Definition *InformationProduction,BankingIndustryStructureandCreditAllocation.*

G. Chiesa

Dipartimento di Scienze Economiche

Introduction

Does the removal of intra-state entry barriers increase welfare? Will all banks survive? Will it lead to a consolidation of the banking industry? The experience of credit market deregulation has not been always successful. Credit market liberalisation, via the removal of entry barriers, of limitations of activity and of markets for funding is generally recognised at the origin of banking crises like the American Savings & Loans, the Scandinavian countries' banking crises at the end of the 80's, the Japanese crises not yet resolved and more recently and on a large scale the East-Asian countries. However there are doubts about whether these problems arose because of liberalisation per se or rather because of failure of prudential regulation in disciplining banks whose lenders were protected by deposit insurance and State guarantees. Financial markets integration in Europe is still at too an early stage to evaluate, but certainly the very clear process that has emerged is a consolidation of the banking system. Moreover, such a process appears to take place worldwide and somehow be the response to anticipation of future competition. Berger, Kashyap and Scalise (1995) documents a striking amount of consolidation in the American banking industry over the 1979-94 period and attribute this consolidation to the relaxation of intra-state entry barriers, moreover consolidation was favoured by regulators through the easing of the process for approving mergers. footnote Does a concentrated banking industry dominate a fragmented one? This is the question this paper seeks to address.

We shall consider banks as limited-liable intermediaries that borrow form investors and lend to firms, and concentrate on bank's (incentive-compatible) information production about borrowers, that is on bank's (incentive-compatible) screening activity. We shall show that when perfectly diversified credit portfolios cannot be constructed, credit allocation depends on aggregate bank capital and on the number of banks that can operate in the same market. In particular: i) a concentrated banking industry, one where aggregate bank capital is held by few banks, leads to credit allocation closer to the first-best optimum; and ii) in the absence of banking industry consolidation, the removal of intra-state entry barriers reduces welfare. Moreover, not all independent banking organisations that were viable in formerly protected national markets remain so when markets are integrated.

One reason that makes fully diversified credit portfolios unfeasible is macroeconomic (systematic) risk. In the economy considered in this paper, a project return depends both on the intrinsic quality of the project – firm – and on overall economic performance. A firm, regardless of its quality, is more likely to succeed in a boom than a recession, and firms that manage to succeed in a recession are the fit ones. By (privately) engaging in costly information production about firms' quality, that is by screening loan applicants, a bank can assess the quality of borrowing firms but not its asset-portfolio return which depends also on overall economic performance. The bank pursues the interests of its shareholders (insiders), it may then have an incentive to take bad risk: it may be tempted to bet that macroeconomic factors will support firms' performance, avoid costly search for good type firms – financing *de facto* negative net-present value projects – and shift resulting losses on (outsiders) investors. At equilibrium projects that are undertaken have positive net-present value (that is, agents' participation constraints are satisfied): banks that are active, find it profitable to engage in information production about borrowers. The questions are: (i) what is a bank's incentive constraint , (ii) how is this affected by banking industry and market structures, and what the implications on credit allocation.

The analysis of a bank's profit maximising choice provides the answer to question (i). The bank finds it incentive-compatible to engage in information gathering (that is search for good-type firms) only if the amount of capital that it invests in its asset portfolio does not fall below a threshold level which is decreasing in its profit margin (the wedge between bank's lending and borrowing rates) and increasing in the cost of searching for good type firms. The lower the average quality of loan applicants, the higher this cost. The bank's incentive-compatibility constraint thus sets a ceiling on the amount of outside financing that the bank can raise and hence on the amount of lending it can make. The higher bank capital and the profit margin and the better the quality of its loan applicants, the higher its lending volume.

With regard to (ii), the key observation is that when several banks are active in the same market, a rejected borrower does not exit he applies to a second bank and if rejected he tries again with a third bank.... The screening and re-application process acts to worsen banks' pool of applicants, it then acts to increase banks' costs of searching for good-type borrowers and thereby to lower the amount of (incentive-compatible) lending, for any given amount of bank capital. The higher the number of banks that can operate in the same market, either because the banking industry is more fragmented and/or intra-state entry barriers are removed, the higher the short-fall of aggregate lending from the first-best optimum and the lower the average quality of borrowing firms. Moreover, at equilibrium, an active bank (that is a bank that engages in lending) is necessarily viable, i.e. it finances firms whose projects have positive net-present value. This sets an upper bound on the number of banks that are active at equilibrium: banks that would be viable in protected national (or local) markets are non-viable when intra-state entry barriers are removed.

These results follow from two crucial observations: (i) the information externality of screening and its adverse pool effects; (ii) the size of a bank (the amount of lending it undertakes) is endogenously constrained: it cannot exceed a ceiling positively linked to its capital, to its (endogenous) profit margin and to the quality of its loan applicants, where the latter depends on banking industry and market structures.

Observation (i) has been made and explored before by Broecker (1990), Riordan (1993) and Nakamura (1996) for given statistical properties' of banks' screening technologies, and by Gehrig (1997) for endogenous screening intensity from a continuum of potential levels. These papers abstract from banks' agency problems by assuming that banks finance firms with their own (unlimited) capital. The size of a bank is unconstrained and the welfare effects of competition depend on the specificity of the model with regard to whether lenders can react to each other (Broecker (1990)) and whether they observe a continuous signal and base their lending decisions on a chosen signal's cut-off (Riordan (1993)); Gehrig (1997) provides conditions under which screening efforts are reduced by competition.

Observation (ii) is specific to this paper, it follows from banks being limited liable and subject to delegation – agency – problems. It provides a role for banking industry structure, independent of competition issues. We do indeed find that banks compete more aggressively — banks' profit margins shrink — when the industry structure is more fragmented. This reinforces the conclusion that aggregate lending and its average quality are lower with a fragmented industry structure than with a concentrated one.

Banks' agency problems have been examined earlier. The financial intermediation literature has stressed the role of risk diversification in solving the conflict of interests between the bank and investors and noted that if all risk could be diversified, agency problems would disappear (Diamond 1984; Ramakrishnan and Thakor 1984; Bhattacharya and Thakor 1993). Boot and Greenbaum (1993) suggests that bank's reputational concerns could ameliorate moral hazard problems and substitute rents in incentivating monitoring of ongoing projects. Bernanke and Gertler (1987) are the first to point out that bank capital affects the scale of banking. In their model, investors' payoffs cannot be conditioned on banks' risky asset returns since these are not public information, bank capital (perfectly) collateralises bank liabilities. Chiesa (1997) analyses banks' agency problems and regulatory issues when risk diversification is limited and banks monitor firms' choices of projects. In contrast to the previous literature, this paper concentrates on banks' ex ante information production.

The paper is structure as follows. Section 2 presents the model. Section 3 derives the incentive-compatibility constraint for bank information production in the (simplest) case of a single bank. Section 4 analyses credit allocation in a one-bank economy. This can be one where there is a single (monopoly) bank, or, equivalently, one with several (monopoly) banks operating in protected national – local – markets. Section 5 analyses credit allocation in a two-banks economy. This economy differs from the previous one either because the banking industry is more fragmented, or equivalently because intra-state entry barriers are removed. Section 6 concludes.

The Model

We consider a one-period credit market consisting of M entrepreneurs or firms, I investors and n banks. Agents are risk neutral, limited liable and maximise their end-of-period expected wealth. Each entrepreneur can undertake an investment project which requires one unit of resources. Each investor is endowed with one unit of resources, which can either be stored at the gross return R_d or deposited in a bank. Investors are large in number ($I > M$). A bank, i, lends to entrepreneurs and borrows from investors. It acts on behalf of its shareholders (insiders), whose equity holdings constitute the bank's endowment of (inside) capital. This capital can be used to finance entrepreneurs or stored at the gross rate R_d per unit. We shall refer to R_d as to the risk-free (gross) interest rate. Banks are endowed with an amount of aggregate capital, A, that satisfies $A > 0$, and is symmetrically distributed across banks. The higher n , the more fragmented the banking sector.

Project and Screening Technologies

Bank lending consists of project financing. A project requires one unit of resources at the beginning of the period and delivers a (random) return at the end of the period. The realisation of this return depends on the macro-state realisation at the end of the period $\theta \in \left\{\underline{\theta},\overline{\theta}\right\}$, where $\overline{\theta}$ occurs with probability p, and the project type $h \in \{g,b\}$. A good (type g) project delivers an observable and verifiable return of x both in (the good macro-state) $\overline{\theta}$ and in (the bad macro-state) $\underline{\theta}$; a bad (type b) project delivers an observable and verifiable return of x in $\overline{\theta}$ and of zero in $\underline{\theta}$. The proportion of type g projects – firms – is λ.

*Assumption A*1

(*an un-screened project has negative net present value*:)

$$[p + (1 - p)\lambda]x < R_d \qquad \#$$

No self-selection devices are available. However, a bank has access to a (costly) screening technology that allows to determine, albeit imperfectly, the credit-worthiness of a firm. The use of this technology imposes a non-pecuniary effort cost $f > 0$ to the bank and produces a (privately observed) signal which can be either G (the test result is a success) or B (the test result is a failure). The screening technology is completely described by:

$$\Pr(G|g) = \Pr(B|b) = \eta \qquad \#$$

Let us consider a bank examining a firm randomly taken from the original population. The probability according to which the bank observes G (the test result is a success) is $\Pr(G)$:

$$\Pr(G) = \eta\lambda + (1 - \eta)(1 - \lambda) \qquad \#$$

the (posterior) probability that the firm is of type g , given the observation of G is then $\Pr(g|G)$:

$$\Pr(g|G) = \frac{\eta\lambda}{\Pr(G)} \qquad \#$$

*Assumption A*2

(*examining a project once and undertaking it only if the test result is a success is a positive net-present value activity*:)

$$\Pr(G)\{[p + (1-p)\Pr(g|G)]x - R_d\} - f > 0 \qquad \#$$

The expression {} in (5) is the net-present value of a project conditional upon examining the project once and the test result being a success. The left-hand side of (5) is then the expected value of the overall surplus generated by screening a project once and undertaking it only if the test result is a success – the signal observed is G.

By contrast, a firm that failed the test should exit. The (posterior) probability that a firm is of type g conditional upon being examined twice, failed one test and passed the other is $\Pr(g|G,B)$:

$$\Pr(g|G,B) = \frac{\eta(1-\eta)\lambda}{\eta(1-\eta)\lambda + \eta(1-\eta)(1-\lambda)} = \lambda \qquad \#$$

and therefore a firm that failed the credit-worthiness test has a negative-net present value project (by (1)). However, a firm that failed the test wishes to take further chances with other banks:

*Assumption A*3
Signals – test results – are independently distributed

*Assumption A*4
(*bank capital is insufficient to fund all socially valuable projects*)

$$A < \Pr(G)M \qquad \#$$

In what follows we shall refer to information-gathering (credit-worthiness examination) as monitoring.

Our assumptions imply that (i) monitoring is value-increasing, unobservable and costly to the bank, (ii) the average loan return is uncertain (its realisation depends on the macro-state realisation), and (iii) a firm that failed to borrow from a bank (that failed the test with a bank) may succeed in obtaining a loan from a different bank (may pass the test with a different bank). (i) is a necessary condition for a delegation – bank-moral hazard – problem; (ii) prevents the solution of this problem through diversification and drives the paper's result of a bank incentive-based lending quality positively linked to bank capital, to bank profit margin and to the average quality of loan applicants; (iii) implies that the average quality of loan applicants decreases as the number of banks that are active in the same market increases. It drives our results about banking fragmentation and intra-state entry barriers removal.

The Credit Game

The game proceeds as follows.

Stage 1: each bank $i = 1,..n$ announces its lending and borrowing rates; these offers are observed by all. Firms and investors apply to banks. *Stage* 2: i chooses its lending policy (L_i, τ_i), $\tau_i \in \{1,0\}$, where L_i is i's lending volume and $\tau_i \in \{1,0\}$ is the proportion of lending that i allocates to successfully screened firms — if $\tau_i = 1$, then i monitors lending: the firms that i finances are those screened successfully; if $\tau_i = 0$, then i does not monitor: i allocates its lending to un-screened firms thus avoiding information-gathering costs altogether.

A bank could arguably monitor a fraction of its lending, i.e. $0 < \tau_i < 1$. However, i's strategy of monitoring some of its loans — that is allocate $\tau_i L_i > 0$ loans to successfully screened firms and $(1-\tau_i)L_i > 0$ to un-screened firms — is strictly dominated by either monitoring all loans, $\tau_i = 1$, or none, $\tau_i = 0$. This will be clarified in Section 3 (see footnote 3).

The benchmark: full liability

Let R^{FB} denote the lending rate that allows to recoup in expected value the resources invested in examining a firm once and funding it only if the test result is a success; R^{FB} is implicitly given by:

$$\Pr(G)\{[p + (1-p)\Pr(g|G)]R^{FB} - R_d\} - f = 0$$

i.e.

$$R^{FB} = \frac{R_d + \frac{f}{\Pr(G)}}{[p + (1-p)\Pr(g|G)]} \qquad \#$$

where $\Pr(G)$ and $\Pr(g|G)$ are given by (3) and (4), respectively.

Proposition *Let banks be fully liable, then the First Best optimum is attained: i) only one bank is active. The active bank lends the socially optimal amount $L^{FB} = Pr(G)M$ allocated entirely to successfully screened firms — the bank's profit-maximising choice is $\tau^* = 1$; ii) all financed firms borrow at the zero profit rate R^{FB} whenever the number of banks, n, satisfies $n \geq 2$.*

[Proof] If a bank is fully liable, then it internalises the consequences of its lending policy, and therefore if it chooses to be active then it examines firms' credit worthiness and lends to successfully screened firms (because of assumptions A1-A2). When there are more than one bank, at Stage 1 banks offer the interest rate under which a single bank would make zero profits, i.e. R^{FB}, and at Stage 2, $n-1$ banks exit (because of (6) and assumption A1) [End Proof]

The crucial point is that full liability eliminates delegation – moral-hazard – problems, a bank is unconstrained in the lending it can make and this implies that it can cover the whole market, financing all firms whose test result is a success. Let i be the bank that does not exit, then if $j \neq i$ would choose not to exit, the firms that it would end up lending to would be those rejected by i and these firms have negative net-present value projects (by (6) and assumption A1). Proposition 1 would continue to hold with limited liable banks, (only) if project returns were independently distributed, i.e. if the only possible macro-state realisation were $\underline{\theta}$ (that is if $p = 0$). In fact, if project returns are uncorrelated, a fully-diversified credit portfolio whose return is certain can be constructed thus eliminating bank-moral hazard (by the same reasoning in Diamond (1984)).

Incentive-Based Lending Constraint

Banks are limited liable and a credit-portfolio outcome depends on overall economic performance (project returns are correlated): the bank may find it profitable to bet that macroeconomic factors will support the performance of firms (bet that the macro-state realisation will be the good one $\overline{\theta}$), thus avoiding costly information gathering (financing *de facto* negative net-present value projects) and shifting resulting losses on investors. This section focuses on this problem in the (simplest) case of a single bank. The key result we shall derive is that the bank finds it profitable to engage in (costly) information gathering – monitor – only if its lending volume does not exceed a critical value which is positively linked to its capital, to its profit margin (the wedge between lending and borrowing rates) and to the average quality of its loan applicants.

We first observe that an un-screened (successfully screened) project has negative (positive) net present value (by assumptions A1-A2), and an equilibrium cannot be one where agents are worse off with respect to their status-quo payoffs:

Proposition *At any equilibrium, firms that are financed are necessarily those that have been screened successfully: the bank's profit-maximising choice is $\tau^* = 1$.*

[Proof] Directly from assumptions A1-A2. [End Proof]

The single bank's pool of loan applicants coincides with the original population of firms whose share of good (type g) firms is λ. The probability according to which the bank examining a firm observes the (good) signal G (the test result is a success) is then $\Pr(G)$ as given by (3), and the (posterior) probability that the firm is of type g, given the observation of G is $\Pr(g|G)$ as given by (4). A corollary to Proposition 2 is then that the bank's (gross) lending rate, R, satisfies:

$$x \geq R \geq R^{FB} \qquad \#$$

i.e. R allows to recoup in expected value the cost of resources invested in examining a firm once and funding it only if the test result is a success. The second corollary is that if at equilibrium the bank's lending volume is L, then the number of firms screened by the bank is $\frac{L}{\Pr(G)}$ and the bank's monitoring costs are $f\frac{L}{\Pr(G)}$. footnote We shall refer to $\frac{f}{\Pr(G)\Pr(g|G)} \equiv \frac{f}{\eta\lambda}$ as the cost of searching for a (good) type g firm. The better the pool of applicants, i.e. the higher the proportion of type g firms, λ, the lower this cost. The third corollary is that at equilibrium the lending volume cannot exceed the First Best optimum level $L^{FB} \equiv \Pr(G)M$, that is the lending volume that would be attained by screening the whole population of firms and funding the firms whose test result is a success.

Proposition 3 below derives the bank's monitoring-incentive constraint, i.e. the constraint that is necessarily satisfied in an equilibrium (by Proposition 2). The key observation is that the bank is limited liable and a monitored lending portfolio, $(L, \tau = 1)$, performs better than an unmonitored one, $(L, \tau = 0)$, in the (bad) macro-state $\underline{\theta}$. Therefore:

(**i**) if the bank's profit-maximising choice is $\tau^* = 1$ (which is possible only if $L \leq L^{FB}$), that is if the bank finds it optimal to engage in information gathering so as to allocate its lending to successfully screened firms, then the bank is solvent with probability one and the risk-free interest rate R_d satisfies the investor's participation constraint. Let the bank's deposit rate be R_d, then its expected profits conditional upon monitoring being the profit-maximising choice are:

$$\begin{aligned}\pi(R, R_d|(L, \tau^* = 1)) &= p[RL - R_d(L-A)] + (1-p)[\Pr(g|G)RL - R_d(L-A)] + \\ &\quad - \frac{f}{\Pr(G)}L - AR_d \qquad \# \\ &= \left\{[p + (1-p)\Pr(g|G)]R - \left(R_d + \frac{f}{\Pr(G)}\right)\right\}L \qquad ;\end{aligned}$$

(**ii**) if the bank's profit-maximising choice is $\tau^* = 0$, then it is necessarily the case that its lending volume, L, is sufficiently large with respect to capital, A, that the bank is insolvent in $\underline{\theta}$ (if it were solvent then it would suffer all the consequences of financing negative net present value projects and its profit-maximising choice would necessarily be monitoring, by A1-A2 and $R \geq R^{FB}$). Hence if $\tau^* = 0$, then the bank is solvent *only if* the macro-state realisation is the good one $\overline{\theta}$, i.e. the state in which loans perform independently of whether they have been granted to successfully screened firms (monitored) or not. In state $\underline{\theta}$, the bank loses its capital so its expected profits are footnote :

$$\pi(R, R_d|(L, \tau^* = 0)) = p(R - R_d)L - (1-p)AR_d \qquad \#$$

Then, the bank's profit-maximising choice is $\tau^* = 1$ only if:

$$\pi(R, R_d|(L, \tau^* = 1)) \geq \pi(R, R_d|(L, \tau^* = 0)) \qquad \#$$

that is iff:

$$\frac{A}{L} \geq 1 - \Pr(g|G)\frac{R}{R_d} + \left[\frac{1}{(1-p)\Pr(G)R_d}\right]f \qquad \#$$

In summary, the bank engages in information gathering — lends to successfully screened firms – only if the amount of capital that it invests in its asset portfolio, i.e. $\frac{A}{L}$, does not fall below the threshold level $c(R, R_d, \lambda)$:

$$c(R, R_d, \lambda) = 1 - \Pr(g|G)\frac{R}{R_d} + \left[\frac{1}{(1-p)\Pr(G)R_d}\right]f \qquad \#$$

Note that $c(R, R_d, \lambda)$ falls below one (by assumptions A1-A2 and $R \geq R^{FB}$), it is decreasing in the bank's profit margin, $\frac{R}{R_d}$, and it is increasing in the cost of searching for a type g firm, i.e. the higher $\frac{f}{\eta\lambda}$ the higher $c(R, R_d, \lambda)$.

Define:

$$L^c(R, R_d; A, \lambda) \equiv \begin{cases} \frac{A}{c(R, R_d, \lambda)} & , \quad if \quad c(R, R_d, \lambda) > 0 \\ \\ \infty & , \quad if \quad c(R, R_d, \lambda) \leq 0 \end{cases} \qquad \#$$

$L^c(R, R_d; A, \lambda)$ is the maximum amount of lending that satisfies the bank's information gathering-incentive constraint — $L^c(\quad)$ satisfies the incentive constraint (12) (or equivalently, (13)) at equality. The higher bank capital and the lower the cost of searching for a type g firm, the higher $L^c(\quad)$.

The above discussion is summarised in the following proposition

Proposition *If the amount of lending that the bank undertakes satisfies*:

$$L \leq min[L^c(R, R_d; A, \lambda), L^{FB}] \qquad \#$$

then the bank's profit maximising choice is to engage in information gathering about borrowing firms, $\tau^* = 1$, *the bank is solvent with probability one and the risk-free interest rate* R_d *satisfies the investor's participation constraint. By contrast if inequality* (16) *fails to hold then the bank's profit maximising choice is to abstain from information gathering,* $\tau^* = 0$.

A corollary to Propositions 2-3 is that the maximum amount of funds that (rational-unprotected) investors are willing to offer the bank is D^m:

$$D^m = \min[L^c(R, R_d; A, \lambda), L^{FB}] - A \qquad \#$$

that is the bank's loanable funds satisfy the (incentive) condition (16). Note that $D^m > 0$ (because $c(R, R_d, \lambda) < 1$, and $L^{FB} > A$). The second corollary is that bank's unit cost of funding is the risk-free interest rate R_d and its expected profits are given by (10).

Discussion

The bank finds it incentive compatible to engage in costly information gathering about the firms to lend to – monitor – only if the stake it takes in its lending is sufficiently high: the amount of capital it invests in its asset portfolio does not fall below the threshold level $c(R = x, R_d, \lambda)$ as given by (14). This also implies that the bank is solvent with probability one and hence its unit cost of funding is the safe-interest rate R_d. For any given amount of bank capital, there is then a maximum amount of outside financing that the bank can raise subject to the constraint that it will find it incentive-compatible not to exploit investors: engage sufficiently in information gathering (monitoring) so as to lend to firms whose projects have positive-net present value. This maximum is D^m as given by (17), and it is also the maximum amount of funds that unprotected investors are willing to provide, given their rational beliefs about the bank (monitoring) choice. The lower the return that the bank gets by engaging in monitoring, i.e. the lower the wedge between the bank's lending and borrowing rate and the higher the cost of searching for a (good) type g firm, the higher the stake it has to take in order to be a credible monitor, consequently the lower the amount of outside financing that the bank can raise and the amount of lending it can make, for any given level of capital A. Lending is then increasing in bank capital and in the bank's profit margin, $\frac{R}{R_d}$, and it is decreasing in the cost of searching for a (good) type g firm. The lower the average quality of loan applicants, i.e. the lower λ, the higher this cost.

We assumed that the bank raises outside financing by issuing deposits (i.e. debt). If the bank were to raise outside financing by issuing equity, then it would share the return of its costly-information gathering activity (monitoring) with investors and the bank-incentive problem would be exacerbated. The optimality property of debt financing when there is (hidden) provision of costly effort has been firstly noted by Innes (1990) and applied to banking by Gehrig (1996).

One-Bank Economy (Intra-State Entry Barriers)

We shall now derive the equilibrium for a one-bank economy. This could be (literally) one where there is a single monopoly bank holding capital *A* and facing the whole firm population, or one with *n* symmetrical (local) markets – States – each with a monopoly bank endowed with $\frac{A}{n}$, that is an economy with intra-state entry barriers. The two economies are by all means equivalent. For expositional convenience we shall refer to the one-bank economy in its literal sense.

The bank's expected profits are increasing in the lending rate and in the lending volume, the monopoly bank thus offers firms the (monopoly) lending rate *x* and raise the maximum amount of funds that (rational-unprotected) investors are willing to provide, i.e. $D^m = \min[L^c(R = x, R_d; A, \lambda), L^{FB}] - A$. Let L^m denote the equilibrium lending volume of the monopoly bank, $L^m = D^m + A$, and let λ^m denote the proportion of borrowing firms that are of type *g* (i.e. have positive net-present value projects).

Proposition *Aggregate lending in a one-bank economy is*:

$$L^m = min[L^c(x, R_d; A, \lambda), L^{FB}] \qquad \#$$

and the proportion of borrowing firms whose projects have positive-net present value is:

$$\lambda^m = Pr(g|G) = \frac{\eta\lambda}{\eta\lambda + (1-\eta)(1-\lambda)} \qquad \#$$

[Proof] Directly from Propositions 2-3 [End Proof]

The better the pool of applicants, that is the higher λ, the higher the proportion of borrowing firms that are of type *g* and the higher aggregate lending L^m.

The equilibrium deviates from the first-best optimum, that is $L^m < L^{FB}$, whenever $\frac{A}{M}$ satisfies:

$$\frac{A}{M} < c(x, R_d; A, \lambda)\,\mathrm{Pr}(G) \qquad \#$$

the higher the cost of searching for a type *g* firm and the lower the bank's profit margin, $\frac{x}{R_d}$, the higher the right-hand side of (20). We thus have that the short-fall of aggregate lending from the first-best optimum level is decreasing in bank capital/over aggregate-lending-demand, and in the bank's profit margin, and it is increasing in the cost of searching for a (good) type *g* firm.

We claimed that the one-bank economy (in its literal sense) is equivalent to one with with intra-state entry barriers, that is an economy with *n* symmetrical (local) markets – States – each with a monopoly bank endowed with $\frac{A}{n}$. This claim is indeed true: the cost of searching for a type *g* firm in each (local) market is exactly the same as that in the one-bank economy (because there is a single bank in each market), and each bank is endowed with $\frac{A}{n}$, the equilibrium lending volume in each market is then the fraction $\frac{1}{n}$ of L^m. Credit allocations are the same.

Discussion

Aggregate lending and its short-fall from the first-best optimum level vary along the business cycle. In a down-turn (in state $\underline{\theta}$) firms whose type has been mis-identified, that is (bad) type *b* firms incorrectly classified as good firms, do not repay their loans, the bank suffers loan insolvencies and decumulates capital. An adverse macroeconomic shock thus leads to a reduction of bank capital and to a lower lending volume, relative to the first-best optimum level.

We simplified by assuming that investment projects – firms – have the same size. Given this simplification, the cost of information gathering – screening – per lending unit is the same for all projects. Maintaining the (reasonable) assumption that screening costs are (mainly) fixed costs and allowing for projects'size etherogeneity, i.e. for the coexistence of "large" and "small" projects, implies that "smaller" projects have larger screening costs per lending unit. The bank then finds it optimal to give priority to large projects and if it rations (i.e., if $L^c(\quad) < L^{FB}$) , then it denies credit to small projects. This is consistent with the empirical evidence that following an adverse macroeconomic shock, small firms experience reduced access to credit, relative to larger firms (Bernanke, Gertler and Gilchrist 1996).

It should also be noted that in the economy we have considered there is no deposit insurance. If depositors were protected by deposit insurance, then prudential regulation would substitute the market (i.e. unprotected investors) in disciplining banks. The (optimal) regulatory scheme would impose the minimum incentive-based capital requirement $c(R = x, R_d, \lambda)$ as given by (14). A failure in prudential regulation would imply that the bank over-lends, does not engage in information gathering and thus finances negative-net present value projects. This does not imply that the bank will necessarily be insolvent, in fact it will be solvent in the lucky event that favorable macroeconomic conditions support firms' performance, i.e. the macro-state realisation is $\overline{\theta}$. Interpreting the state $\overline{\theta}$ as a price-bubble state, and the state $\underline{\theta}$ as the state in which the bubble bursts, may allow for an explanation of the recent Asian experience.

Banking Industry Fragmentation (No Intra-State Entry Barriers)

We now examine the effects of intra-state entry barriers' removal (or, equivalently banking-industry fragmentation). Will all banks survive? Will welfare increase?

To answer these questions, we shall examine a two-banks economy whose aggregate bank capital *A* is (for simplicity) held symmetrically by two banks. This economy differs from the one analysed so far, *only* for the number of banks that can operate in the same market. It could be generated by the integration of two formerly (monopoly) local markets, alternatively it could simply be an economy where the banking industry is more fragmented.

To allow for an explicit (pure strategy) equilibrium solution, we shall assume that at the price-setting stage banks move sequentially. Let bank 1 be the first mover, bank 2 observes the rate offered by bank 1 and chooses whether to be active and which rate to offer. A firm applies to one bank at a time. On the basis of the rates on offer, it chooses the bank to apply to and if rejected it applies to its second choice.

Suppose that both banks are active. Let *i* be firms' first choice and let *j* be firms' second choice, i.e. a firm applies to *j* only if it did not succeed in obtaining a loan from *i*. If banks offer different interest rates, then the lowest-rate setting bank will be firms' first choice, i.e. it will be *i*, with with probability one; if both banks offer the same rate then firms will randomise and a bank will be *j* with probability 1/2. The point being that with two active banks, no matter banks' lending rates, one of the two banks will be (firms' second choice) *j*.

The set of firms faced by *i* coincides with the original population whose share of good (type *g*) firms is λ. Therefore: (a) the minimum amount of capital that *i* must invest in its asset portfolio in order to find it incentive-compatible to engage in information gathering is exactly the same as that of the single bank, i.e. it is $c(R_i, R_d, \lambda)$ as given by (14), i' s (incentive-based) lending is then:

$$L_i \equiv \min\left[L^c(R_i, R_d; \frac{A}{2}, \lambda), L^{FB} \right] \qquad \#$$

lower than that of the monopoly bank (by $\frac{A}{2} < A$ and $R_i \leq x$), and; (b) the average quality of *i*'s borrowing firms is exactly the same as that of the single bank.

Now let us consider *j*. The pool of firms faced by *j* consists of the set of firms that could not borrow from (their first choice) i, and it is made up of two subsets. One subset consists of firms that have not been screened by *i* because of i' s incentive-based lending constraint, the other subset consists of firms that have been examined by *i* but have have failed the test. Let θ denote the probability according to which a *j*'s loan applicant has not been examined by *i*; the proportion of type *g* firms in *j*'s pool of applicants is λ_j :

$$\lambda_j = \theta\lambda + (1-\theta)\lambda_B \qquad \lambda_B \equiv \frac{(1-\eta)\lambda}{(1-\eta)\lambda + \eta(1-\lambda)} < \lambda \qquad \#$$

with probability θ a *j*'s loan applicant has not been examined by *i*, in which case its type is (the good one) *g* with probability λ (the proportion of type *g* firm in the original population); with the residual probability it has been examined and rejected by *i*, in which case its type is *g* with probability $\lambda_B \equiv \Pr(g|B) < \lambda$. By the same reasoning at Propositions 2-3, at equilibrium, a bank lends to firms whose projects have non-negative net-present value, that is *i* engages in information-gathering and lends to successfully screened firms and therefore θ is given by:

$$\theta = \frac{M - \frac{L_i}{\Pr(G)}}{M - L_i} \qquad L_i \equiv \min\left[L^c(R_i, R_d; \tfrac{A}{2}, \lambda), L^{FB}\right] \qquad \#$$

where, $M - \frac{L_i}{\Pr(G)}$, at the numerator of (23), is the number of firms that have not been examined by *i*; $M - L_i$, at the denominator of (23), is the number of firms that did not obtain a loan from *i* , i.e. the number of *j*'s loan applicants. The higher L_i, the higher the probability that a firm that has not been financed by *i* has been examined and rejected by *i*, and therefore the lower θ.

The proportion of type *g* firms in *j*'s pool of applicants, λ_j, is then necessarily lower than the corresponding one for *i* (by 22) and the more so the higher *i*'s (incentive-based) lending (by (23)). Moreover, it may be that *j*'s pool of applicants is so adversely selected that there is no room for a second bank. We derive below the condition for this to be the case.

Suppose that at equilibrium only one bank is active. Then necessarily: a) the active bank lends at the single-bank-zero profit rate, that is $R_i = R^{FB}$ (by standard undercutting arguments); and b) a second bank is not viable – the projects it would finance have negative net-present value. This is true if:

$$ES_j \equiv [p + (1-p)\Pr(g|G)_j]x - (R_d + \frac{f}{\Pr(G)_j}) < 0 \qquad \#$$

where

$$\Pr(g|G)_j \equiv \frac{\eta\lambda_j}{\Pr(G)_j} \quad ; \quad \Pr(G)_j \equiv \eta\lambda_j + (1-\eta)(1-\lambda_j) \qquad \#$$

ES_j, the expression at the left-hand side of inequality (24), is the overall expected surplus generated by examining a firm belonging to *j*'s pool of loan applicants and financing the firm if it passed the test, i.e. the signal received is *G*. If ES_j is negative, then there is no lending rate such that a second bank's expected profits are non-negative.

Lemma *If at equilibrium only one bank is active then necessarily*: *a*) *the active bank lends at the* (*single-bank-zero profit*) *rate* R^{FB}, *and its lending volume is* L_i :

$$L_i = min\left[L^c(R^{FB}, R_d; \frac{A}{2}, \lambda), L^{FB}\right] \quad ; \qquad \#$$

and b) *if a second bank were active*, *then the projects it would finance have negative net-present value*: $ES_j < 0$.

ES_j is increasing in λ_j , i.e. the proportion of type *g* firms in *j*'s pool of loan applicants, and λ_j is decreasing in L_i (by (22)-(23)), ES_j is thus (monotonically) decreasing in aggregate bank capital A : the higher *A*, the higher *i*'s lending volume, L_i, and the number of firms examined by *i*.

Therefore, the higher *A*, the higher the probability that a firm that has not been financed by *i* (that is a *j*'s loan applicant) has been examined and rejected by *i*, and consequently the lower λ_j and ES_j. Let $\underline{A}$ denote the level of *A* such that $ES_j = 0$. Condition (24) holds only if $A > \underline{A}$.

Lemma *There exists a threshold level* $\underline{A}$, *such that if* $A > \underline{A}$ *then at equilibrium only one bank is active. The first-mover is the active bank.*

If $A > \underline{A}$, then condition (24) holds, at equilibrium only one bank is active, the lending rate is the (single-bank) zero profit rate, aggregate lending is L_i as given by (26), lower than the level attained with a monopoly bank, whereas the average quality of borrowing firms is the same. By contrast, if $A \leq \underline{A}$ then a second bank is viable, at equilibrium both banks are active. Both aggregate lending and the average quality of borrowing firms are lower than the ones attained with a monopoly bank.

Lemma *If* $A \leq \underline{A}$ *, then at equilibrium both banks are active. The first mover sets its rate to a value* R_i *that satisfies* $R^{FB} \leq R_i < x$, *where* $R_i = R^{FB}$ *only if* $A = \underline{A}$, *and lends the amount* $L^c\left(R_i, R_d; \lambda, \frac{A}{2}\right)$; *the second mover sets its rate to the monopoly level and lends the amount* $L^c\left(x, R_d; \lambda_j, \frac{A}{2}\right)$.

[Proof] See the Appendix

The results at Lemmas 1-3 are summarised in the following proposition.

Proposition *The higher the number of banks that can operate in the same market (either because intra-state barriers are removed and/or the banking industry is more fragmented), the higher the short-fall of aggregate lending from the First-Best optimum level and the lower the average credit-worthiness of borrowing firms. Moreover, an equilibrium may be one where not all banks are active.*

When more than one bank is active, a rejected borrower does not exit he applies to a second bank and if he is still rejected he tries again with a third bank ... The screening and re-application process acts to worsen the average pool of loan applicants and the more so the higher the number of active banks. Thus the higher the number of active banks, the higher the cost of searching for a (good) type *g* firm and the higher the amount of capital that a bank must put at stake to be a credible monitor, and therefore the lower the amount of lending that a bank undertakes, for any given amount of capital. However, at equilibrium, a bank that is active is necessarily viable, i.e. it finances firms whose projects have non-negative net-present value (by Proposition 2). This sets an upper bound on the number of banks that are active at equilibrium. Depending on *A* and *n*, this upper bound may fall short of *n*. For $n = 2$, one of the two banks is inactive whenever $A > \underline{A}$. More generally, the higher *A* the lower the number of banks that are active at equilibrium, for any given *n*. The point being that if at equilibrium n' banks are active, then: i) each of these banks is viable; and ii) a further bank would not be viable, its (would be) applicants' pool is of sufficiently low quality that the overall surplus generated by financing a successfully screened firm is negative. The higher n' (i.e. the higher the number of banks that a firm can approach) and the higher *A*, the lower the quality of applicants that a further bank would face. Indeed the higher *A*, the higher the probability that a loan applicant to the (would be) further bank has been denied credit by the n' active banks because of un-favourable information and not because of lending constraints. Thus the higher *A*, the lower n' that satisfies the (equilibrium) condition ii).

Three main conclusions can be drawn. A concentrated banking industry dominates a fragmented one: the short-fall of aggregate lending from the First-Best optimum level is lower and the average quality of borrowing firms is higher. In the absence of banking industry reorganization, the removal of intra-state entry barriers reduces welfare. Moreover, not all independent banking organizations that were viable in formerly protected national markets remain so when markets are integrated.

Conclusions

This paper has shown that when banks' main specific function is to engage in information production about borrowing firms, then limitations to risk diversifiability imply that credit allocation depends on banking industry and market structures. Specifically: i) a concentrated banking industry, one where aggregate bank capital is held by few banks, leads to credit allocation closer to the first best optimum ; and ii) in the absence of banking industry consolidation, the removal of intra-state entry barriers reduces welfare. Moreover, not all independent organisations that were viable in formerly protected national (local) markets remain so when markets are integrated.

These results relay on two crucial observations, namely : (a) the information externality of screening and its adverse pool effects; and (b) the size of a bank (the amount of lending it undertakes) is constrained.

Observation (a) follows from the screening and re-application process that takes place when several banks are active in the same market. This process acts to worsen banks' pool of applicants and thereby to increase banks' costs of searching for good firms. These costs may outweight the benefits of providing further chances to good-type firms. If banks would share information about loan applicants then this process would not take place, by contrast there would be social benefits arising from credit-worthiness's assessments based on pooled information. However, there are serious incentive-compatibility constraints to information sharing. A bank would rather wish to misreport firms' type so as to pursue the objective of earning monopoly rents on "captured" good-type borrowers. Most likely, the only way to effectively pool information is bank consolidation. This may be particularly relevant when the formerly independent organisations have specific skills in information gathering with regard to different lines of business or different (possibly, formerly local) markets.

With regard to (b), if bank size were unconstrained then an increase in the number of banks that can operate in the same market would simply make banking competition fiercer and this would not necessarily be detrimental. Indeed, as noted at Proposition 1, when banks are unconstrained and react to each other (i.e. observe each other's lending rate and then each bank chooses whether to be active), competition leads to the first-best-optimal credit allocation. Under the assumed screening technology, this is one where a single bank produces information and lends to successfully screened firms, competition drives the bank's lending rate to the zero-profit value. In contrast, when bank size is limited, at equilibrium, more than one bank is active. Because of observation (a), credit allocation departs from the social optimum and the more so the higher the number of banks that can operate in the same market (Proposition 5). In the economics of this paper, bank-size constraints result from conflict of interests (agency problems) between the bank's (insiders) equity holders and (outsiders) investors. These conflicts of interest imply that the amount of outside financing risen by the bank, and therefore bank size, is constrained by insiders' equity holdings (bank capital) and that outside financing is risen by issuing debt (as noted in Section 3). This makes the analysis especially relevant for institutional environments where investors', and particularly outside-equity holders', protection is limited. But even in more developed environments with well-functioning financial markets, capital appears costly to raise (see Smith 1986 for a survey of evidence) and most of bank financing is risen by issuing debt. Bank capital will then still constrain bank's financing and hence bank size, and therefore the paper's conclusions about banking industry and market structures will still be relevant.

References

Berger, A.N., A.K. Kashyap and J.M. Scalise (1995), ”The Transformation of the US Banking Industry: What a Long Trip It’s Been”, *Brookings Papers on Economic Activity* 2, pp. 55-201.

Bernanke, B., M. Gertler and S. Gilchrist (1996), ”The Financial Accelerator and the Flight to Quality”, *The Review of Economics and Statistics* 78 (LXXVIII), pp.1-15.

Bernanke, B. and M. Gertler (1987), ”Banking and Macroeconomic Equilibrium”, in W.Barnett and K. Singleton (eds.) *New Approaches to Monetary Economics* , Cambridge University Press, Cambridge, US.

Bhattacharya, S. and A. Thakor (1993), ”Contemporary Banking Theory”, *Journal of Financial Intermediation* 3, pp.2-50.

Boot, A.W. and S. Greenbaum (1993), ”Bank Regulation, Reputation and Rents”, in *Capital Markets and Financial Intermediation* (C.Mayer and X.Vives eds.), Cambridge University Press, Cambridge, US.

Broecker, T. (1990), ”Credit-worthiness tests and Interbank Competition”, *Econometrica* 58, pp. 429-452.

Chiesa, G. (1997), ”Incentive-Based Lending Capacity, Competition and Regulation in Banking”, mimeo Università degli Studi di Bologna.

Diamond, D. (1984), ”Financial Intermediation and Delegated Monitoring”, *Review of Economic Studies* 51, pp.393-414.

Gehrig, T. (1996), ”Market Structure, Monitoring and Capital Adequacy Regulation”, *Swiss Journal of Economics and Statistics* 132, pp.685-702.

Gehrig, T. (1997), ”Screening, Cross-Border Banking, and the Allocation of Credit”, *Research in Economics*, forthcoming.

Innes, R.D. (1990), ”Limited Liability and Incentive Contracting with Ex-ante Action Choices”, *Journal of Economic Theory* 52, pp.45-67.

Nakamura, L. (1996), ”The Informational Impact of Competition in Loan Screening”, mimeo Federal Reserve Bank of Philadelphia.

Ramakrishnan, R. and A. Thakor (1984), ”Information Reliability and a Theory of Financial Intermediation”, *Review of Economic Studies* 52, pp.415-432.

Rhoades, S.A. and J. Burke (1990), ”Economic and Political Foundations of Section 7 Enforcement in the 1980s”, *Antitrust Bulletin* 35, pp.373-446.

Riordan, M. (1993), ”Competition and Bank Performance: a Theoretical Perspective”, in *Capital Markets and Financial Intermediation* (C.Mayer and X.Vives eds.), Cambridge University Press, Cambridge, US.

Smith, C. (1986), ”Raising Capital: Theory and Evidence”, *Journal of Applied Corporate Finance* 4, pp. 6-21.

APPENDIX – Proof of Lemma 3

Let $A \leq \underline{A}$, and let R_i be defined by:

$$\left\{[p+(1-p)\Pr(g|G)]R_i - (R_d + \tfrac{f}{\Pr(G)})\right\}L^c\left(R_i,R_d;\tfrac{A}{2},\lambda\right) \equiv \left\{[p+(1-p)\Pr(g|G)_j]x - (R_d + \tfrac{f}{\Pr(G)_j})\right\}L^c\left(x,R_d;\tfrac{A}{2},\lambda_j\right) \qquad \#$$

where $\Pr(g|G)_j$ and $\Pr(G)_j$ are given by (25) and $(22)-(23)$, for $L_i = L^c\left(R_i,R_d;\frac{A}{2},\lambda\right)$ — the right-hand side of (27) is decreasing in R_i.

The expression at the left-hand side of (27) gives the expected profits of (firms' first choice) i ; the expression at the right-hand side gives the expected profits of (firms' second choice) j , for $R_j = x$ and $L_i = L^c\left(R_i,R_d;\frac{A}{2},\lambda\right)$.

Lemma *Let $A \leq \underline{A}$, then R_i satisfies $x > R_i \geq R^*$, where $R_i = R^*$ only if $A = \underline{A}$.*

[Proof] The left-hand side of (27) is continuous and increasing in R_i ; the right-hand side is continuous and decreasing in R_i ; if $R_i = R^*$, then the left-hand side is nought, whereas the left-hand side is nought only if $A = \underline{A}$, and it is strictly positive if $A < \underline{A}$ (because for $R_i = R^*$, $ES_j > 0$ only if $A < \underline{A}$ and $ES_j = 0$ only if $A = \underline{A}$, where ES_j is given by (24)). [End Proof]

Proving Lemma 3 consists in proving the following Lemma.

Lemma *Let $A \leq \underline{A}$, then the first mover sets its rate to R_i as defined by (27) , the second mover offers the monopoly rate x.*

[Proof] The first mover is firms' first choice (by $R_i < x$, Lemma 4), and the second mover's expected profits are identical to the first mover's profits (by (27)). Moreover, the expected profits of firms' second choice are decreasing in the lending rate of firms' first choice. Therefore:

i) the second mover would be worse off by undercutting the first mover, that is by setting a rate lower than R_i . He would be firms' first choice, but his expected profits would be lower than the value of the left-hand side of (27) and hence lower than the one attained at the monopoly rate x. Obviously, he would be worse off by setting its rate to R_j that satisfies $R_i \leq R_j < x$.

ii) The first mover would be worse off by setting a rate above R_i. If he were to do so, then he would be undercutted by the second mover, he would be firms' second choice and his profits would be lower than the value of the right-hand side of (27) and hence lower than the ones attained at R_i. [End Proof]