptimal separating equilibria dominate the corresponding pooling equilibria whereas the remaining separating equilibria are dominated by one of the corresponding pooling equilibria. In these situations we have the so-called *semi-separating suboptimal equilibria* (cf. Stiglitz and Weiss 1992); and, if the pooling equilibrium is such as to imply under-financing of one of the low-quality firms, this under-financing can be identified with type I credit rationing.

As stated above, the separating “second-best” equilibria derived from the optimal equilibria with observable H do not wear out the set of possible suboptimal separating equilibria under non-observable H . For we have also to refer to the equilibria which derive from the suboptimal separating equilibria obtained under perfect information about technology. As far as these latter contracts are concerned, it happens that either they can also obtain under asymmetric information about technology, or they are no longer feasible (being dominated by pooling equilibria of various sorts), or they lead to sub-sub-optimal separating equilibria.

Let us refer to this last alternative by means of Figures 9 and 10. With an observable technology ($H=E$), the supply and indifference curve configuration depicted would have led to a sub-optimal separating equilibrium (like that in Figure 3b). However, the simultaneous “overlap” of the supply and the indifference curves for the $H=I$ case (similar to the first-best equilibrium in Figure 2a) could lead to a sub-suboptimal separating equilibrium provided the pooling equilibrium is not dominant, i.e. provided that the low-quality preferred indifference curves are not dominated by the high-quality E -type’s indifference curve

which is tangent to the corresponding pooling supply curve³⁷. In that case (see Figure 9), the high-quality E -firm would choose the crossing point (N) between the low-quality preferred indifference curve $V^*(q_l, D)$ and the bank's supply curve to the same high-quality E -firm. Under these circumstances, the high-quality E -firm would get even *less* than in the under-financing equilibrium reached when technology was observable (see Figure 3b). Consider now Figure 10. With an observable technology ($H=D$), the supply and indifference curve configuration depicted would have led to a sub-optimal separating equilibrium (like that in Figure 2b). However, the simultaneous "overlap" of the supply and the indifference curves for the $H=E$ case (similar to the first-best equilibrium in Figure 3a), could lead to a sub-suboptimal separating equilibrium, provided the pooling equilibrium is not dominant. In that case, in fact, the high-quality I -firm would choose the crossing point (Q) between the low-quality preferred indifference curve $V^*(q_l, E)$ and the bank's supply curve to the same high-quality I -firm. The high-quality I -firm would thus get even *more* than in the *over-financing* equilibrium reached when technology was observable (see Figure 2b).

If no optimal, sub-optimal "second-best" or even "third-best" separating equilibria obtain, several different pooling equilibria can occur. Therefore we have shown that, even if technology is non-observable, under mild suitable assumptions, there can still be separating equilibria with either *over-financing for the low-quality E-firm* or else *over-* or *under-financing* or both *for the high-quality firms*³⁸. Conversely, we maintain that, when pooling equilibria emerge, high-quality firms will get either *under* or *over-financed* while low-quality firms will get either *over* or *under-financed*. It is impossible to analyze here the rich set of different pooling equilibria which our model could determine. However, if we only refer to the pooling equilibria sketched out above, we can say that while the high-quality I -firm and the low-quality E -firm are under-financed, the high-quality E -firm and the low-quality I -firm are over-financed. A significant sub-set of the *under-financing* pooling equilibrium can be identified with *type I credit rationing*.

We can thus state the following propositions.

PROPOSITION G. When quality and technology are non-observable, under condition (A1), two non-observable quality types and two non-observable technology types (for which assumption (A4) and either assumption (A2) or (A3) hold), the equilibrium set of four separating contracts has the properties:

- (i) the lowest-quality firms that are financed get *first-best* contracts;
- (ii) each contract breaks even for the bank;
- (iii) the first two contracts are decreasing in q : if $\{z_{q+1}, L_{q+1}\} < \{z_{q+1}^*, L_{q+1}^*\}$ (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;

³⁷ It is also necessary to assume that the low-quality preferred indifference curve $V^*(q_l, D)$ does not lie above the low-quality preferred indifference curve $V^*(q_l, E)$. If this was not the case, it would be at the most possible to obtain semi-separating equilibria (see also above).

³⁸ In particular, we get *over-financing* for high-quality scale-economy internalizing firms and *under-financing* for high-quality scale-economy externalizing firms.

(iv) the second two contracts are increasing in q : if $\{z_{q+1}, L_{q+1}\} > \{z_{q+1}^*, L_{q+1}^*\}$ (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.

PROPOSITION 8. When quality and technology are non-observable, under condition (A1), two non-observable quality types and two non-observable technology types (for which assumption (A4) and either assumption (A2) or (A3) hold), the equilibrium pooling contracts have the properties:

- (i) no single contract breaks even for the bank but the different average contracts do;
- (ii) no firm that is financed gets a first-best contract.

4. Summary and Conclusions

In this paper we have analyzed the different equilibria that can emerge in a competitive credit market in which optimizing risk neutral firms sign standard debt contracts with a risk neutral bank in order to finance their production processes whose output is uncertain. We have assumed that firm output was also dependent on the firm's specific quality and technological endowment. The asymmetry of information between banks and firms was due to the inability of the former to observe firms' quality and, possibly, firms' production technology. In the first part of the paper, we have built the basic model in which the technological parameter of each firm is "common knowledge"; in the second part of the paper, we have removed this simplification by making this parameter non-observable by the banks.

The main results, obtained in the basic model, can be summarized in three points. In the first place, our approach encompasses the main features of the previous models on credit rationing, such as S-W and D-W, which analyze loans of fixed size and can, thus, be considered as particular cases of a more general model referring to loans of variable size. In the second place, our approach is able to include the different features, assumed by the observable technological parameter, in a unifying framework of analysis. In so doing, we generalize M-R's contribution which is able to refer to the same features of the technological parameter only by resorting to three different and unrelated "cases", as well as Innes (1991)'s model which is limited to a specific feature of the same parameter. In the third place, our approach adopts a Wilson construction of the contracting game different from the Riley construction chosen by M-R, thus allowing for a richer set of different equilibria. The latter include under-financing separating equilibria, as well as under-financing and type I credit rationing pooling equilibria.

The important refinement, based on the new assumption that the technological parameter is non-observable by the banks, has not only confirmed the results reached in the basic model but has also greatly enriched the set of possible equilibria. The following table summarizes the relations between the two sets of equilibria.

Equilibria when H is observable:	can become, when H is non-observable:
Optimal separating equilibria	Optimal separating equilibria Sub-optimal separating equilibria Pooling equilibria
Sub-optimal separating equilibria	Sub-optimal separating equilibria Sub-sub-optimal separating equilibria Pooling equilibria
Pooling equilibria	Pooling equilibria

Despite these achievements, our paper is still constrained by several simplifications, the main being the exogenous determination of the standard debt contract and the exclusive reference to a one-period horizon and to the finance of firms' "working capital". The first simplification is not difficult to overcome, since the analytical structure of our model is perfectly compatible with Innes (1993) approach, to show that the optimal form of the contract is just the standard debt contract. The second simplification is more important. In our model no role is left for a multiperiod analysis and for the related problem of firms' fixed capital accumulation. As a consequence, no debt accumulation is allowed. These are drawbacks which are common to the whole literature on the topic. Hence, further research is called to analyze the problems relating to capital accumulation through debt (and equity) financing under asymmetric information in a multiperiod horizon. This will certainly be the next step in our research agenda.

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.
- Bester, H. (1987), "The Role of Collateral in Credit Markets with Imperfect Information", *European Economic Review*, 31: 887-99.
- 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.
- Coase, R. (1937), "The Nature of the Firm", *Economica*, 4: 233-61.
- De Meza, D., and D. Webb, (1987), "Too Much Investment: A Problem of Asymmetric Information", *Quarterly Journal of Economics*, 102: 281-92.
- De Meza, D., and D. Webb, (1989), "Efficient Credit Rationing", *European Economic Review*, 36: 1277-90.
- Grossman, S., and O. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94: 691-719.
- Hellwig, M. (1987), "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection", *European Economic Review*, 31:319-25.

- Kreps, D., and R. Wilson (1982), "Sequential Equilibria", *Econometrica*, 50: 863-94.
- Innes, R. (1990), "Limited Liability and Incentive Contracting with Ex-ante Action Choices", *Journal of Economic Theory*, 52: 45-67.
- Innes, R. (1991), "Investment and Government Intervention in Credit Markets When There Is Asymmetric Information", *Journal of Public Economics*, 46: 347-81.
- Innes, R. (1992), "Adverse Selection, Investment, and Profit Taxation", *European Economic Review*, 36: 1427-52.
- Innes, R. (1993), "Financial Contracting under Risk Neutrality, Limited Liability and Ex Ante Asymmetric Information", *Economica*, 60: 27-40.
- Jaffee, D.M., and T. Russell, (1976), "Imperfect Information, Uncertainty, and Credit Rationing", *Quarterly Journal of Economics*, 90: 651-66.
- Leland, H., and D. Pyle, (1977), "Informational Asymmetries, Financial Structure and Financial Intermediation", *Journal of Finance*, 32: 371-87.
- 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*, 12: 380-91.
- Milgrom, P., and J. Roberts (1988), "An Economic Approach to Influence Activities and Organizational Responses", *American Journal of Sociology*, April.
- Miyazaki, H. (1977), "The Rat Race and Internal Labor Markets", *Bell Journal of Economics*, 8: 394-416.
- Riley, J. (1979), "Informational Equilibrium", *Econometrica*, 47: 331-59.
- Rothschild, M., and J.E. Stiglitz (1970), "Increasing Risk: I. A Definition", *Journal of Economic Theory*, 2: 225-43.
- Rothschild, M., and J.E. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics*, 90: 628-49.
- Shavell, S. (1979), "Risk Sharing and Incentives in the Principal and Agent Relationship", *Bell Journal of Economics*, 10: 55-73.
- Stiglitz, J.E., and A. Weiss (1986), "Credit Rationing in Markets with Imperfect Information", *American Economic Review*, 71: 393-410.
- Stiglitz, J.E., and A. Weiss (1986), "Credit Rationing and Collateral", in: J. Edwards, J. Franks, C. Mayer, and S. Schaefer (eds.), *Recent Developments in Corporate Finance*, Cambridge: Cambridge University Press.
- Stiglitz, J.E., and A. Weiss (1992). "Asymmetric Information in Credit Markets and its Implications for Macro-Economics", *Oxford Economic Papers*, 44: 694-724.
- Wilson, C. (1977), "A Model of Insurance Markets with Incomplete Information", *Journal of Economic Theory*, 16: 167-87.

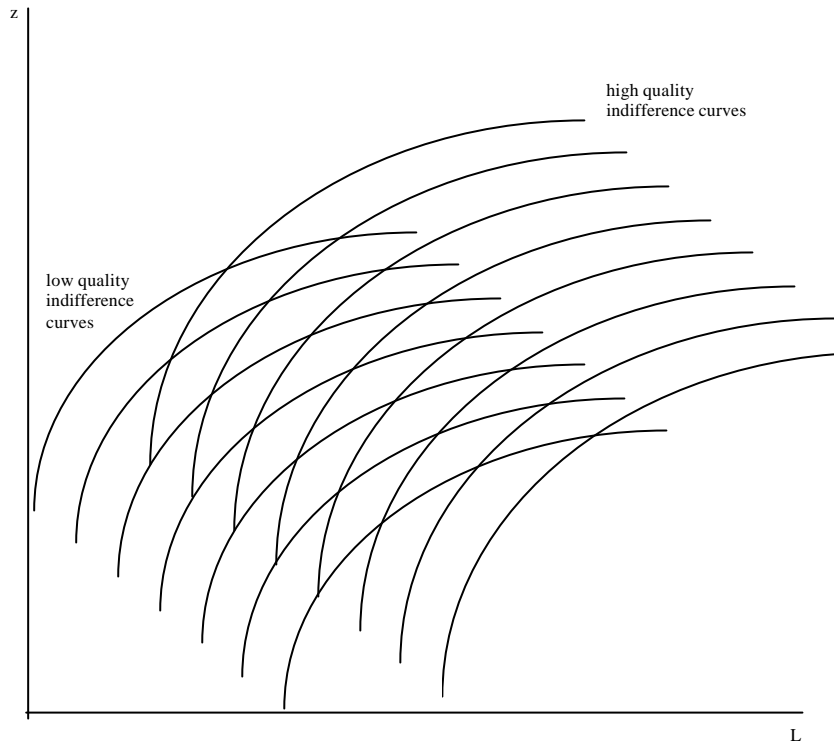
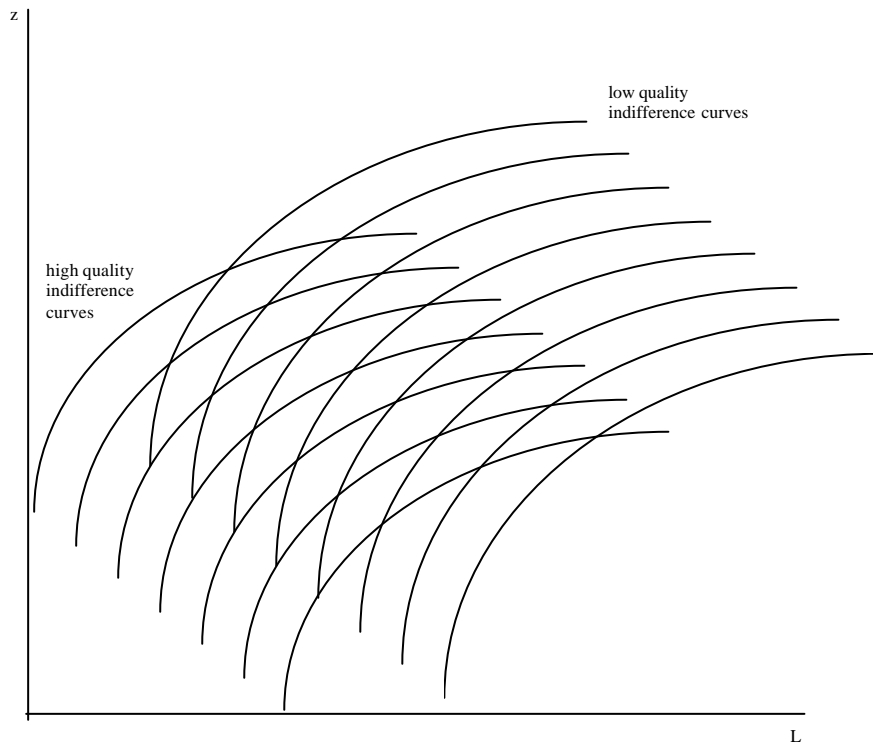


Figure Erreur ! Argument de commutateur inconnu. **a. Firms' indifference curves ($H=I$)**



Erreur ! Argument de commutateur inconnu. **b. Firms' indifference curves ($H=E$)**

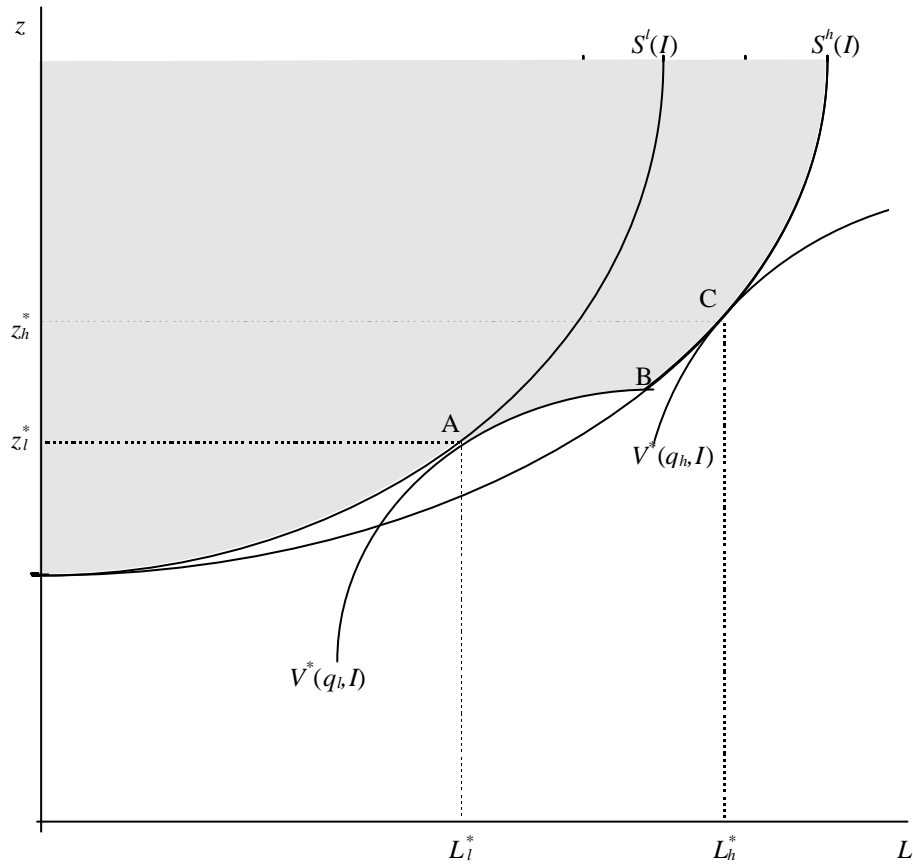
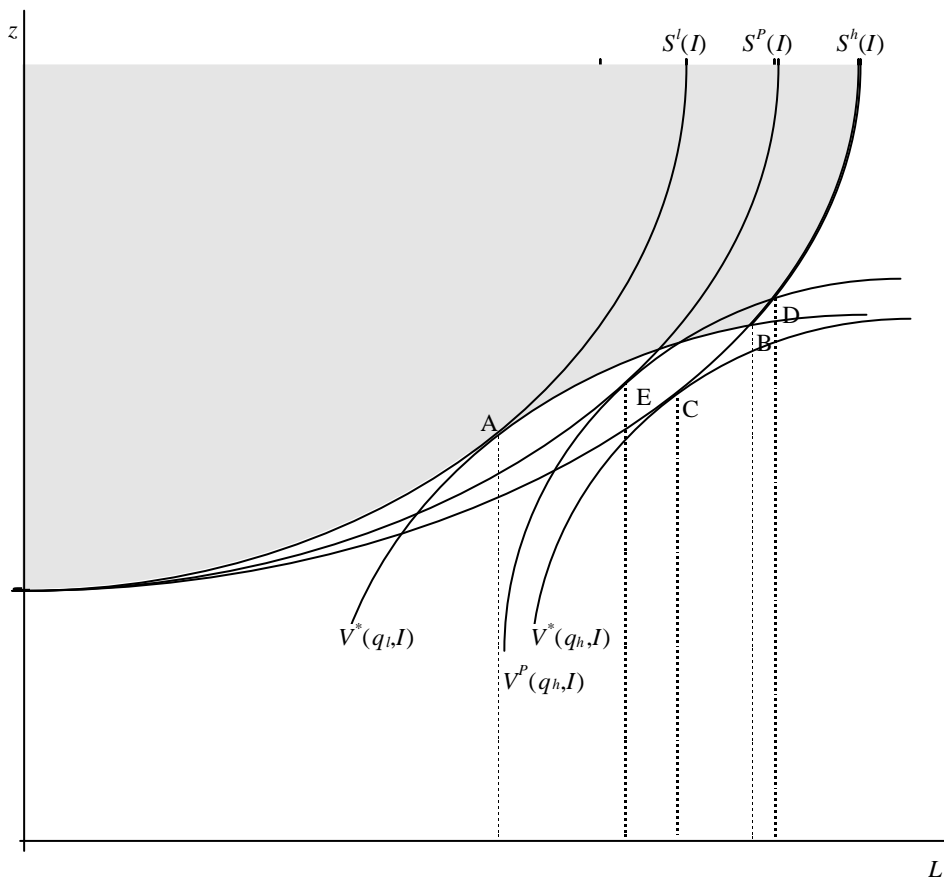
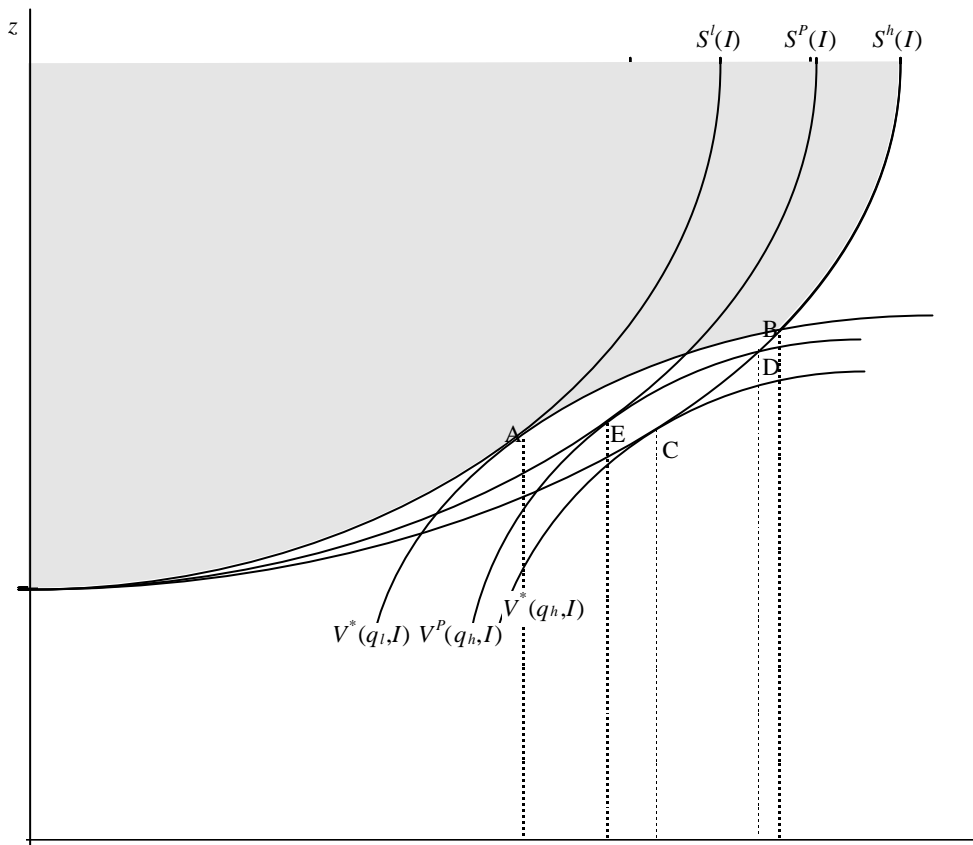


Figure Erreur ! Argument de commutateur inconnu. **a. Optimal separating equilibrium ($H=I$)**



Erreur ! Argument de commutateur inconnu. **b. Sub-optimal separating equilibrium ($H=I$)**



Erreur ! Argument de commutateur inconnu.c. Pooling equilibrium ($H=I$)^L

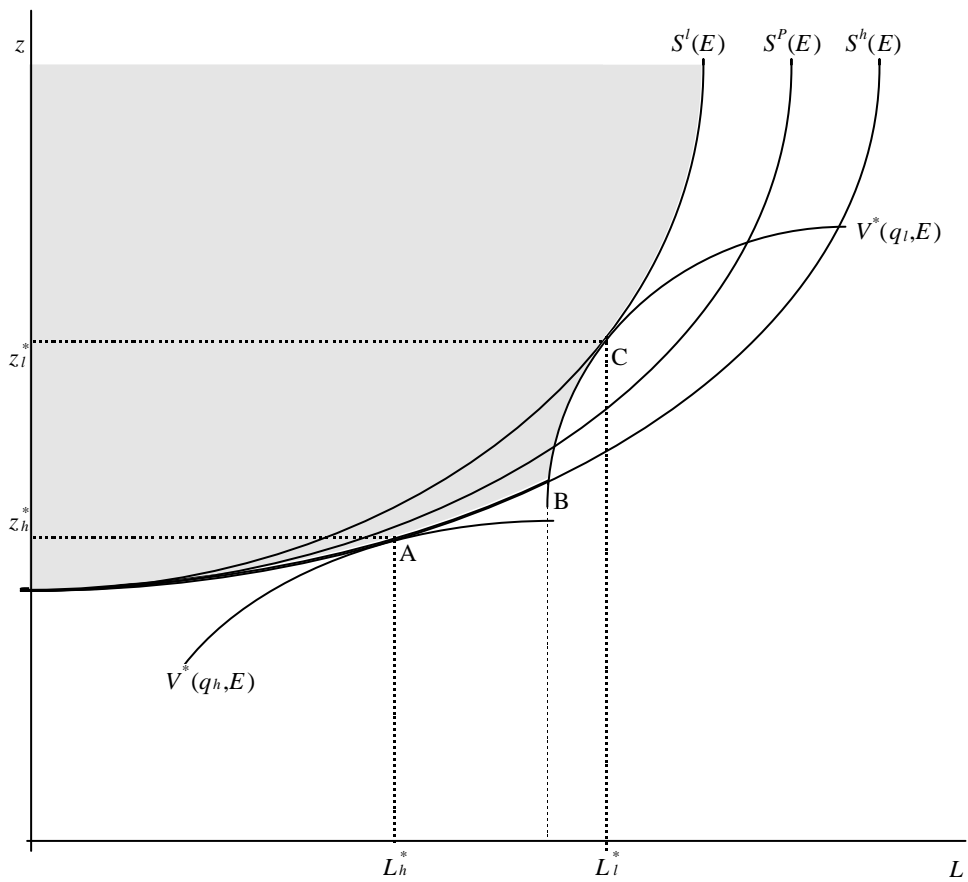
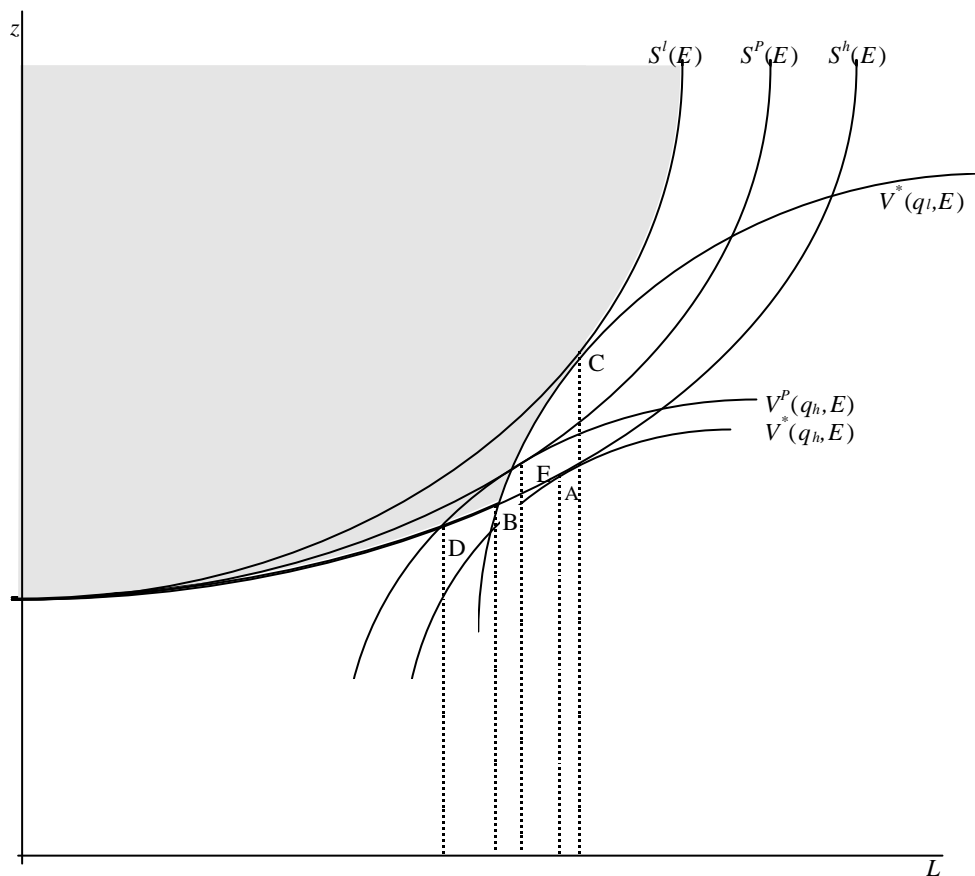
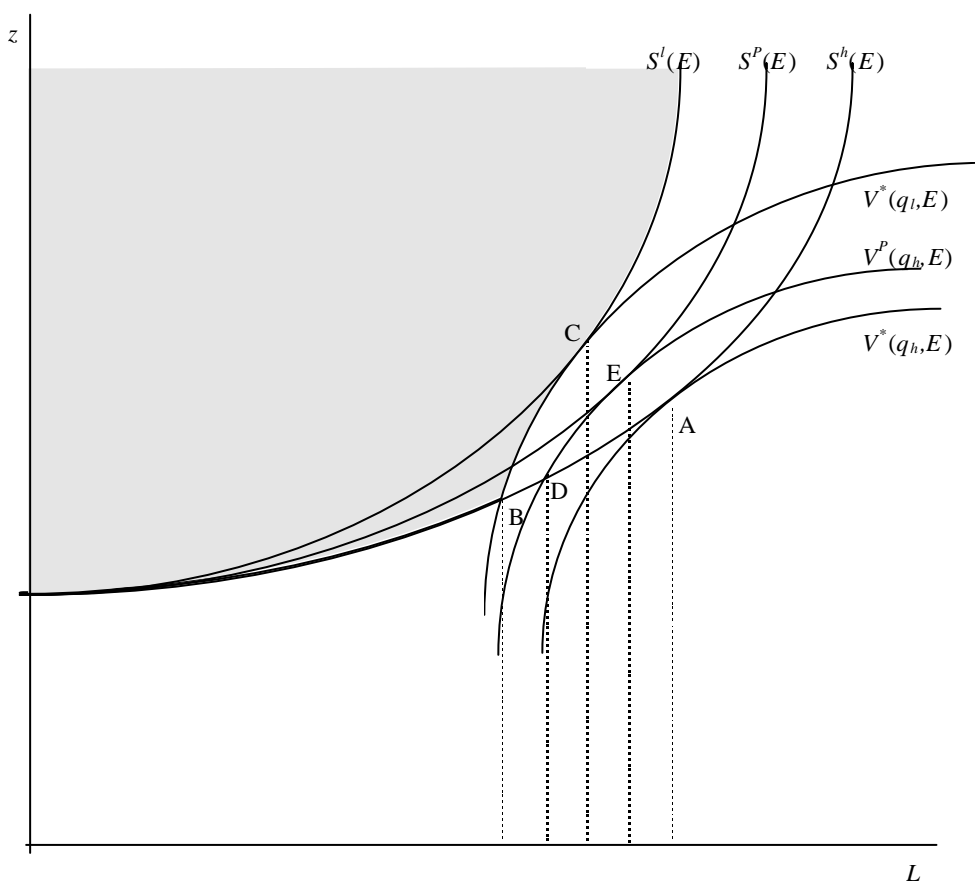


Figure Erreur ! Argument de commutateur inconnu.a. Optimal separating equilibrium ($H=E$)



Erreur ! Argument de commutateur inconnu. **b. Sub-optimal separating equilibrium ($H=E$)**



Erreur ! Argument de commutateur inconnu. **c. Pooling equilibrium ($H=E$)**

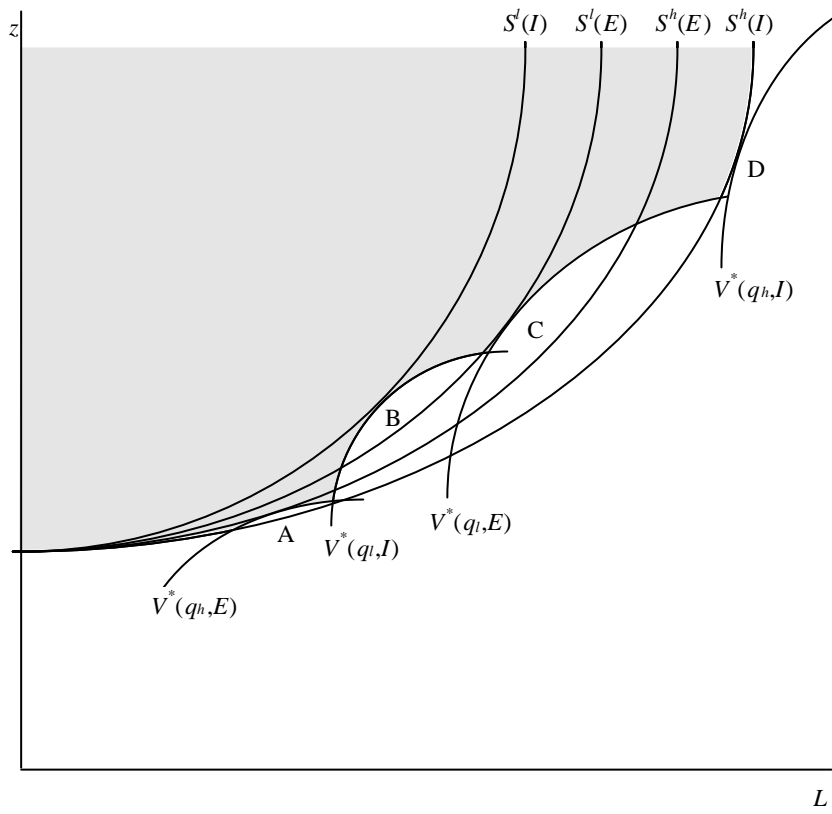


Figure Erreur ! Argument de commutateur inconnu.. **Optimal separating equilibria with non-observable technology**

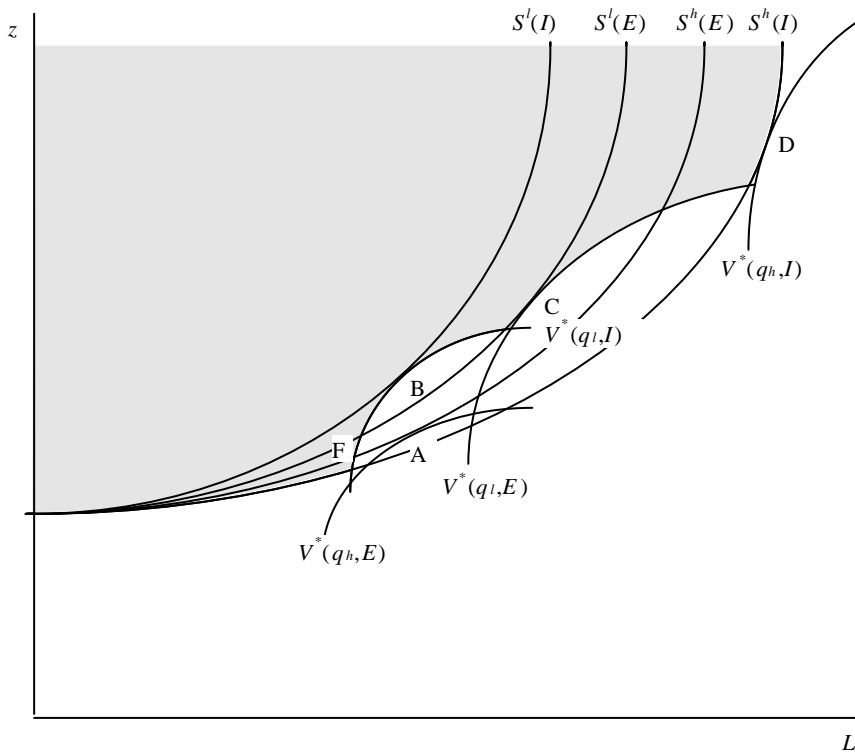


Figure Erreur ! Argument de commutateur inconnu.. **Separating equilibria that are optimal when technology is observable and that become sub-optimal (for high-quality *E*-firms, point F) when technology is non-observable**

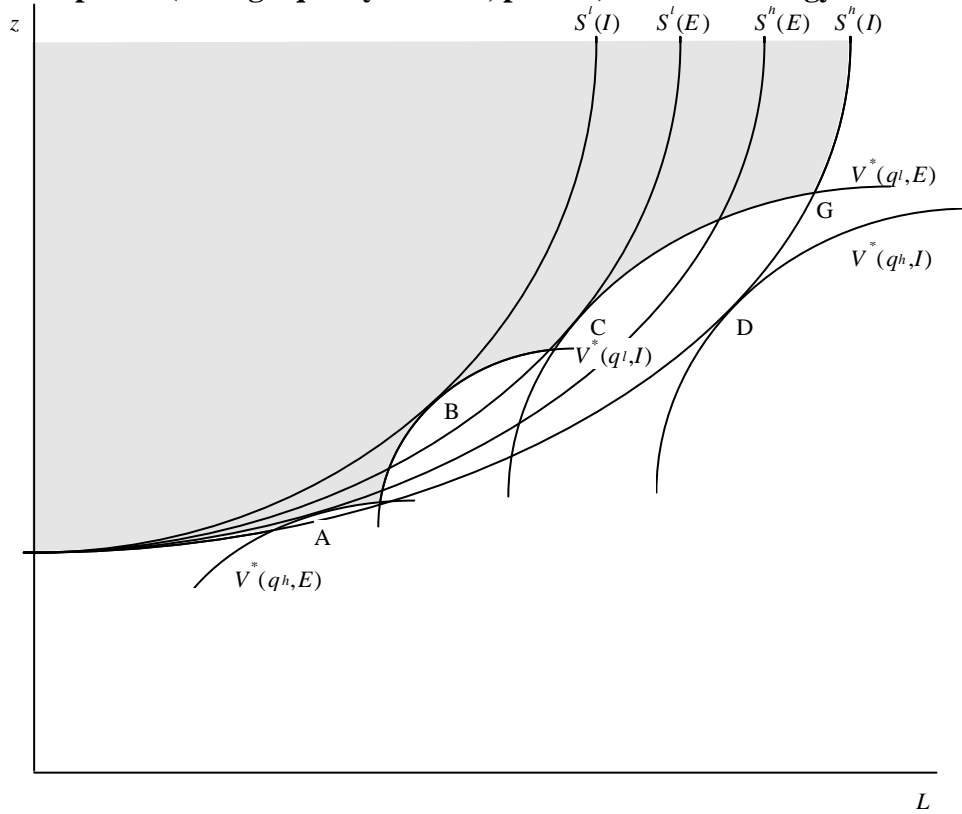


Figure Erreur ! Argument de commutateur inconnu.. **Separating equilibria that are optimal when technology is observable and that become sub-optimal (for high-quality *I*-firms, point G) when technology is non-observable**

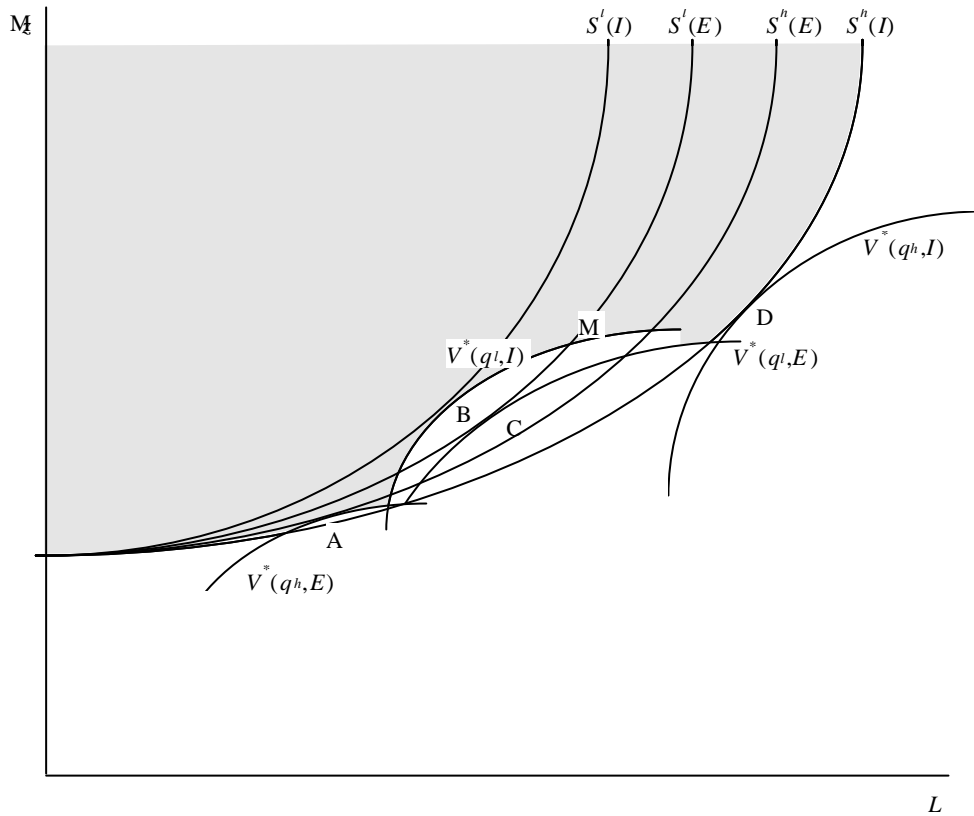


Figure Erreur ! Argument de commutateur inconnu.. **Separating equilibria that are optimal when technology is observable and that become sub-optimal (for low-quality E-firms, point M) when technology is non-observable**

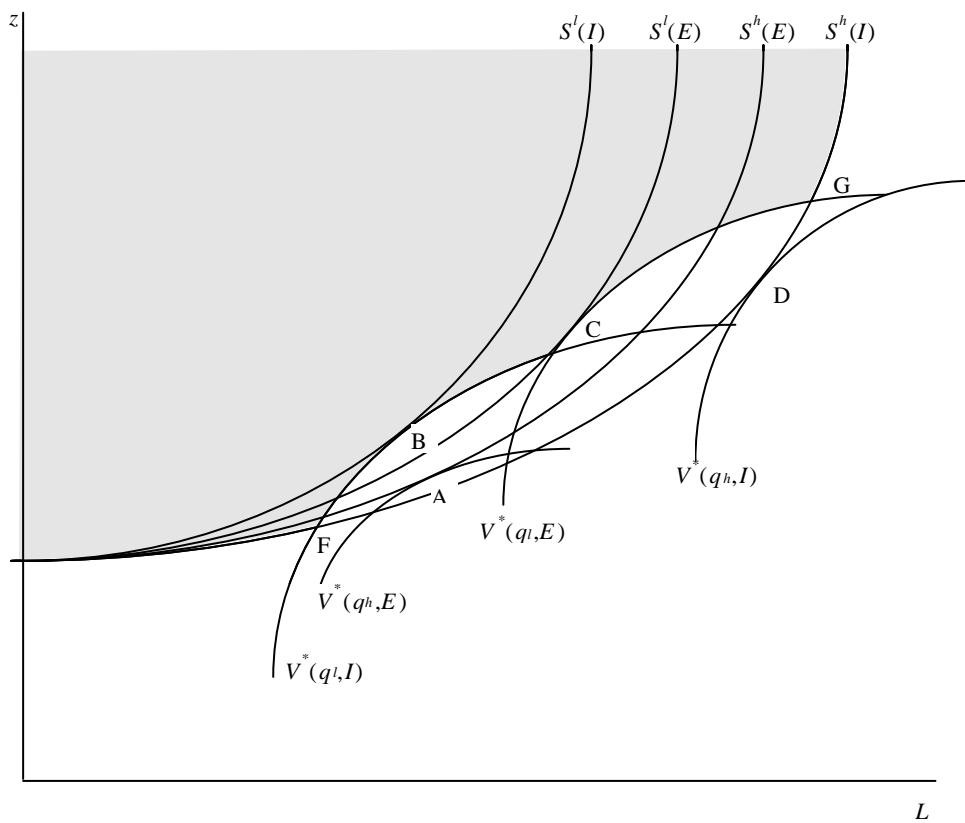


Figure Erreur ! Argument de commutateur inconnu.. **Separating equilibria that are optimal when technology is observable and that become**

sub-optimal (for both high-quality *I*-firms and *E*-firms, points F and G) when technology is non-observable

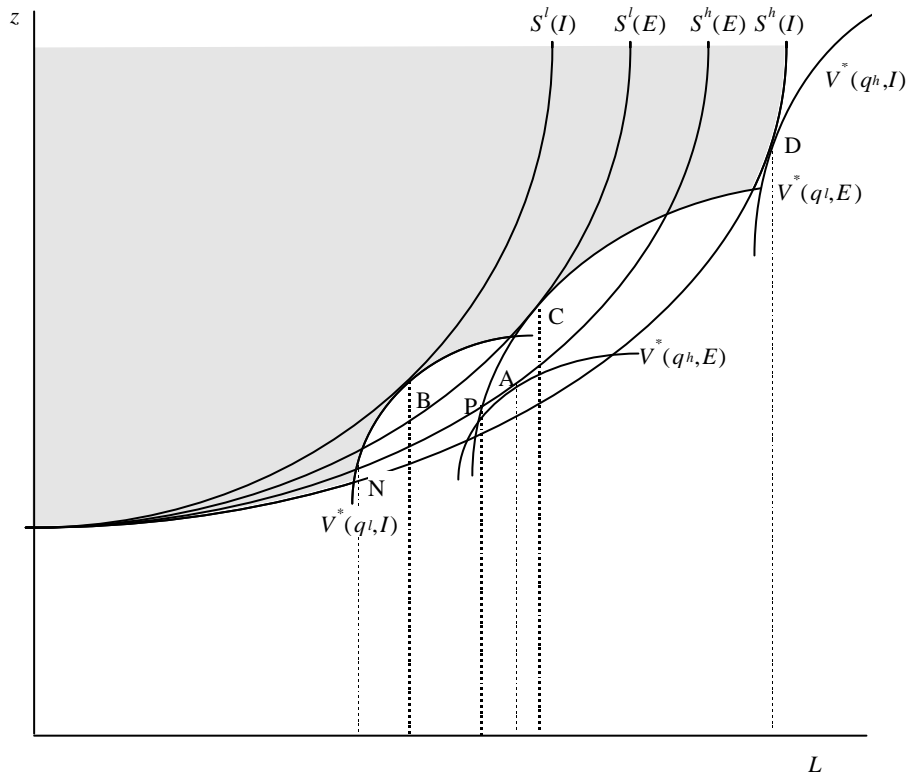


Figure Erreur ! Argument de commutateur inconnu.. Separating equilibria where one is sub-optimal (for high-quality *E*-firms) when technology is observable (point P), and is sub-sub-optimal (point N) when technology is not observable

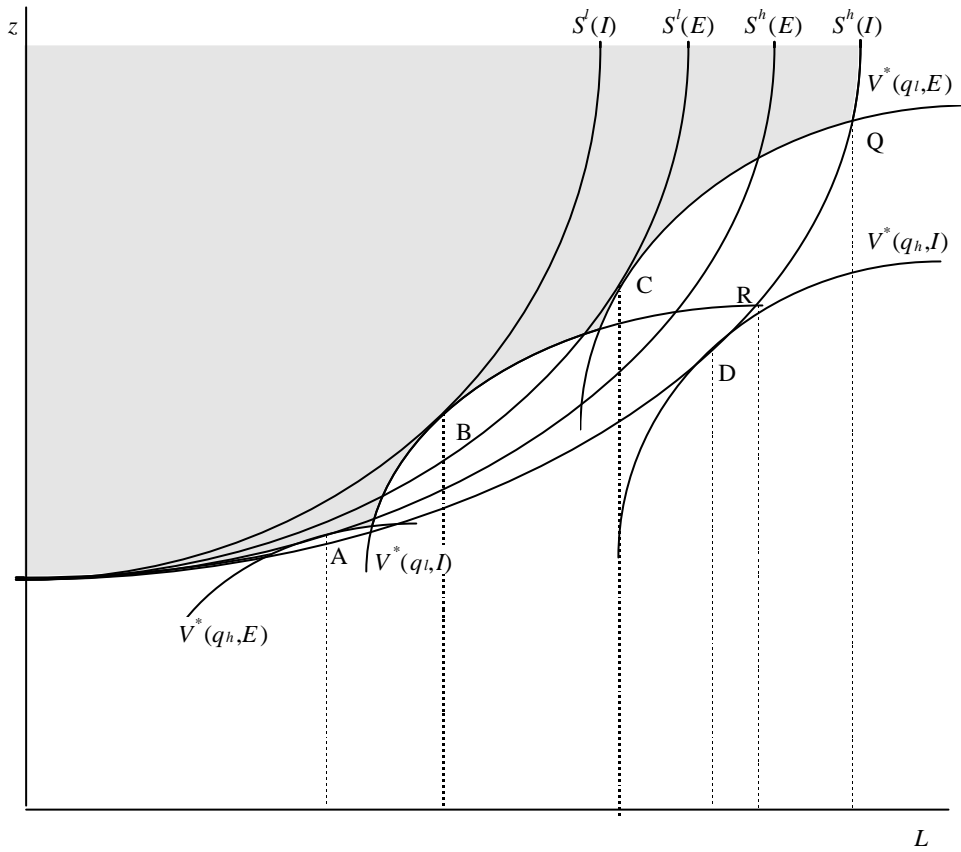


Figure Erreur ! Argument de commutateur inconnu.. **Separating equilibria where one is sub-optimal (for high-quality I -firms) when technology is observable (point R), and is sub-sub-optimal (point Q) when technology is not observable**