ELSEVIER

Journal of Banking & Finance 20 (1996) 917–948

Journal of BANKING & FINANCE

The design of financial systems: An overview

Anjan V. Thakor *

Finance School of Business, Department of Finance, Indiana University, Bloomington, IN 47405, USA

Received 15 February 1995; accepted 15 May 1995

Abstract

I examine the multifaceted aspects of financial system design, focusing on the real effects of this design. My exploration of the key issues and my review of the related literature pertain to three dimensions of financial system design: (i) the permissible scope of activities for banks and other depository financial intermediaries, (ii) the regulations dictating the structure of the banking industry, and (iii) information disclosure requirements in the financial market. I address a diverse set of issues such as borrowers' choices of financing source and how these are affected by financial system design, the impact of financial system design on the capital structure and corporate control decisions of nonfinancial firms, the relationship between financial system architecture and the liability claims of banks, the issues surrounding the desired permissible scope of banking and bank industry structure, and the overall design of a financial system.

JEL classification: G1; G2; G32; G34

Keywords: Financial systems ; Markets and institutions

1. Introduction

There has recently been a surge of academic interest in the design of financial systems. In this paper, I offer some thoughts on the multifaceted aspects of

* Tel.: 812-855-8568; fax: 812-855-8679.

SSDI 0378-4266(95)00033-X

financial system design, examining along the way the burgeoning literature on this subject. My discussion focuses on:

- the key policy-related questions in financial system design;
- the analytical issues involved in theoretical explorations of these questions;
- the manner in which research on this issue is likely to draw upon our knowledge in a variety of seemingly disparate subfields, not only potentially generating unifying themes but also pushing the frontiers of our comprehension of contracts, institutions and markets;
- the potential of future research to illuminate policy initiatives concerning the regulation of institutions and markets;
- the implications of financial system design for the real and financial decisions of firms; and
- possible paths along which financial systems may evolve in the future.

In principle, financial system design encompasses a myriad of institutional and market details, regulations, and disclosure requirements. For simplicity, I shall focus on three of the many aspects of financial system design: (i) the permissible scope of activities for banks and other depository financial intermediaries, (ii) regulations dictating the structure of the banking industry, and (iii) information disclosure requirements in the financial market. These aspects clarify the exogenous instruments that can be used to influence financial system design. Different financial system designs will manifest themselves in different divisions of activities between financial institutions and markets.

My focus on these three aspects of financial system design enables me to restrict the scope of this paper to the seven questions listed below. Prominent by their absence are considerations related to the design of securities exchanges and related market microstructure issues, details of the bankruptcy code, and bank regulation except that concerned with industry structure and banking scope.

Question 1: Why do we care about financial system design? This is perhaps the most obvious policy-related question on this subject. There are many reasons why a systematic examination of financial system design is important. First, as reported by King and Levine (1992), the size of the financial system is strongly correlated with the level of economic development; see Fig. 1. Their paper finds that the citizens of the richest countries hold more of their annual incomes in liquid assets beyond their monetary liabilities than their counterparts in poorer countries. The 'traditional' interpretation of this data would be that rich countries have larger financial systems because these countries are further along on the economic development curve, i.e., economic performance drives financial scope. Such an interpretation relies on the assertion that financial systems are outcomes of the real requirements of an economy, not the drivers of its performance. The details of financial system design – particularly the division of financial activity between intermediated and nonintermediated sources – is of little relevance.

The strength of the *recent* research on this issue, however, is in the observation that the causality may often be reversed, and the design of the financial system

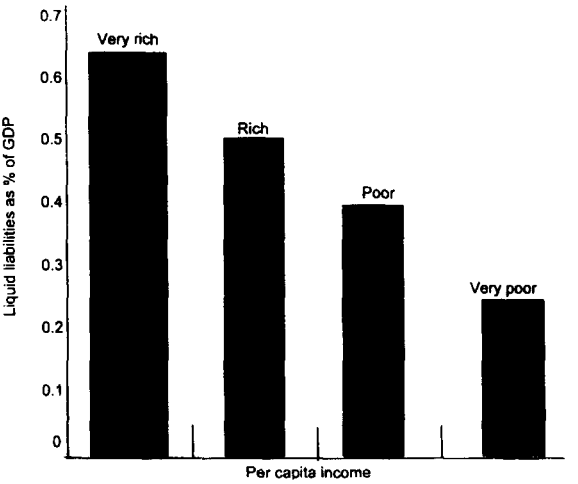

Fig. 1. Financial size and real per capita income, 1985. Source: King and Levine (1992).

could impinge on real activity and economic development. There is a variety of ways in which this could happen; these mechanisms are explored in King and Levine (1992), Bernanke (1988), Gale (1992), and Boot and Thakor (1996), among others. Pagano (1993) summarizes some of these mechanisms. In particular, he points out that financial development can raise the proportion of savings allocated to investment, increase the social marginal productivity of capital, and influence the private savings rate. This viewpoint, therefore, sheds new light on the implications of the growing importance of financial services in developed economies, as depicted in Fig. 2.

Question 2: Why have financial systems in different countries historically been so diverse in design? Economists have been groping for a satisfactory answer to this question for some time. There are many dimensions along which financial systems differ. For example, Mayer (1988) points out that in France, Germany, Japan and the U.K., stock markets were relatively unimportant as a source of funds for corporations during 1970–85. By contrast, U.S. firms relied relatively heavily on bond financing in the capital market during this period; see Fig. 3. This evidence is merely part of a larger body of evidence suggesting that the intermediated/institutional segment of the financial system is more important in Europe and Japan than in the U.S., and financial markets are more important in the U.S. than in Europe and Japan. An example of such evidence appears in Frankel and Montgomery (1991), and is shown in Fig. 4. Apart from 1985–89, when companies were borrowing to repurchase shares, securities have been much more important in the U.S. than elsewhere. We would like to know why.

Question 3: How do firms choose their source of (external) financing and how might this choice be affected by information disclosure requirements in financial

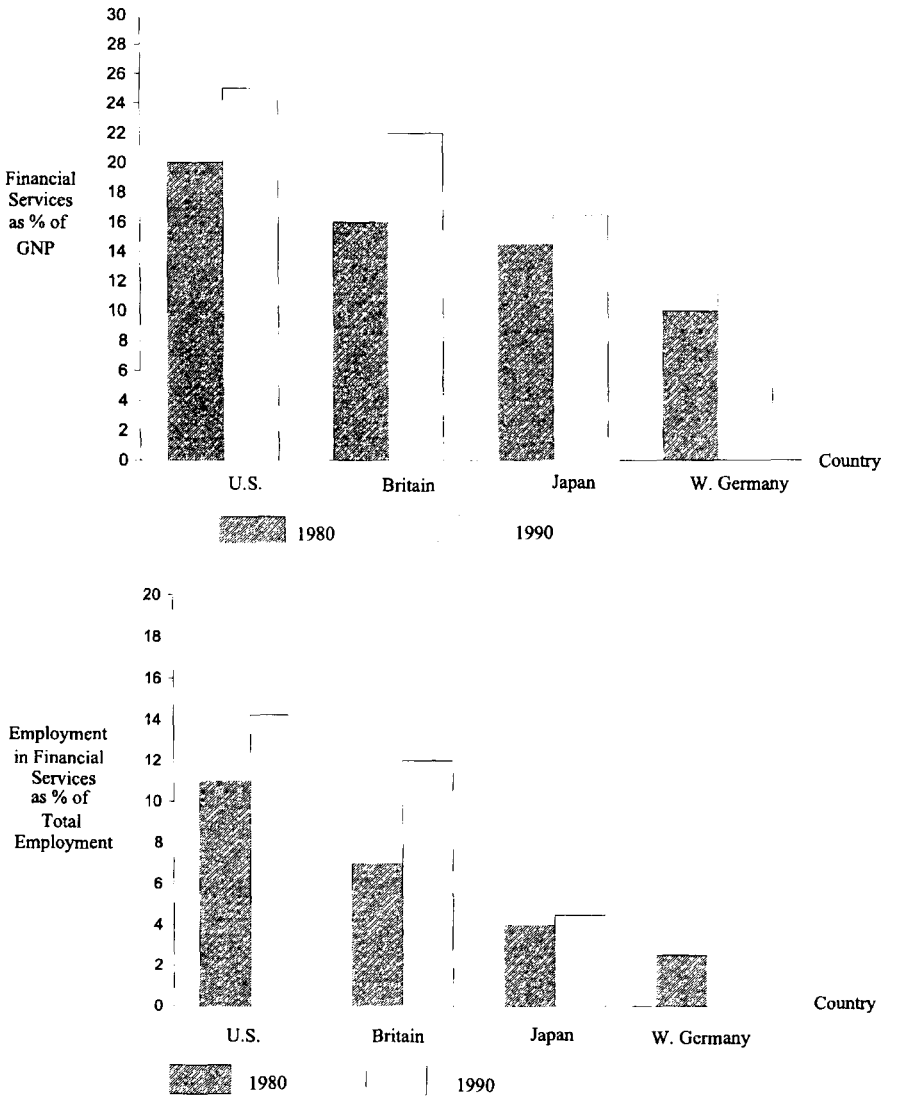

Fig. 2. The growth of financial services in selected countries. Source: Greenbaum and Thakor (1995).

markets and by the development of the financial system? It is now widely believed that firms not only care about their capital structures, but also about *where* they borrow from, conditional on the decision to acquire debt. How is this choice affected by the information disclosure requirements that exchange-listed securities must abide by? There has recently been considerable research that has sharpened our understanding of this aspect of corporate decisionmaking (see, e.g., Diamond,

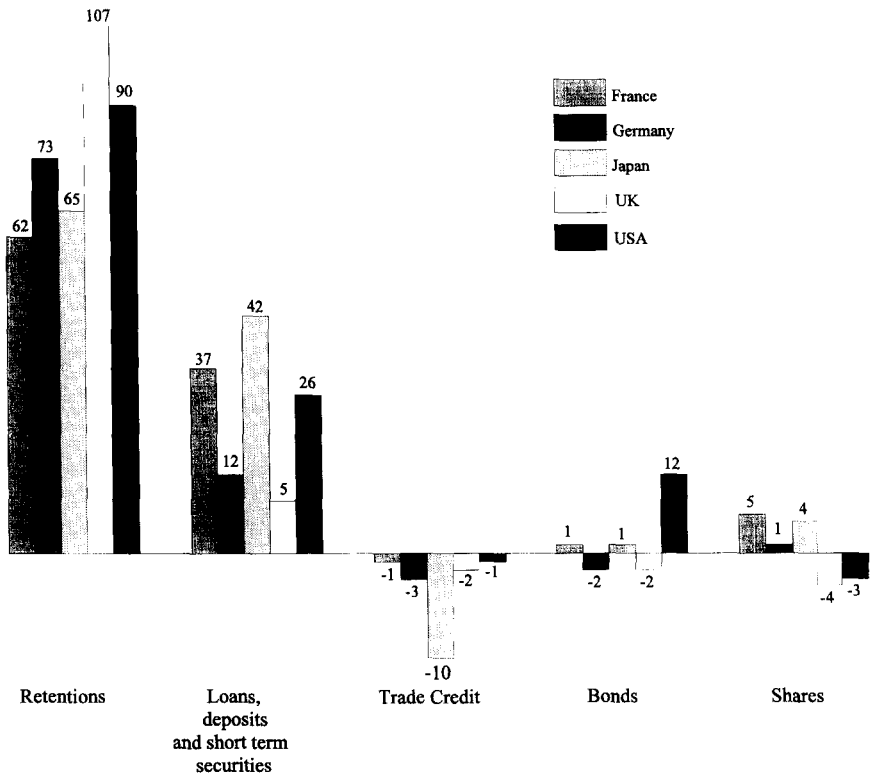

Fig. 3. Net financing of private physical investment by companies in France, Germany, Japan, the U.K. and the U.S. for the period 1970–85. Source: Mayer (1988).

1991a). Yet much remains to be done on the question of how financial system design affects the borrower's choice.

Question 4: Are capital structure and corporate control decisions of nonfinancial firms potentially influenced by financial system design? It has recently come to be understood that a firm's choice of capital structure can influence future contests for control of the firm. Moreover, the role of the capital market is likely to be different from that of banks in determining the outcomes of corporate control contests. Thus, there is reason to suspect that financial system design can impact capital structure and merger/takeover decisions.

Question 5: How are the liability claims of banks likely to be affected by the architecture of the financial system? We would like to know if the variety, nature and volume of deposits used by banks are impacted by how the financial system is configured. The issue here is one of liquidity creation (Diamond and Dybvig, 1983).

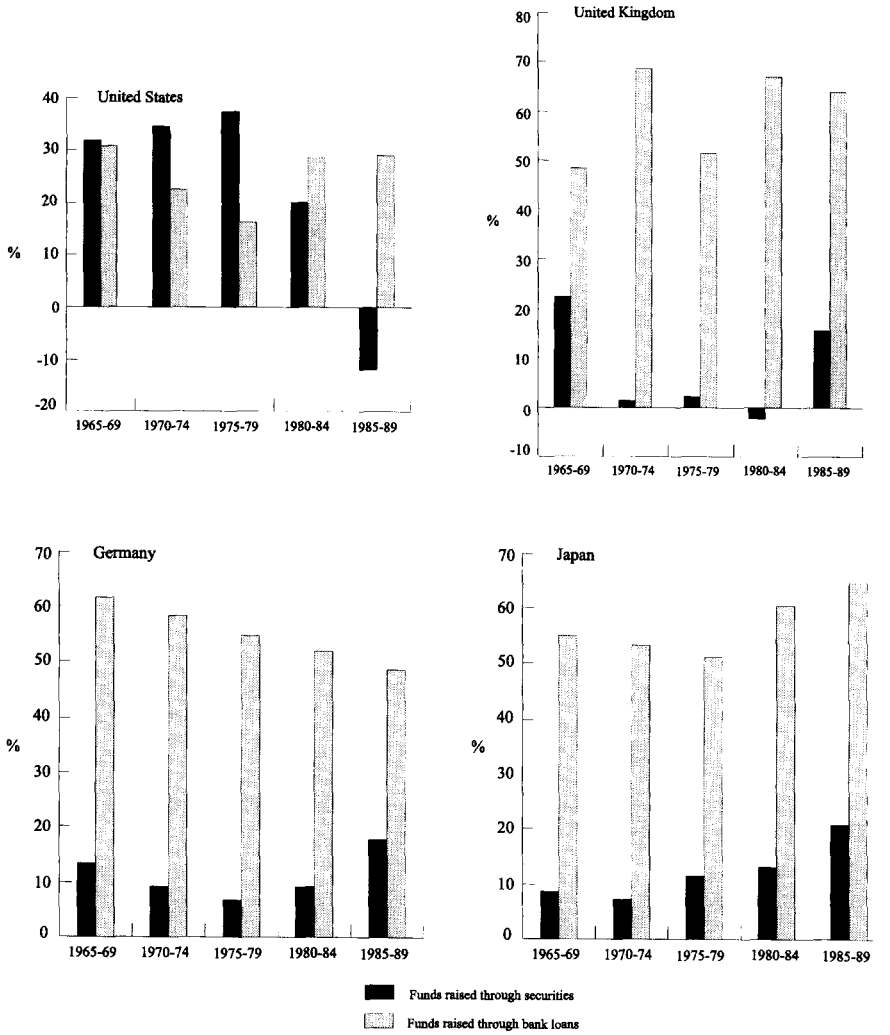

Fig. 4. Percentage of total business funds raised through securities and bank loans, 1965–89. Source: Frankel and Montgomery (1991).

Question 6: What should be the permissible scope of banking? Since the permissible scope of banking and the regulation of banks and financial markets can potentially influence the evolution of a financial system, learning about financial system design may shed new light on optimal regulatory policies. For example, much of the contemporary literature on universal versus restricted commercial banking deals with a particular slice of the set of issues pertinent to

financial system design, such as potential conflicts of interest in universal banking, (see, e.g., Rajan, 1991; Kroszner and Rajan, 1994a; Berlin et al., 1994; Kanatas and Qi, 1993; Puri, 1994). Thus, financial system design is an important issue even in highly developed economies such as the U.S.

Question 7: What is the significance of banking industry structure and how should a new financial system be designed? There is great diversity in banking industry structures in different countries. We would like to understand the significance of this diversity. Moreover, the design of new financial systems is a key policy issue in the emerging market-based economies in Eastern Europe. As Boot and Thakor (1996) point out, the financial systems currently in place in these countries are best viewed as *interim* arrangements designed to facilitate transition to systems that are less dependent on centrally planned capital allocation (see also Catte and Mastropasqua, 1993; Checchi, 1993). There are many who would like to know how these nascent financial systems should be designed and what the stability implications of different designs are likely to be.

The rest of the paper is organized as follows. Section 2 examines the interaction between real and financial decisions and the different mechanisms by which the financial system could drive economic performance (Question 1). Section 3 addresses the issue of the observed diversity of financial systems (Question 2). Section 4 takes up the question of how firms choose their source of financing, particularly in light of the interactions highlighted in Section 2 (Question 3). Section 5 focuses on issues of capital structure and corporate control (Question 4). Section 6 considers the manner in which bank liabilities are affected by financial system design (Question 5). In Section 7, I turn to the role of the permissible scope of banking in financial system design and its potential impact on financial innovation incentives (Question 6). In Section 8, I attempt to synthesize the insights of previous sections to draw conclusions about financial system design (Question 7). Section 9 concludes with thoughts about the future evolution of financial systems. I also discuss at this stage the importance of the political environment and the political economy of financial system design. Wherever appropriate, I note the major unresolved issues in the existing literature.

2. Real and financial decisions

Since an important reason for the interest in financial system design is that this design may influence real decisions, I consider that issue in this section. Given the large and growing literature on this topic, I will not attempt an exhaustive discussion of the ways in which the financial system affects real decisions. Rather, I will focus on six interesting mechanisms by which financial system activity may be propagated to the real sector. These are depicted in Fig. 5.

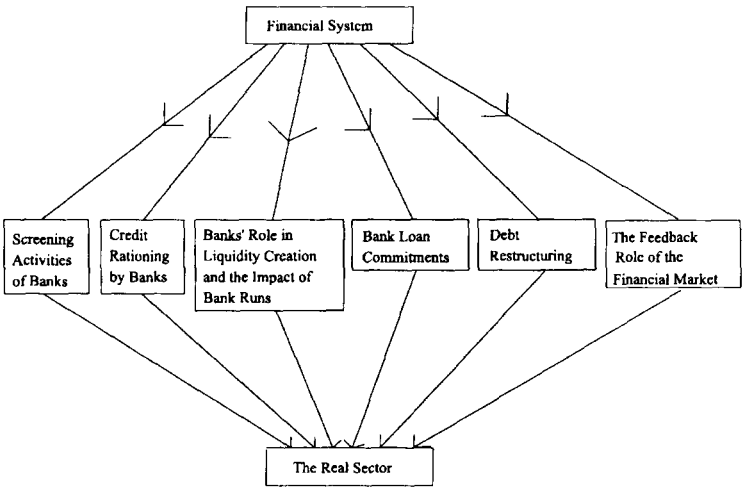

Fig. 5. The ways in which the financial system may affect the real sector.

2.1. Screening and monitoring activities of banks

A significant economic function served by financial intermediaries is to screen potential borrowers to generate signals about their creditworthiness. [1] Ramakrishnan and Thakor (1984) first formalized this intuition to provide a theory that rationalized the endogenous formation of perfectly diversified nondepository financial intermediaries whose principal function is to certify borrowers. [2] Contemporaneously Diamond (1984) rationalized perfectly diversified banks that monitored their borrowers' cash flows in a costly state-verification framework. Later, Boyd and Prescott (1986) exploited the screening role of banks to show how these institutions could enhance aggregate investment in socially beneficial projects by generating signals that facilitate a reduction in the investment capital diverted to inferior projects. By examining the role of banks in altering aggregate investment patterns, Boyd and Prescott thus brought to light an important link between the real and financial sectors. [3]

A similar point was made by King and Levine (1992), who joined the insights of the contemporary financial intermediation literature with those of new growth theory. [4] King and Levine observe that many capital investments that elevate productivity involve intangible capital goods whose value is often difficult to

[1] See Allen (1990) and Bhattacharya and Thakor (1993).
[2] Millon and Thakor (1985) rationalized imperfectly diversified certification agencies.
[3] See also Chan (1983).
[4] See also the review by Mayer and Vives (1992).

establish. In this setting, as in Boyd–Prescott, financial intermediaries arise endogenously as a market mechanism for screening entrepreneurs, but unlike Boyd–Prescott, intermediaries help finance intangible assets. Economies possessed with better financial systems – in particular, better financial intermediaries – are more adept at evaluating assets whose values are difficult to establish. This implies that a better developed financial system leads to higher productivity and growth through superior physical capital accumulation and more efficient investments in intangible and human capital.

An even more dramatic potential impact of screening on real activity was derived by Gale (1992) in the context of a model in which banks can acquire a costly screening technology. Banks have constraints on their capacity to screen borrowers, and it is costly for borrowers to apply for loans. Gale shows the existence of multiple equilibria, some involving (significant) bank lending and others involving little or no bank lending. The latter class of equilibria – called 'financial collapse' by Gale – arises from an adverse selection effect. If all banks except one decline to lend, then the sole lender will attract a relatively high proportion of bad borrowers who find it profitable to exploit the debt contract by investing in high-risk projects, assuming that their participation is masked by the presence of at least some good borrowers. A key to the model is that bad borrowers benefit more from bank loans because their projects have higher variance and therefore provide greater risk-shifting benefits. Given loan application costs, each borrower makes its application decision based on a tradeoff between the cost of applying and the expected value of a bank loan; the latter is declining in the probability of being rationed. Since for any given rationing probability, the value of a bank loan is lower to good borrowers than bad, the good borrowers exit earlier than the bad when the rationing probability rises. Thus, when there is only one bank that is willing to lend, the rationing probability is high, causing good borrowers to exit and causing the pool of borrowers to contain a higher proportion of bad borrowers than when all banks are willing to lend. The sole lender recognizes this, and its screening capacity constraints then lead it to reject a relatively high proportion of credit applicants. Application costs for borrowers now further deter good borrowers from applying because they perceive a low likelihood of receiving credit. This process continues until no good borrowers remain, which then leads banks to decide not to screen and simply reject all applicants. [5]

Cessation of lending is not the only equilibrium, however. If all but one bank decide to lend, then it will be profitable for the sole deviant to screen and lend as well. Thus, financial collapse potentially arises from a coordination failure.

[5] Although the description makes it appear as if these events are sequential, they are not – everything happens simultaneously.

2.2. Credit rationing by banks

In an attempt to explain why profit-maximizing banks would tolerate an excess demand for loans rather than raising the price for credit, Stiglitz and Weiss (1981) developed a model in which a bank's expected profit is not monotonically increasing in the loan interest rate. The nonmonotonicity can stem from adverse selection or moral hazard or both. Adverse selection arises from the assumption that the bank faces borrowers who are observationally indistinguishable but are heterogeneous in credit risk. When the bank raises its loan interest rate, the safer borrowers drop out of the credit market before the riskier borrowers do because the safer borrowers' greater likelihood of repaying their loans makes them more sensitive to the loan interest rate. Consequently, the bank recognizes that an increase in its loan interest rate causes the borrower pool to become riskier, and it may discover that its expected profit peaks at an interest rate that is less than that needed to clear the market; credit rationing is the result.

Raising the loan interest rate could also exacerbate asset-substitution moral hazard. If borrowers could choose between safe and risky assets and the bank was unable to observe which asset was chosen, Stiglitz and Weiss show that a sufficiently large increase in the loan interest rate could cause the borrower's preference to switch from the safe to the risky project. Hence, the bank may again find its expected profit peaking at an interest rate at which there is an excess loan demand.

A weakness of the Stiglitz and Weiss analysis is that the bank is merely a passive conduit for the transfer of funds from savers to borrowers – it provides no screening or monitoring services. [6] Papers by Thakor and Callaway (1983) and Thakor (1996) focus on noisy screening by banks and derive stochastic credit rationing as an equilibrium outcome. Thakor (1996) further shows that risk-based capital requirements – such as the Bank for International Settlements capital guidelines – tend to increase such rationing.

If bank loans are somehow unique and not readily replaceable by alternative credit sources, then credit rationing by banks will depress capital investment and retard economic growth. Moreover, if financial markets, with price-taking agents competing to provide funds, are less prone to ration credit, one might be tempted to conclude that encouraging the development of the financial market as a *viable* alternative to banks could result in higher aggregate investment. But such a conclusion would be premature since this literature has not explicitly considered the incentives that lead to the endogenous determination of markets and institutions within a financial system. I will have more to say on this later.

[6] Another weakness is that the rationing in their model is heavily dependent on the assumption that only a debt contract can be used by the bank in extending a loan. However, Williamson (1987) crafts a model in which the debt contract arises endogenously, and one encounters the Stiglitz–Weiss kind of credit rationing.

2.3. Bank's role in liquidity creation and the impact of bank runs

Banks provide a host of *qualitative asset transformation* (QAT) services. An important QAT service is *liquidity creation*. Banks create liquidity by issuing liquid deposit claims against illiquid loans. Diamond and Dybvig (1983) formalized this role of banks in a model in which the superiority of bank financing over financial market funding is due to the superior risk sharing provided by the bank. However, a coordination failure can lead to a 'bad' Nash equilibrium in which 'long-lived' depositors prematurely withdraw their deposits because they believe others will as well. [7] This run on the bank is disruptive since the only way the bank can satisfy depositors' demand is to liquidate loans. Given the assumptions that premature loan liquidation is costly and loans are illiquid – they cannot be sold in the market – it follows that bank runs adversely impact the real sector. [8]

2.4. Loan commitments and other long-term contracts

Over 80% of commercial and industrial lending by U.S. banks is done via loan commitments. A bank loan commitment is a promise by the bank to lend up to a predetermined amount at predetermined terms. As explained by Thakor et al. (1981), when a borrower acquires a bank loan commitment, it purchases a put option.

With borrower risk aversion, it is not obvious why customers demand loan commitments. Various papers have provided explanations. For example, Boot et al. (1987, Boot et al. (1991), show that a loan commitment can attenuate both effort-aversion and asset-substitution moral hazard. [9] The intuition is as follows. Under a loan commitment, a bank can set the promised loan interest rate low enough to ensure that the borrower will choose the first-best effort or project. When the second-best spot lending outcome is distorted away from the first-best, the bank will expect to make a loss on the loan itself. However, it can recoup this loss by charging a commitment fee at the time that it sells the commitment. Since this fee is treated as a sunk cost by the borrower at the time that it makes its action/project choice, it does not affect that choice. Hence, the first best is attained.

An important prediction of this theory is that bank loan commitments result in improved investment decisions. Moreover, the loan commitment contract usually has a 'material adverse change' clause which permits the bank to not honor the commitment if *it* judges the borrower's financial condition to be unacceptable at the time of commitment takedown. [10] The economic contribution of this clause

[7] In Diamond–Dybvig's model, such beliefs arise as sunspot phenomena.

[8] Bernanke (1988) discusses the disruptive effects of bank failures during the 1930s in the U.S.

[9] See also Houston and Venkataraman (1994).

[10] Boot et al. (1993) rationalize this discretion.

means that the loan commitment outcome cannot be replicated by exchange-traded put options, and the financial market cannot provide the same investment efficiency as bank financing for those borrowers who demand loan commitments.

There are some who believe that a key difference between the financial market and banks is that the latter can make long-term commitments with discretion regarding when to honor them, whereas investors in the financial market cannot. [11] It is an interesting research question to ask why. Although I have not seen a paper that specifically addresses this, Boot et al. (1993) suggest a possible explanation. In particular, they argue that the discretionary aspect of the bank loan commitment facilitates a more rapid development of bank reputation than is possible without that discretion. Moreover, it is the desire to develop/protect this reputation that induces the bank to honor its commitment even when it is not legally obligated to do so. By contrast, since there are numerous small investors typically providing funding to any borrower in the financial market, reputational concerns are of little consequence, and it is unlikely that discretionary commitments would be honored in the financial market. Consequently, such commitments are unlikely to be deployed in a financial market setting.

A somewhat different perspective appears in Mayer (1988) and Shleifer and Summers (1988). Both papers emphasize the importance of long-term relationships. They start with the premise that explicit contracting is essentially incomplete, and this makes it desirable for firms to enter into long-term implicit contracts with various stakeholders – employees, suppliers, creditors, etc. These implicit contracts (or long-term commitments) produce significant ex ante gains. For example, Jaggia and Thakor (1994) formally show that implicit contracts by firms that translate into long-term employment commitments encourage employees to invest more in nonmarketable, firm-specific human capital than would be possible with only explicit wage contracts. Of course, these are 'discretionary promises' in the sense of Boot et al. (1993) in that firms are not legally obliged to honor them. Moreover, it is costly ex post to honor such contracts. The key point then is that the financial market provides numerous opportunities – more so than bank financing does – to violate these implicit contracts. Jaggia and Thakor (1994) demonstrate that formal bankruptcy provides one mechanism for invalidating implicit long-term wage commitments. Shleifer and Summers (1988) point to the possibility of takeovers of publicly traded firms as another mechanism. The possibility that these implicit contracts will be violated ex post then weakens their ex ante (desirable) incentive effects. One advantage of bank-oriented financial structures is that this problem is less serious. As discussed earlier, reputational incentives deter banks from not honoring implicit contracts.

Much more needs to be done on this issue, however. From the loan-commitments or long-term contracting perspective, there are real effects associated with

[11] See, for example, Bhattacharya and Thakor (1993) and Sabani (1992).

changes in the relative proportions of bank and financial market financing. Thus, comprehending key differences in the abilities of banks and financial markets to credibly precommit to future outcomes is likely to prove very useful in contemplating overall financial system design.

2.5. Debt restructuring

Different financial systems may differ in the manner in which they restructure the debts of financially distressed borrowers. To the extent that debt restructuring has real effects, this means that optimal financial system design must also be conditioned on the real effects of that design.

This viewpoint is suggested by the work of Sabani (1992), who develops a model in which banks accumulate borrower-specific proprietary information in the post-lending stage of the relationship. Such information provides incentives for the incumbent bank to protect the relationship by restructuring the debt of its financially distressed borrower. Sabani shows that this framework produces Pareto-ranked multiple equilibria and that there is less restructuring by banks in competitive, market-oriented economies of the Anglo-Saxon mode than in bank-oriented systems of the German–Japanese mode. Thus, there is a potential tension between competition and commitment (to restructure). [12]

A quite different approach appears in Berlin and Mester (1992). They posit that coordination difficulties make it more difficult for loan covenants to be renegotiated in the financial market than by banks (see also Wilson, 1994). Such renegotiation may be necessitated by random shocks to the lender's information about the borrower. Berlin and Mester conclude that banks will tend to negotiate loan contracts that have stringent covenants but will be willing to renegotiate these covenants if warranted. Thus, borrowers who pose a relatively onerous asset-substitution moral hazard to the lender and also display significant information volatility prefer to borrow from banks. [13]

It is natural to assume that the borrower's anticipation of whether a loan will be restructured will affect its ex ante choice of investment project. In this case, investment choices will be partly driven by whether the borrower has access to bank or financial market financing. Moreover, the analysis of Sabani (1992) suggests that borrowers' investment choices are also likely to be influenced by the degree of competition in banking. Thus, details of financial system design are likely to be highly pertinent to the real sector.

[12] Dewatripont and Maskin (1990) suggest that bank-oriented economies persist with bad projects too long, whereas market-oriented economies cut off good projects too early.

[13] Thakor and Wilson (1995) show that risk-based bank capital requirements will reduce the extent of restructuring by banks and hence reduce bank loan demand.

2.6. *The feedback role of the financial market*

There have been numerous papers that have suggested that the financial market can emit signals that contain payoff-relevant information not available in profits or other reported summary measures. The additional information conveyed by the financial market can then be used for a variety of purposes, with significant real effects.

One strand of this literature suggests that stock prices provide signals for the efficient allocation of investment (see, e.g., Grossman, 1976, Grossman, 1978; Grossman and Stiglitz, 1980). This theme is also developed in Allen (1992), who takes a somewhat broader view that stock market signals can lead to improved real decisionmaking in general; I discuss this paper in greater detail in the next section.

Another strand of this literature relies on the dependence of managerial incentive contracts on stock prices. Ramakrishnan and Thakor (1984) derived optimal incentive contracts for managers in publicly traded firms who choose effort that affects real payoffs. The contracts they derive provide optimal risk sharing and effort incentives, and depend heavily on stock-market information. Thus, the feedback role of the financial market is in providing information regarding managerial performance that then affects effort input decisions of managers. This issue is also confronted by Holmstrom and Tirole (1993), who develop a model in which stock prices contain information not available in the firm's own information set, and therefore permit more efficient contracting and superior real decisions.

3. Financial system diversity

As I mentioned in the Introduction, there is a great deal of variety in financial systems across different countries. A brief summary of these distinctions appears below. [14]

- The U.S. has far more banks than any other country. Moreover, the ten largest U.S. banks account for a significantly smaller portion of total lending in the U.S. than their counterparts do in Japan and Europe. That is, the U.S. banking industry is much more fragmented than that in other countries.
- Western European and Japanese banks have traditionally had more powers (wider scope) than U.S. banks. For example, while Japanese banks have, like their U.S. counterparts, been barred from universal banking, they have had a greater ability to hold equity in borrowing firms and exercise corporate control. German, Swiss and Dutch banks have been permitted to be 'universal' and

[14] See Allen (1992) and Greenbaum and Thakor (1995).

engage in investment banking and insurance in addition to their usual commercial banking functions.

- Long-term relationships between banks and their borrowers are much more extensive in countries like Japan and Germany than in the U.S.
- In the U.S. there is considerable competition between depository financial intermediaries and financial markets. This competition is much more limited in other countries.
- The U.S. has far more publicly listed firms than any other country.
- There is considerably more information about publicly listed firms that is available in the public domain in the U.S. than in most other countries.
- The market for corporate control is significantly more active in the U.S. than in other countries.
- In general, financial markets in the U.S. are more highly developed than those in other countries. For example, the futures and options markets in the U.S. display far more liquidity than those in Germany.

While there are many conjectures about the reasons for this diversity in the configuration of financial systems, I do not believe we have anything close to an integrated theory that accommodates all of the relevant stylized facts related to this diversity and also generates additional testable predictions. Perhaps the beginning of the contemporary literature on this subject can be traced to Sah and Stiglitz (1986), who explored the project selection ramifications of different ways of organizing economic systems. They suggest that there will be more Type-II project selection errors in market-oriented economies and more Type-I project selection errors in bureaucracy-oriented economies. As for the architecture of financial systems in particular, a preliminary step was taken by Allen (1992) in a descriptive essay on comparative financial systems. Allen considers a model in which the manager's information set does not include all that is of relevance for optimal real decisions. The more complex the technology employed by the firm and the greater the frequency with which payoff-germane new information arrives in the market, the less complete will be the manager's information. Since the decisions the manager makes are conditioned on his own information, optimality in decision-making will be sacrificed due to the incompleteness in the manager's information. This not only makes the manager value the information of others that he does not possess, but it also creates a divergence of opinions about how the firm should be run, leading in turn to a role for corporate control contests.

In this framework, the financial market is viewed as a mechanism for aggregating many diverse opinions and hence providing information about optimal decision rules in corporations that is superior to that attainable through bank borrowing. The reason is that the bank provides a 'single check' as opposed to the 'multiple checks' of the financial market. Multiple checks are provided by the financial market through a variety of mechanisms – market prices, trading volume, takeover attempts, etc. It follows from this that capital will be allocated in the financial market when optimal decision rules are hard to formulate, and there is

little consensus on how the firm should be run; for example, when information decays rapidly and new information arrives almost constantly. Allen cites biotechnology as an example of such an industry. In contrast, banks are desirable institutions for allocating resources in situations in which there is consensus on decisionmaking, and the main problem is monitoring firms.

Allen uses this framework to rationalize the divergent patterns of financial market development in different countries. He points out that his theory is consistent with the importance of the stock market in the U.K. during the 19th century when it was the first country to go through the Industrial Revolution. It is also consistent with the fact that the U.S. has relied most heavily on stock markets in the 20th century when it was the first country to go through the post-Industrial Revolution. And the prediction is that as the U.S. goes through the 'Information Revolution', the relative importance of the financial market will grow.

Allen also extracts the implications of his theory for the development of financial institutions in Europe. Stock markets will need to develop further in the advanced economics of Western Europe if new technologies and industries are to emerge. In Eastern Europe, on the other hand, the task is to build basic industries where the technology is familiar and optimal decision rules are well known. [15] The development of banks should therefore be given priority over the development of financial markets.

Allen's framework is appealing because it is simple and seems to explain quite a bit. However, because it is only an initial step, it leaves numerous questions unanswered. For example, if there are diminishing marginal returns to adding more agents to the pool of agents whose opinions are being aggregated, then how many agent's opinions do we need to aggregate before we come close to mimicking the aggregation efficiency of the financial market? If this number is around say seven or eight, then it is difficult to see what advantage the financial market offers to large firms relative to bank financing, since each large U.S. company has relationships with seven to eight banks on average. Moreover, given the monitoring advantage of banks, if the optimal number exceeds eight, why do not we observe each firm approaching a larger number of banks rather than going to the financial market? Also, why is it that the financial market is unable to provide the monitoring services that banks provide?

I believe that we will begin to understand these issues only when formal models are crafted to deal with them. Recent papers by Allen and Gale (1994a) and Boot and Thakor (1996) represent initial stabs at formal modeling. I discuss these papers in Section 7.

[15] These rules are well known to managers in developed countries. Whether this is true in Eastern Europe is questionable.

4. Borrower's choice of financing source

How does a borrower decide which source of credit to use at any given time, conditional on having decided to acquire debt? This question was formally articulated by Diamond (1991a) in the context of a model in which a borrower can choose between a bank and the financial market. In Diamond's model, some borrowers can substitute risky assets for safe ones to the lender's detriment. Banks specialize in resolving this asset-substitution moral hazard problem through monitoring. However, the longer that a borrower keeps repaying its loan – with the probability of repayment in any given period enhanced by the choice of the safe project in that period – the better is its credit reputation and the lower are the interest rates on future loans. This means that a borrower with a better reputation will perceive a higher *present value* of its net payoffs from future bank loans. Under the assumption that a borrower who defaults in any given period is then forever excluded from the capital market or bank borrowing, it follows that borrowers with better reputations attach a lower value to choosing the risky project over the safe one. Indeed, a reputational cutoff is reached such that once the borrower's credit reputation exceeds that cutoff, the borrower will always prefer the safe project over the risky project. Given this, bank monitoring is unnecessary. Thus, Diamond's model produces the prediction that borrowers with nascent credit reputations approach banks, whereas more mature borrowers with well-established credit reputations access the capital market directly. This prediction seems roughly consistent with the stylized facts.

Numerous authors have since examined various facets of the borrower's choice of financing source. Rajan (1992) focuses on the intertemporal accumulation of proprietary borrower-specific information by relationship lenders like banks. He shows that this leads to an informational monopoly for an incumbent bank and distorts the borrower's effort choice away from the first best. Thus, Rajan's hypothesis is orthogonal to the earlier research on the value of long-term bank-borrower relationships, in that it highlights the distortionary nature of such relationships rather than its earlier-mentioned benefits. This leads Rajan to conclude that borrowers who perceive significant *future* rents from their projects will wish to avoid bank financing at present in order to preclude the future expropriation of significant portion of these rents by informationally privileged banks.[16]

A different approach was suggested by Campbell (1979). He focused on the tension between shareholders and bondholders. Assuming that the terms of debt are renegotiable, Campbell argued informally that shareholders, represented by

[16] The notion that bank lending is profitable and often a substitute for capital market funding is provided empirical support by Benveniste et al. (1993), who find that the commercial banking industry experienced positive wealth effects when Drexel Burnham Lambert – a major underwriter of junk bonds (a close substitute for bank loans) – failed.

management, may prefer to borrow from a bank in order to conceal favorable news from bondholders. Campbell was therefore the first to raise the possibility that confidentiality concerns could tilt the borrower's choice in favor of bank financing.

This intuition was enriched and formalized by Yosha (1995) and Bhattacharya and Chiesa (1995). Yosha examines firms' choices between bilateral (bank) and multilateral (public) financing after firms have acquired some proprietary information. If a firm chooses multilateral financing, it faces a greater probability that its proprietary information will be leaked to competitors. But if it chooses bilateral financing, its competitors may infer that the firm has something to hide and react in a way that reduces the firm's profit. Yosha shows that in equilibrium, the high-quality firms (that have most to lose from their information being leaked) choose bilateral financing, and that the cost differential between the two financing source choices keeps competitors from unambiguously inferring that these firms are hiding information. Bhattacharya and Chiesa also focus on information-sharing/disclosure issues in the context of bilateral (single bank) and multilateral (multiple banks or the financial market) financing regimes. The novelty of their approach is in the conclusion that the choice of financing arrangement can serve as a precommitment to enforce an ex ante desired information-sharing rule that may otherwise prove difficult to implement ex post. This can affect ex ante incentives to make investments that generate future information. Thus, the design of the financial system can significantly impact R&D investments.

None of the above papers address *intrafirm* incentive problems faced by the borrower, and the potential impact of the choice of financing source on the resolution of these problems. Wilson (1994) develops a model in which divisional managers have a propensity to overinvest in some projects due to private benefits. They are able to overinvest because they know more about their project a priori than their supervisors do. In the interim stage, when more financing is needed to sustain the project, additional information may arrive that reduces the informational gap between divisional managers and their supervisors. But by this time, the initial investment is a sunk cost and therefore is ignored by senior management. Additional funding at the interim stage is therefore provided if the *subgame* NPV (ignoring the initial investment) is nonnegative. Anticipating this, the divisional manager sometimes invests in projects that produce private benefits for him but have negative ex ante NPV for the firm.

Wilson observes that one way to ameliorate this moral hazard is to have a liquidity constraint on interim financing that effectively raises the hurdle rate for that kind of financing. This ensures that threats by senior management to choke off interim financing for ex-ante-negative-NPV projects are credible. Now, if there is a bank that is providing financing, the firm faces only a 'soft' budget constraint since the relatively easy renegotiability of bank financing means that the firm can acquire whatever additional funds it needs to maximize NPV at the interim stage. In this case, it is not subgame perfect for the firm to deny interim financing to

Table 1
Real effects associated with borrower's choice of financing source

Paper	Role of bank	Role of financial market
Diamond (1991a)	resolves asset-substitution moral hazard through monitoring	provides residual funding for borrowers with sufficiently high credit reputations
Rajan (1992)	exploits ex post informational monopoly due to its relationship with borrower and thereby depresses borrower's effort-supply incentives	provides 'arms-length' financing that avoids effort-avoidance moral hazard
Bhattacharya and Chiesa (1995)	provide greater confidentiality of R&D information generated by borrower	may help to implement ex ante efficient information-sharing arrangements
Wilson (1994)	provides financing to borrowers that do not face significant intrafirm incentive problems	helps attenuate intrafirm incentive problems that lead to overinvestment

projects with negative ex ante NPVs. On the other hand, the firm faces a 'harsh' budget constraint with financial market financing – whatever constraint is imposed ex ante on the firm's interim funds cannot be relaxed. The firm can choose the constraint ex ante in such a way that moral hazard is efficiently dealt with.

An interesting commonality among all of these papers is that the firm's choice of financing source has real effects. Table 1 summarizes the real effects predicted by each of these papers. What this means is that how the financial system is configured will matter to the real sector because of its impact on the borrower's choice of financing source. For instance, if the financial market is relatively underdeveloped, Rajan (1992) and Wilson (1994) would predict that effort-avoidance and overinvestment moral hazard are likely to be exacerbated. On the other hand, an underdeveloped banking system will mean that there will be an excessive reliance on borrower reputation and a relatively poor resolution of asset-substitution moral hazard (Diamond, 1991a).

5. Corporate control and capital structure

The design of the financial system can affect information aggregation and the outcomes of corporate control contests. This in itself can influence the firm's capital structure. There is also another way in which the design of the financial system can affect capital structure. It has recently been shown that the availability of a bank loan commitment can affect the firm's choice of capital structure (see Shockley, 1995). Both these effects imply real effects of financial system design since it is well known that capital structure can affect investment decisions.

5.1. Information and corporate control

We have already seen how information aggregation can be of value to the firm when decisionmaking is complex. There are many instances in which firms engage voluntarily in direct exchanges of information. Examples are trade associations, cartels and research joint ventures. Vives (1984) shows that if goods are substitutes and there is Bertrand competition or if goods are complements and there is Cournot competition, information sharing is optimal.

There is also a pretty extensive literature on the formation of cartels and information exchange. Although many papers have shown that information sharing will not be optimal in many situations (e.g., Roberts, 1985; Cramton and Palfrey, 1990), Kihlstrom and Vives (1996) develop a model in which information is exchanged.

Research joint ventures are another way to exchange information. Katz (1986) has explored the incentives firms have to share the costs of and knowledge created by research. Bhattacharya et al. (1990) examine the optimal level of information sharing and its implementation. Bhattacharya and Chiesa (1995) analyze the impact of the configuration of the financing source on information sharing incentives.

If direct voluntary exchange of information is not sufficient, other mechanisms may come into play. One such mechanism is provided by the market for corporate control. Manne (1965) has suggested that an important attribute of market-based economics is the fact that different management teams can compete for the control of assets. As Allen (1992) observes, proxy fights, direct share purchases and mergers represent three prominent market-based mechanisms for the transfer of control. These three mechanisms can be interpreted in the context of the Allen (1992) model described earlier.

Suppose a raider has access to better information than the incumbent management of the firm, and is therefore in a position to implement better decisions. However, it is likely to be difficult to convince shareholders that the raider will make better decisions than the existing management. This implies that, with many shareholders and fragmented ownership of the firm, proxy fights are likely to be difficult for raiders to win, an observation consistent with the stylized facts. Consequently, it may be necessary for the raider to simply offer target shareholders a sufficiently high premium over the current stock price to induce them to sell out. This can be achieved with a tender offer to purchase shares at a premium. [17]

[17] Of course, this will raise the usual free-rider problem highlighted by Grossman and Hart (1980). However, as pointed out by Bagnoli and Lipman (1988), this problem can be avoided with nonatomistic shareholders. Grossman and Hart (1980) suggested that 'exclusionary devices' – such as permitting the raider to exclude from sharing in the post-takeover gains those shareholders who do not tender their shares to the raider – could eliminate the free-rider problem. Shleifer and Vishny (1986) propose that a large minority shareholder could also help overcome the free-rider problem.

It is, of course, possible that when opinions are sufficiently diverse, the raider will not be able to overcome the holdout problem raised by Grossman and Hart (1980). Efficient information sharing may then be possible only with a (negotiated) merger.

The point is that the design of the financial system can affect the efficiency with which information relevant to corporate decisions is shared/aggregated. A well-developed financial market facilitates the information aggregation process by making corporate control contests easier. This is consistent with the stylized fact that the U.S., with its relatively more complex real technologies, also has a better developed financial market.

The observation that corporate control contests are facilitated by a greater degree of development of the financial market has other implications as well. In particular, Harris and Raviv (1991) have shown that a firms's capital structure may depend on the tension between two control considerations. If the firm has too little debt, then its incumbent management risks losing control to a raider who can increase firm value simply through an elevation of the firm's leverage. On the other hand, if the firm has too much debt, then its incumbent management risks losing control to the bondholders if bond covenants are violated. Thus, the optimal capital structure balances the probability of losing control to a corporate raider against the probability of losing control to bondholders.

An implication of this logic is that the expected control cost of equity is likely to be higher in better developed financial markets because the threat of a hostile takeover is greater. This means that we should expect a heavier reliance on debt in such markets. The empirical prediction then is that leverage ratios of firms should be higher in the U.S. than in, say, Germany.

An offsetting consideration is that economies in which banks are dominant are also characterized by long-term bank–borrower relationships to a greater extent than economies in which banks play a lesser role. These relationships can help mitigate moral hazard and encourage greater use of debt. Boot and Thakor (1996) develop an infinite-horizon model of banking in which long-term contracts improve efficiency by reducing reliance on collateral to resolve effort-avoidance moral hazard. To the extent that such moral hazard reduces optimal leverage, economies dominated by banks are also likely to have firms that borrow more from banks and have *higher* leverage ratios.

What is missing, of course, is an integrated theory that accounts for *all* of these effects. The Harris–Raviv model does not consider the value of banks or relationships between borrowers and lenders. I think this is a particularly interesting open research question that deserves attention.

5.2. Banks and corporate capital structure

The fact that banks can make loan commitments that involve discretionary promises to lend in the future at predetermined terms has an important implication

for the capital structures of nonfinancial firms. Shockley (1995) has recently shown in an interesting model that a firm that purchases a bank loan commitment will acquire more *nonbank* debt than a firm that does not have a loan commitment. The intuition is that a loan commitment attenuates the underinvestment moral hazard associated with risky debt (see Myers, 1977) and thus improves the terms of nonbank debt.

Again, the implication is that bank financing is associated with higher leverage ratios for those who borrow from banks. Moreover, the result that bank loan commitment financing resolves underinvestment moral hazard implies improved decisions related to a particular financial system design.

6. Liquidity, bank liabilities and financial system design

As mentioned in Section 2, an important QAT service provided by banks is that of liquidity creation. In the process of creating liquidity, banks end up mismatching their balance sheets as they typically finance (illiquid) long-maturity assets with (liquid) liabilities of varying but shorter maturities. Diamond (1991b) has examined the liability maturity decision and shown that liquidation incentives and control considerations play significant roles in this decision.

What is particularly interesting about liquidity creation by banks is that it typically involves a demand deposit contract constrained by a *sequential service constraint* (SSC). In the Diamond and Dybvig (1983) model, the SSC plays a pivotal role in generating bank runs. Thus, it forces the bank to cope with risk in the process of creating liquidity. There are some who believe that the SSC is just part of the technology of banking. For example, in continuous time, serving withdrawals sequentially until funds were exhausted would be a natural way of doing business; the SSC in Diamond–Dybvig's model can then be viewed merely as a way to approximate this continuous-time phenomenon in discrete time. However, there are others who believe that the SSC is a contract design parameter that must either be rationalized endogenously or be excluded from the demand deposit contract. This is the approach taken by Calomiris and Kahn (1991), who propose that the SSC solves a free-rider problem in depositor monitoring. In their model, uninsured depositors monitor bank management to prevent activities like fraud and excessive risk taking. Without the SSC, however, depositors have an incentive to free-ride on the costly monitoring performed by others, and the monitoring mechanism consequently breaks down. The SSC prevents free-riding because monitoring depositors precipitate a run on the bank when they detect untoward behavior, and it is the monitoring depositors who are first in line to withdraw. A free-riding depositor would get a lesser expected payoff upon withdrawal due to his later position in the withdrawal queue.

In the Calomiris–Kahn framework, bank runs are socially desirable. This makes it difficult to juxtapose this insight about the SSC with its role in the

Diamond–Dybvig model in which runs are productively disruptive. Moreover, it makes it virtually impossible to visualize a meaningful role for deposit insurance. These issues were resolved by Peters (1995). In his model, bank runs are often *ex post inefficient* but *ex ante* effective in disciplining bank management. The SSC increases the marginal payoff to a depositor from becoming an informed monitor and thus leads to a high probability of a bank run. However, this probability may be higher than that needed to discipline management and thus leads to excessive ex post inefficiencies. Peters shows that partial deposit insurance reduces the marginal payoff to a depositor from becoming an informed monitor and thus reduces the probability of a bank run. The level of deposit insurance coverage can be calibrated to ensure that the probability of a bank run is no more than that needed to align the incentives of bank management with those of depositors.

The connection of these liability management issues with banking scope and financial system design was recently highlighted by Peters and Thakor (1995). They craft a model in which depositors must be incented to monitor the asset choices of bank managers and the bank's managers must be incented to monitor borrowers. Their key result is that it is impossible to provide *both* of these incentives at their first-best levels in a single bank. They prescribe *functionally separating* the bank into: (i) a pure liquidity-creation bank that has access to insured deposits and is permitted to invest only in assets whose payoffs are invariant to bank monitoring, and (ii) an uninsured bank that can invest in any asset (a sort of universal bank).

A somewhat different functional separation prescription emerges from the analysis of Craine (1995). Craine shows that various distortions can be diminished by separately chartering banks that invest in 'private-information' assets – those about which the bank has proprietary information – and those that invest in public-information assets. In a separating equilibrium, the public-information banks raise funds through insured deposits and the private-information banks raise funds through uninsured liabilities.

These functional separation prescriptions are reminiscent of the 'narrow bank' proposal originally made by Simons (1934, Simons (1935), but are different in their underlying logic and details. In particular, these contemporary proposals focus on the impact of informational frictions and organizational design issues. Moreover, they do *not* necessarily prescribe a narrow bank that produces no net liquidity; this prescription is at the heart of the original Simons proposal.

Policy discussions of the merits of these proposals often revolve around the uniqueness of *insured* banks in providing liquidity. If we limit the scope of federal deposit insurance to a smallish slice of the banking industry *and* if deposit insurance is critical to creating net liquidity – as it is in Diamond and Dybvig (1983) – then narrow banking could be deleterious to liquidity creation. However, it has been suggested that liquidity can be effectively created by means other than deposit insurance. For example, cash flow partitioning creates liquidity outside the banking sector in Boot and Thakor (1993a); see also Gorton and Pennacchi (1990).

Indeed, the financial innovation analysis of Boot and Thakor (1995) suggests that there may be *more* liquidity created in financial-market-dominated economies than in bank-dominated economies due to the stronger incentives for financial innovation in the former. On the other hand, the Allen and Gale (1994a) argument is that the *nature* of financial instruments available for liquidity management differ across the two types of economies. In particular, cross-sectional risk sharing is likely to be better in market-dominated economies, whereas intertemporal risk sharing is likely to be better in bank-dominated economies.

7. Scope of banking and financial innovation

In casual discussions about financial system design, one hears the argument that it is likely to be influenced by regulatory restrictions on the scope of banking. For example, it is claimed that the 1933 Glass–Steagall Act, that has separated commercial and investment banking in the U.S. for over six decades, has had a major impact on the development of banking and financial markets in the U.S., and provides a plausible explanation for the differences in financial systems across different countries (see Saunders and Walter, 1994).

Numerous authors have recently examined whether the Glass–Steagall restrictions make sense anymore (e.g., Rajan, 1993; Puri, 1994). Traditional arguments against universal banking have focused on conflicts of interest arising from the dual role of universal banks in commercial lending and securities underwriting (see Kanatas and Qi, 1993). These conflicts may be manifested in misrepresentation of the financial condition (or quality) of a borrower whose capital market issue is being underwritten by the lender or a delay in taking a borrower to the financial market because profits on lending exceed those on underwriting.

While it is accepted that banking regulation affects the development of the financial market, we do not understand very well all of the richness in the interplay between bank regulation and the evolution of the financial system. For example, if conflicts of interest between the commercial and investment banking parts of a universal bank are serious, why would the market itself not provide the necessary signals to discipline the universal bank? In particular, if these conflicts were sufficiently serious, it would be optimal for the universal bank to provide *credible* separation between its commercial and investment banking activities. Or put a little differently, if a universal bank were really inefficient from a conflict-of-interests standpoint, market forces would make it unnecessary to have regulation to outlaw universal banking. Indeed, the empirical evidence provided by Kroszner and Rajan (1994a) supports the market discipline argument and militates against the notion that these conflicts are of significance. They compare the performance of securities underwritten by commercial and investment banks prior to the Glass–Steagall Act and find no significant difference in the qualities of issues underwritten by commercial and investment banks. In a subsequent paper,

Kroszner and Rajan (1994b) find that the market performance of underwritten securities in the 1920s was better when underwriting was done by a separately capitalized securities affiliate than when it was done by a securities division of the universal bank. Consistent with this, commercial banks during this time almost uniformly moved towards the affiliate structure, which involved greater separation between lending and underwriting.

7.1. Financial innovation

In a series of papers, Allen and Gale have explored the economic incentives underlying financial innovation (see, e.g., Allen and Gale, 1994b). The basic premise of these papers is that financial innovation improves the risk-sharing opportunities available to investors. Since successful innovation often leads to the emergence of new markets, this perspective is useful in understanding the development of futures and options markets.

A somewhat different approach is taken by the earlier-mentioned Boot and Thakor (1993a) paper. Boot and Thakor explain financial innovation in terms of the incentives of investors to acquire costly information in an adverse-selection economy. They show that innovation that produces more information-sensitive partitions of existing securities stimulates greater information acquisition by investors. This enhances the liquidity of the firm's securities and leads to more efficient market pricing.

Whether innovation leads to the emergence of new markets that expand risk-sharing opportunities or improves liquidity in existing markets, it is likely to have real consequences. Consider improved risk sharing first. It has been shown that corporate hedging and other risk-reduction initiatives like conglomerate mergers may improve investment decisions (see, e.g., Ramakrishnan and Thakor, 1991). An intuitive way to think about this is to visualize a principal-agent framework in which managerial incentive contracts make the manager's payoff sensitive to idiosyncratic risk to cope with effort-aversion moral hazard. In this case, the greater the variance of the idiosyncratic risk, the lower will be welfare (see Ramakrishnan and Thakor, 1984), so that hedging this risk can be welfare-improving.

Consider next more efficient market pricing. As shown by Myers and Majluf (1984), asymmetric information about firm valuation can result in the firm eschewing positive-NPV projects. Thus, financial innovation that reduces the impact of informational asymmetries can help diminish underinvestment distortions.

An open research question that interests me is whether financial innovation incentives are affected by financial system design. Are bank-dominated financial systems more or less likely to innovate than market-dominated systems? Does the permissible scope of banking influence how much financial innovation occurs? This issue has recently been addressed by Boot and Thakor (1995). They show

that a financial system with a *concentrated* universal banking industry exhibits stochastically lower innovation than a functionally separated financial system in which commercial and investment banking are kept distinct by regulatory fiat. This provides a perspective on the greater rate of financial innovation in the U.S. than in continental Europe.

8. Overall financial system design

I have discussed a wide variety of issues related to financial system design. The ultimate goal, of course, is to join the insights of these seemingly disparate strands of the literature to craft a unified theory of financial system design. This has not happened yet. But two recent papers have taken modest initial steps.

Allen and Gale (1994a) build a theory that seeks to explain the welfare implications of two prominent designs – the market-dominated U.S. financial system and the bank-dominated German financial system. They argue that bank-dominated systems provide better intertemporal risk sharing than market-dominated systems. The reason is that banks are better equipped to make long-term commitments that are essential to attain desired intergenerational wealth transfers and risk sharing. In the Allen–Gale framework, such commitments are unavailable in the financial market because they are unraveled by competition in a manner similar to the competitive unraveling of the optimal (from a risk-sharing standpoint) pooling contract in the Rothschild and Stiglitz (1976) insurance model.

On the other hand, market-dominated systems provide better cross-sectional risk sharing at any point in time than bank-dominated systems. The reason is that there are more instruments available to economic agents for hedging their risk exposure in the financial market than with banks.

There is thus no unambiguous answer to the question of which financial market design is better from a welfare standpoint. Moreover, the Allen–Gale approach suggests that the German financial system achieves a better allocation of risk across generations, whereas the U.S. financial system achieves a superior intragenerational allocation of risk.

A different aspect of financial system design is explored by Boot and Thakor (1996). They start with assumptions about primitives – the types of agents in the economy and their endowments of wealth and information. They show that agents who choose to specialize in monitoring borrowers to deter asset-substitution moral hazard will prefer to coalesce and form banks. On the other hand, agents who choose to acquire payoff-relevant information about firms that is unavailable to the managers of these firms will prefer to trade independently in the financial market.

In addition to rationalizing the emergence of banks and financial markets from primitives, Boot and Thakor also show that borrowers with high observable credit qualities will opt for financial market financing, and those with relatively low observable credit qualities will opt for bank financing. Thus, they provide a theory

of the borrower's choice of financing source that is predicated on observable differences in credit attributes. An important feature of the Boot–Thakor analysis is that *both* bank financing and capital market funding have real effects. The real effect associated with bank financing comes from the fact that the bank is more effective in enforcing a choice of the socially preferred project by the borrower. The real effect associated with financial market funding arises from the 'feedback' provided by the market to the firm's manager that induces him to improve his decisions with respect to investments in enhancing project payoffs. That is, the manager does not possess all of the information he needs for optimal decisions; he obtains some of this information by observing financial market trading, and this leads to more efficient decisions.

The importance of industry structure in financial system design should also be appreciated. The Boot and Thakor (1996) analysis shows that the demand for bank credit is lower when the banking industry is concentrated than when it is fragmented. The reason is that the price of bank credit is higher in concentrated banking industries. Moreover, problems of relationship-specific informational monopolies are also likely to be more acute in concentrated banking systems (Rajan, 1992). On the other hand, Boot and Thakor (1995) show that the more concentrated the *universal* banking industry, the lower will be the rate of financial innovation, and hence the slower will be the pace at which the financial market evolves. Thus, industry structure can profoundly influence how the financial system configures itself.

As we consider overall financial system design, can we say anything about the relationship between the design of the financial system and its stability? Recently, Allen and Gale (1995) have suggested that bank-dominated economies in which banks do not face considerable competition from financial markets have a greater ability to smooth asset returns over time than market-dominated economies. They present an example of an economy in which the incompleteness of financial markets leads to underinvestment in reserves, and for a broad class of welfare functions, the optimum requires the holding of large reserves in order to smooth asset returns over time. They argue that a long-lived intermediary may be able to implement the optimum. However, it would be premature to conclude from this that bank-dominated financial systems are more stable. The Allen–Gale analysis does not deal with a multibank financial system with the possibility of bank runs and contagion-induced panics. The question of which system is more stable remains open at this point.

9. Conclusion

While there has been a great deal of work done in related areas, we have a lot to learn when it comes to overall financial system design. I believe that this topic represents perhaps the most exciting research possibility in the area of institutions

and markets. What I suspect it will teach us is that we need to rethink the way we visualize institutions and markets. We tend to view institutions and markets as distinct and competitive. Yet, as Merton (1993) has pointed out, they also complement each other since financial market innovations often open up new business opportunities for banks in the risk management arena.

No discussion of the design of financial systems would be complete without a discussion of the political environment and the political economy of design. It is apparent that the recent discussions of repealing the Glass–Steagall Act in the U.S. and adopting interstate branching have been greatly politicized. For example, the securities and insurance industries have been opposed to Glass–Steagall repeal, whereas small banks have been opposed to interstate branching. Unfortunately, there has been little analysis of these issues. Exceptions are papers by Boot and Thakor (1993b), Campbell et al. (1992), and Kane (1989, Kane (1990).

The distinctions between institutions and markets will continue to become even more blurred, and market forces will press on to make national regulators less important and regulatory restrictions such as Glass–Steagall obsolete. If research on financial system design yields a good harvest, we may have something to say about how financial systems will evolve in the future and what the welfare implications of this evolution are likely to be.

Acknowledgements

Prepared initially for a keynote address at the 7th Annual Australian Finance and Banking Conference.

I thank an anonymous referee, David Hirshleifer and Patty Wilson for helpful comments and Kathleen Petrie for research assistance. Only I am responsible for any infelicities, however.

References

Allen, F., 1990, The market for information and the origin of financial intermediation, Journal of Financial Intermediation 1, 3–30.

Allen, F., 1992, in: C. Mayer and X. Vives, eds., Stock markets and resource allocation in capital markets and financial intermediation (Cambridge University Press, Cambridge).

Allen, F. and D. Gale, 1994a, A welfare comparison of the German and U.S. financial systems, Working paper (Wharton School).

Allen, F. and D. Gale, 1994b, Financial innovation and risk sharing (MIT Press, Cambridge, MA).

Allen, F. and D. Gale, 1995, Financial markets, intermediaries, and intertemporal smoothing, Rodney L. White Center for Financial Research WP #5-95 (Wharton School).

Bagnoli, M. and B.L. Lipman, 1988, Successful takeovers without exclusion, Review of Financial Studies 1, 89–110.

Benveniste, L.M., M. Singh and W.J. Wilhelm, Jr., 1993, The failure of Drexel Burnham Lambert: Evidence on the implications for commercial banks, Journal of Financial Intermediation 3, 104–137.

Berlin, M. and L.J. Mester, 1992, Debt covenants and renegotiation, Journal of Financial Intermediation 2, 95–133.

Berlin, M., K. John and A. Saunders, 1994, Bank equity stakes in borrowing firms and financial distress, Working paper (Indiana University).

Bernanke, B., 1988, Monetary policy transmission through money or credit?, Business Review (November/December), 3–11.

Bhattacharya, S. and G. Chiesa, 1995, Financial intermediation with proprietary information, Journal of Financial Intermediation 4, 328–357.

Bhattacharya, S., J. Glazer and D. Sappington, 1990, Sharing productive knowledge in internally financed R&D contests, Journal of Industrial Economics 39, 187–208.

Bhattacharya, S. and A.V. Thakor, 1993, Contemporary banking theory, Journal of Financial Intermediation 3, 2–50.

Boot, A.W. and A.V. Thakor, 1993a, Security design, Journal of Finance 48, 1349–1378.

Boot, A.W. and A.V. Thakor, 1993b, Self-interested bank regulation, American Economic Review 83, 206–212.

Boot, A.W. and A.V. Thakor, 1996, Financial system architecture, Working paper (Indiana University).

Boot, A.W. and A.V. Thakor, 1995, Banking scope, financial innovation, and the evolution of the financial system, Working paper (Indiana University).

Boot, A.W., A.V. Thakor and G.F. Udell, 1987, Competition, risk neutrality and loan commitments, Journal of Banking and Finance 11, 449–472.

Boot, A.W., A.V. Thakor and G.F. Udell, 1991, Credible commitments, contract enforcement problems and banks: Intermediation as credibility assurance, Journal of Banking and Finance 15, 605–632.

Boot, A.W., S. Greenbaum and A.V. Thakor, 1993, Reputation and discretion in financial contracting, American Economic Review 83, 1165–1183.

Boyd, J.H. and E.C. Prescott, 1986, Financial intermediary-coalitions, Journal of Economic Theory 38, 211–232.

Calomiris, C.W. and C.M. Kahn, 1991, The role of demandable debt in structuring optimal banking arrangements, American Economic Review 81, 497–513.

Campbell, T., 1979, Optimal investment financing decisions and the value of confidentiality, Journal of Financial and Quantitative Analysis 14, 913–924.

Campbell, T., Y.-S. Chan and A.M. Marino, 1992, An incentive based theory of bank regulation, Journal of Financial Intermediation 2, 255–276.

Catte, P. and C. Mastropasqua, 1993, Financial structure and reforms in central and eastern Europe in the 1980's, Journal of Banking and Finance 17, 785–817.

Chan, Y.-S., 1983, On the positive role of financial intermediation in allocation of venture capital in a market with imperfect information, Journal of Finance 38, 1543–1568.

Checchi, D., 1993, Creation of financial markets in (previously) centrally planned economies, Journal of Banking and Finance 17, 819–847.

Craine, R.B., 1995, Fairly priced deposit insurance and bank charter policy, Working paper (University of California at Berkeley).

Cramton, P. and T. Palfrey, 1990, Cartel enforcement with uncertainty about costs, International Economic Review 31, 17–47.

Dewatripont, M. and E. Maskin, 1990, Contract renegotiation in models of asymmetric information, European Economic Review 34, 311–321.

Diamond, D.W., 1984, Financial intermediation and delegated monitoring, Review of Economic Studies 51, 393–414.

Diamond, D.W., 1991a, Monitoring the reputation: The choice between bank loans and directly placed debt, Journal of Political Economy 99, 689–721.

Diamond, D.W., 1991b, Debt maturity structure and liquidity risk, Quarterly Journal of Economics 106, 709–737.

Diamond, D.W. and Ph.H. Dybvig, 1983, Banks runs, deposit insurance, and liquidity, Journal of Political Economy 91, 401–419.

Frankel, A.B. and J.D. Montgomery, 1991, Financial structure – an international perspective, Bookings Papers on Economic Activity (April), 191–231.

Gale, D., 1992, Information capacity and financial collapse, in: C. Mayer and X. Vives, eds., Capital markets and financial intermediation (Cambridge University Press, Cambridge).

Gorton, G. and G. Pennacchi, 1990, Financial intermediaries and liquidity creation, Journal of Finance 45, 49–71.

Greenbaum, S. and A.V. Thakor, 1995, Contemporary financial intermediation (Dryden Press, Fort Worth, TX).

Grossman, S., 1976, On the efficiency of competitive stock markets where trades have diverse information, Journal of Finance 31, 573–585.

Grossman, S., 1978, Further results on the informational efficiency of competitive stock markets, Journal of Economic Theory 18, 81–101.

Grossman, S.J. and J.E. Stiglitz, 1980, On the impossibility of informationally efficient markets, American Economic Review 94, 543–566.

Grossman, S.J. and O.D. Hart, 1980, Takeover bids, the free-rider problem, and the theory of the corporation, Bell Journal of Economics 11, 42–64.

Harris, M. and A. Raviv, 1991, The theory of capital structure, Journal of Finance 46, 297–355.

Holmstrom, B. and J. Tirole, 1993, Market liquidity and performance monitoring, Journal of Political Economy 101, 678–709.

Houston, J.F. and S. Venkataraman, 1994, Information revelation, lock-in and bank loan commitments, Journal of Financial Intermediation 3, 355–378.

Jaggia, P. and A.V. Thakor, 1994, Firm specific human capital and optimal capital structure, International Economic Review 35, 293–308.

Kanatas, G. and J. Qi, 1993, Should commercial banks be prohibited from underwriting corporate securities?, Working paper (University of South Florida).

Kane, E.J., 1989, Changing incentives facing financial-services regulators, Journal of Financial Services Research 2, 263–272.

Kane, E.J., 1990, Principal-agent problems in S&L salvage, Journal of Finance 45, 755–764.

Katz, M., 1986, An analysis of cooperative research and development, Rand Journal of Economics 17, 527–543.

Kihlstrom, R. and X. Vives, 1996, Collusion by asymmetrically informed firms, Journal of Economics and Management Strategy, Forthcoming.

King, R. and R. Levine, 1992, Financial intermediation and economic development, in: C. Mayer and X. Vives, eds., Capital markets and financial intermediation (Cambridge University Press, Cambridge).

Kroszner, R.S. and R.G. Rajan, 1994a, Is the Glass–Steagall Act justified? A study of the U.S. experience with universal banking before 1933, American Economic Review 84, 810–832.

Kroszner, R.S. and R.G. Rajan, 1994b, The role of firewalls in universal banks: Evidence from commercial bank securities activities before the Glass–Steagall Act, Working paper (University of Chicago).

Manne, H., 1965, Mergers and the market for corporate control, Journal of Political Economy 73, 110–120.

Mayer, C., 1988, New issues in corporate finance, European Economic Review 32, 1167–1182.

Mayer, C. and X. Vives, eds., 1992, Introduction, in: Capital Markets and Financial Intermediation (Cambridge University Press, Cambridge).

Merton, R.C., 1993, Operation and regulation in financial intermediation: A functional perspective, in: P. England, ed., Operation and regulation of financial markets (The Economic Council, Stockholm).

Millon, M. and A.V. Thakor, 1985, Moral hazard and information sharing: A model of financial information gathering agencies, Journal of Finance 40, 1403–1422.

Myers, S.C., 1977, Determinants of corporate borrowing, Journal of Financial Economics 5, 147–175.
Myers, S.C. and N.S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, Journal of Financial Economics 13, 187–221.
Pagano, M., 1993, Financial markets and growth: An overview, European Economic Review 37, 613–622.
Peters, S.R., 1995, Why are demand deposits sequential-service-constrained and partially insured?, Working paper (University of Connecticut).
Peters, S.R. and A.V. Thakor, 1995, A rationale for the functional separation of qualitative asset transformation services in banking, Working paper (Indiana University).
Puri, M., 1994, The long-term default performance of bank underwritten security issues, Journal of Banking and Finance 18, 397–418.
Rajan, R.G., 1991, A theory of the costs and benefits of universal banking, CRSP working paper 346 (University of Chicago).
Rajan, R.G., 1992, Insiders and outsiders: The choice between informed and arm's-length debt, Journal of Finance 47, 1367–1400.
Rajan, R.G., 1993, Conflict of interest and the separation of commercial and investment banking, Working paper (University of Chicago).
Ramakrishnan, R.T.S. and A.V. Thakor, 1984, Information reliability and a theory of financial intermediation, Review of Economic Studies 51, 415–432.
Ramakrishnan, R.T.S. and A.V. Thakor, 1991, Cooperation versus competition in agency, Journal of Law, Economics and Organization 7, 248–283.
Roberts, K., 1985, Cartel behavior and adverse selection, Journal of Industrial Economics 33, 33–45.
Rothschild, M. and J. Stiglitz, 1976, Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, Quarterly Journal of Economics 90, 629–650.
Sabani, L., 1992, Market oriented versus bank oriented financial systems: Incomplete contracts and long term commitments, Unpublished manuscript (Dipartimento di Economia Pubblica, Universita' di Roma La Sapienza and Trinity College, Cambridge).
Sah, R.D. and J.E. Stiglitz, 1968, The architecture of economic systems: Hierarchies and polyarchies, American Economic Review 76, 716–727.
Saunders, A. and I. Walter, 1994, Universal banking in the United States (Oxford University Press, New York).
Shleifer, A. and L. Summers, 1988, Break of trust in hostile takeovers, in: A. Auerback, ed., Corporate takeovers: Causes and consequences (University of Chicago Press, Chicago) 33–56.
Shleifer, A. and R.W. Vishny, 1986, Large shareholders and corporate control, Journal of Political Economy 94, 461–488.
Shockley, R., 1995, Theory and evidence on the leverage implications of bank loan commitments, Journal of Financial Intermediation 4, 272–301.
Simons, H.C., 1934, A positive program for laissez faire, Public Policy Pamphlet No. 15 (University of Chicago).
Simons, H.C., 1935, Rules versus authorities in monetary policy, Journal of Political Economy (February), 1–30.
Stiglitz, J. and A. Weiss, 1981, Credit rationing in markets with imperfect information, American Economic Review 71, 393–410.
Thakor, A.V., 1996, Capital requirements, monetary policy and aggregate bank lending: Theory and evidence, Journal of Finance 51, 279–324.
Thakor, A.V. and R. Callaway, 1983, Costly information production equilibria in the bank credit market with applications to credit rationing, Journal of Financial Quantitative Analysis 18 (June), 229–248.
Thakor, A.V. and P. Wilson, 1995, Capital requirements, loan renegotiation and the borrower's choice of financing source, Journal of Banking and Finance 19, 693–712.

Thakor, A.V., H. Hong and S.I. Greenbaum, 1981, Bank loan commitments and interest rate volatility, Journal of Banking and Finance 5, 497–510.

Vives, X., 1984, Duopoly information equilibrium: Cournot and Bertrand, Journal of Economic Theory 34, 71–94.

Williamson, S., 1987, Costly monitoring, loan contacts, and equilibrium credit rationing, Quarterly Journal of Economics 102, 134–145.

Wilson, P., 1994, Public ownership, delegated project selection and corporate financial policy, Working paper (Indiana University).

Yosha, O., 1995, Information disclosure costs and the choice of financing source, Journal of Financial Intermediation 4, 3–20.