Tilburg University

Banking

van Damme, E.E.C.

Published in:
Oxford Review of Economic Policy

Publication date:
1994

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
van Damme, E. E. C. (1994). Banking: A survey of recent microeconomic theory. *Oxford Review of Economic Policy*, 10(10), 14-33.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

BANKING: A SURVEY OF RECENT MICROECONOMIC THEORY

ERIC VAN DAMME

*CentER for Economic Research, Tilburg University*¹

I. INTRODUCTION

In the last decade, economic theorists have again turned their attention to the questions of how financial institutions influence the allocation of funds for investment and the possibilities for economic growth. Although the relationships between the real and the financial sectors of the economy had been extensively discussed during and immediately after the Great Depression—mainly on the basis of Fisher's (1933) debt deflation theory—interest in this topic had waned in the three decades following the Second World War. Two reasons can be advanced for why the theory then largely ignored the relations between the real and financial sectors of the economy:

- (i) The macroeconomic literature following Keynes' General Theory stressed outside money as the financial variable most relevant to aggregate economic behaviour.² The main support for the importance of the monetary mechanism was provided by the Friedman and Schwarz (1963) study of the historical relationship between money and output, a study that took emphasis away from the significance of all other aspects of the financial system.
- (ii) The microeconomic literature relied almost entirely on the Walrasian paradigm of perfect competition on markets in which exchange is costless. In this Arrow/Debreu world with a complete set of markets there are no trade

¹ This paper resulted from discussions among a group of researchers that met at CentER to study the recent developments in banking theory. I thank the members of the group, Svend Albæk, Helmut Bester, Piet Moerland, Henk Oosterhout, Eduardo Siandra, Jacques Sijben, Dale Stahl, and Jonathan Thomas, for their contributions. I am grateful to Geert Almekinders, Sylvester Eijffinger, and Martin Fase for their comments on an earlier version. The usual disclaimer applies.

² In the section on financial markets, Blanchard and Fischer (1989) write, 'Despite the complexity and sophistication of financial markets, they are typically represented in macroeconomic models by only two variables: the money stock and an interest rate' (p. 478).

frictions and institutional details can be neglected.³ There is, hence, no link between the real and financial sectors of the economy—the financial sector is just a veil—and financial structure is irrelevant for investment decision and the value of the firm, as the famous Modigliani and Miller (1958) proposition shows.

Although some economists, such as Gurley and Shaw (1955), argued that financial intermediaries were important since they alleviated the trade frictions between borrowers and lenders, these ideas were not incorporated into economic theory since they could not (yet) be formalized. Similarly, although economists were certainly aware of the restrictiveness of the assumptions underlying the Modigliani and Miller irrelevance theorem, the investigation of the robustness of this result had to wait until the trade frictions associated with an incomplete set of markets could be formally modelled in a satisfactory way. The latter happened only in the 1970s, when work in information economics showed that inefficiencies in trade can arise when one of the transacting parties has superior information concerning the value of the trade and that, in such circumstances, intermediaries have a real role to play (Akerlof, 1970; Jensen and Meckling, 1976; Leland and Pyle, 1977). The aim of this paper is to survey the recent developments in this theory of imperfect information in so far as they relate to banking and financial intermediation.⁴ We restrict ourselves to microeconomic issues. For discussions about the links between the financial system and aggregate economic behaviour, we refer to Bernanke (1983), Gertler (1988), Kindleberger (1978), and Mishkin (1990).⁵

We start in section II by reviewing the main functions of banks that are described in the literature. In section III we focus on the asset side of the bank's balance sheet, i.e. on the relation between a bank and the firms to which it provides funds. We discuss three types of information problems and three corresponding roles for banks: the screening of projects, the monitoring of project choices, and

the monitoring of project outcomes. We take a closer look at the latter role, that of a bank as a delegated monitor (Diamond, 1984). Attention is shifted to the relation between a bank and its lenders in section IV. In particular, following the influential paper by Diamond and Dybvig (1983), we discuss the possibility of bank runs and show how such panics can be avoided by deposit insurance schemes. Section V is devoted to discussing competition in the banking sector. As banks compete simultaneously at the input and output markets, the natural question that arises is whether aggressive behaviour on one of markets allows a bank to monopolize the other market (Yanelle, 1991). In section VI we turn to a comparison of different financial systems. In particular we touch upon the comparison between the US/UK 'market' system with the Japanese/German 'banking' system, i.e. we investigate whether long-term relationships between firms and banks foster economic growth. Section VII concludes with some remarks concerning the (de-)regulation of the banking sector.

II. FINANCIAL INTERMEDIATION

Despite the important role played by intermediaries in many markets, the theory of intermediation is still in its infancy. Intermediation can be profitable only in markets in which trade frictions exist, but such frictions do not exist in the standard market models that populate economics textbooks. Actually, until recently, theoretical analysis relied almost exclusively on the Walrasian paradigm of anonymous frictionless markets and in this model the actual process of trading and price formation is left unmodelled. To address the question of why intermediaries exist is equivalent to asking why bilateral transactions between buyers and sellers are not efficient, i.e. to ask what the trade frictions are and how intermediaries can alleviate them. The traditional banking literature focuses on the imperfection that arises from maturity mismatch. While lenders prefer to hold liquid assets, borrowers pre-

³ The so-called 'equivalence theorems' justify the restriction to Walrasian equilibria, given complete markets, see Aumann (1964).

⁴ Another recent survey is Bhattacharya and Thakor (1993).

⁵ Blanchard and Fischer (1989) conclude their discussion about the role of financial markets with 'We believe that the recent work attempting to account for certain features of the financial markets from the viewpoint of asymmetric information is extremely important and that it will be increasingly integrated in complete macroeconomic models.' (p. 489).

fer to issue loans that have a time to maturity that is not too short (Gurley and Shaw, 1955). The more recent literature stresses imperfections because of asymmetries of information: borrowers know more about their projects than prospective lenders. Both types of imperfection imply that banking may be profitable and efficiency-enhancing. Once the public holds deposits with banks it is efficient to let banks provide the payments between depositors. Hence, the three main functions of banks described in the literature are:

(i) *Transformation*

Banks transform the illiquid primary securities issued by firms into secondary liquid assets desired by consumers.

(ii) *Delegated screening and monitoring*

Banks screen potential borrowers and monitor actual borrowers on behalf of the depositors.

(iii) *Payments*

Banks facilitate payments between consumers by providing a payments system.

Gurley and Shaw (1955) focus on the transformation role of banks. Financial intermediaries alleviate the imperfections arising from maturity mismatch. They issue the (low-yield) liquid liabilities that are preferred by lenders and invest a proportion of the funds thus received into higher-yielding earning assets of the type that borrowers prefer to issue. Their theory assumes that the transformation technology is non-convex (for example, there exist indivisibilities and/or economies of scale in transactions in financial markets), and hence that there are limits to the amount of risk-sharing and diversification that are feasible (profitable) through direct transactions, so that intermediaries enable cost savings to be obtained by allowing transactions in bulk. To a certain extent this approach is incomplete, since the existence of transaction costs is assumed rather than derived. It is not explained why direct finance through markets would not allow full risk diversification.

The more recent literature argues that transaction costs arise from imperfect information and the

peculiar character of credit markets, on which goods (money) that are delivered today are traded against *promises* to deliver goods (money plus interest, or profit shares) at some date in the future. Obviously, such a market can function only if the promise is sufficiently credible. This literature (which follows the seminal papers by Leland and Pyle (1977) and Jensen and Meckling (1976)) views intermediaries as institutions that can reduce or eliminate these information and credibility problems.

The literature starts from the assumption that asymmetries of information will prevent markets from being complete. Borrowers typically will have much better information about their future prospects and, if the informational gap is sufficiently wide, no lender will be willing to trade with a borrower because he fears being deceived, and markets cannot be established (Akerlof, 1970). Although, the borrower might engage in signalling activities to reduce the information gap (e.g. a trustworthy auditing firm might be hired), frequently (e.g. if the borrower is a recently established firm or a consumer without a track record) the signalling instruments available will not be sufficient to allow markets to open. In such cases a bank might step in and screen the borrower.⁶ As screening reduces the information gap between the bank and the borrower, profitable bilateral transactions may take place in case the market could not be established. Hence, banks may reduce the transactions costs resulting from adverse selection. Furthermore, by engaging in monitoring activities during the contract execution stage, banks can also reduce the transactions costs resulting from moral hazard. These monitoring activities by banks may also generate external benefits. Renewal of short-term bank loans provides a positive signal to other market participants and, hence, avoids duplication of information costs (Fama, 1985). In effect, the firm borrows the bank's reputation. Finally, note that such bilateral transactions provide the bank with inside information that might be worthwhile in future interactions with the borrower. Hence, the inside bank acquires a competitive advantage and it is natural for long-term relations to develop. Consequently, financial intermediation might pro-

⁶ Schumpeter (1934, pp. 72–4) calls bankers the gatekeepers of capitalist economic development and describes their strategic function as to screen potential innovators and advance credit to the most promising.

vide a mechanism of commitment to a long-term relationship (Mayer, 1988; Hellwig, 1991). In section VII we consider the (dis-)advantages of such longer-term relations.

Since the bank has superior information about the true value of the loans of its borrowers and cannot easily communicate this information to the market, it follows that a large part of the bank's assets are not marketable. This non-marketability of assets might be said to constitute the essential difference between a bank and another type of financial intermediary.⁷ It might create problems since the liabilities of banks consist, for a large part, of liquid deposits whose returns are not tied to the bank's performance. Consequently, a bank always faces the risk that it underestimates the number of withdrawals and is unable to fulfil its obligations. Hence, a bank might be subject to runs and panics. Since such runs might be contagious and might result in much vital information being lost to the economy, it is important that unnecessary runs be avoided. In section IV, which is based on Diamond and Dybvig (1983), we will investigate runs in more detail and discuss various institutions that can prevent them.

Besides providing credit for risky projects and making fixed-interest loans to the public, banks, of course, provide many other facilities that improve the allocation of resources. For example, they facilitate transactions by providing payment services. That banks provide such services is very natural: for consumers, bank accounts are more attractive than holding cash, and payment through the banking system is cheaper on average than personal payment. Nevertheless, it might be argued (e.g. Goodhart, 1989, p. 118) that this service is inessential to banking and that it might just as well, or perhaps even better, be carried out by other financial intermediaries. (Payment services might be offered on the basis of marketable assets and thereby runs and panics automatically avoided.) Be that as it may, it seems very unlikely that banks will give up this service since it is so useful to them. When a client has an account with a bank, the bank can easily monitor the financial transactions of the client and in this way it obtains vital information about the client's credit-worthiness, reliability, pref-

erences, etc. Hence, the bank has a comparative advantage when making loans to its own depositors. Indeed, opening a checking account is often a requirement for obtaining a loan. Since this information is not only valuable to the bank itself, but also to other financial intermediaries, such as insurance companies, and since the transfer of this information is easier within one firm than between two firms, one might expect that the deregulation of the banking and insurance industry will lead to insurance companies trying to get access to this information by merging with banks.

III. THE BANK/BORROWER RELATION

Imagine an entrepreneur who wants to raise investment capital in order to carry out a project that might possibly yield substantial returns. In most cases, entrepreneurs know considerably more about their business than anyone else and, hence, even in the case where it is efficient to carry out the project, there seem to be three reasons why entrepreneurs might not be able to raise the capital from investors:

(i) *The adverse-selection problem*

Entrepreneurs might not be able to convince the investors that the project really is as profitable as claimed. Since, also, entrepreneurs with low quality projects can possibly gain by claiming that their project is of good quality, the claim that the investment will be profitable cannot be taken at face value.

(ii) *The moral-hazard problem*

Entrepreneurs might not be able to convince the investors that they will really invest money in the project proposed, rather than in some more risky project or in a project that yields higher personal benefits. Since, obviously, 'entrepreneurs' who want to misuse the money can raise funds only if they claim that the money will be put to good use, a claim of the latter type is not necessarily credible.

(iii) *The ex post verification problem*

They might not be able to convince the investors that, if the project is successful, they will indeed pay the dividends that they promised.

⁷ Dewatripont and Tirole (1993) argue that the essential difference between a bank and an ordinary firm is that the bank's liabilities consist mainly of debt (i.e. deposits) and that this debt is held by many small depositors.

Since any money that is not returned to investors remains in the hands of the entrepreneur, they might, *ex post*, have an incentive to claim to have been less successful than they actually were, in order to minimize the pay-out.

Simple examples illustrate these three cases further.

Initial public offerings provide a good example of the adverse-selection problem. Investors have to sort out whether the firm is offered simply for reasons of risk-sharing or whether the entrepreneur wants to get rid of the firm before the bad news comes out in the open. A simple numerical example (based on Akerlof, 1970) is the following. Suppose the entrepreneur needs to raise an investment of I in order to carry out a project that with probability x yields a return of \$100, and with probability $1 - x$ yields a return of \$0. It is efficient to carry out the project if $I < 100x$. Assume that the entrepreneur knows the exact value of x but that investors only know the portfolio of projects, say they consider x to be uniform on $[0,1]$. Assume further that, *ex post*, it can be verified whether or not the project was successful. If investors cannot screen projects, they will be willing to invest only if the average revenue of the portfolio exceeds the investment, i.e. if $I < 50$. Consequently, if $I > 50$, then no projects will be carried out, even though it would be profitable to carry out the projects with $I < 100x$. It is easy to see how in this case an intermediary who has the opportunity to monitor x at a cost of C could induce a more efficient outcome. Monitoring results in the funding of all profitable projects and it pays⁸ if $C < 50 - I + I^2/200$. Note that since information is a public good (monitoring by a second person does not add new information), it is efficient if just one participant engages in monitoring activities. Furthermore, because of this public good character, there exists a free-rider problem and an investor will only be willing to acquire information if this gives certain exclusivity rights. Namely, if outsiders observe that the

monitor is willing to invest, they rationally infer that the project is profitable and, since they have not incurred the monitoring costs, they are able to underbid the insider.⁹ Consequently, only an investor that can finance the project on its own has an incentive to monitor. If there are only small investors, nobody will acquire information (cf. Grossman and Stiglitz, 1980).

In their seminal paper, Stiglitz and Weiss (1981) discuss a slight modification of the above example to construct a theory of credit rationing. In the paper it is assumed that there is a portfolio of projects, x , distributed on $[0,1]$ according to a distribution function, F . Project x needs an investment of I , it yields a return, $R(x)$, with probability x and a return 0 with probability $1 - x$. It is assumed that all projects have the same expected return, hence $xR(x) = E$ for all x . The entrepreneur knows x , but lenders only know the portfolio. In addition, *ex post*, it is only revealed whether or not the project failed, the actual return remains unknown to borrowers. Hence, the repayment obligation R cannot be conditioned on x . Given R , an entrepreneur with project x only decides to apply¹⁰ for a loan of I if $R(x) \geq R$, hence $x \leq E/R$. Consequently, raising the repayment obligation R has adverse effects: it raises the overall riskiness of the pool of applicants, there is adverse selection. If this effect is sufficiently strong, then the lender's profit,

$$\int_0^{E/R} (xR - I) dF(x) \quad (1)$$

need not be monotonic in R . Hence, equilibrium may occur at an interest rate $r = R/I - 1$ at which there is credit rationing. Even though some lenders who are denied credit would be willing to pay a higher interest rate, lenders are unwilling to increase the interest rate since this would lower their profits.

Stiglitz and Weiss (1981) also discuss the moral-hazard problem. This problem arises, for example, when, once entrepreneurs have received the de-

⁸ If the intermediary monitors, it will decide to finance only those projects with $100x \geq I$, i.e. for which the expected returns exceeds the investments. The aggregate expected return, therefore, is the expected value of $\max(0, 100x - I)$ and it pays to pay the monitoring cost if this expectation exceeds C .

⁹ Fama (1985) argues that renewal of short-term bank loans provides a positive externality to other market participants. If the bank renews, other investors are willing to step in as well. Hence, the bank privately supplies a public good. If the benefits are sufficiently high, such private provision is indeed optimal.

¹⁰ This property follows if there are (vanishing) costs for the entrepreneur associated with bankruptcy.

sired funds, they invest them not according to the business plan they proposed, but rather in a more risky project. For example, they might start expanding on a larger scale knowing that, once they have raised some initial investment, the investors will be more likely to invest more, fearing that otherwise they might lose their initial investment. Stiglitz and Weiss (1981) discuss the following numerical example. The entrepreneur, who has raised I for investment, has a portfolio of investment projects available. Project x yields the return $\$R(x)$ with probability x and $\$0$ with probability $1 - x$. The choice of x is not observable to investors: they just learn whether the project was successful or not, they do not get to know the actual realization of the returns. Consequently, the contract between investors and the entrepreneur specifies that the latter will pay back a fixed amount, R , in case the project is successful, and nothing otherwise. The efficient project is the one that maximizes the expected return $xR(x)$. The entrepreneur, however, does not have the incentive to maximize social surplus, but, rather, wants to maximize the expected net return $x(R(x) - R)$, and in most cases the project that is optimal from the entrepreneur's point of view will not be efficient. Typically,¹¹ the entrepreneur will choose a project that is too risky (i.e. one with the probability of success being small but with a high pay-off in the case of success). Consequently, if the agency problem is sufficiently severe, then investors may be unwilling to invest in the first place, even though the efficient project is profitable. As in the above adverse-selection example, credit rationing may be the resulting equilibrium outcome. Lenders are unwilling to raise R since this could lead to more risky project choices by the entrepreneur and to lower profits for lenders.

The problem of *ex-post* verifiability is known from bankruptcy cases or, more generally, from the dissolving of partnerships. If the entrepreneur can divert money or goods from the firm to private sources and if it cannot be verified whether such divestiture has occurred, the entrepreneur will engage in such activities until the level where they are optimal from his personal point of view and this level is excessive from society's point of view. A simple numerical example is as follows. The entrepreneur's project yields a random return, X , where

X has a cumulative distribution function, F , with $F(0) = 0$ (i.e. there are no negative returns); the entrepreneur observes the realization x of X but investors only know F . In this case, the entrepreneur can always (and will always) credibly claim to have had extremely bad luck, that the project failed, hence $x = 0$, and the investors will make losses on their investment. Rational investors realize that the promise to repay is not credible, hence, they will not be willing to invest. Gale and Hellwig (1985) and Diamond (1984) show that the picture is not so gloomy if the lenders can monitor the project returns or if non-pecuniary penalties can be levied upon the entrepreneur in case of default (see also Townsend, 1979). In these cases, the entrepreneur can be given the right incentives to pay back. The remainder of the section illustrates this by discussing the model from Diamond (1984). The conclusion from that article is that banks can play an important role as delegated monitors.

The basic situation analysed in Diamond (1984) is one where a firm has to borrow money to finance some investment. Diamond restricts himself to the *ex-post* verifiability problem, there are no other moral-hazard problems, nor adverse-selection problems. Each potential lender has a limited amount of wealth so that the firm needs the funds of, say, m lenders. The investment yields some random return, X , which realization an outside lender can observe only after having spent a monitoring cost, C . Still the other lenders do not observe x as a result. This means that if all the lenders want to be informed about x , the total monitoring cost is mC . One possible way of solving the contracting problem between the m lenders and the firm involves each lender indeed monitoring the firm. The firm promises each of the lenders some part of its future profit, X , in return for obtaining their wealth and this solution is feasible if the expected return on the project $E(X)$ exceeds the sum of the investment cost and the monitoring cost, mC . The total deadweight-loss with this solution equals mC .

An alternative solution does not involve any monitoring cost but relies on the construction of an 'incentive contract'. If the firm repays amount z , a non-pecuniary punishment, $\phi(z)$, is inflicted upon it. The punishment $\phi(z)$ can be interpreted as bank-

¹¹ This holds if $2R'(x) + xR''(x) \leq 0$ for all x .

ruptcy penalties or the firm's loss of 'reputation' in bankruptcy. Given this penalty function, $\phi(\cdot)$, the firm chooses its payment to maximize its net profit $x - z - \phi(z)$ subject to $z \leq x$. Diamond shows that the optimal incentive contract, $\phi^*(\cdot)$, takes the following simple form. As long as the firm pays some fixed amount $\bar{z}$, it is not punished, i.e. $\phi^*(z) = 0$ for $z \geq \bar{z}$. If it pays some amount $z < \bar{z}$, then the punishment is $\phi^*(z) = \bar{z} - z$. Given this punishment scheme, the firm will choose to pay $\bar{z}$ as long as $x \geq \bar{z}$. If the project returns, x , are not sufficient to repay $\bar{z}$, the firm pays $z = x$ and incurs the punishment $\bar{z} - x$. Hence, this outcome is very much like a standard debt contract with a fixed repayment obligation, $\bar{z}$. Whenever the firm fails to pay $\bar{z}$, it is punished by bankruptcy and is not allowed to keep any of its profits, x . Of course, this contract also leads to some deadweight loss in the form of bankruptcy costs. These costs are equal to the expected non-pecuniary penalty, $E\{\phi^*(z)\} = E\{\phi^*(X)\}$. The cost of contracting determines which kind of contract is optimal. Clearly the incentive contract is superior to the monitoring solution if and only if $E\{\phi^*(z)\} < mC$.

The second part of Diamond's analysis uses these insights to construct a model of intermediation. Consider a bank which contracts with mN lenders to receive deposits for funding N firms. Assume the firms' project returns are independent and identically distributed. Also assume $E\{\phi^*(z)\} > mC$ so that the lenders would optimally select the monitoring solution if they were to lend their money directly to the firms. Of course, if the bank obtains the funds from m lenders and performs the monitoring, it incurs only the cost, C . Since $C < E\{\phi^*(z)\}$ the bank will sign a contract under which it monitors the outcome, x , of each firm it finances. In comparison with direct lending and borrowing this reduces the total monitoring cost per firm by $(m - 1)C$. As a result, it appears optimal to let the bank monitor the firms on the behalf of the depositors. The contract between a firm, i , and the bank then specifies some fraction, $g_i(x_i)$, of profits, x_i , that firm i has to pay to the bank in case its realized profits are x_i . The bank uses the total sum it receives, $\sum_{i=1}^N g_i(x_i)$, to make its payments to the depositors.

This raises the obvious question of how the bank is monitored. The bank might cheat and claim that all

firms had negligible profits so that it cannot pay back anything to the lenders. Clearly, this problem cannot be solved by letting each lender monitor the bank because this would actually increase the monitoring cost per project to $(m + 1)C$. Diamond's solution is that the lenders sign an incentive contract with the bank. Under this contract the bank is obliged to pay some fixed amount, $\bar{z}_B$, of money per deposit unit. Should the bank fail to fulfil this obligation, it is punished by bankruptcy. Hence, this optimal incentive contract is just like an ordinary deposit contract. It is optimal for the bank to pay out $\bar{z}_B$ whenever it can do so, but, since with some probability the bank may actually be unable to make the promised payment, there is a deadweight loss of $D_N = E\phi^*_B(\sum_i g_i(X_i))$. These costs might be termed the 'cost of delegation'. Diamond's important observation is that these costs become smaller and smaller as the number of firms, N , is increased. The reason is the 'law of large numbers'. When the firms' profits are independently distributed, then the average profit per firm is not very risky when N is large enough. Therefore, given $\bar{z}_B < E(X)$, the probability that the bank cannot pay $\bar{z}_B$ is rather small if it can diversify among a large number of independent investments. Hence, $D_N \rightarrow 0$ as $N \rightarrow \infty$ and, as a result, delegated monitoring may save resources. The net gain from financial intermediation amounts to $(m - 1)C - D_N$, which is certainly positive when N is large enough.

The key to Diamond's theory, therefore, is the diversification of the bank that allows the incentive problems to be resolved. The bank finances and monitors a large number of firms, so that by the law of large numbers, the bank's returns are approximately deterministic and the agency problem between the bank and its lenders vanishes. Hence, the bank can offer depositors a fixed-interest deposit scheme. At this point it is natural to ask whether the incentive problems are really as severe as the theory wants us to believe and whether all important moral-hazard problems have been taken into account. Recall that Diamond's model abstracts away from the portfolio-selection problem of banks: all projects are assumed identical. If this assumption is dropped, there might be an additional moral-hazard problem in that bank managers might have an incentive to select projects that are too risky. This is particularly so if government insures deposits using a flat rate, see Merton (1977). We return to

my remains in autarky. Each agent invests in the technology initially, but agents of type 1 withdraw their investment immediately after they have learned their type. Hence, type 1 agents consume 1 unit at time 1, type 2 agents consume R units at time 2 and agents remain uninsured. This market outcome is inefficient and a Pareto improvement can be obtained if consumers get together and form a mutual bank at date 0. The bank pools resources, allows type 1 consumers to consume c_1 at date 1 (hence it keeps a fraction tc_1 of the deposits in reserves) and it invests $1 - tc_1$ into the technology. The optimal contract maximizes the consumer's *ex ante* expected utility, hence, it solves

$$\begin{aligned} \max & \quad tu(c_1) + (1 - t)u(c_2) \\ \text{st.} & \quad (1 - t)c_2 = R(1 - tc_1). \end{aligned} \quad (2)$$

In the optimum, $u'(c_1^*) = Ru'(c_2^*)$, hence $c_2^* > c_1^*$ so that the contract is incentive compatible: type 2 consumers do not have an incentive to claim to be type 1 if they indeed get c_2^* at $t = 2$. Furthermore, if the coefficient of relative risk-aversion exceeds 1, the optimum involves risk-sharing, i.e. $1 < c_1^* < c_2^* < R$.

One way for the bank to realize the optimal allocation is by means of a simple demand deposit contract. The bank offers each withdrawer at time 1 a return of r for each unit deposited at time 0. At time 2 the bank is dissolved and the assets remaining in the bank are shared equally among all consumers that have not withdrawn at time 1. The efficient allocation is realized if the bank puts $r = c_1^*$ and if only type 1 consumers withdraw at date 1. Furthermore, since $r < c_2^*$, it seems that type 2 consumers have no incentive to withdraw early, hence, the banks should not face any liquidity problems. However, closer investigation reveals that the answer is not that straightforward. The optimal choice of a type 2 consumer depends on his expectations about what the other type 2 individuals will do. Assume that the type 2 consumers conjecture that a fraction s of the population will withdraw at time 1 (where, naturally $s \geq t$). As long as $rs \leq 1$ the bank will be able to fulfil its contractual obligations and it will be optimal for a type 2 consumer not to withdraw as long as

$$s \leq (R - r) / r(R - 1). \quad (3)$$

(Note that the right-hand side in this expression exceeds t .) If $rs > 1$, then the bank cannot fulfil its obligations: even if the bank withdraws all funds from investment, funds are not sufficient to satisfy the demand for withdrawals at date 1. If we assume that consumers get served on a 'first-come first-served' basis then, if a consumer withdraws at time 1, there is at least a chance of getting something, while the consumer is bound to get nothing if he waits till time 2. Hence, in this case all consumers will choose to withdraw early. Consequently, if the expected s satisfies (3), then the realized s is equal to t . If the expected s does not satisfy (3), then $s = 1$. Hence, if the bank offers a demand deposit contract with return $r > 1$ such that (3) holds for $s = t$, then there are two possible conjectures that are consistent (self-fulfilling). Consequently, there are two (rational expectations, or Nash) equilibria:

- (i) the good equilibrium, $s = t$: only the type 1 consumers withdraw at time 1;
- (ii) the bank-run equilibrium, $s = 1$: everybody withdraws early.

Note that in this latter run equilibrium, every consumer has an expected utility of $u(r)/r$, hence the expected utility is less than $u(1)$ so that this equilibrium is strictly worse than autarky. Of course, the good equilibrium is strictly better than autarky. A case might be made that, under the conditions of the model, only the good equilibrium is viable. After all, if people expected the bank-run equilibrium to occur, they would not deposit their money in the bank in the first place. However, simple extensions of the model make this argument less compelling (see, for example, Chari and Jagannathan, 1988). If consumers start out with optimistic expectations but then, in the course of time, become more pessimistic because of bad news of the economy in general (a bad earnings report, a bank run at another bank, a stock-market crash, outbreak of a war, the death of a president, or maybe sunspots), the bank-run equilibrium might become a very real possibility.

Let us now focus on the question of how the bank-run equilibrium can be prevented. First of all, note that reserve requirements cannot avoid runs, unless the requirement is 100 per cent, but then we are back to autarky. One way the bank can defend itself

this problem below. On the other hand, a little reflection reveals that the domain of applicability of Diamond's theory might be restricted to situations with small, relatively unknown firms and large banks. If a firm is large and needs to make huge investments, then one bank will probably not be able to diversify the risk sufficiently. Furthermore, in such a case the firm can possibly put up enough collateral so that incentive problems vanish (Bester, 1985), or it might even be able to finance its investments from retained earnings and, as the latter involves no agency costs, it is actually preferable (Myers and Majluf, 1984). Finally, large firms have credit ratings which allow them to borrow directly, by issuing bonds or commercial paper on the market. Since they may want to borrow repeatedly, and since their actions today influence their credit-ratings and, hence, the borrowing terms that they will be offered in the future, they will want to keep their good reputation. The incentive constraints that such firms face are, therefore, different from the one discussed in the above one-period world. Following the seminal work on reputations in markets by Kreps and Wilson (1982), Diamond (1991) constructs a model of repeated borrowing by firms and he allows that firms with high credit-ratings will rely on the market while those with ratings that are more to the middle of the spectrum will rely on bank loans. This theory, however, is dismissed by Davis and Mayer (1991), who provide empirical evidence that firms often continue to obtain bank finance *after* they have issued bonds. We will return to this issue of bank finance versus market finance in section VI.

IV. THE BANK/LENDER RELATION

Bank runs are a common feature of the extreme crises that have played a prominent role in monetary history (Kindleberger, 1978). Bernanke (1983) has shown that much of the economic damage in the Great Depression was caused directly by bank runs, the essential reason being that much vital information was lost as a consequence of them (see also Mishkin, 1990). Consequently, it is important to

understand why bank runs occur and how they can be prevented. In this section we study these issues by reviewing the seminal paper by Diamond and Dybvig (1983)¹² and some of the literature that has followed that paper.

The Diamond and Dybvig paper sees the role of the bank as providing insurance against random shocks to consumers' preferences for the *timing* of consumption. As the authors want to concentrate on the liability side of the bank's balance sheet, they choose to model the economy's technology in a very rudimentary and simple way. There is one productive, illiquid, riskless technology that involves constant returns to scale. In return for 1 unit of input today (date 0) it yields $R > 1$ units of output in the future (date 2). Investments can be withdrawn without costs (but also without profits) tomorrow (date 1).¹³ Consumers also have access to an unproductive, liquid, storage technology. Consumers start out in identical positions today, each having 1 unit to invest. All consumers face the risk of needing to consume tomorrow (date 1), i.e. at a point in time when the returns of the productive technology have not yet been realized. A certain fraction, t , of the consumers will have to consume at date 1, and at date 1 each consumer will know whether he belongs to this unlucky group or not. We will refer to the unlucky group as 'type 1 consumers' and to the others as 'type 2 consumers'. Consumers are assumed to be risk-averse with utility function, u , and they would like to insure themselves against the risk of being of type 1.

Diamond and Dybvig make the important assumption that a consumer's need for liquidity tomorrow is private information, i.e. that tomorrow it cannot be verified whether a consumer really belongs to the type 1 group or whether he just claims this to be so. This lack of observability of agent's types rules out a complete market of Arrow/Debreu state contingent claims, hence, only markets for uncontracted claims can be established. Since agents are identical and the technology involves constant returns to scale, it is easily seen that the equilibrium on these markets involves no trade, hence, the econo-

¹² Bryant (1980) is another early paper making a similar point. A distinction between these papers is that Bryant focused on information-based runs (i.e. some investors know that the bank actually is insolvent), while Diamond/Dybvig consider the case of pure panics (see below).

¹³ If it is common knowledge that investments cannot be withdrawn until date 2, then the 'bad equilibrium' (i.e. the bank run) does not exist.

against runs is by simply refusing to allow further withdrawals after a certain number have been made, i.e. there is suspension of convertibility. In this simple model the bank could refuse to accept withdrawals after a fraction, t , of the population has withdrawn. This policy forces the expectations of the type 2 consumers to fulfill $s \leq t$. Consequently they find it optimal not to withdraw and the suspension of convertibility will not occur in equilibrium. Of course, the only reason that this policy works so well is that the model is so highly stylized: the bank knows exactly what fraction of the population has a need for liquidity. With less information the bank might underestimate the demand for liquidity and suspension of convertibility may imply that fewer consumers are served in the first period than is required by efficiency. Diamond and Dybvig show that, if the fraction of consumers that has liquidity needs is stochastic, *no* bank contract can achieve optimal risk-sharing. However, in that case efficiency can be obtained by means of demand deposit contracts backed by government deposit insurance. Essentially, the Diamond and Dybvig deposit-insurance mechanism amounts to a tax that is levied on the early withdrawals and that is conditioned on the total amount of early withdrawals. In effect the government's promise to provide insurance is accompanied by the threat to the public to levy high taxes if there are too many withdrawals and this threat discourages type 2 consumers from withdrawing early. Since the government can tax *ex post* (i.e. reduce withdrawals after they have been made), the government is able to provide better risk sharing and Diamond and Dybvig conclude that there is a potential benefit from government intervention into banking markets. Note that the power to tax gives a second advantage to the government: there is no need to keep reserves. Diamond and Dybvig, however, also point out that deposit insurance has drawbacks. If deposit insurance is provided, then banks have an incentive to take on interest-rate risk: they might mismatch maturities of assets and liabilities, or they might accept too risky investments, or they might simply neglect their task of monitoring the performance of their borrowers (see also Merton, 1977). As the 'Savings & Loan' (S&L) crises in the US have shown, these concerns are real.

Several researchers have sought to investigate the robustness of the Diamond and Dybvig results by

analysing variations of the basic model outlined above. Jacklin (1987) argues that, given the simple technology assumed by Diamond and Dybvig, one can question the assumption that banks can only offer deposit contracts. He shows that dividend-paying traded equity contracts can achieve the same efficient allocation while preventing bank runs from occurring. However, he also shows that this equivalence depends on the restrictive nature of the preferences assumed in Diamond and Dybvig. With a richer specification, demand deposits typically outperform equity contracts. Other papers allow for a richer production sector. For example, the returns of the profitable technology might be stochastic and, over time, investors' information about the distribution of the returns might improve. In this case, besides pure panics, information-based runs might occur. Some agents who do not inherently want to consume early might obtain information of likely future low returns of the technology and might prefer to withdraw early. Runs might arise if uninformed agents, perhaps mistakenly, interpret long queues at the bank as a signal that others have unfavourable information and withdraw early as well. Postlewaite and Vives (1987), Jacklin and Bhattacharya (1988), and Chari and Jagannathan (1988) construct models of this type in which bank runs occur in the *unique* equilibrium. In these models of information-based runs, deposit insurance plays a somewhat different role than in the simple 'panic' model discussed above. Deposit insurance reduces the incentive to gather information about the solvency of the bank. This might lead to fewer runs, but it might also have the adverse effect of increasing moral-hazard problems with the bank.

Bhattacharya and Gale (1987) extend the Diamond and Dybvig model to a more macro setting in which there are multiple, geographically dispersed banks. Each bank is a local monopolist that finds itself confronted with the Diamond and Dybvig problem. It is assumed that there exists an interbank market which allows a bank that faces unexpected high liquidity demand at time 1 to borrow from other banks that have relatively few withdrawals. Hence, the bank is not forced to liquidate the investment in the profitable technology, but it can trade on the interbank market instead. Bhattacharya and Gale show that this market induces a free-rider problem. Each bank manager counts on the others to provide

liquidity and overinvests funds in the profitable technology, hence there is under-investment in liquid assets. As a consequence, the incidence of bank runs is increased and establishing the inter-bank market leads to less efficient outcomes. Bhattacharya and Gale conclude by suggesting that a central bank may have an important role in overcoming this free-rider problem. By monitoring the banks' transactions, it may allow better risk-sharing to be achieved. Again, therefore, a role for government intervention is suggested. However, the availability of a lender of last resort might also lead to perverse incentives for banks to take on risks, hence, to bank insolvency. Of course, Bagehot (1873) was already well aware of this fact. We return to this issue in section VII.

V. COMPETITION AMONG BANKS

Competition for depositors has traditionally been linked to instability problems, runs, excessive risk-taking and inferior economic performance (Friedman and Schwarz, 1963). In many countries, the banking sector has been exempted from anti-cartel laws and sometimes governments have even encouraged collusive agreements among banks in the belief that this would promote stability (Vives, 1991). In this section we address the question of whether competition in banking is viable and whether competition makes the banking sector more fragile or unstable. We will see that the theory of competition in financial intermediation is still at an unsatisfactory stage, and so many interesting questions will remain unanswered. Even though intermediaries are involved in 'double' competition, i.e. they compete simultaneously for deposits and loans, we first consider some aspects of 'single' competition. To start with, we focus exclusively on the competition for deposits, next we move to the competition in the loan market, while at the end of the section we discuss aspects of 'double' competition.

Let us first consider the problem of the consumer who has to choose among competing banks who all offer standard, fixed-interest demand deposit contracts. (We refer to section III for a justification of this restriction.) Assume that there exists a pool of independent, identically distributed risky projects

in which banks can invest the funds that they receive from depositors. (This assumption simplifies the problem by excluding the bank's moral-hazard problem associated with excessive risk taking.) Each project needs a substantial investment to get started. The independence assumption, together with the assumption of lumpy investment, imply that a larger bank is better diversified, hence, there are economies of scale. The question is at which bank the consumers should open deposits if they want to maximize their expected profits and what terms the bank should offer. A consumer's expected profit depends on two key variables: (i) the interest rate, r , offered by the bank and (ii) the probability, p , that the bank will fulfil its obligations, i.e. that it does not go bankrupt. There are two factors that complicate the consumer's problem: (i) *ceteris paribus*, an increase in the interest rate, r , makes it less likely that the bank will be able to fulfil its obligations (an increase in r is accompanied by a decrease in p) and (ii) because of the economies of scale, the probability, p , that the bank is solvent is increasing in the number of depositors the bank attracts and, *ceteris paribus*, this number increases with r . The more depositors a bank attracts, the more projects it can finance, and the safer it becomes. Hence, the larger the bank, the more attractive it becomes for each single depositor. There exist important network externalities (just as, for example, for computers or for telecommunication networks). Consequently, consumers face a coordination problem and multiple equilibria exist, including equilibria in which all consumers go to the bank that offers an interest rate that is strictly lower than that of any other bank. Since each consumer expects, for one reason or another (e.g. because this bank has the highest reputation, or for one of the reasons mentioned in Schelling, 1960) that all other customers go to that bank, he finds it optimal to go to that bank himself and the expectation that all people will take deposits at that bank is self-fulfilling. There is indeterminacy about which banks are active in equilibrium and what interest rates the banks offer.

Matutes and Vives (1992) study a variant of the above model that allows for product differentiation among the banks. Their paper integrates essential elements from the financial intermediation theories discussed in the previous sections into a

Hotelling-type model known from the industrial organization literature. Depositors are assumed to be uniformly located on the line segment $[0,1]$, while at each of the two end-points a bank is located. Each depositor wants to deposit one unit of funds and faces linear transportation costs to reach the bank. Hence, there is horizontal differentiation. Depositors have (homogeneous) prior expectations about the reliability of each bank, *a priori* they believe that bank i will honour its obligations with a probability p_i^e . Hence, the model allows for vertical differentiation as well (we can have $p_i^e \neq p_2^e$). Knowing these prior expectations, banks simultaneously offer deposit rates, while consumers decide thereafter with which bank to deposit. Bank i then invests all its deposits in a risky project with random per-unit return, $R_i(n_i)$, with positive expected value. For simplicity, it is assumed that this expected value is independent of n_i . However, it is assumed that there are economies of diversification: the variance of $R_i(n_i)$ decreases with the market share, n_i , of the bank. As a consequence, a larger bank is more likely to be able to meet its obligations. (Since each bank is subject to an optimal 'incentive contract', à la Diamond (1984), the bank will indeed pay out if it can.) In the basic model, Matutes and Vives assume that a depositor of bank i does not receive any payment if the bank defaults, hence, the expected pay-off of a depositor located at a distance t from bank i is $p_i(n_i)r_i - t$ where $p_i = P(R_i(n_i) \geq r_i)$. The authors focus on rational expectations equilibria, i.e. all agents optimize and the depositors' prior expectations are fulfilled ($p_i(n_i) = p_i^e$). Not surprisingly, the network externality effect ($\partial p_i / \partial n_i > 0$) implies that in general there is a multiplicity of equilibria. A bank which is perceived to be safer can afford to offer lower deposit rates while still acquiring a larger market share, and both these aspects contribute to the bank's safety and to making the beliefs self-fulfilling. (Note that, in general, if there exists a risk-free investment opportunity yielding positive returns, there is always an equilibrium in which no bank is active. Depositors believe that nobody will deposit, resulting in $n_i = 0$ and $p_i(n_i) = 0$, making it indeed optimal not to deposit with the bank, but to invest in the riskless asset instead.)

Matutes and Vives then study the effects of public intervention in this banking market. They first

consider deposit insurance with fair (risk-based) pricing. The deposit insurance fund guarantees the depositor the return r_i also in case of bankruptcy, and it is filled with taxes that are levied on the bank and that depend on the bank's position (r_i, n_i) . In this case, depositors perceive both banks as perfectly safe ($p_i^e = 1$), this eliminates the network externality and, hence, results in a unique equilibrium à la Hotelling. However, such an insurance scheme need not necessarily increase welfare. Consumers may spread themselves between banks, preventing the full economies from diversification from being obtained. Also, deposit insurance may make banks more aggressive, hence, it might lead to higher deposit rates and, as a result, higher probabilities of bankruptcy. Consequently, the model suggests that, in addition, ceilings on deposit rates might be needed to improve welfare.

Let us now turn to the competition for borrowers. To focus exclusively on the asset side of banks' balance sheets, assume that a bank can raise arbitrary amounts of deposits at a given rate. Broecker (1990) studies the consequences of adverse selection, resulting from unobserved borrowers' quality, for the competition among banks. He assumes that before a bank decides on providing a loan, it conducts a credit-worthiness test on the borrower. However, the testing technology is imperfect, so that there is residual uncertainty about the client's riskiness. All banks make use of the same imperfect technology. If there is just one bank, then the pool of applicants to this bank is simply an unbiased sample from the entire population and the risk premium charged will reflect the percentage of bad loans in the overall pool. If there are multiple banks, then borrowers that have been rejected at one bank can apply for loans at other banks. Consequently, each bank perceives that the pool of loan applicants that it gets has on average lower quality than the population as a whole has; after all the bank gets previously untested as well as previously rejected candidates. As a consequence, in order not to fall prey to the winner's curse, the bank should be more conservative and charge a higher risk premium. As the winner's-curse problem is more severe if there are more banks, we see that a larger number of banks may be accompanied by a higher interest rate on loans. In extreme cases it might even happen that the market breaks down as soon as there is more

than one bank, hence, only a monopoly is viable. An alternative way to avoid the adverse-selection problem is to allow the exchange of information about which borrowers they have already rejected. Hence, banks might be allowed to form an 'information-exchanging' cartel. Of course, from a regulatory point of view the question is how can one ensure that banks exchange information without simultaneously colluding on prices.

The question of whether banks will set up a credit bureau to exchange information about the credit-worthiness of their customers is addressed in Pagano and Japelli (1991). In the Pagano and Japelli model, banks are local monopolies in different cities, i.e. there is one bank per city. Consumers want to borrow in order to smooth consumption and it is assumed that the bank in a customer's home town has perfect information about this customer's credit-worthiness. It is also assumed that consumers who migrate from one town to another have to cancel their original loans, apply for new loans in the new town, and cannot take evidence of their credit history with them. Pagano and Japelli assume that only fixed costs are involved in setting up the credit agency. The benefits of the agency increase with the mobility of the population and the volume of the loan market. Hence, the model predicts that at most one agency is set up and that it is, indeed, set up if volume is large or mobility is high. Pagano and Japelli indeed confirm this prediction by international and historical evidence. Note, however, that since it is assumed that each bank is a local monopoly in its own town, the model is not really one of competition. In an extension, Pagano and Japelli allow banks to compete but they retain the assumption that each bank has a natural advantage in its home market. In this case, the exchange of information also has a negative effect. It reduces the informational gap between a bank and its competitors, allowing the other banks to compete more aggressively and, consequently, reduces the bank's monopoly rents. This effect reduces the expected gains from information-sharing and may conceivably deter banks from sharing information altogether. In their analysis, Pagano and Japelli also abstract away the moral-hazard problem that is associated with information revelation: a bank might join the credit bureau only in order to receive information from its competitors,

it might choose to provide no (or worthless) information about its own customers. Such moral hazard might also prevent a sharing agreement from being feasible. (See Jaynes, 1978, for an early analysis of this problem.)

Let us now move on to discuss aspects of 'double' competition, i.e. simultaneous competition for both deposits as well as loan contracts. As early as 1770, James Stewart stressed that 'double' competition is fundamentally different from the traditional 'single' competition. He argued that *a priori* it is unclear which competition, for inputs or for outputs, will be fiercer and that double competition may lead to non-market clearing prices on both the input and output markets. In modern language, Stewart's first observation may be paraphrased as 'the outcome may depend on the rules of the game and *a priori* it is unclear which rules govern the competition'. His second observation is particularly interesting as it suggests that intermediation does not necessarily lead to efficient outcomes. Hence it raises the question as to what extent the analysis from section III has to be modified if strategic behaviour between intermediaries is taken into account. Recall that Diamond's approach was contract theoretic. The essential argument was that only a constellation with a single bank was individual rational and (incentive) efficient. The important implicit assumption in Diamond's analysis is that competition will indeed produce this efficient outcome. Whether this assumption is justified, or whether perhaps we should instead take seriously the cautious words of Stewart, has recently been investigated in Yanelle (1989, 1991) (see also Stahl, 1988). To illustrate that there might be reason for caution, note that it is not implausible that one intermediary tries to corner the input market by offering depositors an excellent deal, a position which would then allow him to make monopoly profits in the loan market. Or perhaps the story is with the roles of the input and output markets reversed, but also in this case, the intermediary might acquire a monopoly position.

To analyse these issues, Yanelle (1991) studies two different models. In both of them, there are borrowers, lenders, and prospective banks. The total amount lenders are willing to lend exceeds the amount required by borrowers. As in Diamond (1984), each

borrower has a project of which the outcome is privately observed. In the first model, the timing is as follows. First, banks simultaneously make loan offers to borrowers. Second, borrowers decide which offer, if any, to accept. In the third stage, those agents who need money for investment (i.e. the banks who have contracts with borrowers and the entrepreneurs who do not have contracts with banks) make offers to lenders. Each of these offers specifies the maximum amount the agent is willing to borrow plus the return per unit of deposit. In the final stage lenders decide whether and where to deposit their money. The second game that is analysed reverses the order of the moves: first there is competition for deposits, next there is competition to grant loans. Yanelle shows that the first game admits a continuum of equilibria. There exist (efficient) equilibria in which only one bank is active, but there are also equilibria with multiple active banks. The reason for this multiplicity is the same as the one encountered when discussing the Matutes and Vives model. Non-convexities in the technology give rise to a coordination problem and this can be resolved in many ways. Hence, the outcome suggested by Diamond is only one of the many equilibria of game 1 and, as Yanelle argues, this equilibrium outcome is very risky for all agents in the game. Considerations of risk-dominance (Harsanyi and Selten, 1988) would favour the selection of a different equilibrium. To justify the selection of the Diamond outcome, criteria of collective rationality have to be invoked. The second game also has a continuum of equilibria, but there is only one equilibrium in which there are no coordination failures and Yanelle focuses attention on this 'coalition-proof' equilibrium. She shows that it leads to quite different outcomes than the one considered by Diamond. First, for some parameter values, intermediation will not occur in the equilibrium. Second, if intermediation occurs, it does not result in the competitive outcome, since both the buying and selling prices will exceed the competitive level. Third, preventing banks from operating may actually result in a Pareto superior outcome, hence, intermediation need not minimize transaction costs and it may actually be socially bad. Consequently, it seems wise to heed Stewart's

warnings, and not to take for granted that competition in intermediation automatically leads to efficient outcomes. Of course, it might be that yet another specification of the competition between intermediaries yields conclusions that are in full agreement with those of Diamond. Perhaps the negative conclusions reached by Yanelle are caused by the static nature of the game and the (coordination) problems vanish in (more realistic?) dynamic models in which individuals have the time to search until a stable outcome is reached. Further development of dynamic models is, therefore, called for. Recently, some work in this direction has begun. We look into some of these aspects in the next section.

VI. LONG-TERM OR ARM'S-LENGTH RELATIONS?

In a thought-provoking paper that addresses the question of how well different financial systems have performed in funding physical investment, Colin Mayer uncovers the paradox that the most advanced and most competitive financial markets in the world, those of the US and the UK, appear to be most deficient at funding their industries. In the UK and the US, net capital investment is entirely self-financed by industry, while, for example, in Japan, bank loans have contributed approximately 40 per cent of corporate investment. Despite being the most highly developed stock markets in the world, the UK and US markets have made the lowest net funding contribution, in fact they have made negative contributions to the financing of capital formation (Mayer, 1988). Mayer's interpretation of the data is that intensified competition can be seriously detrimental over a longer horizon since it may discourage incentives for capital formation and risk taking. In this section we address the question of whether a system of 'house banks' (such as in Germany and Japan) that is based on long-term relationships between banks and firms necessarily performs better than a system that is based on intensive competition and arm's-length relationships, such as those in the US and in the UK.¹⁴

¹⁴ Corbett and Jenkinson (1994) challenge the conventional wisdom that the UK and US are 'market-based' and that Germany and Japan are 'bank-based'. They use national income accounts data to argue that only 10 per cent of finance was externally raised in Germany, the UK, and the US and that the proportions of bank finance were roughly equal in these countries. This proportion, however, is higher in Japan.

Mayer's central claim is that competition in financial markets undermines the ability of a firm and investors to commit to each other and thereby may discourage incentives for risk taking and risk sharing. Long-term, bilateral, exclusive bank/firm relationships allow such commitments and this may be advantageous, as is illustrated by corporate rescues. As Hoshi *et al.* (1989) document, if a Japanese firm is in trouble, bank managers step in and help with managerial restructuring, rescheduling loans and, more generally, acting as risk-sharers by being willing to sustain current losses. In contrast, if the firm that threatens to get into trouble has multiple creditors, then all of them will rush in in order to make the most of a hopeless situation. Since the first-come first-served element is important, there might exist runs and panics even in cases where the firm is basically healthy, for the same reasons as such runs exist in the model discussed in section IV (see also Weskamp, 1989). The reason that the Japanese bank is willing to help in the rescue operation is that it expects that its current losses will be compensated in the future. After all, the firm is committed to its house bank. In a competitive situation, no creditor expects to reap positive profits from a rescue operation since the firm cannot commit itself. Promises that the firm will be loyal to the rescuing bank and that it will give this bank a favourable treatment in the future are not credible, since, *ex post*, when the rescue operation has been successful, the firm will want to renege on its promise and will do business with the bank that offers the best deal. Foreseeing this, no bank will be willing to rescue the firm in the first place. In short, there is a time-inconsistency problem.

Mayer argues that long-term relations are not only beneficial in times of trouble but that they also give managers the appropriate incentives for risk-taking behaviour. For a manager to be induced to undertake risky investments, there must be the expectation of positive rewards in case the investments are successful. To induce the investments, investors would like to commit themselves to rewarding the manager in case of success but, *ex post*, there will be the incentive to renege, for example by accepting an attractive takeover offer. Hence, takeovers might discourage efficient investment (see also Stein, 1988). Since having a main bank can dis-

courage the takeover, a firm that is tied to such a bank might undertake more efficient investment.

An interesting example where intensive firm/bank relations indeed induce firms to undertake more efficient investment is described in Von Thadden (1990). He considers a situation in which there are two types of firms ('good' and 'bad') who can choose between two types of projects ('short-term' and 'long-term') that require identical amounts of funds. All projects of bad firms are unprofitable, projects of good firms are profitable, and the long-term project is ultimately more profitable but is associated with short-term losses. Investors cannot observe the type of the firm (i.e. there is adverse selection) and in the absence of monitoring they will interpret low first-period earnings as a signal that the firm is bad and discontinue funding. As a consequence, without monitoring, good firms will invest in short-term, inefficient projects. A bank that monitors the firm, however, is able to determine whether low first-period earnings are due to bad luck or bad management. In this case a good quality firm is assured that additional funds will be provided for the long-term project and the efficient investment will be realized. The crucial element in this story is that the monitoring bank acquires an informational advantage over the other investor. This advantage yields the bank an information rent and, hence, gives the bank an incentive to spend money on monitoring activities in the first place. Exactly the same forces are at work in Fischer's (1990) formalization of Mayer's thesis that Japanese banks are more willing to engage in rescue operations. The house bank obtains an informational advantage over its competitors that reduces the strength of competition from outside financiers; this allows the house bank to earn monopoly rents and provides incentives for the bank to support the rescue operation. Sharpe (1990) discusses a similar model. He argues that long-term relationships arise endogenously since a bank that actually lends to a firm learns more about this firm than any other bank, and, hence, gains a competitive advantage. Sharpe also points out that since banks make positive profits on high quality firms that are 'informationally captured', there will be excessive competition for firms that are not yet captured. As a result, too much capital may be allocated to new,

low quality firms. From a welfare point of view, long-term relationships need not be optimal.

The fact that firms typically maintain relations with multiple banks suggests that a long-term bank/firm relationship must involve costs as well as benefits. Indeed such costs might arise from a house bank not being able to restrain its monopoly power: it may use its informational edge to squeeze the firm's profits. Rajan (1992) shows that such monopoly power may induce under-investment on the part of the firm. If a successful firm has a weak bargaining position *vis-à-vis* its bank, it will have to give a large share of the surplus to the bank, hence, it may under-invest. By keeping a second bank informed, the firm's *ex-post* bargaining power is increased, and it will get a large share of the surplus. Since such a shift in bargaining power might induce more efficient investment on the part of the firm, such 'double sourcing' might actually also be beneficial for the house bank itself. Another beneficial aspect of such multiple bilateral relationships is that they may make renegotiations more difficult in times of trouble, which may be desirable from an *ex-ante* point of view. That is to say, under incomplete information, *ex-ante* efficient contracts usually involve some kind of *ex-post* inefficiency. If renegotiation cannot be prevented, then such contracts are not credible. Parties know that inefficiencies will be renegotiated away and, as a consequence, the *ex-ante* efficient outcome is not attainable. If, however, multiple parties are involved at the renegotiation stage, the transaction costs increase and it is not so clear that the *ex-post* inefficiency will be negotiated away. Hence, in this case, the *ex-ante* optimal outcome might be obtained.

In a fairly general model, Dewatripont and Maskin (1990) have also investigated whether or not bank finance leads to better allocation of credit than market finance. The authors note that a decentralized market-based system is very successful at deterring poor projects (as the managers of 'bad' firms know that their projects will be quickly terminated anyway—the market will not continue funding after low earnings have been reported, they do not even start requesting the funds). The flipside of this coin is that, as a consequence, the system may also fail to identify and implement the really efficient projects. On the other hand, the German/

Japanese house-bank system is more likely to implement the really good projects, but at the same time it is liable to the 'soft budget constraint' problem of persistence with unprofitable projects. For example, returning to Von Thadden's model, if the monitoring bank cannot perfectly identify the type of the firm after low earnings have been reported in period 1, it might continue funding the project in the hope that eventually the initial losses will be recouped, even though the firm will never yield positive profits. Hence, both systems have their advantages with the market system leading to a portfolio of investment projects with lower variance than the house-bank system, and the latter probably leading to a somewhat higher average return per project.

VII. CONCLUSION

What are the policy implications that follow from the theory as discussed in this paper? How should the banking industry, and the financial sector in general, be regulated in order to guarantee that it functions efficiently and contributes optimally to the growth of the real economy? In this concluding section we briefly review some of the main aspects from this survey with these questions in mind. We can draw one conclusion straight away: there are no easy answers. As the paper has amply shown, we are still far from having a satisfactory theoretical model of financial intermediation. We have a large variety of simple partial models and a lot of work remains to be done to see how robust the conclusions of these models are with respect to several of the crucial assumptions involved. Care has to be exercised in drawing strong policy conclusions on the basis of these simple models, in particular since policy that may be defended on the basis of one model may result in disastrous outcomes in sensible elaborations of the same model. Consequently, it is not (yet) possible to give detailed advice about how the financial sector should be regulated. All that can be said at this general level is that, since the banking sector is imperfectly competitive and since regulation has to be based on imperfect information, the principle of second-best theory applies. One cannot be sure that regulation is not free from side-effects; it is difficult to guarantee that regulation of one type of inefficiency does not introduce

another inefficiency somewhere else in the system.

The two main goals that regulation of the banking sector aims to achieve are (i) to protect individual investors and (ii) to ensure the stability of the financial system. Regulation is needed because of possible market failures owing to information problems, externalities, or problems of market power. The primary reason why the individual investor has to be protected is that the financial sector is based on trust—the system only works because enough people have sufficient confidence that it works—and that fraud is always a possibility. Regulation might also protect the depositor from abuse of market power by banks, and the deregulation and liberalization of financial markets may make this issue more important. Banks are holders of information and some of this information might be of vital importance to other market participants. Consequently, we may expect liberalization and deregulation to lead banks to merge with companies that can profit from such information. For example, since banks provide payment services, they have knowledge about the consumer's expenses on health care which might be of relevance to insurance companies. Regulation ('Chinese Walls') might be needed to protect the consumer.

Because of the network externalities involved in banking (a larger bank being better able to diversify risk and, hence, being able to offer a larger return per unit of deposit), the answer to the question as to what extent competition in banking is viable and/or desirable is not straightforward. Certainly, more competition (for example, as a result of (European) financial integration) is not always beneficial. More competition may aggravate the winner's-curse problem. It may make each bank more prudent, while at the same time it may result in more bad quality projects being financed. Competition may also erode the benefits that can be obtained from long-term relationships, and the monopoly rents associated with the latter might be necessary to induce banks to make investments that are socially profitable. Efficiency might be enhanced if banks are allowed to exchange information about their clientele. However, allowing such exchange of informa-

tion might make it more difficult to prevent banks from colluding on prices. Furthermore, Hellwig (1991) makes the important observation that since a bank holds information about multiple firms, the bank might provide these firms with information about each other. Hence, banks might weaken the moral-hazard problems associated with cartel behaviour by making secret deviations from the cartel agreement easier to detect. Cartelization might, therefore, be easier where banks are present.

The fundamental liquidity transformation function that banks perform makes them liable to runs. Since such panics might be contagious and do not discriminate between good and bad, they may result in viable information being destroyed. This externality provides the second main motivation for the regulation of the banking sector. We showed that intervention, by providing deposit insurance or lender-of-last-resort facilities, reduces the incidence of bank runs and panics. At the same time we argued that such insurance possibilities induce a moral-hazard problem: banks may have an incentive to assume too much risk or to offer too high deposit rates. It is now universally agreed that these problems contributed significantly to the deepening of the S&L crises in the US (see Jaffee, 1989). Hence, more involvement by regulators (central banks) may give more scope for moral hazard and may make matters worse. By providing liquidity at times of distress, the central bank may induce more serious insolvency problems. Of course, Bagehot (1873) was aware of this fundamental dilemma. He insisted that the lender of last resort should help only solvent banks and let the others fail. This, of course, raises the questions of whether illiquidity problems can be distinguished from insolvency problems and whether a regulator will be able to commit to such a policy.

Capital adequacy requirements might mitigate the incentives to take on excessive risk¹⁵ and following the suggestions of the 1988 G10 meeting in Basle, it is now (as of 1 January 1993) required in the EC that a bank's Cooke ratio, i.e. the ratio of capital to risk-weighted assets (both on- and off-balance), be at least 8 per cent. As soon as a bank fails to meet this requirement, it should recapitalize or else it

¹⁵ Rochet (1992) argues that capital adequacy requirements may distort a bank's investment behaviour in such a way as actually to make insolvency more likely!

should be closed. Dewatripont and Tirole (1993) (see also Tirole, 1994) rationalize such a procedure by building on the literature on the optimal governance structure of firms. The starting point of this literature is that external pressure or interference is a main component of managerial incentive schemes. Since managers dislike interference, the best incentives are provided by interfering after, and only after, bad performance. In the case of a bank, poor performance corresponds with a low Cooke ratio, hence, the control model requires intervention in case this ratio drops below a certain threshold. Since equity holders have no incentive to intervene in case of impending insolvency—they want to gamble for resurrection—only the threat of transferring control to the debtholders can induce bank managers to limit risk. However, since depositors are small and numerous, they generally will lack the information and power to intervene on an individual basis. Consequently, Dewatripont and

Tirole argue that a regulator is needed to perform their task. Hence, a regulator should not only be involved in crisis management (i.e. when banks face liquidity problems) but also in the active monitoring of banks *ex ante*. Such active monitoring will presumably also allow the regulator to better distinguish illiquidity problems from insolvency problems.

Of course, the important question remains as to how to ensure that the regulator has the proper incentives. If regulatory action is (or might be) interpreted as a signal that the regulatory agency did not perform properly in the past, the agency has an incentive to 'cover up' and to be inactive when action is called for (Boot and Thakor, 1993). Kane (1989) argues that such forbearance—allowing insolvent institutions to continue to operate in the hope that they will recover, is exactly what happened in the US S&L crisis.

REFERENCES

- Akerlof, G. (1970), 'The Market for Lemons: Quality Uncertainty and the Market Mechanism', *Quarterly Journal of Economics*, **84**, 488–500.
- Aumann, R. J. (1964), 'Markets with a Continuum of Traders', *Econometrica*, **32**, 39–50.
- Bagehot, W. (1873), 'Lombard Street', republished in *Collected Works*, vol. 9, London, *The Economist*, 1986.
- Bernanke, B. S. (1983), 'Non-Monetary Effects of the Financial Crisis in the Propagation of the Great Depression', *American Economic Review*, **73**, 257–76.
- Bester, H. (1985), 'Screening vs Rationing in Credit Markets with Imperfect Information', *American Economic Review*, **75**, 850–5.
- Bhattacharya, S., and Gale, D. (1987), 'Preference Shocks, Liquidity, and Central Bank Policy', in W. A. Barnett and K. J. Singleton (eds), *New Approaches to Monetary Economics*, Cambridge, Cambridge University Press.
- Thakor, A. (1993), 'Contemporary Banking Theory', *Journal of Financial Intermediation*.
- Blanchard, O., and Fischer, S. (1989), *Lectures on Macroeconomics*, Cambridge, MA, MIT Press.
- Boot, A., and Thakor, A. (1993), 'Self-Interested Bank Regulation', *American Economic Review*, **83**, 206–12.
- Broecker, T. (1990), 'Credit Worthiness Tests and Interbank Competition', *Econometrica*, **58**, 429–52.
- Bryant, J. (1980), 'A Model of Reserves, Bank Runs, and Deposit Insurance', *Journal of Banking and Finance*, **4**, 335–44.
- Chari, V. V., and Jagannathan, R. (1988), 'Banking Panics, Information, and Rational Expectations Equilibrium', *Journal of Finance*, **43**, 749–61.
- Corbett, J., and Jenkinson, T. (1994), 'The Financing of Industry, 1970–89: An International Comparison', CEPR Discussion Paper 948.
- Davis, P., and Mayer, C. (1991), 'Corporate Finance in the Euromarkets and the Economics of Intermediation', CEPR Discussion Paper 570.
- Dewatripont, M., and Maskin, E. (1990), 'Credit and Efficiency in Centralized and Decentralized Economies', Harvard Institute of Economic Research Discussion Paper 1512.
- Tirole, J. (1993), 'Banking: Private Governance and Regulation', in C. P. Mayer and X. Vives (eds), *Financial Intermediation in the Construction of Europe*, Cambridge, Cambridge University Press, 12–35.
- Diamond, D. W. (1984), 'Financial Intermediation and Delegated Monitoring', *Review of Economic Studies*, **51**, 393–414.

- Diamond, D. W. (1991), 'Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt', *Journal of Political Economy*, **99**, 689–721.
- Dybvig, P. (1983), 'Bank Runs, Deposit Insurance, and Liquidity', *Journal of Political Economy*, **91**, 401–19.
- Fama, E. F. (1985), 'What's Different About Banks?', *Journal of Monetary Economics*, **15**, 29–39.
- Fischer, K. (1990), 'Hausbankbeziehungen als Instrument der Bindung zwischen Banken und Unternehmen: Eine theoretische und empirische Analyse', University of Bonn Doctoral Dissertation.
- Fisher, I. (1933), 'The Debt Deflation Theory of Great Depressions', *Econometrica*, **1**, 337–57.
- Friedman, M., and Schwarz, A. J. (1963), *A Monetary History of the United States, 1867–1960*, National Bureau of Economic Research, Princeton, Princeton University Press.
- Gale, D., and Hellwig, M. (1985), 'Incentive-Compatible Debt Contracts: The One-Period Problem', *Review of Economic Studies*, **52**, 647–63.
- Gertler, M. (1988), 'Financial Structure and Aggregate Economic Activity: An Overview', *Journal of Money, Credit and Banking*, **20**, 559–88.
- Goodhart, C. A. E. (1989), *Money, Information and Uncertainty* (2nd edn), London, Macmillan.
- Grossman, S., and Stiglitz, J. (1980), 'On the Impossibility of Informationally Efficient Markets', *American Economic Review*, **70**.
- Gurley, J., and Shaw, E. (1955), 'Financial Aspects of Economic Development', *American Economic Review*, **45**, 515–38.
- Harsanyi, J. C., and Selten, E. (1988), *A General Theory of Equilibrium Selection in Games*, Cambridge, MA, MIT Press.
- Hellwig, M. (1991), 'Banking, Financial Intermediation and Corporate Finance', in A. Giovanni and C. Mayer (eds), *European Financial Integration*, 35–72, Cambridge, Cambridge University Press.
- Hoshi, T. A., Kashyap, A., and Scharfstein, D. (1989), 'The Role of Banks in Reducing the Costs of Financial Distress in Japan', Massachusetts Institute of Technology.
- Jacklin, C. (1987), 'Demand Deposits, Trading Restrictions and Risk Sharing', in E. Prescott and N. Wallace (eds), *Financial Intermediation and Intertemporal Trade*, Minneapolis, MN, University of Minnesota Press.
- Bhattacharya, S. (1988), 'Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications', *Journal of Political Economy*, **96**, 658–92.
- Jaffee, D. M. (1989), 'Symposium on Federal Deposit Insurance for S&L Institutions', *Journal of Economic Perspectives*, **3**, 3–6.
- Jaynes, G. D. (1978), 'Equilibria in Monopolistically Competitive Insurance Markets', *Journal of Economic Theory*, **19**, 394–422.
- Jensen, M., and Meckling, W. (1976), 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure', *Journal of Financial Economics*, **3**, 305–60.
- Kane, E. (1989), 'The High Cost of Incompletely Funding the FSLIC Shortage of Explicit Capital', *Journal of Economic Perspectives*, **3**, 31–47.
- Kindleberger, C. (1978), *Manias, Panics and Crashes: A History of Financial Crises*, New York, Basic Books.
- Kreps, D. M., and Wilson, R. (1982), 'Reputation and Imperfection Information', *Journal of Economic Theory*, **27**, 253–79.
- Leland, H., and Pyle, D. (1977), 'Informational Asymmetries, Financial Structure, and Financial Intermediation', *Journal of Finance*, **32**, 371–87.
- Matutes, C., and Vives, X. (1992), 'Competition for Deposits and Risk of Failure in Banking', University of Barcelona Discussion Paper.
- Mayer, C. (1988), 'New Issues in Corporate Finance', *European Economic Review*, **32**, 1167–89.
- Merton, R. (1977), 'An Analytical Derivation of the Cost of Deposit Insurance Loans Guarantees', *Journal of Banking and Finance*, **1**, 242–56.
- Mishkin, F. S. (1990), 'Asymmetric Information and Financial Crises: A Historical Perspective', NBER Working Paper 3400, Cambridge, MA.
- Modigliani, F., and Miller, M. (1958), 'The Cost of Capital, Corporation Finance and the Theory of Investment', *American Economic Review*, **48**, 261–97.
- Myers, S., and Majluf, N. (1984), 'Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have', *Journal of Financial Economics*, **13**, 187–221.
- Pagano, M., and Japelli, T. (1991), 'Information Sharing in the Consumer Credit Market', University of Bocconi Discussion Paper.

- Postlewaite, A., and Vives, X. (1987), 'Bank Runs as an Equilibrium Phenomenon', *Journal of Political Economy*, **95**, 485–91.
- Rajan, R. (1992), 'Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt', *Journal of Finance*, **47**, 1367–400.
- Rochet, J.-C. (1992), 'Capital Requirements and the Behaviour of Commercial Banks', *European Economic Review*, **36**, 1137–70.
- Schelling, T. (1960), *The Strategy of Conflict*, Cambridge, Cambridge University Press.
- Schumpeter, J. (1934), *The Theory of Economic Development*, Cambridge, MA, Harvard University Press (first published in German in 1911).
- Sharpe, S. (1990), 'Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships', *Journal of Finance*, **45**, 1069–87.
- Stahl, D. O. (1988), 'Bertrand Competition for Inputs and Walrasian Outcomes', *American Economic Review*, **78**, 189–201.
- Stein, J. (1988), 'Takeover Threats and Managerial Myopia', *Journal of Political Economy*, **96**, 61–80.
- Stewart, J. (1770), *An Inquiry into the Principles of Political Economy: Being an Essay on the Science of Domestic Policy in Free Nations*, Basel, J. J. Tourneisen.
- Stiglitz, J. E., and Weiss, A. (1981), 'Credit Rationing in Markets with Imperfect Information', *American Economic Review*, **71**, 393–411.
- Tirole, J. (1994), 'On Banking and Intermediation', *European Economic Review*, **38**, 469–88.
- Townsend, R. M. (1979), 'Optimal Contracts and Competitive Markets with Costly State Verification', *Journal of Economic Theory*, **21**, 1–29.
- Vives, X. (1991), 'Regulatory Reform in European Banking', *European Economic Review*, **35**, 505–15.
- von Thadden, E. L. (1990), 'Bank Finance and Long-Term Investment', WWZ Discussion Paper No. 9010, University of Basel.
- Weskamp, A. (1989), 'Die Auswirkungen der Sicherungsrechte auf die effiziente Ausgestaltung von Kreditbeziehungen', University of Bonn Doctoral Dissertation.
- Yanelle, M.-O. (1989), 'The Strategic Analysis of Intermediation', *European Economic Review*, **33**, 294–301.
- (1991), 'The Strategic Analysis of Intermediation', Delta (Paris) Discussion Paper.