Czech Economic Review 8 (2014) 102–124
Acta Universitatis Carolinae Oeconomica

Relationship Lending, Bank Competition and Financial Stability

Juha-Pekka Niinimäki*

Received 6 March 2014; Accepted 28 November 2014

Abstract This paper explores the effects of relationship lending on bank stability under perfect bank competition. Relationship banking generates profitable old lending relationships, which ease the stress to search and monitor new borrowers, and create great charter values for banks even under bank competition. These everlasting charter values mitigate risk-taking incentives improving financial stability. We find the maximum growth speed for banks so that they will avoid risk taking, and discover the optimal equity capital requirement: new banks and rapidly growing banks should have relatively more capital.

Keywords Relationship lending, bank competition, financial stability, charter value, bank capital requirements
JEL classification G21, G22, G28

1. Introduction

This paper investigates several themes of bank regulation—de novo banks, entry to new markets, maximal risk-free growth in lending, charter value, risk shifting and optimal capital requirements—in the model of relationship lending with perfect bank competition.

To begin, the theory of strategic bank competition has been fundamentally advanced by insights of relationship banking (e.g. Rajan 1992; Baas and Schrooten 2006; Dell'Arriccia et al. 1999; Bouckaert and Degryse 2004, 2006). Bank's lending to a borrower creates a relationship that provides ex post monopoly power for the inside bank and allows it to profit from its locked-in old borrowers.[1] Our paper shows how profits from old lending relationships mitigate the risk-taking incentives of banks. The existence of old lending relationships also lessens the required effort to search and monitor

* University of Turku, Department of Economics, 20014 Turku, Finland. Phone: +3580233351, E-mail: juhnii@utu.fi.

[1] As to empirical research on relationship lending we focus on the most recent findings. Neuberger et al. (2008) explore the number of bank relationships hold by small and medium-sized entrepreneurs (SMEs) in Switzerland. Firm and industry structure have the largest explanatory power, while banking market structure and conduct play a minor role. Relationship lending tends to enhance the concentration of banking relationships. Hernandez-Canovas and Martinez-Solano (2010) discover that SMEs with longer bank relationships have enhanced access to loan, but at the same time they bear a higher cost for their debts. Cotugno et al. (2013) find that relationship lending variables are significant contributory factors to the loan portfolio quality. The findings support our results. Boot (2000) and Elyasiani and Goldberg (2004) survey research on relationship lending. For a survey on financial intermediation, see Gorton and Winton (2002).

new loan applicants, thereby making the monitoring strategy more profitable.[2]

Second, according to pathbreaking findings (e.g. Keeley 1990) competition shrinks bank profits, thus eroding their charter value and thereby driving them to take more risk, and dramatically increasing the probability of bank failures. We show that even under perfect competition an established relationship lending bank enjoys profits period after period, forever. Even if a single lending relationship lasts for two periods, overlapping lending relationships create an everlasting and expanding charter value to the bank thereby mitigating incentives to take risks.[3]

Third, abundant empirical evidence reveals that de novo banks are risky (e.g. Gunther 1990; Hunter et al. 1996; DeYoung 2003).[4] Hunter et al. (1996, p. 237), for instance, document:

> "We find that 37 percent of the 353 de novo institutions initiating operations between January 1, 1980, and December 31, 1986 failed by the end of 1990. Reflecting the lower capital requirements in existence for institutions chartered before 1984, de novos that initiated operations before 1984 had a failure rate of 53%, as compared to 22 percent for those entering under more stringent capital requirements with more capital in the post-1984 environment. *We interpret this as evidence of the positive effects of higher capital requirements and improved monitoring."*

Hunter et al. (1996) also find that de novo institutions most likely to fail were those with rapid asset growth and with low capital levels. Empirical evidence confirms the severe failure rate of rapidly growing banks (e.g. White 1991; Hunter et al. 1996; Logan 2000; Altunbas et al. 2011; Jin et al. 2011; and Berger et al. 2013). More precisely, regarding the S&L crises in the U.S.A., White (1991) finds that the thrifts which were liquidated or acquired during 1986–1989 grew by 101% over 1982–1985, whereas the remainder of the industry grew only by 49% during the same period. Hunter et al. (1996) study de novo S&Ls, which were chartered during 1980–1986. The De novo bank group with annual growth rate above 100% had a failure rate of 60%, compared to a failure rate of 32% for their slower growth counterparts. Logan (2000) investigates the small banks' crisis of the early 1990s in Great Britain. In 1988 almost 40% of banks

[2] Empirical evidence reveals that *switching costs* are important in loan markets. Barone et al. (2011) report significant evidence on switching costs in Italy. Banks price discriminate between new and old borrowers by charging lower interest rates to the former. The discount amounts to about 44 basis points and is equal to 7% of the average interest rate. These findings are supported by Ioannidou and Ongena (2010). They discover that a loan granted by a new (outside) bank carries a loan rate that is significantly lower than the rates on comparable new loans from the firm's current (inside) banks. The new bank initially decreases the loan rate but eventually ratchets it up sharply. These results are consistent with the existence of hold-up costs in bank-firm relationships. Zhao et al. (2013) study British banks during 1993–2008. They find evidence on high switching costs in the latter part of the sample period. Calem et al. (2006) document evidence on switching costs in the market of credit cards.

[3] Empirical evidence confirms that a bank with a substantive charter value takes less risks, e.g. Keeley (1990), Gan (2004), Ghosh (2009) and Lee and Hsich (2013).

[4] Gunther (1990) explores the experiences of Texas banks during the 1980s; 39% of de novo banks failed in comparison to 21% for established banks. Many de novo banks invested in high-risk assets as soon as they opened. DeYoung (2003, p. 742) documents: "... de novo banks and established banks tended to fail for similar reasons. Imprudent lending practices (e.g. aggressive lending, ...)."

within the highest loan growth quartile went to fail compared with 17% or less of banks in the lower growth quartiles. Altunbas et al. (2011), Jin et al. (2011) and Berger et al. (2013) explore the 2007–2009 global financial crises. Altunbas et al. (2011) focus on listed banks operating in European Union and Jin et al. (2011) and Berger et al. (2013) analyze US banks. Jin et al. (2011) find that growth in commercial loans, growth in real estate loans and overall loan growth are reliable predictors of a bank failure. Our paper suggests a theoretical explanation for these empirical regularities: Both rapidly growing banks and de novo banks expand their lending to novel markets in terms of geographic areas, sectors and borrowers where the bank has no earlier experience. The paper derives the *optimal growth path* for a bank so that it will not take excessive risks. The bank is allowed to expand its operations but it should not expand too rapidly. The more severe the problem of asymmetric information, the slower the permitted growth speed is.

Fourth, imposing capital requirements for banks has proved to be a fairly effective regulatory tool, but because equity capital is expensive the requirements must be carefully designed to economize capital usage. The Basle Capital Accord, for example, classifies loans to different risk categories. The risk classification principles have encountered criticism, because they are considered to constitute one of the main causes of the subprime crisis. This paper derives the optimal incentive compatible capital ratio, which recommends a high capital requirement for de novo banks and rapidly growing banks.

Fifth, the paper shows how the overlapping structure of loans can be used to mitigate the incentive problem existing between the bank and its depositors. This new method can be more commonly utilized to improve incentives between lenders and borrowers.

Finally, we will detail links to the relationship lending literature (e.g. Rajan 1992, Baas and Schrooten 2006). In this literature, information advantage provides ex post monopoly power for the inside bank and allows it to profit from its locked-in old borrowers. Our model is more general, because several reasons (not only information advantage) can generate the lock-in effect. Besides, the relationship lending literature examines a single lending relationship which takes for two years. Our paper explores a bank with infinite number of overlapping relationship lending contracts and the bank operates forever. Moreover, the relationship lending literature focuses on loan markets and banks are risk free. Our paper extends this literature to bank failures and bank regulation. The paper finds out the maximal risk-free growth path for banks and derives rules for the optimal incentive compatible equity capital requirement. In addition, the paper is related to the literature on financial intermediation. As in Diamond (1984), our bank offers intermediation services to depositors. Diamond (1984) eliminates the incentive problem between the bank and depositors by assuming that the bank's loan portfolio is fully diversified which makes the bank risk free. Our paper investigates the most difficult incentive problem between a bank and depositors, because loan risks are completely correlated. The paper shows how relationship lending and equity capital together eliminate the incentive problem.

The paper proceeds as follows. Section 2 introduces an economy and Section 3 characterizes equilibrium loan interest rates. Sections 4 and 5 present a bank with the overlapping structure of loans. The main findings are in Section 6, Section 7 provides with an example and Section 8 concludes.

2. Economy

Consider a model of perfectly competitive credit markets with infinite number of periods and three types of risk-neutral agents: owner-managed banks, investors (depositors) and owner-managed firms (borrowers). Firms and banks are protected by limited liability and they maximize the expected wealth of their owners, whereas investors maximize their personal wealth. A new generation (continuum) of firms and investors is born in each period. An investor lives only for a period whereas a firm operates for two periods. Banks are identical. A bank operates for ever if it does not take excessive risks and fail. As standard in the credit models of hidden characteristics, the type of each firm (borrower) is given (good, bad) and private information. A bank can learn firm type by monitoring it. The NPV of an average firm project is negative without monitoring but under bank monitoring the NPV of a financed project is positive. Unfortunately, limited liability causes the problem of moral hazard: the bank has incentives to take excessive risks and neglect monitoring. Only the bank and its borrowers know the true strategy of the bank (monitoring, non-monitoring). A bank regulator, who knows the economic environment, can infer the true strategy by gathering information on bank capital and bank growth. The regulator supervises that the bank pursues the monitoring strategy. The main part of the supervision process consists of the calculation of bank profits under both strategies. This requires that the regulator figures out the magnitude of profitable old lending relationships and returns from these relationships period after period.[5]

Next, the paper details the regulator's supervision process. Sections 2.1–2.2 illustrate the project types. Section 3 defines loan interest rates under relationship banking. Thereafter, it is possible to solve profits from both strategies (monitoring, non-monitoring) and infer the constraint for bank monitoring in Section 4.

2.1 Borrower types

This section introduces firms and projects. A new generation of firms is born in every period. Period t, $t \in \{0,1,2,\ldots\}$, begins at time t and ends at time $t+1$. This is also the starting point of a new period, period $t+1$, which ends at $t+2$. A firm operates for two periods and it can undertake a project in both periods. A firm, which is born in period t, is a *new firm* during period t and an *old firm* in period $t+1$. The age of a firm is observable. Everyone knows in period $t+1$ whether an old firm received a loan in period t.

New firms: In its first operating period a firm seeks for a loan. The problem of hidden characteristics is present. The type of a firm is either bad or good and it is firm's private information. A project of a *good firm* succeeds with certainty and has positive NPV,

[5] Alternatively, it is possible that deposits are uninsured and depositors supervise lending strategies.

$Y > r$, where r indicates gross interest rate of the economy. The project of a *bad firm* succeeds with probability p and has negative NPV, $pY < r$. An average project has negative NPV, $(g+(1-g)p)Y < r$, where g denotes the share of good firms in the economy.

Old firms: Recall that each firm seeks for a loan in period 1. If a bank exerts effort in monitoring, only a good firm receives a loan in period 1. Hence, the lending decision of period 1 reveals the firm type to market participants. Consider a firm, which received a loan in period 1. The firm bears a switching cost Φ in period 2, if it changes a bank after period 1. The *switching cost* is so small that it does not make a good project unprofitable, $Y > r+\Phi$.[6]

An investor is endowed with a unit of capital, which can either be stored at the interest rate of the economy, r, or deposited in a bank.

2.2 Information acquisition: monitoring

Market power is such that banks are driven to their reservation utility levels and a firm can seize the project surplus. Since the problem of hidden characteristics is present, banks must screen good and bad firms by exerting effort in monitoring. Let m_1 indicate the cost of monitoring a new firm. Monitoring displays the type (good or bad). Each firm has time to seek a loan from one bank only. A bank monitors firms until it finds a good one and the expected monitoring costs towards a loan amounts to $M_1 = m_1/g$ units. The project of a good firm is assumed to have positive NPV even with the costs of monitoring, $Y > M_1 + r$. As is standard in relationship lending models, pre-commitments to two-period contracts are unenforceable, and projects are financed using single-period loans. We also assume, again obeying the tradition of the relationship lending literature, that the owner of each firm consumes the project surplus at the end of first period. We make the following assumption:

Assumption 1. *The risks and returns of bad borrowers are completely correlated.*

Assumption 1 causes the most difficult incentive problem between a bank and the regulator (or uninsured depositors), because diversification does not mitigate the incentive problem as in Diamond (1984). If we can develop a method which eliminates the incentive problem under Assumption 1, the method eliminates the incentive problem when correlation is incomplete.

3. Loan interest rates

This section finds out loan interest rates. Recall that borrower's age is observable. In addition, it is public information whether a borrower receives a loan in his first operating period. Hence, a bank is able to price discriminate between its locked-in old borrowers, old borrowers locked-in to a rival bank and the fresh loan applicants of the newborn generation (new borrowers). The equilibrium loan interest rates prove to be

[6] Barone et al. (2011) list several reasons for switching costs in loan markets.

similar to standard models of relationship lending: a new firm is unprofitable to a bank, but an old one yields profit. To begin, we make the following assumption:

Assumption 2. *Each unit of equity capital entails $r+c$ units costs for a bank.*

Assumption 2 motivates a bank to minimize capital ratio. Competition drives each bank to maintain the same capital ratio. First we investigate a monitoring bank. Thereafter, we focus on the non-monitoring strategy. An investor deposits his endowment in a bank, which monitors firms at the start of each lending relationship. If a bank is indifferent whether or not to grant a loan, it grants the loan. Loan interest offers are public information.

Consider the second operating period of a firm, which received a loan in the first operating period. First, the inside bank makes it loan interest offer. Then, outside banks make sequentially public their offers. After this, it is not possible to change the offers. The firm chooses the bank with the lowest interest rate offer. If it is indifferent between the inside bank and an outside bank, it favors the inside bank. Since the lending decision of the first operating period reveals the type of the firm, no problem of asymmetric information appears in the second operating period. The switching cost Φ favors the original lending relationship. The loan interest rate offer of the inside bank satisfies $R_2 = (1-e)r + e(r+c) + \Phi$, or

$$R_2 = r + ec + \Phi. \tag{1}$$

Here e is the capital ratio of the bank. The zero profit offer of an outside bank is $R_2 - \Phi$. Given the switching cost, the old firm is indifferent between the loan interest rate of the inside bank and the offer of the outside bank. Thus, an old firm stays with its original bank.

More precisely, suppose that the inside bank offers $R_2 + \varepsilon$, $\varepsilon \geq 0$. If $\varepsilon > 0$, an outside bank profits if it later bids, e.g. $R_2 - \Phi + \varepsilon/2$. Given this bid, the old firm switches from the inside bank to the outside bank and the inside bank losses a profitable borrower. Thus, the inside bank optimally offers R_2. It looses the old borrower if $\varepsilon > 0$ in $R_2 + \varepsilon$. On the other hand, the inside bank will not offer $R_2 - \varepsilon$. This kind of offer would reduce its loan interest income.

Since the inside bank has an advantage during the second operating period of a firm, the inside bank enjoys profit Φ from each old borrower. To gain profitable old borrowers, the banks compete fiercely for new borrowers. The banks must sacrifice profits because they must offer a lower introductory loan interest, R_1 to attract new borrowers so that the expected life-time profits from a new borrower are zero

$$R_1 - r - M_1 - e_t c + \delta(R_2 - r - e_{t+1} c) = 0, \tag{2}$$

where $\delta = 1/r$. Given (1), (2) provides

$$R_1 = r + e_t c + M_1 - \delta\Phi. \tag{3}$$

In (3) the return from a new borrower is negative, $-\delta\Phi$ units, but the return from an old borrower is positive, Φ units. No old borrower will actually switch a bank. Loan interest rates depend on capital ratio but the profit from an old firm and the loss from a new firm are independent of it. Since loan interest rates are public information, a non-monitoring bank must offer the same loan interests as a monitoring bank.[7] A deviant loan interest rate offer would reveal the non-monitoring strategy to the regulator (or uninsured depositors, who would not save in the bank).[8]

The regulator knows bank profit (loss) from a single old (new) borrower. The regulator can infer the shares and volumes of new and old borrowers in a bank's loan portfolio. He can also infer the total profits (losses) of the bank from old (new) borrowers. Given this information and the capital ratio, he can calculate bank profits with monitoring and without monitoring. In next we find out the conditions such that the bank optimally monitors.

4. Banking with overlapping structure of loans and equity capital

Sections 4–6 indicate how overlapping structure of loans (lending relationships) together with equity capital eliminate the incentive problem to neglect monitoring.[9] This section derives *incentive constraint* for the bank. It exerts effort in monitoring only if the constraint is satisfied. The incentive constraint includes information on new and old borrowers and equity capital. At the end of the section, the characteristics of the incentive compatible bank behavior are summarized in three claims.

Each bank is formed in period 0 and it has a continuum of borrowers $[0, S_t]$, where S_t denotes bank size in period t. A banker is endowed with a fixed amount of funds, E. He injects the endowment into his bank as equity capital. The banker does not receive the endowment at once. He receives $\frac{1}{2}E$ units at the start of period 0, when he sets up the bank, and the rest, $\frac{1}{2}E$, at the start of the next period. The bank funds the rest of the loans by attracting deposits and pays risk-free interest r on them.

Assumption 2, equity capital is more expensive than deposits, motivates the bank to minimize capital ratio, E/S. Since the amount of equity capital is fixed, E, Assumption 2 drives the bank to maximize its size, S, by growing as fast as possible. In this model, equity capital represents the basic instrument for mitigating the incentive problem between the bank and depositors, and the incentive compatible equity ratio measures the magnitude of the incentive costs.

The regulator supervises banks. Given the negative NPV of non-monitored projects, it is not socially optimal to operate a bank which neglects monitoring. The regulator does not directly observe whether a bank monitors. Fortunately, he can infer the presence of the monitoring strategy by keeping track of the capital ratio and the growth

[7] We assume that a non-monitoring bank must charge the same loan interest rates as a monitoring bank. The assumption is not critical. An alternative assumption creates a bit different model.

[8] A non-monitoring bank is not interested to attract old borrowers from another bank, because they generate zero profit to an outside bank.

[9] The overlapping structure of lending relationships requires that a bank grants loans to new firms in each period. As a result, the bank has both new borrowers and old borrowers in each period, $t \geq 1$. The generations of borrowers are thus overlapping. We demonstrate how the overlapping structure of lending relationships can be utilized to mitigate the incentive problem. Alternatively, suppose that the bank grants loans to new loan applicants only in every second period: $t, t+2, t+4, \ldots$ Now the bank has new borrowers in periods t, $t+2, t+4, \ldots$, and old borrowers in periods $t+1, t+3, \ldots$ The borrower generations are not overlapping.

path of the bank. The growth path reveals implicitly information on new and old borrowers. The regulator knows the characteristics of the economy and the bank sizes of the previous periods $S_0, S_1, \ldots, S_{t-1}$. The sizes must have been such that the incentive constraint has been satisfied in every period and the bank has monitored borrowers. If the bank selects its size S_t in the current period so that the incentive constraint is not satisfied, the regulator closes the bank.[10] A rational bank knows this and selects its size S_t so that the incentive constraint is satisfied. Thereafter, the bank announces how many deposits $S_t - E$ it will then allow and depositors make their deposits.[11] The bank grants S_t loans. At the end of the period, project outputs materialize and the loans are repaid. The bank pays interest r on deposits and the bank's owner then receives the remainder of the return. Consequently, the bank cannot make commitments regarding its size in the future. The bank's optimal actions can be specified using three claims.

Claim 1 (Binding incentive constraint). *The incentive constraint is binding in every period.*

The incentive constraint is satisfied when the bank prefers the monitoring strategy to the non-monitoring strategy or is indifferent between them. The incentive constraint determines the maximal size of the bank in every period. If the incentive constraint were non-binding, the bank could grow, thereby reducing its capital ratio and the costs of funding. If the bank is too large, its capital ratio is too low and the banker prefers the non-monitoring strategy. Thus, the bank size at the maximum level is such that the expected returns of the monitoring strategy, π_t^m, and non-monitoring strategy, π_t^{nm}, are equal in every period t.

$$\sum_{i=0}^{\infty} \delta^i \pi_{t+i}^m = \sum_{i=0}^{\infty} (\delta p)^i \pi_{t+i}^{nm} \qquad \forall t,\ t \in \{0,1,2,\ldots\}. \tag{4}$$

If the bad risk materializes, bad loans default and this becomes public. We will observe later that the bank fails due to the materialization of the bad risk. We can restate (4) as

$$\pi_t^m + \delta \sum_{i=1}^{\infty} \delta^{i-1} \pi_{t+i}^m = \pi_t^{nm} + \delta p \sum_{i=1}^{\infty} (\delta p)^{i-1} \pi_{t+i}^{nm}. \tag{5}$$

The binding incentive constraint in period $t+1$ will be

$$\sum_{i=1}^{\infty} \delta^{i-1} \pi_{t+i}^m = \sum_{i=1}^{\infty} (\delta p)^{i-1} \pi_{t+i}^{nm}. \tag{6}$$

[10] Alternatively, we learn in Section 6 that the regulator determines a capital requirement which depends on the bank's age and growth.

[11] Consider an economy without the regulator. Uninsured depositors acquire information on the bank. Depositors make their deposits only if they know that the incentive constraint is satisfied. Then the bank monitors loan applicants and the bank is risk-free.

Inserting (6) in (5) gives

$$\pi_t^m + \delta(1-p)\sum_{i=1}^{\infty}\delta^{i-1}\pi_{t+i}^m = \pi_t^{nm}. \tag{7}$$

On the R.H.S., π_t^{nm} is the return on the non-monitoring strategy in period t. On the L.H.S., the first term represents the returns of the monitoring strategy in period t. The second term is the difference in the present value of future profits from period $t+1$ onwards for the monitoring and non-monitoring strategies. If the bank chooses the non-monitoring strategy, it risks failure and loss of these profits. Hence, future profits are greater for the banks that monitor. The exact magnitude of future profits is solved in Appendix A.

Claim 2 (**Future profits**). *Future profits in period $t+1$ and thereafter add up to $\sum_{i=1}^{\infty}\delta^{i-1}\pi_{t+i}^m = V_t\Phi + Er$, if the current period is t, $t \in \{1,2,3,\ldots\}$, and $V_0\Phi + \frac{1}{2}Er$ in period $t=0$.*

Here, V_t denotes the volume of new borrowers in period t. Intuitively, in period t the bank invests in its V_t new borrowers by monitoring them. These borrowers will have the second project in period $t+1$. Hence, the future profits from $t+1$ on, $V_t\Phi$, arise from the profitable old borrowers in period $t+1$. In addition, future profits include a compensation for the equity capital injection, Er. In period 0, the amount of injected equity capital is $\frac{1}{2}E$, but at the start of the next period the banker injects more capital into the bank thereby increasing the compensation term from $\frac{1}{2}Er$ to Er.

It is necessary to investigate bank returns when the bad risk materializes. Does a non-monitoring bank fail? Two scenarios occur. First, consider a de novo bank without old loans. If it neglects monitoring, the share of good loans is g and the rest are bad in its loan portfolio. Suppose now a downturn and the bad risk materializes. It is possible to show that the bank fails. That is, the value of the bank assets does not cover the value of deposits.[12] The idea is self-evident. Since the NPV of the financed projects is negative without monitoring, the non-monitoring strategy can be profitable—and the incentive constraint can be binding—only if the bank benefits from limited liability when the bad risk materializes.

Second, consider now a bank which operates in period t, $t \geq 1$, and has pursued the monitoring strategy so far. It inherits old borrowers with good projects from the previous period. The loans to these yield profits with certainty. Suppose that the bank does not monitor new borrowers. The share g of these are good and the rest are bad. What happens if the bad risk materializes? Is the bank solvent, when loan portfolio includes both bad borrowers (share $1-g$ of new borrowers) and good borrowers (all old borrowers, part g of new borrowers)? The incentive constraint can be binding if $\pi_t^m < \pi_t^{nm}$ (recall (7)). This is possible only if the bank benefits from limited liability when it neglects monitoring.[13] Thus, the bank fails when the bad risk materializes. The cases can be summarized as follows.

[12] The proof is omitted due to the lack of space.

[13] Detailed proof that $\pi_t^m < \pi_t^{nm}$ is omitted.

Claim 3 (**Limited liability**). *When the incentive constraint is binding, the optimal bank size (the fraction of new borrowers) is so large that a non-monitoring bank fails if the bad risk materializes.*

When the incentive constraint is binding, the optimal share of new borrowers is large. Since a high fraction of these are bad, the bank fails when the risk materializes. Claim 3 simplifies the analysis considerably when the incentive constraint is binding. With probability p , each of the bad loans is repaid and a non-monitoring bank makes the same monetary returns as a monitoring bank. Yet, it avoids the non-monetary cost of monitoring. With probability $1-p$, bad loans default and the non-monitoring bank fails. The fact that the loan portfolio includes both good and bad loans has no effect on the risks and returns of the non-monitoring bank. It makes the same returns as if the loan portfolio had included only bad loans.

A bank does not ever pursue a "temporary non-monitoring strategy" such that a monitoring bank turns to the non-monitoring strategy for some periods and then returns to the monitoring strategy. The expected profit from the return is always lower than if a non-monitoring bank retains the non-monitoring strategy. As a result, a non-monitoring bank will also keep its old borrowers who pay high interest R_2 on their loans.

We have now characterized the incentive constraint. It abstracts information on profits from old borrowers, losses from new borrowers, capital ratio, etc. The regulator knows the incentive constraint and can use to ensure that the bank pursues the monitoring strategy in every period. In equilibrium banks always monitor borrowers.

5. Bank's optimal size in periods 0, 1, 2, ...

We have determined the incentive constraint and summarized the bank's operations using Claims 1–3. Given this information, it is possible to find out the optimal bank size (=maximal size) in each period such that the incentive constraint is satisfied and the bank pursues the monitoring strategy. The regulator observes the bank size and the amount of equity capital and infers that the incentive constraint is satisfied. Subsection 5.1 derives the optimal size in period 0 and Subsection 5.2 discovers it in period t, $t \geq 1$. Thereafter, it is possible to express the optimal size of period t as a function of the initial size.

5.1 The de novo bank

This section identifies the optimal size in period 0, S_0. Given (7), in period 0 the bank maximizes its size so that the incentive constraint is binding, i.e.

$$\pi_0^m + \delta(1-p)\left[S_0\Phi + \frac{1}{2}Er\right] = \pi_0^{nm}. \tag{8}$$

On the L.H.S., the first term displays the return from the monitoring strategy in period 0. The second term indicates the present value of future profit. The R.H.S. contains the expected return without monitoring in period 0. Appendix B shows that (8) simplifies to

$$0 = pS_0M_1 - (1-p)\frac{1}{2}Er. \quad (9)$$

On the R.H.S., the first term is the benefits of neglecting monitoring. With probability p, loans succeed under both strategies. Without monitoring the bank earns higher returns than with monitoring by avoiding the costs of monitoring. As to the second term, with probability $1-p$ the bad risk materializes, these loans default and the bank fails, thereby losing its equity capital. Recall that the banker's endowment is only $\frac{1}{2}E$ in period 0. Three findings follow. First, without equity capital, the bank would neglect monitoring since $0 < pS_0M_1$. Second, a fully equity-funded bank, $S_0 = \frac{1}{2}E$, would monitor. To see this, note that the R.H.S. of (9) can then be written as

$$S_0\left[pM_1 - (1-p)r\right] < S_0\left[p(Y-r) - (1-p)r\right] = S_0(pY - r) < 0.$$

Third, since the bank maintains a fixed amount of equity capital, the non-monitoring strategy is profitable if the size, S_0, is sufficiently large. This can be observed from (9). There is a maximal (=optimal) size, which can be determined from (9),

$$S_0^* = \frac{(1-p)\frac{1}{2}Er}{pM_1}. \quad (10)$$

If the bank were smaller, the capital ratio would be higher. The costs towards a loan unit would be higher and the bank should charge more interest on loans. Yet, this is impossible. Other banks, which choose the optimal bank size and thereby a lower capital ratio, can charge less interest on loans.[14] The representative bank can succeed in competition only by expanding lending so that it can reduce costs towards a loan unit and thereby offer lower loan interest rates.

If the bank were bigger, the incentive constraint would not be binding, and the bank would neglect monitoring. A conclusion follows.

Lemma 1 (De novo bank). *In period 0, the bank has no profitable old lending relationships and its monitoring incentives are created entirely by equity capital. Without equity capital the bank would not monitor. Given the fixed amount of equity capital, the bank's optimal size is S_0^*.*

5.2 The bank in periods $t = 1, 2, 3, \ldots$

This section determines the optimal size of the bank in period t when the bank has pursued a monitoring strategy thus far. Then, the bank has V_{t-1} old good borrowers and

$$V_t = S_t - V_{t-1}, \quad t \in \{1, 2, \ldots\},$$

new borrowers, for example $V_0 = S_0$, $V_1 = S_1 - V_0$, $V_2 = S_2 - V_1$, ... Obviously, the number of new borrowers is equal to the difference between the size of the bank and

[14] We study a representative bank and assume that the sector consists of identical banks, which are set up in period 0.

the number of old borrowers (new borrowers in the previous period). Given (7), in period t the incentive constraint is

$$\pi_t^m + \delta(1-p)\left[V_t\Phi + Er\right] = \pi_t^{nm}. \tag{11}$$

On the L.H.S., the first term denotes the return from the monitoring strategy in period t. The second term shows the present value of future profits. The R.H.S. includes the expected returns from non-monitoring in period t. Appendix C indicates that (11) can be restated as

$$(S_t - V_{t-1})pM_1 - (1-p)V_{t-1}\Phi - (1-p)Er = 0. \tag{12}$$

The first term represents the extra return from the non-monitoring strategy compared to the monitoring strategy. With probability p the loan interest income and payments on deposits are identical under both strategies. Yet, under the non-monitoring strategy, the bank avoids the costs of monitoring. This return depends on the amount of new borrowers, $S_t - V_{t-1}$. The second term is the profit from old borrowers. Without monitoring, the bank risks failure and the loss of this profit. The third term is the risk of losing equity capital without monitoring. Thus, the first term weakens monitoring incentives, while the last two terms strengthen these. Note that old borrowers V_{t-1} reduce the incentive problem in two ways. The first term reveals that the existence of old borrowers decreases the need to invest in monitoring. The second term reveals that old borrowers induce profits, because they are locked in the initial bank. The optimal size can be solved from (12),

$$S_t = 2S_0^* + GV_{t-1}, \text{ where } G = \left[\frac{(1-p)\Phi}{pM_1} + 1\right], \tag{13}$$

or $S_t = 2S_0^* + G(S_{t-1} - S_{t-2}) + GV_{t-3}$. This reveals two points. First, (13) shows that the size grows in parallel with the number of old borrowers V_{t-1}. Second, the faster the growth of the previous period, $S_{t-1} - S_{t-2}$, the larger the size in the current period. A conclusion follows.

Lemma 2 (Established bank). *In periods* $1, 2, \ldots$, *both equity capital and old lending relationships motivate the bank to monitor. Old relationships reduce monitoring costs and provide profits.*

Profits from old lending relationships mitigate risk-taking incentives, even through the profits are not injected in the bank as additional equity capital. The intuition is that these profits boost dividends and the banker will not risk this income by neglecting monitoring. This incentive effect is based on the overlapping structure of lending relationships. The bank grants new loans when it still has ongoing old lending relationships. The old borrowers will yield profits after a period only if the new loans do not default.[15] This motivates the bank to monitor new borrowers. The incentive effect

[15] In period 0, size S_0^* ensures that the bank monitors. Through monitoring, the bank creates profitable lending relationships for period 1. In period 1, both these old relationships and equity capital encourage the

does not occur without the overlapping structure, that is, if the bank reinvests its funds in new loans only in every second period. The bank's optimal size in period 1 can now be solved from (13),

$$S_1^* = 2S_0^* + GS_0^*, \tag{14}$$

since $V_0 = S_0$. The bank grows from period 0 to period 1. Unfortunately, equation (13) becomes impracticable in later periods. For example, in period 100 the size is $2S_0^* + GV_{99}$. This is obscure when the number of old borrowers, V_{99}, is unknown. Fortunately, the optimal size can be expressed as a function of the bank's initial size. Proposition 1 is proved in Appendix D by induction.

Proposition 1 (Bank's optimal size). *A bank's optimal size in period t is*

$$S_t^* = \left[2 + 2G\sum_{i=1}^{t-1}(G-1)^{i-1} + G(G-1)^{t-1}\right]S_0^*, \quad G = 1 + \frac{(1-p)\Phi}{pM_1}. \tag{15}$$

We have solved the incentive compatible bank size in every period. Using (15) it is possible to solve the optimal growth speed for the bank, the optimal capital ratio, etc. in the following sections.

6. Findings

6.1 Optimal growth speed

It is easy to obtain the following result from (15).

Proposition 2 (Growth speed). *The optimal growth is* $S_{t+1} - S_t = (G-1)^{t-1}G^2S_0^*$.

Using Proposition 2, the following results are proved in Appendix E.

Corollary 1 (Growth factors). *Growth speeds up with Φ and g, but slows down with p and m_1. If $G > 2$, growth speeds up over time and approaches infinity in eternity. If $G < 2$, growth slows down over time and approaches zero in eternity.*

Intuitively, the larger the switching cost, Φ, the more profitable a single old lending relationship. Obviously, great profits from old lending relationships reduce risk-taking incentives. When g is high, the economy has plenty of good loan applicants. This decreases the costs of screening and makes the monitoring strategy more worthwhile. High monitoring cost, m_1, erodes monitoring incentives. (Here m_1 is likely to be high in novel markets or industries. In many emerging economies weak accounting information increases monitoring costs.) The larger p, the higher the profitability of the non-monitoring strategy. This weakens monitoring incentives and the bank can grow only slowly. When $G > 2$, growth speeds up over time; through growth, the bank creates sufficiently new profitable lending relationships that enable it to grow even more

---

bank to monitor. Hence, the bank can grow further. Size S_1^* ensures that the bank monitors in period 1. By monitoring in period 1, the bank creates profitable lending relationships for period 2. In period 2, both the old relationships from period 1 and equity capital encourage monitoring. The bank can grow further. This growth process continues for ever from one period to the next. See Niinimäki (2001).

rapidly in the next period. We have $G > 2$ if p is small and ratio Φ/M_1 is high. When $G < 2$, growth slows down over time. Growth creates profitable lending relationships, but lowers the capital ratio. In contrast to the case of $G > 2$, the value of lending relationships is insufficient to fully compensate for the declining capital ratio. Hence, the incentive problem worsens and the rate of growth needs to slow down. Since $G > 2$ is rather unrealistic, only the alternative $G < 2$ is explored in the following sections.[16]

6.2 Optimal size in eternity

Growth slows down to zero in eternity and the optimal size settles to a steady-state level.

Proposition 3 (Size in eternity). *When $G < 2$, the bank's size in eternity approaches the steady-state level $S_t^* = 2S_0^* + \frac{2G}{2-G}S_0^*$.*

Proof. The bank's optimal size in Proposition 1 can be rewritten as

$$S_t^* = 2S_0^* + 2G\left[1 + (G-1) + (G-1)^2 + \dots\right] S_0^* + G(G-1)^{t-1}S_0^*.$$

As t increases without limit, the first term is fixed, the second term can be expressed as a sum of an infinite geometric series, $2S_0^*G/(2-G)$, and the third term approaches zero. □

The steady-state level increases with the switching cost, Φ, and with the share of good loan applicants, g, but decreases with the success probability of a non-monitored loan, p, and the costs of monitoring a new borrower, m_1.

6.3 Charter value

The bank has a positive charter value in the steady state. To observe this, recall the number of new borrowers, $V_t = S_t - V_{t-1}$. Let V_{ss} denote the number of new borrowers and S_{ss} the bank size in the steady state. Then, V_{ss} and S_{ss} are fixed, $V_t = V_{t-1} = V_{ss}$ and $S_t = S_{ss}$. Inserting these into $V_t = S_t - V_{t-1}$ gives the number of new borrowers as a function of the bank's size $V_{ss} = \frac{1}{2}S_{ss}$. To identify the profits in the steady state, recall that in every period the bank has V_{ss} old borrowers. Their loans provide total profits $V_{ss}\Phi$. The rest of the loan portfolio, $S_{ss} - V_{ss}$, consists of unprofitable loans by new borrowers. These loans entail a loss $(S_{ss} - V_{ss})\delta\Phi$ (recall (3)) and the bank makes returns $V_{ss}\Phi - (S_{ss} - V_{ss})\delta\Phi$. Given $V_{ss} = \frac{1}{2}S_{ss}$, the returns can be restated as $\frac{1}{2}\Phi S_{ss} - (S_{ss} - \frac{1}{2}S_{ss})\delta\Phi$ or

$$\frac{1-\delta}{2}S_{ss}\Phi.$$

[16] The alternative of $G < 2$ seems more practical than that of $G > 2$. When $G > 2$, growth speeds up over time and approaches infinity in eternity. The size of the bank then approaches infinity as well. Thus, the bank is a natural monopoly. This is problematic, since the banking sector is assumed to be fully competitive. Further, the assumption that an infinite rate of growth is possible ignores any demand-side considerations. It is likely that the quality of the financed projects declines as the rate of growth approaches infinity.

The bank makes similar returns in each period of the steady state; that is, forever! The present value of the returns is $\frac{1}{2}S_{ss}\Phi$. This is the bank's charter value. The following conclusion can be drawn.

Proposition 4 (Charter value). *Although the banking sector is fully competitive, the bank makes profits in every period of the steady state (forever), since it has profitable old lending relationships. The charter value of the bank is* $\frac{1}{2}S_{ss}\Phi$.

If we sum the charter value and the compensation for the equity capital investment, Er, we obtain the future profits in Claim 2.

In his classic article, Keeley (1990) advances the notion that during earlier decades various anti-competitive restrictions endowed banks with market power and created positive charter values, which reduced risk-taking incentives. Strict competition later eroded charter values and made risk- taking profitable, thereby adding to bank failures. In broad agreement with the visions of Keeley, this paper suggests that a bank may have a considerable charter value even when the banking sector is fully competitive, if the bank has profitable old lending relationships. We emphasize two points:

(i) Charter value is not based on a huge start-up cost in period 0. In contrast, the bank builds its charter value gradually over a long time span by growing slowly and at the same time investing in lending relationships.

(ii) The bank has positive charter value for ever, although the charter value is based on lending relationships, which last for two periods. In each period, the bank invests in new lending relationships, which creates valuable old borrowers for the next period.

Each bank monitors and earns zero life-time returns. The optimal growth path determines the loan interest rates in each period. Suppose that a bank does not grow as rapidly as the optimal growth path allows. This policy yields negative life-time returns. The bank takes the loan interest as given by other banks. At the same time its operating costs exceed the costs of the other banks due to the higher equity capital ratio. Hence, slow growth is not optimal. A bank maximizes it returns (i.e. earns zero life-time returns) only if it grows as rapidly as the optimal growth path allows. It is possible to say that perfect competition forces banks to invest as much as possible (according to the optimal growth path) to monitoring. The life-time returns are zero only if the investment in monitoring is at the upper limit in each period. If a bank invests less in growth the life-time returns are negative. Hence, perfect competition drives to the maximal investment in each period.

6.4 Dynamic capital ratio

A bank has a fixed amount of equity capital, E, and it grows according to the optimal growth path. Propositions 2 and 3 give the following result.

Corollary 2 (Optimal capital ratio). *Optimal capital ratio declines as the bank matures. Established banks maintain lower capital ratios than de novo banks. In eternity, the capital ratio approaches the steady-state level,* E/S_{ss}.

Given the incentive problem—the bank may neglect monitoring—the bank needs to maintain some equity capital in order to be able to commit to monitoring. Equity capital and charter value are substitutes in reducing the incentive problem. The greater the charter value, the lower the needed capital ratio and vice versa. A de novo bank has no charter value and it needs to maintain a high capital ratio. As the bank matures and grows, it gradually gains more and more charter value. Thus, the bank can gradually lower its capital ratio. Finally, the bank achieves the steady state. The bank's size, charter value and capital ratio produce the fixed levels of steady state. Since the charter value is at its maximal level, the needed capital ratio is at its minimal. That is, since an established bank (a bank that has achieved the steady state) has relatively great charter value, it needs to maintain relatively low capital ratio in order to ensure its monitoring incentives.

Equity capital entails an incentive cost Ec/S_{ss} per a loan unit. If G approaches 2, S_{ss} is infinite. The incentive cost approaches zero. The incentive problem can be eliminated at no cost in the steady state using the overlapping structure of loans even if the incentive problem is very difficult, because loan risks and returns are completely correlated.

7. Example

Suppose that a banker has one million Euros to inject into a bank as equity capital. The economy has the following characteristics: $g = 0.3$, $p = 0.85$, $r = 1.05$, $Y = 1.1$, $M_1 = 0.03$, $\Omega = 0.03$.

Using Proposition 1, we can solve the bank size: 3.09 in period 0, 9.8 in period 1, and 15 million Euros in eternity. As for the capital ratio, it is 16.2% tier 1 capital for de novo banks $(0.5/3.09 = 0.162)$ and 10.2% after a period $(1/9.8 = 0.102)$. For an established bank in the steady state, the capital ratio is 6.7% $(1/15)$. In the steady state, charter value is 0.225 million Euros. It is rather small. Yet, the overlapping structure of lending relationships decreases strongly the incentive compatible capital ratio.

7.1 Rapid growth

The section determines the capital ratio when the bank hopes to grow more rapidly than the optimal growth path would allow. Alternatively, an established bank may operate in the steady state and then start to grow again. What is the incentive compatible capital ratio of the bank? The model framework is a bit different than above, because the bank does not follow the optimal growth path.

In order to explore this, suppose that a banker gains new wealth, which he can invest in the bank. The new capital induces the banker to expand the size of the bank. The desired size is denoted by $\hat{S}$. Recall that (13) implies that the bank's maximal size with the given amount of equity capital and number of old borrowers

$$S_t^* = (1-p)\hat{E}r/pM_1 + GV_{t-1}, \tag{16}$$

where $\hat{E}$ is the new amount of equity capital. The incentive compatible amount of equity capital can be determined from (16) as a function of the desired size

$$\hat{E} = \frac{pM_1\hat{S} - GV_{t-1}M_1p}{(1-p)r}.$$

Dividing this by $\hat{S}$ gives the incentive compatible capital ratio

$$\hat{e} = \frac{pM_1 - \frac{GV_{t-1}M_1p}{\hat{S}}}{(1-p)r}. \tag{17}$$

Here $\hat{e}$ increases with the desired size $\hat{S}$. When a bank grows without limit, i.e. $\hat{S}$ approaches infinity, the incentive compatible capital ratio approaches

$$\hat{e} = \frac{pM_1}{(1-p)r}.$$

Hence, the capital ratio of a rapidly growing bank is equal to the capital ratio of a de novo bank in period 0, E/S_0 (see (10)). Therefore, a bank that has followed the optimal growth path up to period t, and has thus monotonously lowered its capital ratio, must raise its capital ratio back to the initial high level. Intuitively, when a bank grows rapidly the share of profitable old lending relationships declines in the loan portfolio (see (17)). In the extreme case, this share approaches zero when the bank's size, $\hat{S}$, grows without limit. The positive incentive effect of old lending relationships then disappears and monitoring incentives must be created exclusively with equity capital. This case is identical to that of the de novo bank. Thus, the bank must raise its capital ratio to the initial high level. Let us continue the example. Consider a bank in the *steady state*. Given Proposition 1 and the steady state, (17) simplifies to

$$\hat{e} = \frac{17}{105}\left[1 - \frac{\frac{10}{17}}{1+\%}\right]$$

Here % indicates the growth speed. If it is zero, the capital ratio is at the steady state level, 6.7%. If the bank grows 30% (100%) in a period, the incentive compatible capital ratio is 8.9% (11.4%). It is possible to draw the following conclusions.

Proposition 5 (**Rapid growth**). *If an established bank plans to grow (or if a de novo bank plans to grow more rapidly than the optimal growth path allows), it must attract more equity capital and raise its capital ratio. The faster the growth, the higher the incentive compatible capital ratio is.*

8. Conclusion

This paper contributes to the research on relationship lending (e.g. Rajan 1992; Baas and Schrooten 2006). Relationship loans generate profitable lending relationships thereby creating a positive charter value to the bank even under perfect competition.

Charter value improves the banks' incentives to lend safely and monitor loan applicants carefully. The existence of old lending relationships also lessens the stress to search for and screen new loan applicants. As a result, established banks, which have plenty of old lending relationships, will lend safely even when they operate with a low capital ratio. In contrast, de novo banks and rapidly growing banks that lack old lending relationships, must maintain more equity capital to signal that they lend safely.

Acknowledgment The article is based on the discussion paper no. 16 of the Bank of Finland, Helsinki. I am grateful to the Bank of Finland for a nice working environment.

References

Altunbas, Y., Manganelli, S. and Marques-Ibanez, D. (2011). Bank Risk during the Financial Crisis: Do Business Models Matter? Frankfurt am Mein, European Central Bank, ECB Working Paper No. 1394.

Baas, T. and Schrooten, M. (2006). Relationship Banking and SMEs: A Theoretical Analysis. *Small Business Economics*, 27, 127–137.

Barone, G., Felici, R. and Pagnini, M. (2011). Switching Costs in the Local Credit Markets. *International Journal of Industrial Organization*, 29, 694–704.

Berger, A., Imbierowicz, B. and Rauch, C. (2013). The Roles of Corporate Governance in Bank Failures during the Recent Financial Crisis. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2021799.

Boot, A. (2000). Relationship Banking: What Do We Know? *Journal of Financial Intermediation*, 9(1), 7–25.

Bouckaert, J. and Degryse, H. (2004). Softening Competition by Inducing Switching in Credit Markets. *Journal of Industrial Economics*, LII, 27–52.

Bouckaert, J. and Degryse, H. (2006). Entry and Strategic Information Display in Credit Markets. *Economic Journal*, 116, 702–720.

Calem, P., Gordy, M. and Mester, L. (2006). Switching Costs and Adverse Selection in the Market for Credit Cards: New Evidence. *Journal of Banking and Finance*, 30, 1653–1685.

Cotugno, M., Stefanelli, V. and Torluccio, G. (2013). Relationship Lending, Default Rate and Loan Portfolio Quality. *Applied Financial Economics*, 23, 573–587.

Dell'Arriccia, G., Friedman, E. and Marquez, R. (1999). Adverse Selection as a Barrier to Entry in the Banking Industry. *Rand Journal of Economics*, 30, 515–534.

DeYoung, R. (2003). De Novo Bank Exit. *Journal of Money Credit and Banking*, 35(5), 711–728.

Diamond, D. (1984). Financial Intermediation and Delegated Monitoring. *Review of Economic Studies*, 51(3), 393–414.

Elyasini, E. and Goldberg, L. (2004). Relationship Lending: A Survey of the Literature. *Journal of Banking and Finance*, 56, 315–330.

Gan, J. (2004). Banking Market Structure and Financial Stability: Evidence from the Texas Real Estate Crisis in the 1980s. *Journal of Financial Economics*, 73, 567–601.

Ghosh, S. (2009). Charter Value and Risk-taking: Evidence from Indian Banks. *Journal of Asia Pacific Economy*, 14(3), 270–286.

Gorton, G. and Winton, A. (2002). Financial Intermediation. Philadephia, University of Pennsylvania, The Wharton School, Working Paper No. 02-28.

Gunther, J. (1990). Risk and Failure Among Newly Established Texas Banks. Dallas, Federal Reserve Bank of Dallas, Financial Industry Studies Working Paper No. 90-6.

Hernandez-Canovas, G. and Martinez-Solano, P. (2010). Relationship Lending and SME Financing in the Continental European Bank-based System. *Small Business Journal*, 34, 465–482.

Hunter, W., Verbrugge, J. and Whidbee, D. (1996). Risk Taking and Failure in De Novo Saving and Loans in the 1980s. *Journal of Financial Services Research*, 10, 235–271.

Ioannidou, V. and Ongena, S. (2010). Time for a Change: Loan Conditions and Bank Behaviour when Firms Switch Banks. *Journal of Finance*, LXV(5), 1847–1877.

Jin, J., Kiriradan, K. and Lobo, G. (2011). Ability of Accounting and Audit Quality Variables to Predict Bank Failure during the Financial Crisis. *Journal of Banking and Finance*, 35, 2811–2819.

Keeley, M. (1990). Deposit Insurance, Risk, and Market Power in Banking. *American Economic Review*, 80(5), 1183–1200.

Lee, C.-C. and Hsich, M.-F. (2013). Beyond Bank Competition and Profitability: Can Moral Hazard Tell us More? *Journal of Financial Services Research*, 44, 87–109.

Logan, A. (2000). The Early 1990s Small Banks Crisis: Leading Indicators. *Financial Stability Review*, December 2000.

Neuberger, D., Pedergnana, M. and Räthke-Döppner, S. (2008). Concentration of Bank Relationships in Switzerland: The Result of Firm Structure or Banking Market Structure. *Journal of Financial Services Research*, 33, 101–126.

Niinimäki, J.-P. (2001). Intertemporal Diversification in Financial Intermediation. *Journal of Banking and Finance*, 25(5), 965–991.

Rajan, R. (1992). Insiders and Outsiders: The Choice Between Informed and Arm's Length Debt. *Journal of Finance*, XLVII(4), 1367–1399.

White, L. (1991). *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation*. New York and Oxford, Oxford University Press.

Zhao, T., Matthews, K. and Murinde, V. (2013). Cross-Selling Switching Costs and Imperfect Competition in British Banks. *Journal of Banking and Finance*, 37, 5452–5462.

Appendix A

The expected return of the monitoring strategy from $t+1$ onwards can be written as

$$\begin{aligned}\sum_{i=1}^{\infty}\delta^{i-1}\pi_{t+i}^{m} = \quad & (S_{t+1}-V_t)(R_1-M_1)+V_tR_2-rS_{t+1}-Ec+\\ & \delta\left\{(S_{t+2}-V_{t+1})(R_1-M_1)+V_{t+1}R_2-rS_{t+2}-Ec\right\}+\\ & \delta^2\left\{(S_{t+3}-V_{t+2})(R_1-M_1)+V_{t+2}R_2-rS_{t+3}-Ec\right\}+\\ & \delta^3\left\{\ldots+V_{t+3}R_2\ldots\right\}+\ldots+Er.\end{aligned}$$

Some manipulation gives

$$\begin{aligned}\sum_{i=1}^{\infty}\delta^{i-1}\pi_{t+i}^{m} = \quad & V_t(R_2-r-ec)+\\ & +(S_{t+1}-V_t)\left[R_1-r-M_1-ec\right]+\delta V_{t+1}(R_2-r-ec)+\\ & +\delta\left\{(S_{t+2}-V_{t+1})\left[R_1-r-M_1-ec\right]+\delta V_{t+2}(R_2-r-ec)\right\}+\\ & +\delta^2\left\{(S_{t+3}-V_{t+2})\left[R_1-r-M_1-ec\right]+\delta V_{t+3}(R_2-r-ec)\right\}+\\ & +\ldots+Er.\end{aligned}$$

Since $V_{j+1}=S_{j+1}-V_j\ \forall j,\ j\in\{1,2,\ldots\}$, we have

$$\begin{aligned}\sum_{i=1}^{\infty}\delta^{i-1}\pi_{t+i}^{m} = \quad & V_t\Phi+(S_{t+1}-V_t)\left[R_1-r-M_1-ec+\delta\Phi\right]+\\ & \delta(S_{t+2}-V_{t+1})\left[R_1-r-M_1-ec+\delta\Phi\right]+\\ & \delta^2(S_{t+3}-V_{t+2})\left[R_1-r-M_1-ec+\delta\Phi\right]+\ldots+Er.\end{aligned}$$

Given (2), all sums in the square brackets are equal to zero and the R.H.S. simplifies to $V_t\Phi+Er$. In period 0, value of the expected returns is $V_t\Phi+\frac{1}{2}Er$, because the banker needs to inject $\frac{1}{2}Er$ units of fresh equity capital in the bank at time 1. □

Appendix B

This Appendix shows how (8), $\pi_0^m+\delta(1-p)\left[V_0\Phi+\frac{1}{2}Er\right]=\pi_0^{nm}$, simplifies to (9), $0=pS_0M_1-(1-p)\frac{1}{2}Er$. To begin, (8) consists of three parts.

(i) The part π_0^m is the returns from the monitoring strategy in period 0:

$$\pi_0^m=S_0(R_1-r-M_1)+\tfrac{1}{2}E(r-1)-\tfrac{1}{2}Ec. \tag{B1}$$

(ii) The part π_0^{nm} is the expected returns from non-monitoring in period 0:

$$\pi_0^{nm}=pS_0R_1-pS_0r+p\tfrac{1}{2}E(r-1)-p\tfrac{1}{2}Ec. \tag{B2}$$

The bank fails when the bad risk materializes. It is possible to restate (B2) as

$$\pi_0^{nm}=p\pi_0^m+pS_0M_1. \tag{B3}$$

With probability p, bad loans succeed and the non-monitoring strategy is more profitable than the monitoring strategy, since the bank avoids the monitoring cost, pS_0M_1.

(iii) The part $\delta(1-p)\left[V_0\Phi+\frac{1}{2}Er\right]$ denotes future profits. Here, we have $V_0=S_0$. Inserting $V_0=S_0$, (B1) and (B3) into (8) gives

$$S_0(1-p)\left[R_1-r-M_1-\tfrac{1}{2}cE/S_0+\delta\Phi\right]=pS_0M_1-(1-p)\tfrac{1}{2}Er. \qquad \text{(B4)}$$

The term in the square brackets equals zero, since a borrower provides zero returns for the bank during his lifetime. Hence, (B4) simplifies to $0=pS_0M_1-(1-p)\frac{1}{2}Er$. □

Appendix C

The incentive constraint of period t, $\pi_t^m+\delta(1-p)\left[V_t\Phi+Er\right]=\pi_t^{nm}$ in (11), is shown to simplify to (12), $0=(S_t-V_{t-1})pM_1-(1-p)V_{t-1}\Phi-(1-p)Er$. Here (11) has three parts.

(i) The part π_t^m is the return from the monitoring strategy in period t:

$$\pi_t^m=(S_t-V_{t-1})(R_1-M_1)+V_{t-1}R_2-S_tr+E(r-1)-Ec. \qquad \text{(C1)}$$

On the R.H.S., the first (second) term shows the interest income from new (old) borrowers.

(ii) The part π_t^{nm} is the return from the non-monitoring strategy in period t:

$$\pi_t^{nm}=p\left\{(S_t-V_{t-1})R_1+V_{t-1}R_2-S_tr+E(r-1)-Ec\right\}. \qquad \text{(C2)}$$

A non-monitoring bank mimics a monitoring bank by charging the same loan interest rates. Given Claim 3, the non-monitoring bank fails when the bad risk materializes. When it does not materialize, with probability p, the loan interest income is the same as if the bank had followed the monitoring strategy but the bank avoids the cost of monitoring. We can restate (C2) as

$$\pi_t^{nm}=p\pi_t^m+p(S_t-V_{t-1})M_1. \qquad \text{(C3)}$$

The second term shows the returns from neglecting costly monitoring.

(iii) The part $\delta(1-p)\left[V_t\Phi+Er\right]$ is the future profits. Inserting $V_t=S_t-V_{t-1}$, (C1) and (C3) into (11), gives a rewritten incentive constraint

$$\begin{aligned}(S_t-V_{t-1})(1-p)\left[R_1-r-M_1-ec+\delta\Phi\right]&=\\ =(S_t-V_{t-1})pM_1-(1-p)V_{t-1}\Phi-(1-p)Er.&\end{aligned} \qquad \text{(C4)}$$

The term in the square brackets is equal to zero, since an average borrower provides zero expected returns for the bank over his lifetime. Thus, (C4) simplifies to (12). □

Appendix D

We show that equation (15) is true through induction. The proof has two steps.

Step 1. It is shown that when (15) is true in period t, it will be true also in period $t+1$. Given (13), a bank's size in period $t+1$ can also be expressed as $S^*_{t+1} = 2S^*_0 + V_t G$.

Recalling that $V_t = S_t - V_{t-1}$, it is possible to rewrite $S^*_{t+1} = 2S^*_0 + V_t G$ as $S^*_{t+1} = 2S^*_0 + GS^*_t - GV_{t-1}$ or

$$S^*_{t+1} = 2S^*_0 + GS^*_t - S^*_t + S^*_t - GV_{t-1}. \tag{D1}$$

Recalling from (13) that $S^*_t = 2S^*_0 + GV_{t-1}$, (D1) simplifies to

$$S^*_{t+1} = 4S^*_0 + (G-1)S^*_t. \tag{D2}$$

Inserting S^*_t from (15) to (D2) gives

$$\begin{aligned} S^*_{t+1} &= 4S^*_0 + \left[2(G-1) + 2G(G-1)\sum_{i=1}^{t-1}(G-1)^{i-1} + G(G-1)^{(t+1)-1}\right]S^*_0 \\ &= \left[2 + 2G + 2G\sum_{i=2}^{(t+1)-1}(G-1)^{i-1} + G(G-1)^{(t+1)-1}\right]S^*_0 \\ &= \left[2 + 2G\sum_{i=1}^{(t+1)-1}(G-1)^{i-1} + G(G-1)^{(t+1)-1}\right]S^*_0. \end{aligned} \tag{D3}$$

Now (D3) is fully in accordance with (15). Hence, if (15) is true in period t, it is also true in period $t+1$.

Step 2. We show that (15) is true in periods 1 and 2. When $t = 1$ and $t = 2$, (15) gives

$$S^*_1 = [2+G]\,S^*_0, \quad S^*_2 = [2 + 2G + G(G-1)]\,S^*_0. \tag{D4}$$

It is easy to verify from (14) that S^*_1 in (D4) is true. It is enough to show that S^*_2 is true. To observe this, recall that $S^*_t = 2S^*_0 + GV_{t-1}$ from (13). When $t = 2$, this equals $S^*_2 = 2S^*_0 + GV_1$. If we insert $V_1 = S_1 - S_0$, we have $S^*_2 = 2S^*_0 + GS^*_1 - GS^*_0$. Inserting S^*_1 from (14) into this gives (D4). Hence, (15) is true in periods 1 and 2. It has been shown that (15) states the bank's optimal size correctly in each period. □

Appendix E

The optimal growth speed is

$$\Delta = S_{t+1} - S_t = (G-1)^{t-1}G^2 S^*_0,$$

which implies
$$\frac{d\Delta}{dG} = G(G-1)^{t-2}S_0^*\,[Gt+G-2] > 0.$$
The rate of growth decreases with p
$$\frac{d\Delta}{dp} = \frac{d\Delta}{dG}\frac{dG}{dp} + (G-1)^{t-1}G^2\frac{dS_0^*}{dp} < 0,$$
because $d\Delta/dG > 0$, $dG/dp < 0$, $dS_0^*/dp < 0$. The rate of growth decreases with M_1
$$\frac{d\Delta}{dM_1} = \frac{d\Delta}{dG}\frac{dG}{dM_1} + (G-1)^{t-1}G^2\frac{dS_0^*}{dM_1} < 0,$$
because $d\Delta/dG > 0$, $dG/dM_1 < 0$, $dS_0^*/dM_1 < 0$. Because $M_1 = m_1/g$, we obtain
$$\begin{aligned}\frac{d\Delta}{dm_1} &= \frac{d\Delta}{dM_1}\frac{dM_1}{dm_1} = \frac{1}{g}\frac{d\Delta}{dM_1} < 0,\\ \frac{d\Delta}{dg} &= \frac{d\Delta}{dM_1}\frac{dM_1}{dg} = \frac{-m_1}{g^2}\frac{d\Delta}{dM_1} > 0.\end{aligned}$$
The rate of growth decreases with m_1 and increases with g. The rate of growth increases with Φ
$$\frac{d\Delta}{d\Phi} = \frac{d\Delta}{dG}\frac{dG}{d\Phi} > 0,$$
because $d\Delta/dG > 0$, $dG/d\Phi > 0$.

The rate of growth increases with time if
$$\frac{d\Delta}{dt} = \frac{d}{dt}\,[G-1]^{t-1}\,G^2S_0^* = \ln[G-1]\,[G-1]^{t-1}\,G^2S_0^*.$$
This is positive if $G-1 > 1$, that is $G > 2$. □