

NBER WORKING PAPER SERIES

FINANCIAL INTERMEDIATION

Gary Gorton
Andrew Winton

Working Paper 8928
<http://www.nber.org/papers/w8928>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2002

Forthcoming in Handbook of the Economics of Finance, edited by George Constantinides, Milt Harris and Rene Stulz (Amsterdam: North Holland). The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

© 2002 by Gary Gorton and Andrew Winton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Financial Intermediation
Gary Gorton and Andrew Winton
NBER Working Paper No. 8928
May 2002
JEL No. G0, G2

ABSTRACT

The savings/investment process in capitalist economies is organized around financial intermediation, making them a central institution of economic growth. Financial intermediaries are firms that borrow from consumer/savers and lend to companies that need resources for investment. In contrast, in capital markets investors contract directly with firms, creating marketable securities. The prices of these securities are observable, while financial intermediaries are opaque. Why do financial intermediaries exist? What are their roles? Are they inherently unstable? Must the government regulate them? Why is financial intermediation so pervasive? How is it changing? In this paper we survey the last fifteen years' of theoretical and empirical research on financial intermediation. We focus on the role of bank-like intermediaries in the savings-investment process. We also investigate the literature on bank instability and the role of the government.

Gary Gorton
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367
and NBER
gorton@wharton.upenn.edu

Andrew Winton
Carlson School of Management
University of Minnesota

I. Introduction

Financial intermediation is a pervasive feature of all of the world's economies. But, as Franklin Allen (2001) observed in his AFA Presidential Address, there is a widespread view that financial intermediaries can be ignored because they have no real effects. They are a veil. They do not affect asset prices or the allocation of resources. As evidence of this view, Allen pointed out that the millennium issue of the *Journal of Finance* contained surveys of asset pricing, continuous time finance, and corporate finance, but did not survey financial intermediation. Here we take the view that the savings-investment process, the workings of capital markets, corporate finance decisions, and consumer portfolio choices cannot be understood without studying financial intermediaries.

Why are financial intermediaries important? One reason is that the overwhelming proportion of every dollar financed externally comes from banks. Table 1, from Mayer (1990), is based on national flow-of-funds data. The numbers are percentages, so in the United States for example, 24.4% of firm investment was financed with bank loans during the 1970 - 1985 period. Bank loans are the predominant source of external funding in all the countries. In none of the countries are capital markets a significant source of financing. Equity markets are insignificant. In other words, if finance department staffing reflected how firms actually finance themselves, roughly 25 percent of the faculty would be researchers in financial intermediation and the rest would study internal capital markets.

As the main source of external funding, banks play important roles in corporate governance, especially during periods of firm distress and bankruptcy. The idea that banks "monitor" firms is one of the central explanations for the role of bank loans in corporate finance. Bank loan covenants can act as trip wires signaling to the bank that it can and should intervene into the affairs of the firm. Unlike bonds, bank loans tend not to be dispersed across many investors. This facilitates intervention and renegotiation of capital structures. Bankers are often on company boards of directors. Banks are also important in producing liquidity by, for example, backing commercial paper with loan commitments or standby letters of credit.

Consumers use bank demand deposits as a medium of exchange, that is, writing checks, using credit cards, holding savings accounts, visiting automatic teller machines, and so on. Demand deposits are securities with special features. They can be denominated in any amount; they can be put to the bank at par (i.e., redeemed at face value) in exchange for currency. These features allow demand deposits to act as a medium of exchange. But, the banking system must then "clear" these obligations. Clearing links the activities of banks in clearinghouses. In addition, the fact that consumers can withdraw their funds at any time has, led to banking panics in some countries, historically, and in many countries more recently.

Banking systems seem fragile. Between 1980 and 1995, thirty-five countries experienced banking crises, periods in which their banking systems essentially stopped functioning and these economies entered recessions. (See Demirgüç-Kunt, Detragiache, and Gupta (2000), and Caprio and Klingebiel (1996).) Because bank loans are the main source of external financing for firms, if the banking system is weakened, there appear to be significant real effects (e.g., see Bernanke (1983), Gibson (1995), Peek and Rosengren (1997, 2000)). The relationship between bank health and business cycles is at the root of widespread government policies concerning bank regulation and supervision, deposit insurance, capital requirements, the lender-of-last-resort role of the central bank, and so on. Clearly, the design of public policies depends on our understanding of the problems with intermediaries. Even without a collapse of the banking system, a credit crunch has sometimes been alleged to occur when banks tighten lending, possible due to their own inability to obtain financing. Also, the transmission mechanism of monetary policy may be through the banking system.

Basically, financial intermediation is the root institution in the savings-investment process. Ignoring it would seem to be done at the risk of irrelevance. So, the viewpoint of this paper is that financial intermediaries are not a veil, but rather the contrary. In this paper, we survey the results of recent academic research on financial intermediation.

In the last fifteen years, researchers have made significant progress in understanding the roles of financial intermediaries. These advances are not only theoretical. Despite a lack of data as rich as stock market prices, significant empirical work on intermediaries has been done. All of this work has contributed to a deeper appreciation of the role of banks in the savings-investment process and corporate finance, of the issues in crises associated with financial intermediation, and of the functioning of government regulation of intermediation. We concentrate on research addressing why bank-like financial intermediaries exist, and the implications for their stability. By bank-like financial intermediaries, we mean firms with the following characteristics:

1. They borrow from one group of agents and lend to another group of agents.
2. The borrowing and lending groups are large, suggesting diversification on each side of the balance sheet.
3. The claims issued to borrowers and to lenders have different state contingent payoffs.

The terms “borrow” and “lend” mean that the contracts involved are debt contracts. So, to be more specific, financial intermediaries lend to large numbers of consumers and firms using debt contracts and they borrow from large numbers of agents using debt contracts as well. A significant portion of the borrowing on the liability side is in the form of demand deposits, securities that have the important property of being a medium of exchange. The goal of intermediation theory is to explain why these financial intermediaries exist, that is, why there are firms with the above characteristics.

Others have cited additional important characteristics of bank-like financial intermediaries, but in our view these seem less important. For example, the maturity of the loan contracts is typically longer than the maturity of the debt on the liability side of the balance sheet, but that is essentially the third point above. Also, Boyd and Prescott (1986) assert that financial intermediaries lend to agents whose information set may be different from their own, in particular, would-be borrowers have private information concerning their own credit risk. Although this suggests a clear role for intermediaries, it is not clear that this is a necessary condition.

Empirical observation is the basis for the statement that intermediaries involve large number of agents on each side of the balance sheet and also for the view that the nature of the securities issued to borrowers and lenders are different. On the liability side of the balance sheet, intermediaries often issue a particular security to households, demand deposits, securities that serve as a medium of exchange. On the asset side of the balance sheet, bank loans are not the same as corporate bonds. Moreover, the structure of the bank loans does not mirror the bank's obligations in the form of deposits. Financial intermediaries with the above characteristics correspond most closely to commercial banks, savings and loans, and similar institutions. But, securitization vehicles and conduits also satisfy the above definition, blurring the distinction between intermediated finance and direct finance, a topic we return to below.

There are a number of issues in studying intermediation that are perhaps unique, compared to other areas of finance. First, there are issues of data. While governments often collect an enormous amount of data about banks, for example, in the U.S. there are the Call Reports that provide a massive amount of accounting information about commercial banks, there is a lack of price data. Thus, unlike other areas of finance, there is an almost embarrassing lack of essential information, prices of loans, of secondary loan sales, and so on. Researchers have been creative in finding data, however, as we discuss below. Other periods of history have also been intensively studied. Apparently, more so than other areas of finance, research in financial intermediation is intimately linked with economic history. In addition, other countries offer rich laboratories as banking systems vary across countries to a significant degree.

Second, in the study of financial intermediation, institutions, regulations, and laws are important. Banking systems have been influenced by laws and regulations for hundreds of years and it is difficult to make progress on many issues without understanding the enormous variation in banking system structures across countries and time, which is due to these laws and regulations. This is most apparent in the variety of industrial organization of banking systems around the world and through history. This variation is just beginning to be exploited by researchers and seems a likely area for further work.

Finally, intermediation is in such a constant state of flux that it is not much of an exaggeration to say that many researchers in financial intermediation do not realize that they are engaged in economic history. It is a challenge to determine whether there are important features of intermediation that remain

constant across time, or whether intermediation is being fundamentally altered by securitization, loan sales, credit derivatives, and other recent innovations.

The paper proceeds as follows. We begin in Section II by discussing evidence on the uniqueness of banks and theories that seek to motivate the existence and structure of these financial intermediaries. Key issues include monitoring or evaluating borrowers, providing consumption smoothing and other types of liquidity, combining lending and liquidity provision as a commitment mechanism, and the coexistence of banks and markets.

In Section III we focus on the specifics of interaction between banks and borrowers. Key issues include the pros and cons of dynamic bank-borrower relationships, the relationship between loan structure and monitoring and between banking sector structure and monitoring, “credit cycles” and capital constraints, and the role of “non-traditional” bank activities such as equity investment.

In Section IV we focus on banking panics and the stability of the banking system. Key issues include evidence on the incidence of banking panics internationally and historically, the causes of panics, the role of bank coalitions in forestalling panics, whether banks are inherently flawed. Section V concerns bank regulation, deposit insurance, and bank capital requirements. Government intervention into banking is a fairly recent phenomenon, but has come to be a widely accepted role because of concerns about moral hazard problems emanating from deposit insurance. The paradigm of moral hazard is reviewed, with particular focus on the empirical evidence. Corporate governance in banks, capital requirements for banks, and other issues are also reviewed.

Finally, in Section V we summarize where all of this research leaves us, both in terms of our present understanding and in terms of directions for the future.

II. The Existence of Financial Intermediaries

The most basic question with regard to financial intermediaries is: why do they exist? This question is related to the theory of the firm because a financial intermediary is a firm, perhaps a special kind of firm, but nevertheless a firm. Organization of economic activity within a firm occurs when that organizational form dominates trade in a market. In the case of the savings-investment process, households with resources to invest could go to capital markets and buy securities issued directly by firms, in which case there is no intermediation. To say the same thing a different way, nonfinancial firms need not borrow from banks; they can approach investors directly in capital markets. Nevertheless, as mentioned in the Introduction, most new external finance to firms does not occur this way. Instead, it occurs through bank-like intermediation, in which households buy securities issued by intermediaries who in turn invest the money by lending it to borrowers. Again, the obligations of firms and the claims ultimately owned by investors are not the same securities; intermediaries transform claims. The existence

of such intermediaries implies that direct contact in capital markets between households and firms is dominated. “Why is this?” is the central question for the theory of intermediation.

Bank-like intermediaries are pervasive, but this may not require much explanation. On the liability side, demand deposits appear to be a unique kind of security, but originally this may have been due to regulation. Today, money market mutual funds may be good substitutes for demand deposits. On the asset side, intermediaries may simply be passive portfolio managers, that is, there may be nothing special about bank loans relative to corporate bonds. This is the view articulated by Fama (1980). Similarly, Black (1975) sees nothing special about bank loans. Therefore, we begin with an overview of the empirical evidence, which suggests that there is indeed something that needs explanation.

A. Empirical Evidence on Bank Uniqueness

What do banks do that cannot be accomplished in the capital markets through direct contracting between investors and firms? There is empirical evidence that banks are special. Some of this evidence also attempts to discriminate between some of the explanations for the existence of financial intermediaries, discussed below.

To determine whether bank assets or liabilities are special relative to alternatives, Fama (1985) and James (1987) examine the incidence of the implicit tax due to reserve requirements. Their argument is as follows. Over time, U.S. banks have been required to hold reserves against various kinds of liabilities. In particular, if banks must hold reserves against the issuance of certificates of deposit (CDs), then for each dollar of CDs issued, the bank can invest less than a dollar. The reserve requirement acts like a tax. Therefore, in the absence of any special service provided by bank assets or bank liabilities, bank CDs should be eliminated by nonbank alternatives. This is because either bank borrowers or bank depositors must bear the tax. Since CDs have not been eliminated, some party involved with the bank is willing to bear the tax. Who is this party? Fama finds no significant difference between the yields on CDs and the yields on commercial paper and bankers acceptances. CD holders do not bear the reserve requirement tax and he therefore concludes that bank loans are special. James revisits the issue and looks at yield changes around changes in reserve requirements and reaches the same conclusion as Fama.

Another kind of evidence comes from event studies of the announcement of loan agreements between firms and banks. Studying a sample of 207 announcements of new agreements and renewals of existing agreements, James (1987) finds a significantly positive announcement effect. This contrasts with non-positive responses to the announcements of other types of securities being issued in capital markets (see James (1987) for the references to the other studies). Mikkelsen and Partch (1986) also look at the abnormal returns around the announcements of different type of security offerings and also find a positive

response to bank loans.¹ Table 2 provides a summary of the basic set of results. There are two main conclusions to be drawn. First, bank loans are the only instance where there is a significant positive abnormal return upon announcement. Second, equity and equity-related instruments have significantly negative abnormal returns. James (1987) concludes, "...banks provide some special service not available from other lenders" (p. 234).

The results of James are quite dramatic and many researchers followed up on them. Lummer and McConnell (1989) distinguish between new bank loan agreements and revisions to agreements already in place. Further, they classify announcements concerning existing agreements into announcements containing positive information and those containing negative information. This classification is based on whether the terms of the agreement (maturity, interest rate, dollar value, covenants) are revised favorably or unfavorably (some have both favorable revisions in some dimensions and unfavorable revision in others). They find no abnormal return to announcement of new agreements. Favorable renewals have significantly positive abnormal returns, while negative renewals have significantly negative abnormal returns. The strongest negative response comes when the bank initiates a loan cancellation. The strongest positive response is associated with loan renewals where there was previously public information suggesting the loan was in trouble. The results of Lummer and McConnell suggest that the bank is not producing information upon first contact with a borrower. Rather, the bank either learns information later or takes action later, and this is revealed when a loan is renewed or restructured. The results are consistent with the view that a continuing relationship with a bank can signal changes in value to capital markets.

Best and Zhang (1993) confirm Lummer and McConnell (1989). But, with a revised definition of "new" loans, Billet, Flannery, and James (1995) find no significant differences between initiation of loans and loan renewals. Slovin, Johnson, and Glascock (1992) and Hadlock and James (2000) also find no differences.

Slovin, Sushka, and Polonchek (1993) look at an interesting implication of the result that bank loans are somehow different than other securities. If loans are special, in some sense, then when a borrower's bank fails, does that adversely affect that borrower? To address this they examine share price responses of bank borrowers' shares upon the announcement of the failure of their bank, Continental Illinois. If banks are simply passive investors, and their loans are indistinguishable from bonds, then when there is a bank failure, borrowers simply go elsewhere to borrow funds. However, if there is a "customer relationship," then banks acquire private information about their borrowers and the bank's

¹ Slovin, Sushka, and Hudson (1988) find significantly positive announcement abnormal returns associated with the announcement of standby letters of credit. Preece and Mullineaux (1989) find that the reaction to loan agreements with insurance companies is similar to that for bank loan agreements. Also, see Mullineaux and Preece (1996).

failure would mean that this intangible asset is destroyed, causing borrowers losses. Slovin, Sushka, and Polonchek (1993) find that Continental Illinois borrowers incurred significantly negative abnormal returns (- 4.2% annually) during the bank's impending failure. This evidence is consistent with bank relationships being important, an issue discussed further below. Bernanke (1983) essentially argues that crisis in the U.S. banking system during the Great Depression can be viewed in the same way, causing real adverse effects for borrowers. Gibson (1995) studying the effects of the health of Japanese banks finds that investment is thirty percent lower by firms that have a Japanese bank that is weak.

Another area in which banks appear to be different from bondholders' concerns reorganization of firms in financial distress, though this depends on the characteristics of the particular sample studied. Gilson, John, and Lang (1990) find that the likelihood of a successful debt restructuring by a firm in distress is positively related to the extent of that firm's reliance on bank borrowing. The interpretation is that it is easier to renegotiate with a single bank, or small number of banks, than it is with a large number of dispersed bondholders, in which case there are free rider problems. However, Asquith, Gertner, and Scharfstein (1994), and James (1995), find that for firms with public debt outstanding, banks rarely make unilateral concessions to distressed firms. Franks and Torous (1994) study 45 distressed exchanges and 37 Chapter 11 reorganizations during the period 1983 to 1988. Unlike Gilson, John, and Lang (1990), Franks and Torous find that firms that successfully complete exchange offers do not owe significantly more of their long-term debt to banks. Franks and Torous' firms all have publicly traded debt and tend to be larger than the firms in the Gilson, Lang, and John sample. James (1996) partially reconciles some of these conflicting results. He finds that the higher the proportion of total debt held by the bank, the higher the likelihood the bank debt will be impaired, and so the higher the likelihood it participates in the restructuring. Banks do not act unilaterally when the firm has significant public debt outstanding because banks, as senior lenders, would be transferring wealth to the public debt holders in these cases.

In other countries, banks interact with borrowers in different ways than in the United States. Such examples offer another type of evidence on the ability of banks to provide valuable services that cannot be replicated in capital markets. Hoshi, Kashyap, and Scharfstein (1990a, b, 1991) find that firms in Japan in keiretsu, that is, firms with close ties to banks, are less liquidity constrained compared to firms without such ties. Also, firms with close ties are able to invest more when they are financially distressed, suggesting the importance of a bank relationship. In Germany, Gorton and Schmid (1999) find that bank equity ownership improves the performance of firms. Also, see Fohlin (1998). We review more evidence on "bank relationships" in Section III below.

We conclude that financial intermediaries are producing services that are not easily replicated in capital markets. We turn now to the major theories that have been put forth as explanations for the existence of financial intermediation. These theories are not mutually exclusive.

B. Banks as Delegated Monitors

Diamond (1984) offered the first coherent explanation for the existence of financial intermediaries.² Diamond's intermediaries "monitor" borrowers. Since monitoring is costly, it is efficient to delegate the task to a specialized agent, the bank. The notion of monitoring borrowers has become an influential idea, which subsequent researchers have further developed.

Not only do Diamond's intermediaries contain most of the important elements of a theory of intermediation, discussed above, but he also identifies and solves a fundamental problem at the root of intermediation theory. That problem concerns the fact that whatever problem the intermediary solves to add value with respect to borrowers would seem to imply that lenders to the intermediary would face the same problem with respect to their lending to the intermediary. In Diamond (1984), the intermediary "monitors" borrowers on behalf of investors who lend to the intermediary. But, then it would appear that the lenders to the intermediary have to "monitor" the intermediary itself. How is this problem, which has come to be known as the "monitoring the monitor" problem, solved? Diamond (1984) was the first to recognize and then solve this problem.

In Diamond (1984) borrowers must be "monitored" because there is an ex post information asymmetry in that lenders do not know how much the firm has produced. Only the individual borrower observes the realized output of his project, so contracts cannot be made contingent on the output. Consequently, a lender is at a disadvantage because the borrower will not honor ex ante promises to pay unless there is an incentive to do so. The first possibility Diamond considers to solve this contracting problem is the possibility of relying on a contract that imposes nonpecuniary penalties on the borrower if his payment is not at least a certain minimum. This contract is costly because such penalties are imposed in equilibrium, reducing the utility of borrowers. If, instead, the lender had available an information production technology, then the information asymmetry could be overcome by application of this technology, at a cost. Perhaps this would be cheaper, and hence more efficient, than imposing nonpecuniary penalties. Diamond termed production of information about the borrower's realized output, at a cost, "monitoring."

The notion of "monitoring" in Diamond (1984) appears inspired by Townsend (1979), but there is a critical difference. In Townsend the lender must bear a cost to determine whether the borrower has the resources to repay the loan or not, a decision made after the borrower's project output has been realized and after a payment has been offered to the lender. That is, in Townsend, the decision by a lender to

² For reasons of space, we do not survey the previous transaction cost-based literature. For surveys of this literature, see Benston (1976) and Baltensperger (1980).

monitor a borrower is made *after* the entrepreneur has made a payment to the lender; it is contingent on the amount of the payment. Hence, it is known as “costly state verification.” In Diamond, however, monitoring is not state contingent and the cost must always be borne because, in Diamond, the monitoring cost must be incurred *before* the output realization of the borrower’s project is known to anyone.

This difference between Townsend and Diamond, with respect to monitoring, leads to another difference. In Townsend, the costly state verification problem motivates the form of the contract between a borrower and lender: it is a debt contract (since random monitoring is assumed away; see Boyd and Smith (1994)). In Diamond, the optimal contract between the borrower and the lender is a debt contract in the absence of monitoring, but once monitoring is introduced, the optimal contract is undetermined. It is feasible for the contract to be an equity contract, for example. On the one hand, this does not matter for Diamond’s basic argument, but, on the other hand, it seems potentially important for understanding why agents trading in markets cannot replicate the function of the intermediary, as we discuss further below.

The monitoring solution may dominate the contract that imposes nonpecuniary penalties, but it raises another problem. If a single borrower has many lenders, then each lender will have to bear the cost of monitoring, which in turn will lead to duplication of monitoring costs or free riding problems among individual lenders. This raises the prospect of a third solution. If the task of monitoring were delegated to a single agent, free riding and duplication of monitoring costs problems could potentially be eliminated. But if the lenders were to delegate the task of monitoring, then the same problem would still exist, but at one step removed. That is, the individual lenders would then face the task of monitoring the agent delegated to monitor the borrower(s). This is the problem of “monitoring the monitor.” Diamond (1984) presents the first coherent theory of banking that solves the problem of monitoring the monitor.

To be more precise, the problem of “monitoring the monitor” is this: lenders to the intermediary can reduce monitoring costs if the costs of monitoring the intermediary are lower than the costs of lenders lending directly to borrowers and directly incurring the monitoring costs. Diamond’s fundamental result is to show that as an intermediary grows large, it can commit to a payment to depositors that can only be honored if, in fact, the intermediary has monitored as it promised. If not, then the intermediary incurs nonpecuniary penalties, interpreted by Diamond as bankruptcy costs or loss of reputation.

To see the argument, we follow Williamson’s (1986) presentation of the Diamond result; unlike Diamond, it does not rely on precise contractual specification of nonpecuniary penalties, which is rarely seen in practice. Williamson’s monitoring technology follows Townsend, so Diamond’s result does not depend on the timing of monitoring (that is, whether it is state contingent or not). A brief outline of the essential part of the Williamson model is as follows. Borrowers need resources to invest in their projects. They invest K units of endowment at date 0 and receive $K\tilde{w}$ at date 1, where \tilde{w} is a random variable distributed according to the density $f(w)$. As shown by Gale and Hellwig (1985), the optimal contract

between the borrower and a lender is a debt contract. At date 1 borrower j has a realized return of w_j per unit invested. Borrower j pays the lending intermediary a gross rate of return \bar{R} in a state, w_j , where there is no monitoring and $R(w_j)$ when there is monitoring. Define the set $B = \{w_j: R(w_j) < \bar{R}\}$ and $B^c = \{w_j: R(w_j) \leq \bar{R}\}$. Finally, let r denote the certain market return, required by risk neutral investors.

When the intermediary has m borrowers, each investing K , then the total return to the intermediary (before compensating depositors) is:

$$\pi_m = K \sum_{j=1}^m \min\{R(w_j), \bar{R}\}.$$

By the strong law of large numbers:

$$\lim_{m \rightarrow \infty} \frac{1}{mK} \pi_m = \int_B R(w_j) dw_j + \int_{B^c} f(w_j) dw_j.$$

Consequently, since the intermediary's return must be at least the market return, r , if the following inequality holds:

$$\int_B R(w_j) f(w_j) dw_j + \bar{R} \int_{B^c} f(w_j) dw_j - \left(\frac{c}{K}\right) \int_B f(w_j) dw_j \geq r,$$

then, as the intermediary grows large, it can guarantee a certain return of r to its depositors.

If the intermediary is finite sized, that is, it lends to a finite number of borrowers, then depositors must monitor the intermediary to ensure that the intermediary, in turn, is monitoring the borrowers. Since monitoring is costly, and given the certain market return that must be obtained, the depositors must be compensated for these monitoring costs by the intermediary. Compensating the depositors for monitoring costs incurred, lowers the profitability (utility) of the intermediary. However, the central result of Diamond (1984) applies here, namely, that the depositors need not monitor an infinitely large intermediary because such a firm can achieve r with probability one. In the limit, depositors do not need to monitor the intermediary. The “monitoring the monitor” problem is solved by diversification.

One might object that, in practice, financial intermediaries are not infinitely diversified, and some credit risk is not diversifiable; also, it seems likely that a depositor finds it more difficult to monitor a large bank than to monitor a small bank. Krasa and Villamil (1992a,b) address these concerns. Suppose we modify Williamson (1986) by assuming that larger banks' returns are more costly to verify. If loan returns are stochastically independent of one another, Krasa and Villamil (1992a) apply the Large Deviation Principle to show that, so long as a depositor's cost of monitoring doesn't increase exponentially with bank size, the expected costs of monitoring a sufficiently large bank go to zero. Moreover, they show through examples that even relatively small banks (e.g., 32 loans) get enough gains from diversification to dominate direct lending. If some loan risk is systematic, the chance of bank failure

is bounded away from zero as bank size grows (Krasa and Villamil, 1992b). In this case, since the cost of monitoring banks that fail is increasing in bank size, there is a bank size past which the increase in monitoring costs dominates marginal benefits from additional diversification. Moreover, this optimal size diminishes as the systematic component of loan risk increases.

Winton (1995a) addresses another issue, namely the role of bank capital. Suppose that the banker invests his own funds in the bank as “inside” equity capital. Being junior, such equity absorbs losses first, reducing the probability with which the bank defaults and depositors must monitor. Thus, bank capital is another mechanism for implementing delegated monitoring. Since the bankers’ capital is fixed, it will be most helpful for smaller banks; also, the relative importance of capital versus diversification increases as more loan risk is systematic.³

Of course, Diamond (1984) does not explain all the characteristics of intermediaries. But, he elegantly explains the existence of intermediaries, in particular, as coalitions, of borrowers and lenders, which dominate the alternative of direct investment by investors in securities issued by firms. The securities market fails in the sense that intermediation, centralization of the task of monitoring, is a lower cost solution to the ex post information asymmetry between borrowers and lenders. Diversification is critical to intermediation providing a lower cost solution because diversification is critical to reducing the monitoring the monitor problem. The textbook idea that individual investors can diversify nonsystematic risk on their own does not take into account the role diversification plays in allowing an intermediary to be monitored costlessly (in the limit).

Other papers that study banks as delegated monitors include Gorton and Haubrich (1987) and Seward (1990).

C. Banks as Information Producers

If information about investment opportunities is not free, then economic agents may find it worthwhile to produce such information. There will be an inefficient duplication of information production costs if multiple agents choose to produce the same information. Alternatively, a smaller number of agents could produce the information, becoming informed, and then sell the information to the uninformed agents. This, however, introduces the “reliability problem” originally identified by Hirshleifer (1971): it may be impossible for the information producer to credibly ensure that he has, in fact, produced the valuable information.

³ Winton (1995b) shows that further reductions in monitoring costs are possible if a class of “outside” equity holders is created, who are junior to depositors but senior to the banker.

A related problem concerns resale of the information. If an information producer could credibly produce valuable information, and then sell it to another agent, then there is no way to prevent the second agent from selling it to a third agent, and so on. In other words, purchasers of the information can sell or share the information with others without necessarily diminishing its usefulness to themselves. This is known as the “appropriability problem.” The returns to producing the information could not all be captured by the information producer, possibly making the production of information uneconomic (see Grossman and Stiglitz (1980)). The resale and appropriability problems in information production can motivate the existence of an intermediary.

Leland and Pyle (1977) were the first to suggest that an intermediary could overcome the reliability problem. The intermediary can credibly produce information by investing its wealth in assets about which it claims to have produced valuable information. The starting point for Leland and Pyle (1977) is a single entrepreneur who has private information about an investment opportunity, but who has insufficient resources to undertake the investment. Since outside investors do not observe the entrepreneur’s private information, there is an adverse selection problem. Leland and Pyle show that the entrepreneur’s private information can be signaled by the fraction of equity in the project that the entrepreneur retains, while he sells the remaining fraction to outside investors.⁴ At the end of their paper, Leland and Pyle suggest that financial intermediaries might efficiently solve the reliability and appropriability problems inherent in information production by issuing securities and using the proceeds to invest in a portfolio of securities about which the intermediary has become privately informed. After deriving his delegated monitoring model, Diamond (1984) also derives a Leland and Pyle model in which diversification lowers the intermediary’s signaling costs compared to the entrepreneur’s costs.

Following Leland and Pyle, a number of papers, notably Campbell and Kracaw (1980), also argued that financial intermediaries might exist to produce information about potential investments, information that could not be efficiently produced in securities markets. Campbell and Kracaw (1980) show that appropriability and reliability problems can be eliminated if the information producer has a sufficient minimum amount of wealth to risk if he does not produce the information. To risk his own money requires that the intermediary actually invest on behalf of other agents. The paper, however, that most fully articulates the argument that coalitions of agents should form to produce information *ex ante* about potential investments is Boyd and Prescott (1986).

The underlying problem faced by agents in Boyd and Prescott (1986) is an information asymmetry that occurs prior to contracting and investing, resulting in an adverse selection problem. Agents are of different types and this information is private to each agent. Each agent, however, is

⁴ See also Kihlstrom and Mathews (1990) and Duffie and Demarzo (1999).

endowed with a technology to evaluate projects, that is, the technology can determine agent type. Ex ante information production can alleviate the adverse selection problem. This can be done in a market context, where an agent evaluates his own project, and then issues securities to investors that promise specified returns. Or, a coalition of agents can offer investors a claim on group returns. Financial intermediaries are coalitions of agents that evaluate projects, invest in those determined to be high-value projects, and share the returns from the portfolio of projects.

More specifically, the outline of the model is as follows. Agents live for two periods. Each agent is endowed with a project of unknown type (good or bad). Agents know their own type, so there is no opportunity to enter into contracts before knowing their types. Each project type can have a high or low return (good projects are more likely to realize the high return). An agent can expend his endowment either on producing information about a single project's type or as an investment in a single project, his own if he has not evaluated it or another agent's project. If a project is evaluated, then a noisy signal of true project type is received. Project evaluation and investment are publicly observable and verifiable, as are project returns, evaluation results, consumption outcomes, and contract terms.

An efficient outcome invests in as many good projects as possible. But, the difficulty in accomplishing this is that bad-type agents will want to mimic good-type agents, claiming that they are good, promising the same high return to investors as the good-type agents, and then hoping that their project realizes the high return. Indeed, there is such a securities market equilibrium, but it is one in which some bad-type projects are evaluated, by mimicking agents. This is inefficient.

The alternative is the financial intermediary coalition. The model is one of mechanism design. One interpretation of how to implement the equilibrium with the coalition (given by Boyd and Prescott) is as follows. Coalition members deliver their endowments to the coalition prior to investment. These endowments are used for project evaluation. Depositors are other agents who turn over their endowments to the coalition in exchange for a promised amount of consumption. The depositors give the coalition the right to invest in their project and to receive the entire project output, if the coalition desires. Project owners are promised very high returns if evaluation reveals a good project and if the realized return is high. Otherwise, depositors are promised an amount of consumption which is more than a bad-type agent could achieve on his own, but less than the promised amount for projects with a good evaluation and high realized returns. Members of the coalition are residual claimants and share profits equally.

The coalition's sharing rules induce truthful revelation of agent type. The coalition then evaluates good-type projects and funds each of these projects with a good evaluation. It uses the remaining proceeds to fund bad-type projects without evaluation. This is the critical point. The promised returns separate types, and since good types are relatively scarce, the coalition ends up funding some bad-

type projects, but it does not waste resources evaluating those projects. This is why it dominates the securities market.

The intermediary dominates the securities market because the intermediary coalition can induce agents to truthfully reveal their type and this cannot be achieved in the securities market. Truthful revelation allows the coalition to avoid inefficiently evaluating some bad-type projects. The reason is that, by conditioning returns on the coalition's portfolio returns, rather than on the returns of a single project, the coalition can offer higher returns to bad-type agents, so they will participate in the coalition. The relative proportions of good-types and bad-types are also important. In particular, good-type agents must be scarce. Note also that it is important that a coalition be large because a small coalition may end up with so many good-type projects that they cannot all be funded. In the population, good-type projects are relatively scarce and this must be reflected in coalition membership. Thus, as in Diamond (1984), size of the coalition is critical for the argument.

The equilibrium concept in Boyd and Prescott is based on the core of an economy. That is, an allocation is an equilibrium if no large coalition of agents, with specified fractions of agent types, can achieve a different allocation, satisfying resource, consumption, incentive and other constraints, and make at least some agent type better off without reducing any other type's utility. Deviating coalitions are not allowed to attract higher than population proportions of type- i agents unless it makes them strictly better off. Although the solution of the model is standard in that it relies on the revelation principal, the equilibrium concept is less common in the finance and financial contracting literature. This may account for why this paper has not led to a successor literature in banking per se; instead, it has been more influential in macroeconomics, where the equilibrium concept has been taken up, though see the discussion of Williamson (1988), below.

Boyd and Prescott's intermediary has the characteristics of bank-like intermediaries identified in the introduction. Other researchers have pursued solutions to the problems of reliability and appropriability of valuable private information, but these other solutions do not involve bank-like intermediation. Two settings in particular have been examined. The first considers delegated portfolio management, i.e., a setting where a fund manager may claim to have superior information or superior ability and offers to invest on behalf of investors. The second considers the sale of valuable information about investments when the information producer does not invest on behalf of investors. A theory of intermediation must distinguish between firms that sell information, like rating agencies, firms that are delegated portfolio managers, like mutual funds or hedge funds, and bank-like financial intermediaries.

At the level of casual empiricism there are identifiable differences between these types of arrangements. A bank-like intermediary does not sell information that it produces. Rather, as in Boyd and Prescott, it uses the information internally to improve the returns to coalition members. This is very

different from the case of a firm that sells information to investors, like a rating agency. Firms selling information face problems of reliability and appropriability, but they do not lend money. Purchasers of the information may lend in reliance on the information purchased, but they are then directly lending, not via an intermediary. A portfolio manager, claiming to have superior information, accepts investments from one set of agents and then uses the proceeds to invest in securities. This seems very similar to a bank-like intermediary. One difference is that the claims held by the investors do not have different state contingent payoffs than the payoff on the portfolio of claims chosen by the portfolio manager; essentially, the investors and the portfolio manager all hold equity claims in the portfolio.

In Bhattacharya and Pfleiderer's (1985) model, investors want to hire portfolio managers, but there are two sources of private information that make this difficult. First, investors must hire a manager from pool of managers with heterogeneous abilities. A manager or agent has the ability to receive an informative signal about the risky asset (there is also a riskless asset). Once a manager has been hired, he must be induced to truthfully reveal the signal he has received. However, once the principal has designed the contract and hired a manager, the manager/agent's only role is to transmit the information to the principal. There is no portfolio management by the manager/agent since the principal can directly invest using the information supplied by the manager/agent. There is no intermediary (nor do Bhattacharya and Pfleiderer claim that there is; the purposes of their paper are different).

Allen (1990) presents a model that distinguishes conditions under which information is sold to agents who then use the information to make investments from the case where the buyers of the information then act as intermediaries and resell the information. Essentially, reselling the information allows more of the value of the information to be captured. Because the initial information seller must distinguish himself from potential uniformed mimics, he faces a number of constraints. These constraints limit the amount of profit he can take in from selling the information. This is the basis for information resellers to enter the market; they find it profitable to resell the information rather than use it as a basis for their own investments because they can capture more of the value of the information. Here there is a type of intermediation: there are agents who buy information and then resell it. But, these agents do not invest on behalf of others.

Ramakrishnan and Thakor (1984) consider a setting in which firms issuing new shares to the public can hire an agent to produce information about their quality. Information production requires a costly, and unobservable, effort, so the information producer would like to avoid this cost if he can do so without being detected. There is an ex post noisy indicator of the information producer's effort choice, so compensation for information production can be linked to this indicator. Because information producers are risk averse, they would prefer to avoid the risk that the noise in the indicator prevents them from obtaining compensation for their efforts. The main point of Ramakrishnan and Thakor is that this risk is

mitigated if one infinitely large intermediary is formed since this diversifies the risk associated with the effort indicator. The large intermediary is formed when information producers can costlessly monitor each other's efforts. (Millon and Thakor (1985) extend the analysis to the case where the internal monitoring is costly.) Ramakrishnan and Thakor's intermediary, however, does not accept funds for investment. Rather, it is a pure information seller. In this regard, also see Lizzeri (1999).

In general, the differences in settings where some agents would like valuable, but costly, information produced for investment purposes are subtle. In many models there is no need for the information seller to actually accept the funds that will be invested on the basis of the superior information. In Bhattacharya and Pfleiderer, Allen, and Ramakrishnan and Thakor, the information producer sells the information to investors, but does not need to actually invest the funds of the investors. In Boyd and Prescott the intermediary accepts deposits, produces information, and invests in projects based on the information produced. Only by conditioning the returns on the portfolio that is produced by the coalition can truthful revelation be induced.

A potentially important aspect of information production by banks concerns whether the information is produced upon first contact with the borrower or is instead learned through repeated interaction with the borrower over time. Another strand of the literature on banks as information producers argues that banks acquire (private) information over time through repeatedly lending to a borrower. The acquisition of this private information over time is known as a "customer relationship" and is discussed in Section III below.

D. Banks as Consumption Smoothers

Bryant (1980) and Diamond and Dybvig (1983) develop a role for bank liabilities, without stressing any particular features of bank assets. Bank liabilities do not function as a transactions medium. Rather, banks are vehicles for consumption smoothing; they offer insurance against shocks to a consumer's consumption path.

The Diamond and Dybvig model assumes that the payoffs from the available investment opportunities are inconsistent with the possible consumption paths desired by consumers. In particular, consumers have random consumption needs, and satisfying these needs may require them to prematurely end investments unless they save via intermediation so that they can to some extent diversify these consumption shocks. The model offers a view of the liability side of banking; the right to withdraw from the bank, prematurely ending investment in order to satisfy sudden consumption needs, corresponds with notions of how demand deposits actually work. The model also focuses on banking panics, a separate topic that we discuss in Section IV below.

The outlines of the Diamond and Dybvig model are as follows. There are three dates 0, 1, and 2 and a single good. The available technology allows one unit of investment to be transformed over two periods into $R > 1$ units at the final date. If this investment is interrupted at the interim date, then it just returns the initial one unit. Importantly, the long-term investment only realizes a return over the initial investment if it reaches fruition at date 2. All consumers are identical initially, at date 0, but each faces a privately observable, uninsurable risk with regard to their preferences. At date 1, each consumer learns whether he cares only about consumption at date 1, an “early consumer,” or only about consumption at date 2, a “late consumer.” The problem is evident: consumers would like to insure themselves against the bad luck of being an early consumer. Without being able to write such insurance contracts, because consumer type is not observable, early consumers can do no better than consuming their single unit of endowment, which was invested in the investment technology but which is liquidated early. The lucky late consumers consume $R > 1$.

Diamond and Dybvig (1983) argue that a bank can provide insurance against the risk of being an early consumer. Basically, a bank works as follows. At date 0 the bank opens and accepts “deposits” of endowment. The bank promises a fixed claim of r_1 per unit deposited will be paid out to consumers who withdraw at date 1. The return on a deposit that is not withdrawn at date 1, but is withdrawn at date 2, depends on how much was withdrawn at date 1. Suppose the fraction of consumers who will turn out to be early consumers is fixed and known. Then Diamond and Dybvig show that the return of r_1 can be set to the amount that an early consumer would achieve if there were complete insurance markets. So the bank can support the full-information risk-sharing equilibrium.

The Diamond and Dybvig model has important features of intermediaries and the real world environment. First, it incorporates the idea that consumers have uncertain preferences for expenditure streams, producing a demand for liquid assets. Furthermore, the modeling representation of this uncertainty, the technique of early and late consumers, has been very influential in its own right. Uncertainty about preferences for expenditure streams leads to the bank offering claims that look like demand deposits. This is combined with a second important feature, namely, real investment projects are irreversible, or at least costly to restart once stopped. A third important feature of the model is the idea that individual consumers have private information about the realization of their type, the realization of their preferred consumption stream. There is no credible way to truthfully reveal this information.

We now turn to some details about why insurance or securities markets cannot provide consumption smoothing or insurance against the risk of uncertain preferences for expenditure streams. In Diamond and Dybvig, an intermediary that issues demand deposits allows greater risk sharing than autarky. Diamond and Dybvig assume that demand deposits cannot be traded and do not consider other securities markets. Their model assumes a sequential service constraint, that is, a first-come-first-served

ruled under which at date 1 the bank honors claims to withdraw in the order in which they are received until the bank runs out of resources to honor the claims. The remaining consumers seeking to withdraw receive nothing and the bank fails. We discuss models that motivate the sequential service constraint in Section IV below. Here, we simply note that the idea it attempts to capture is that consumers cannot coordinate to go to any securities market at the same time to trade; they are busy doing other things such as shopping, eating, sleeping, working, etc. Thus, Diamond and Dybvig's assumptions that demand deposits cannot be traded and that no other securities markets are open are not completely without foundation.

This point is important because Jacklin (1987) and Haubrich and King (1990) argue that the existence of Diamond and Dybvig intermediaries requires the restriction that consumers only have nontraded demand deposits available to them. Jacklin (1987) begins by asking, why does a securities market fail in the Diamond and Dybvig model? In order to highlight the importance of trading restrictions and preferences for Diamond and Dybvig's result that intermediation is the best insurance arrangement, he proposes an alternative arrangement that uses traded securities. Suppose that there are firms in Diamond and Dybvig that own the two-period production technology. Each firm raises capital by issuing dividend-paying shares at date 0. Consumers buy the shares, entitling them to set the production policy and to set dividend policy about the amount paid out to share owners at date 1. The "dividends" on Jacklin's equity are set to smooth income in exactly the desired way; they are not just pass-throughs from the firms. Shareholders of record at date 0 receive the dividend at date 1 and then can sell the share in a share market at date 1. At date 1 consumers learn their preferred consumption streams. Early consumers will want sell their shares ex dividend to late consumers. Jacklin shows that the social optimum obtains with this share market in place. Thus, the bank cannot do any better.

Jacklin goes on to show that the result that the intermediary cannot improve upon trading dividend-paying shares is not true in general. Recall that, in Diamond and Dybvig, some consumers find that they must consume early; it is all or nothing. If instead preferences are smooth, so that one type of consumer will learn that he has a stronger preference for earlier consumption than the other type, then it can happen that demand deposits dominate traded equity shares, but only under certain conditions. Furthermore, if demand deposits can be traded, then optimal risk sharing does not occur regardless of preferences. Finally, Jacklin argues that, if new assets can be introduced, individuals will deviate from either the demand deposit arrangement or from the economy with traded dividend-paying shares. These points lead Jacklin to conclude that the Diamond and Dybvig "demand-deposit" intermediary can only exist if trading restrictions limit consumers to the type of demand deposits that Diamond and Dybvig model. This highlights the importance of the sequential service constraint and its interpretation.

Haubrich and King (1990) revisit in detail the issue of financial intermediation in settings where agents are subject to privately observable income shocks. Their main conclusions are similar to Jacklin (1987), namely, that “demand deposits *uniquely* provide insurance only if there are restrictions on financial side exchanges, which may be interpreted as exclusivity provisions or regulations on security markets. If these restrictions cannot be implemented, then our environment does not rationalize banks” (p. 362; emphasis in original). They also make the useful distinction between two separate issues. One is the fact that the available investment technology is illiquid in the sense that no return is earned if the two-period investment is ended early. The other is that risk averse consumers with privately observable income shocks have a demand for insurance. They argue that a securities market is as good as banks in providing liquidity. In their model, the bank’s comparative advantage is in providing insurance against private income shocks rather than providing liquidity per se, but that advantage still depends on trading restrictions.

Hellwig (1994) and von Thadden (1998) examine how banks function when additional considerations are introduced into Diamond and Dybvig’s structure. Hellwig shows that if market returns at the interim (early-consumption) date are subject to systematic “interest-rate shocks,” banks optimally do not provide insurance against such interest rate risk. Von Thadden shows that if depositors can join outside coalitions that engage in market activity, banks’ ability to provide insurance is severely curtailed, and banks are more constrained as long-term investment opportunities are more reversible. Intuitively, ex ante insurance makes the return to holding deposits at the interim date deviate from returns available by directly investing, allowing arbitrage.

Diamond (1997) responds to Jacklin (1987), Haubrich and King (1990), Hellwig (1994), and von Thadden (1998) in a model with both banks and a securities market in which (by assumption) only a limited subset of agents participate in the market. The main focus of the paper is on the interactions between bank provision of “liquidity” and the depth of the market. As more agents participate in the securities market, banks are less able to provide additional liquidity.

Allen and Gale (1997) introduce a different smoothing role for financial intermediaries, namely, that they are unique in providing a mechanism for smoothing intertemporal intergenerational risks. Allen and Gale study the standard overlapping (risk averse) generations model with two assets, a risky asset in fixed supply and a safe asset that can be accumulated over time.⁵ The risky asset lasts forever and pays out a random dividend each period. The safe asset consists of a storage technology. First, consider the market equilibrium in this economy. Perhaps counterintuitively, the safe asset is not a useful hedge

⁵ Freeman (1988) and Qi (1994) introduce Diamond and Dybvig banks into an overlapping generations model, but do not consider intertemporal smoothing of risk.

against the uncertainty generated by the risky asset. Because the risky asset's returns are independently and identically distributed in any period, a representative young agent solves the same decision problem at any date. Old agents supply the risky asset inelastically, so the equilibrium price of the risky asset is constant and nonstochastic. Because the dividend is nonnegative, the safe asset is dominated and is not held in equilibrium.

This market equilibrium is in contrast with the portfolio allocation that would occur for an infinitely lived agent facing the same investment opportunities. Such an individual can self-insure against low dividend periods by holding a buffer stock of precautionary savings in the form of the safe asset. Intuitively, when the dividend is high, the individual saves some of the dividend for a "rainy day" when the dividend is low. In the overlapping generations setting, a social planner can make a Pareto improvement by following the same type of rule.

The market equilibrium in the overlapping generations model cannot achieve the allocation that the social planner could achieve because private agents cannot trade before they are born, while the social planner can, in effect, trade at all dates. In particular, the social planner trades *ex ante*, that is, before the realization of the path of dividends. A representative young agent, however, is born into a world where the dividend has just been realized. There is no willingness to implement insurance once the state is known. For example, suppose the dividend just realized is low. Then, the social planner would like to implement a transfer from the young to the old, to smooth their income. On the other hand, if the dividend just realized is high, then the social planner would like to transfer some of that to the current young. Some excess may be saved for the next period. These transfers insure that each generation receives the expected utility targeted by the social planner.

It is well known that markets are incomplete in overlapping generations models, but the point made by Allen and Gale is that a long-lived financial intermediary may be the institutional mechanism to provide for this intertemporal smoothing. The intermediary would hold all the assets and offer a deposit contract to each generation. After accumulating large reserves, the intermediary offers (almost) all generations a constant return on deposits, independent of the actual dividend realizations. How such an institution would be set up initially, and how it would be maintained when some agents will have incentives to renege on the arrangement, are not clear. Allen and Gale loosely interpret the institution as corresponding to German universal banks.

Consumption insurance that implements smooth patterns of intertemporal consumption plans is at the center of the model of consumer behavior of neoclassical economics. Another central notion concerns the use of "money" to facilitate exchange. The search models of money or models with cash-in-advance constraints attempt to explain why "money" exists. The notion of banks as consumption-smoothing institutions attempts to wed these two ideas. Bank liabilities are seen as claims that facilitate

consumption smoothing. But, there is no notion of exchange in the model, no sense in which transactions are taking place where bank “money” is being used to facilitate the smoothing. Instead, agents are essentially isolated from each other; there is no trade with other agents where “money” buys goods. Rather agents fear missing out on long-term investment opportunities because of possible shocks to their preferences. Agents trade only with the bank.

E. Banks as Liquidity Providers

Bank liabilities function as a medium of exchange. This basic observation leads to ideas and models concerning “liquidity” that are quite distinct and perhaps more natural than viewing bank liabilities as allowing consumption smoothing. A medium of exchange is a set of claims or securities that can be offered to other agents in exchange for goods. Such claims can dominate barter and may dominate government-supplied money. What are the advantages of privately-produced trading claims to be a medium of exchange? One class of these models considers settings where agents cannot contract and trade with each other due their inability to meet at a single location. Without “money” they must barter, and this is clearly inefficient. This generates a need for a payments system, essentially a trading center or bank that can produce and net claims. A second notion of liquidity is related to the information properties of claims that are privately produced as a medium of exchange. The focus is on reducing trading losses that agents who need to consume face when other traders with private information seek to use this information to make trading profits. Yet a third notion of liquidity uses a setting where moral hazard problems limit firms’ ability to borrow to meet unexpected investment needs. Because moral hazard limits the effectiveness of transactions between firms with excess liquidity and firms that need liquidity, a bank that provides contingent liquidity to those that need it can dominate a decentralized market.

The first view of banks as liquidity providers concerns the role of banks in the payments system. Freeman (1996a,b) models an environment where agents are spatially separated and the timing of transactions is such that they cannot simultaneously trade at a central location. The problem in the model is that some agents, buyers, wish to consume goods from other agents, but have no goods that the buyers want to offer these sellers in exchange. Nor do buyers have any money, though later at another location they will be able to sell their goods in exchange for money. So, buyers issue i.o.u.’s – promises to pay at the central location next period with fiat money – to the sellers. Fiat money is used to settle the debts, but money and private debt coexist. Now, at the central clearing location it may happen that all creditors and debtors arrive simultaneously, in which case clearing occurs directly. If arrival is not simultaneous, however, settling can take place through a clearinghouse. The clearinghouse accepts money in payment of i.o.u.’s and pays off i.o.u.’s presented. However, if creditors arrive first, then the clearinghouse must have some means of paying them before the debtors arrive. A basic point of Freeman is develop the

notion of the clearinghouse issuing its own i.o.u.'s, bank notes for example, that can circulate and be redeemed for fiat money later.

Green (1997) builds on the Freeman model, arguing that a clearinghouse “netting by novation” can also achieve the efficiency gain in Freeman’s model. In the same vein, McAndrews and Roberds (1999) model the efficiency gains from introducing banks that allow for centralized netting of claims. A bank can lend to firms via overdrafts. The firms are willing to accept payment in bank funds since the income funds can be used to repay the overdraft loan. Banks can provide “liquidity” to the extent that the payments they are requested to make are offsetting. Williamson (1992) also presents a model in which fiat money and private bank “money” coexist in equilibrium. Cavalcanti and Wallace (1999) study a random-matching model in which some agents, called banks, can produce information about the trading histories of other agents, called nonbanks. The equilibrium is one in which the banks issue and redeem private bank notes.

In these models, banks issue private money to facilitate their role in clearing transactions. This is related to the historical experiences during which banks actually did issue their own private money, notably during the American Free Banking Era, 1838-1863. During this period hundreds of different banks’ monies circulated. Early economic historians and monetary theorists viewed the experience as a failure, arguing that it was marked by “wildcat banking” which justified a role for the government in the provision of a fiat currency. Following earlier work by Rockoff (1974, 1975), Rolnick and Weber (1982, 1983, 1984, 1985) reexamine failure rates over the cross section of states with different banking regimes and conclude that the period was not marked by such episodes. Gorton (1996, 1999) analyzes the prices of private bank notes and concludes that the market for banknotes worked well in pricing the risk of bank failure and in preventing wildcatting. Some experiences of Illinois, New York and Wisconsin during the Free Banking period are studied by Economopoulos (1988, 1990).

The pricing of free bank notes raises another issue concerning the production of liquidity. When offered a bank liability in exchange for goods, the seller of goods must recognize the risk that the bank can fail before the liability is honored. If some agents have private information about the likelihood of bank failure, they may be able to benefit from this when trading bank liabilities. An important property of a medium of exchange may well be that there is little or no such risk; that is, the value of the medium of exchange is independent of such considerations. But then it must be riskless in the sense that its value does not depend on the likelihood of the bank failing. This intuition is developed in the second view of liquidity, exemplified by Gorton and Pennacchi (1990).

Gorton and Pennacchi (1990) begin with a common assumption of financial market models, namely the existence of “noise traders” or “liquidity traders.” Kyle (1985) originally introduced these traders as a reduced-form modeling device, following Grossman and Stiglitz (1980). These models do

not explicitly examine the motives of these noise traders; instead, they are posited to conveniently trade and lose money, making it profitable for other traders to undertake costly information production. Implicitly, these models seem to assume larger settings in which agents face cash-in-advance constraints and suffer shocks to income or preferences, causing them to sell securities. These models assume symmetrically that there are also noise traders who have sudden urges to buy securities, though it is less clear what the source of this urge is exactly. Intuitively, these traders either sell securities at too low a price or buy securities at too high a price because they are uninformed and the prices at which securities are traded are not fully revealing.

Gorton and Pennacchi (1990) observe that these noise traders should recognize their problem, namely, that they lose money when they trade securities with better-informed traders. Consequently, they should demand securities with the property that when they are traded it is not possible for insiders to benefit at the expense of less informed traders. Thus, a security is said to be “liquid” if uninformed traders can sell it (unexpectedly) without a loss to more informed traders. The higher the variance of the value of a security, the greater the potential losses to insiders when uninformed traders must sell. If securities could be valued independently of information known only to the informed traders, then these securities would be highly desirable for trading purposes. Gorton and Pennacchi (1990) argue that splitting the cash flows of an underlying portfolio to create debt and equity can create such “liquid” securities, namely the debt. If the debt is riskless, then there can be no information advantage that other agents could possess. Uninformed agents with unexpected needs to sell securities can sell these securities to satisfy their liquidity needs. Financial intermediaries are the natural entities to create such securities, as they hold diversified portfolios of assets. Consequently, their debt should be used for transactions purposes.

Holmström and Tirole (1998) provide another rationale for intermediaries based on a third notion of “liquidity.” They begin by deriving a demand for “liquidity” that emanates from firms rather than consumers. There are three dates in their model, 0, 1, and 2. At date 0 the entrepreneur running the firm raises outside financing. At date 1 there is a “liquidity shock” requiring the entrepreneur to invest more in the project if it is to obtain a return at date 2. After the realization of the liquidity shock, the decision to continue or not is made, followed by the entrepreneur’s effort choice. If the project is continued, then an outcome is realized at date 2 and contract payments are made. Because there is a moral hazard problem in inducing the entrepreneur to expend effort, outside investors cannot be promised the full social value of the investment. The firm raises less financing than the first-best social optimum. If the firm can store the initial resources, then it faces a dilemma. It can reduce the amount it invests at date 0, to have an amount to hedge against a liquidity shock. Or, it can invest more at date 0, but then have less on hand if it needs more at date 1.

Now, suppose there is no storage and no aggregate uncertainty. The only way to transfer value across time is to use claims issued by firms. In general equilibrium, some firms will need resources at date 1 and some will not. A second-best arrangement would allow firms with large needs for resources at date 1 to utilize the market value of those firms with low needs at date 1. How would this actually work? A firm with a liquidity shock at date 1 cannot meet its needs by selling claims at date 1; it is too late to do that. Could the firm instead hedge against an adverse liquidity shock at date 1 by buying claims on other firms at date 0, and then selling those claims at date 1? There are two problems with this arrangement. First, if the moral hazard problem is severe enough, then a market in firm claims will not supply enough “liquidity.” Second, there is an inefficient distribution of liquid assets (the claims that can be sold at date 1 by firms needing resources). Firms without adverse liquidity shocks end up holding claims at date 1 that they do not need.

An intermediary can provide liquidity by issuing claims to investors at date 0 on its value at date 2. At date 0 it contracts with each firm to provide a line of credit at date 1. The maximum credit line is incentive compatible with the entrepreneur making an effort.⁶ Unlike claims in the financial market, which cannot be made contingent on a firm’s liquidity shock, firms only draw on the credit line at date 1 to the extent that they need resources. If there is aggregate uncertainty, then this arrangement may not work, and there can be a role for a government bond market.

Both Gorton and Pennacchi (1990) and Holmström and Tirole (1998) have intermediaries creating securities that have desirable state contingent payoffs. In Gorton and Pennacchi, the bank creates demand deposits whose value does not depend on the state of the world. This security is in demand because its value is not state contingent and, therefore, uninformed traders will not lose to better-informed traders who know the state of the world. In Holmström and Tirole, the intermediary creates a security, the credit line, which is valuable because it is state contingent; it is only drawn on when a firm needs resources at the interim date. Capital market securities issued by firms cannot replicate this state-contingent payoff.

F. Banks as Commitment Mechanisms

An important question concerns why illiquid bank assets are financed by demand deposits that allow consumers to arrive and demand liquidation of those illiquid assets. Calomiris and Kahn (1991) and Flannery (1994) link the fragility of bank capital structures to the role of banks. These authors begin with the assumption that banks are somewhat opaque institutions, more so than nonfinancial firms.

⁶ Because the entrepreneur pays a fee at date 0 for the credit line, he can borrow at a lower rate at date 1 than if he issued securities at date 1. This lower rate leads to greater effort for any amount of borrowing.

Evidence for this opaqueness compared to nonfinancial firms can be found, for example, in Slovin, Sushka, and Polonchek (1992) and Morgan (2000). Calomiris and Kahn (1991) argue that bank demand deposits include the right to withdraw at anytime at par along with a sequential service constraint in order to control the risk taking activities of bankers. If information about the banker's decisions must be produced at a cost, then individual depositors who expend resources to produce the information will get into line to withdraw at the bank first. Because the sequential service constraint is a first-come-first-served rule, it rewards those depositors in line first, and so information-producing depositors will recover more than other depositors. This argument was the first to suggest that banks' capital structures are deliberately made fragile so as to commit to not engaging in certain activities. From this viewpoint, fragility is a positive attribute of banks. Jean-Baptiste (1999) also argues that the instantly callable feature of demand deposits is necessary as a device to discipline bankers.

Flannery (1994) makes a related argument. He argues that bank creditors cannot effectively control bank asset substitution because of the ease of flexibly altering the bank portfolio, but they can estimate a bank's riskiness at any point in time. To control bankers, short-term debt is used because changes in bank risk will be reflected in financing costs. Again, the basic point is that the capital structure of banks is designed to be fragile, so that it functions as a commitment mechanism. Flannery and Sorescu (1996) show empirically that bank debt prices do reflect bank risk.

Diamond and Rajan (2001) use this idea that fragility is a commitment device to construct a model of bank-like financial intermediation. In their model, entrepreneurs need to raise money from outside investors to finance their projects. The specific abilities of the entrepreneur are important for the project to generate high cash flows; that is, if the entrepreneur refuses to work, then the project is worth less when someone else runs it. Moreover, the entrepreneur cannot commit to stay with the project. A lender, however, can build a relationship by lending to entrepreneur and learning about the project. If this relationship lender "liquidates" the project by separating the entrepreneur from the project, then the project is worth less than it would be worth with the entrepreneur, but more than if it is run by someone other than the relationship lender.

Because the entrepreneur cannot commit to stay with the asset, Diamond and Rajan say that the asset is "illiquid." This "illiquidity" makes it possible for the entrepreneur to hold up the relationship lender. Because potential relationship lenders anticipate this holdup problem, the amount that the entrepreneur can borrow is limited. Lenders also have problems because they may face a realized liquidity shock at an interim date. If a relationship lender needed cash at the interim date, the project would have to be sold to a non-relationship investor in whose hands it is worth even less. The prospect of such a shock makes relationship lending expensive, if not prohibitive.

The consequences of this chain of illiquidity could be mitigated if the relationship lender could borrow against the full value of the loan when faced with a liquidity shock. But this requires that the relationship lender commit to not separate from the project in the future. Diamond and Rajan argue that a bank can achieve such a commitment by designing a fragile capital structure, as follows. If the relationship lender issues demand deposits that are subject to collective action problems among the depositors, then if the relationship lender threatens to withdraw from the project, depositors will run the bank and the relationship lender will receive no rents.

As Diamond and Rajan note, this fragile structure is not first-best if banks face undiversifiable liquidity shocks. In this case, runs may occur because of high liquidity demand rather than because of bank moral hazard. Diamond and Rajan (2000) use this problem to motivate the existence and optimal level of bank equity capital. We return to this point in Section III below.

G. Empirical Tests of Bank Existence Theories

Theories of the existence of bank-like financial intermediaries link banks' activities on the asset side of their balance sheets with the unique liabilities that banks issue on the liability side of their balance sheets. Such a link is important for establishing what it is that banks do that cannot be replicated in capital markets. As we have seen, these arguments take two linked forms. First, the banks' balance sheet structure may ensure that the bank has incentive to act as delegated monitor or information producer. Second, by virtue of holding a diversified portfolio of loans, banks are in the best position to create riskless trading securities, namely, demand deposits.

Two papers, in particular, construct empirical tests of hypotheses about links between the two sides of bank balance sheets. Berlin and Mester (1999) look for a link between bank market power in deposits markets and the types of loan contracts that the bank enters into with borrowers.⁷ "Core deposits" are those deposits, demand deposits and savings deposits, which are mostly interest rate inelastic. To the extent that a bank has such core deposits, it can safely engage in long-term contracts with borrowers; in particular, it can smooth loan rates. Using a large sample of loans from the Federal Reserve's Survey of Terms of Bank Lending to Business, they find that banks that are more heavily funded through core deposits do provide borrowers with smoother loan rates in response to aggregate shocks.

Kashyap, Rajan, and Stein (2001) empirically analyze the link between loan commitments and demand deposits. While demand deposits are liabilities and loan commitments are assets, the two

⁷ Hannan and Berger (1991) and Neumark and Sharpe (1992) provide evidence of bank monopoly power in the retail deposit market.

securities both commit the bank to potentially meet demands for cash. That is, depositors may withdraw their deposits and borrowers may draw on their loan commitments. To be prepared for such contingencies, each of these security types requires the bank to hold liquidity. As long as the demands for cash on loan commitments and on deposits are not perfectly correlated, there are economies of scale to holding cash against both types of contingencies. They find that banks make more loan commitments than other types of intermediaries and that, within the banking sector, banks with high ratios of transaction deposit to total deposits also have high ratios of loan commitments to total loans.

The dramatic increase in loan sales constitutes a challenge, both theoretically and empirically, to arguments concerning bank existence. In a loan sale, the cash flows from a loan on a bank's balance sheet are sold to investors in the capital markets, through issuance of a new security (a secondary loan participation). This seems paradoxical: the borrowing firm could have issued a security directly to the same investor in the capital markets without going to the bank, and yet chose to borrow from a bank. The above arguments for the existence of financial intermediation imply that the bank loan should not be resold because if it can be resold there is no incentive for the bank to screen *ex ante* or monitor *ex post*. Gorton and Pennacchi (1995) explore these issues empirically, testing for the presence of incentive-compatible arrangements that could explain loan sales. One of their main findings is that the bank keeps a portion of the cash flows that is consistent with maintaining incentives. The idea is that the bank faces the same incentives as it would have had the entire loan been kept on its balance sheet. There are now a number of papers on this subject, but the basic paradox of loan sales remains unexplained. Indeed, the paradox is somewhat deepened to the extent that banks can transfer the credit risk of their loans to third parties via credit default swaps. Market participants seem to rely on banks' incentives to maintain their reputations for monitoring, but the efficacy of this mechanism is largely unexplored.

H. Bonds versus Loans

If banks monitor borrowers in ways that cannot be accomplished by dispersed bondholders, or produce information that capital markets investors cannot produce, then how can bonds and loans coexist?⁸ Why don't loans dominate bonds? This poses the question of the existence of bank loans in a different light. A number of authors have addressed this issue, attempting to differentiate between bonds and loans, in terms of their characteristics, but also in such a way that firms will demand both.

Detragiache (1994) presents a model in which firms use both bonds and loans. Bonds (or synonymously "public debt") cannot be renegotiated, while loans (synonymously "private debt") can be

⁸ The idea that dispersed lenders cannot renegotiate effectively compared to a single lender, like a bank, is commonly assumed. Bolton and Scharfstein (1996) provide the theoretical foundation for this notion.

costlessly renegotiated. Loans are senior to bonds. Equity holders face an asset substitution problem at the initial date, and renegotiation with the bank or liquidation may occur at the interim date. In renegotiation only the senior lender, the bank, can forgive debt, so bank debt has a clear advantage. But bonds also have a role. The payoff to equity is decreasing in the amount of bonds and this helps to limit the incentive to engage in asset substitution at the interim date. Renegotiation allows equity to capture some surplus at the interim date, even if creditors are not satisfied, so risky projects can become attractive. But, if some debt is owed to creditors with whom it is impossible to renegotiate, equity's payoff is reduced and the incentive to engage in asset substitution is mitigated.

In Diamond (1991), new borrowers, i.e., young firms, borrow from banks initially. Then later, based on the credit record established while being monitored by a bank, the firm can issue bonds. There are three types of borrowers. Two of these types are fixed at either Good (G) or Bad (B) while the third can choose between Bad and Good (BG). Type refers to the value of the firm's project so that if the investors knew the firm's type, its debt would be priced accordingly. Over time, investors or the bank can learn the type of the firms with a fixed type by observing whether there has been a default. Diamond (1989), with the same model, shows that, over time, a surviving BG type has a "reputation" that is consistent with being a G type, and this makes the cost of funding so low that it always chooses the Good project. In Diamond (1991) reputation effects eliminate the need for future monitoring so G types can issue bonds. Also, B types cannot benefit from monitoring, so they issue bonds that are appropriately priced. The BG types borrow from banks, which then monitor them. This result explains the coexistence of bonds and loans, but not for the same firm.

In Bolton and Freixas (2000) bank loans are valuable to firms because, unlike bonds, bank loans can be renegotiated. The problem, however, is that bank capital is costly, which makes bank loans costly relative to bonds. Firms trade off the benefits of bank loans against the cost. Cantillo (1998) also considers the cost of banks in determining the choice between loans and bonds. Outside investors lend to firms but can only verify the firms realized returns at a cost, which is necessary if there is a default on the debt. Thus, the set-up is one of costly state verification following Townsend (1979). Banks, however, are better at performing the costly state verification; banks take less time to verify than do nonbank lenders. Nevertheless, banks do not dominate nonbank lenders because they too are firms, so their returns also require costly state verification. The more banks lend, for a given amount of equity, the more likely that consumers financing the bank will have to expend costs to verify the state of the bank. Again, this causes firms that are more likely to be in financial distress to choose bank loans, whereas firms that are less likely to be in financial distress choose direct lending.

Chemmanur and Fulghieri (1994) take a different approach. They assume that some firms are more likely to be in financial distress than others. Firm type is private information. Banks are valuable to

these firms if they will commit more resources to evaluating firm type in financial distress. Banks can develop a reputation for doing this. Then firms with a high likelihood of needing a bank select banks that are more committed to adding value to distressed firms because of their reputation. Firms that have a low likelihood of being in financial distress issue bonds because they do not want to pool with the riskier firms. Banks are of high and low cost types in terms of their ability to evaluate firms. The game is repeated so that over time firms learn about bank type, corresponding to a “reputation.”

Cantillo and Wright (2000) empirically investigate two large panel data sets of firms to investigate the choice of firms between bonds and loans. Their evidence is consistent with their model, in which large firms that are less likely to need banks as monitors and reorganizers in the event of financial distress issue bonds. Smaller firms rely more heavily on bank loans. These determinants of choice of lender are most important during downturns.

In the papers just discussed, firms generally choose between bank loans and bonds but do not mix the two. In Section III, we return to this issue in more detail and discuss two extensions – papers that focus on the optimal mix of bank loans and bonds, and papers that explore differences in the contractual features of these two funding sources.

I. Banks versus Stock Markets

The fact that some economies are more bank-dependent and have small or almost nonexistent stock markets raises the question of how these savings and investment organizations differ. Research on the roles of banks has been discussed above. But what function do stock markets perform? Do stock markets perform the same functions as banks, so that banks and stock markets are substitute institutions? These questions are implicitly posed by studies of Germany, for example, where the economy appears to be very successful, but where historically the economy has been organized around banks.⁹ Little research has been done on these questions. In part, more economic history research is needed, but perhaps surprisingly it is also not so clear what role the stock market really performs.

Dow and Gorton (1997) present a model of the stock market in which stock prices serve two roles. First, informative stock prices can lead to efficient executive compensation. But stock prices are only informative if some traders are willing to trade on their information about projects that the firm is considering undertaking. Thus, informative stock prices have a second role: the firm can use information from stock prices in making capital budgeting decisions. In this way, the stock market performs both a screening role for projects and a monitoring role in the sense of performance-sensitive compensation. But

⁹ The distinction between bank-based systems and stock market-based systems is not as stark as it is usually presented. In the case of Germany, for example, see Fohlin (1999).

Dow and Gorton show that a bank can also perform these roles, suggesting that banks and stock markets are alternative institutions in the savings/investment process.

By contrast, Allen (1993) and Allen and Gale (1999) argue that banks and stock markets are fundamentally different in the way that they process information. Stock markets can aggregate diverse opinions, particularly about new technologies, while banks are inherently conservative. The prediction is that stock market-based economies will embrace new technologies, while bank-based economies will be less dynamic. This appears consistent with casual observations about venture capital, and raises interesting questions about the differences between banks and venture capital.¹⁰

Baliga and Polak (2001) address the question of why financial systems with banks and with bonds arise and persist. They argue that an important distinction between the German and Anglo-Saxon financial systems concerns not the distinction between banks and stock markets, but between bank loans and bonds. They ignore equity because equity finance was unimportant quantitatively at the start of the industrial revolution, the period that they have in mind for their model. Their model is one of moral hazard on the part of borrowers. If a borrower is monitored (bank finance), then the first best outcome can be enforced, but at a cost. If the borrower is not monitored (bonds), then only a second best outcome can be achieved, but there is no monitoring cost. There are multiple equilibria that can be Pareto-ranked. Interestingly, the Anglo-Saxon system can only persist if it is efficient, but the economy can get stuck in an inefficient German system.

Levine (2000) and Levine and Zervos (1998) are examples of the latest empirical research to explore questions about bank-based systems and stock market-based systems. The general conclusion of this literature is that the level of financial development is more critical than the relative dominance of banks or stock markets. Nevertheless, these papers focus on questions of overall economic growth rather than cross-sectional effects on different industries or firms. Differential effects such as those predicted by Allen (1993) or Allen and Gale (1999) remain largely unexamined.

III. Interactions Between Banks and Borrowers

As discussed in the previous section, one view of banks as intermediaries focuses on their role as delegated monitors or evaluators of borrowers that hold loans as a way of making their monitoring credible. To make this basic point, however, initial models of delegated monitoring abstracted from a number of realistic complications. In terms of interaction between borrower and lender, monitoring or evaluation is a one-time affair, behavior is fixed ex ante by contract, and optimal loan contracts are quite simple. In terms of interaction among lenders, borrowers use only one lender, and if competition among

¹⁰ See also Boot and Thakor (1997).

lenders is modeled at all, it is assumed that there are infinite numbers of perfectly diversified intermediaries offering identical terms.

Reality is much more complex. Loan contracts often include many pages of terms and conditions (“covenants”), and some banks hold equity claims on borrowers. Terms are often selectively enforced or renegotiated as bank and borrower interact repeatedly over time. Borrowers often obtain financing from multiple sources: even small firms may have trade creditors, and larger firms often use several bank lenders or mix bank finance with funding from dispersed investors. Indeed, many loan terms govern relations between the bank and other claimants. Banks are rarely perfectly diversified and face varying degrees of competition.

In what follows, we survey research on these issues. In order to narrow the field, we emphasize work that has appeared since the beginning of the 1990s; for a survey of some of the earlier work, see Bhattacharya and Thakor (1993). Moreover, we emphasize papers that focus on monitored finance – that is, papers in which some lenders have access to information that the investing public does not have. Because work on banks as underwriters has largely focused on potential conflicts of interest between banks and investors rather than its impact on bank borrowers per se, we reserve this topic for our discussion of regulation in Section V below. For reasons of space, we also abstract from work on how bank lenders may influence borrowers’ industry structure.¹¹

A. Dynamic Relationships and the Pros and Cons of Bank Monitoring

As noted above, early papers on delegated monitoring focused on one-time interaction between banks and borrowers and emphasized the savings from having a single investor monitor. Subsequent research has shown that a dynamic setting introduces additional pluses and minuses to delegated monitoring. On the plus side, long-term relationships between banks and borrowers allow for improved outcomes through implicit contracts enforced by concerns for reputation or future rents. On the minus side, a credible long-term relationship leaves bank and borrower locked in to one another, so the borrower may exploit the bank, the bank may exploit the borrower, or the borrower may find itself without needed funding if the bank suffers difficulties from the rest of its business. If these problems are severe, it may be better to forgo delegated monitoring entirely and rely instead on “arm’s-length” finance from dispersed investors.

¹¹ Examples include how cross-shareholdings or board seats between firms and banks affect tradeoffs between improved industry coordination on the one hand (e.g., Da Rin and Hellmann, 1996) and harmful anticompetitive effects on the other hand (e.g., Cantillo Simon, 1998, Adams, 1999, Arping, 2001). Other papers on how banks influence borrowers’ industry structure include Maksimovic (1990), Yosha (1995), Bhattacharya and Chiesa (1996), Kanatas and Qi (2001), and Stomper (2001).

One of the first papers to recognize the potential gains from long-term interaction between banks and borrowers is Haubrich (1989). In single-period delegated monitoring models, there is some probability each period that the borrower will do poorly enough that the bank must monitor. Haubrich's insight is that, in a repeated relationship between bank and borrower, the bank can simply keep track of reports from the borrower and penalize the borrower if too many reports are bad. If both the bank and borrower are sufficiently patient (have discount rates close to zero), the threat of being penalized and losing funding access in the future is sufficient to make the borrower report truthfully, sparing the bank the need to monitor more closely. This is true even if the reports are not verifiable in a court of law – the relationship is sustained by a tacit (“implicit”) agreement. To the extent that reviewing and tracking reports is still costly, there is still a gain to having a single bank lender rather than multiple lenders.

Boot, Greenbaum, and Thakor (1993) suggest another way in which long-term relationships and implicit contracting may reduce costs compared to explicit, completely specified contracts. They start with the observation that many bank-borrower arrangements give banks a great deal of flexibility: for example, credit lines often have clauses which let the bank renege on the credit line if the borrower's situation undergoes a “material adverse change.” Since many of these lines are used to back commercial paper borrowings and are meant to be used when the borrower is having trouble refinancing its commercial paper, it is not clear why borrowers pay fees for such a credit line – and yet they do.

The key insight of Boot, et al. is that if the bank committed to honor the credit line no matter what, it might be forced to make good on its commitment in situations where its overall situation was weak, further weakening its financial position. The “material-adverse-change” clause allows the bank to renege in such situations. Nevertheless, the bank does not wish to renege all the time – if it did, market participants would not pay for its credit lines, hurting its future profits. By putting its reputation on the line, the bank is able to offer a product that lets it renege when its current situation is so severe that the hit to its reputation and future profits is less costly than honoring the commitment now. Of course, if this arrangement is to work, the bank must be sufficiently patient and must have future rents or quasi-rents that are sufficiently attractive; an increase in competition among banks would reduce such rents and jeopardize such implicit contracting.

There is some evidence that bank relationships do help borrowers through implicit contracting. Petersen and Rajan (1994) find that small firms that have been with their bank for a longer time (controlling for firm age) have greater access to credit, especially if they rely on a single bank rather than multiple banks. Berger and Udell (1995) find that a longer bank relationship (again controlling for firm age) lowers interest rates and collateral requirements on loan commitments. As noted in Section II, Hoshi, Kashyap, and Scharfstein (1990a,b, 1991) find that Japanese firms that are members of a *keiretsu* face lower costs of financial distress than those faced by non-member firms. Elsas and Krahnen (1998) find

that troubled German firms are more likely to get additional financing if they have a “main bank” (*hausbank*) relationship.¹²

Petersen and Rajan (1995) suggest yet another way in which long-term relationships can benefit borrowers. Suppose that banks do not initially know which borrowers are good and which are bad, but do learn this over time as the firm establishes a track record for itself. Initially, firms can also unobservably choose projects with higher risk but lower returns, and their incentive to do so increases in the interest rate they are charged. If banks compete actively for loans, the rate they charge initially will reflect average credit quality, which may in turn be so high that even good firms choose risky projects, which in turn may lead to credit rationing (as per Stiglitz and Weiss, 1981). By contrast, if banks have some market power, then they can choose a lower rate initially, knowing that they can make up any losses by earning monopoly rents on good firms in the future; this in turn may reduce initial risk-shifting incentives and thus initial credit rationing.

Petersen and Rajan test their theory by regressing the rates that small businesses pay on their loans against a number of controls for firm risk and the Herfindahl index of the local banking market. Consistent with their theory, in highly concentrated banking markets, young firms are more likely to receive bank finance, and the rate of interest that firms pay declines more slowly over time, allowing banks to earn rents on survivors.¹³

The common theme of these papers is that long-term relationships increase contracting flexibility. Since the theoretic models rely on future rents or quasi-rents to maintain incentive compatibility, it follows that if increased competition among banks decreases rents, such competition should also undermine relationships. This in turn suggests a possible drawback to relationships: since bank-borrower relationships implicitly rely on lack of competition, they create an environment where the borrower is exposed to the bank’s weakness or outright exploitative behavior, and vice versa.

The first paper to focus on the drawbacks of bank-borrower relationships is Sharpe (1990), who shows that a bank’s monitoring activity may give it informational rents which in turn may distort borrower behavior. To see this, consider a firm’s “inside” bank (one that already has a lending relationship with the firm). Based on the relationship, the inside bank will have some idea of the firm’s

¹² See also the references and discussion in Berger and Udell’s (1998) review article on small firm finance.

¹³ In related work, Fischer (2000) finds that German firms in more concentrated banking markets are less credit-constrained and transfer more information to their lender; also, banks provide more liquidity if they have received such information. Bonaccorsi di Patti and Dell’Ariccia (2001) examine the impact of local bank concentration on rates of firm creation in Italy. They find that less competitive banking markets have lower rates of firm creation on average, but the opposite is true in industries with low fixed assets, which they argue proxies for high asymmetric information. At a “macro” level, Cetorelli and Gambera (2001) examine how cross-country differences in banking sector concentration affects different industries’ growth rates. They find that more concentrated banking sectors lower average growth rates, but do increase the growth rates of industries that have very high external finance needs.

eventual chance of success (“credit quality”); by contrast, banks that do not have a relationship with the firm (“outside banks”) have not monitored and are thus relatively uninformed about the firm’s credit quality. If outside banks were to offer rates that reflected average credit quality, only below-average firms would switch, and the outside banks would lose money on average. To protect themselves from this “Winner’s Curse,” these banks offer higher rates. This in turn lets the inside bank charge higher rates, letting it earn informational rents on above-average quality firms.¹⁴ Higher borrowing rates discourage investment by good firms with established bank relationships. Moreover, even if initial competition for unattached borrowers causes banks to compete away these subsequent rents, this gives too much capital to firms that are unproven.

Rajan (1992) takes this analysis further in several ways. Unlike Sharpe, he explicitly models agency problems between a firm and its investors that may make delegated monitoring attractive. Specifically, a firm’s eventual chance of success is determined by the unobservable costly effort of its entrepreneur. After this initial exertion of effort, an interim private and unverifiable signal reveals whether the firm will be successful or not; if not, it is efficient to liquidate the firm promptly rather than let it continue. If a single investor (“bank”) holds the firm’s debt, it also sees this signal, but free-rider problems are assumed to rule this out when investors are dispersed (“arm’s-length finance”).

First, if the bank holds short-term debt, it can threaten to liquidate the firm regardless of the interim signal, triggering renegotiation of the loan. If the bank’s bargaining power in renegotiation is high, it can hold up the entrepreneur for a high share of any surplus; knowing this will happen, the entrepreneur underinvests in effort. Long-term bank debt removes the hold-up problem by removing the bank’s ability to threaten to liquidate the borrower. The *borrower* can still initiate negotiations if liquidation is efficient, capturing a share of any liquidation proceeds. The drawback to this arrangement is that, *ex ante*, the borrower’s incentive to put effort into assuring good future outcomes is reduced, since the borrower effectively gets some insurance against bad outcomes that lead to liquidation. If the bank has a lot of bargaining power, long-term bank debt is better: renegotiations favor the bank, so short-term debt would lead to excessively high rents whereas long-term debt gives the borrower little insurance against liquidation. When bank bargaining power is low, the opposite is true.

Second, Rajan considers the incentive effects of arm’s-length finance from dispersed lenders. Since arm’s-length lenders are poorly informed, their decisions to renew loans or demand immediate repayment forcing liquidation are not efficient. Moreover, since they will charge a rate to protect

¹⁴ Strictly speaking, Sharpe’s analysis of the Winner’s Curse is incorrect, since he assumes that a pure strategy equilibrium in rate-setting exists. Broecker’s (1990) analysis of bidding by multiple banks with private signals of borrower quality suggests that the equilibrium should involve mixed strategies (randomized rate-setting); this holds in Rajan’s (1992) setting, and von Thadden (forthcoming) shows that this holds in Sharpe’s setting as well. Nevertheless, the informed bank does earn positive rents on average, so Sharpe’s basic intuition is correct.

themselves from inefficient continuation/liquidation decisions, the borrower's initial investment decisions may also be distorted. Nevertheless, for intermediate levels of bank bargaining power, arm's-length finance may dominate both types of bank finance.

Finally, Rajan endogenizes bargaining power by allowing for interim competition between an "inside" bank, which has monitored and knows the firm's situation, and uninformed "outside" banks. As in Sharpe (1990), the "Winner's Curse" lets the inside bank earn rents on average; the greater the information advantage, the greater the rents and thus the bank's effective bargaining power. If a firm's chance of success ("quality") is high, the inside bank's information advantage is small, and there is little difference between bank finance and arm's-length finance. At intermediate quality levels, bank finance dominates, while at lower quality levels, the informed bank's information advantage is so high that the benefits of efficient liquidation are outweighed by the bank's high rents and the entrepreneur's consequent underinvestment in effort.

Dinc (2000) examines how such informational rents affect banks' ability to sustain implicit relationship lending arrangements. In his model, it is *ex ante* efficient for bank lenders to commit to rescue firms that are distressed but not outright failures; however, because much of the benefits from such rescues flow to the entrepreneur, it is not in a bank's interest *ex post* to rescue the firm. Although the bank can be compensated by giving it higher payments when the firm is successful, in one-shot arm's-length lending, the bank always reneges on rescues *ex post*. In a repeated-game setting, so long as banks do not discount future profits too much, they may be able to establish reputations for rescuing distressed firms and so capture relationship rents. Such implicit contracting is impossible if too few banks compete: informational monopoly lets the lender set its rate so high that any additional gains to maintaining a good reputation are too small to prevent renegeing. On the other hand, if too many banks compete, rents from maintaining a good reputation are reduced, and once again banks renege.¹⁵ Thus, relationship lending can only be sustained for an intermediate number of banks.

As noted by Detragiache, Garella, and Guiso (2000), another implication of the "Winner's Curse" is that firms that rely on a bank may be hurt if that bank faces financial distress. Intuitively, a distressed bank may have difficulty supplying its good borrowers with sufficient credit for their needs. If instead these borrowers try to switch to new banks, the "Winner's Curse" problem will be especially severe because the inside bank's mix of loans is known to be worse than average. As a result, good borrowers of a distressed bank may find that additional financing is very expensive or perhaps even nonexistent.

¹⁵ A key difference between Rajan (1992) and Dinc (2000) is that, in Rajan, only the inside bank has private information, so it alone earns informational rents; in Dinc, all competing banks get (costless) private signals, and in equilibrium they earn positive informational rents that decline with the number of banks.

In Section II, we noted that Slovin, Sushka, and Polonchek (1993) find evidence that Continental Illinois Bank's near-failure in 1984 had a significant negative impact on firms for whom Continental was the main bank. Moreover, since their results focus on firms with *some* access to public markets (those with publicly-traded stock), it seems likely that smaller firms without such access were hit even harder. Kang and Stulz (2000) find similar effects when examining how Japan's long-drawn banking crisis of the 1990s affected firms that were dependent on bank finance.

By themselves, these findings do not prove that bank relationships are wholly bad; after all, before the banks in these samples got into trouble, they may have been critical in funding their relationship borrowers. Nevertheless, to the extent that a relationship with a single bank leads to possible hold-up problems or overexposure to that bank's risk, firms may wish to establish multiple bank relationships, breaking a single bank's information monopoly and diversifying exposure to any one bank's risk.¹⁶ Jean-Baptiste (2001) shows how multiple relationships mitigate the hold-up problem. Detragiache, Garella, and Guiso (2000) examine the diversification argument. They find that, when asymmetric information concerns are high, firms opt for multiple banks; the risk of being denied funding if the firm relies on a single bank that gets into trouble is too great. Similarly, the number of relationships should increase as bank fragility increases.¹⁷

Empirical tests of the number of bank relationships per firm yield mixed results. For example, using data on Italian firms, Detragiache et al. generally find support for their model's predictions: if bank liquidity shocks are high or bank size is low, firms are more likely to opt for multiple banks and more likely to have more relationships once they go to multiple banks. By contrast, Foglia, Laviola, and Reedtz (1998) examine Italian data and find that multiple bank relationships tend to be associated with greater borrower risk; however, they do not control for bank fragility. Similarly, Farinha and Santos (1999) examine Portuguese data and find that firms with greater growth opportunities, less liquidity, or greater bank dependence are more likely to switch to multiple bank relationships, all of which is consistent with reducing hold-up problems. Ongena and Smith (2000) examine how the number of bank relationships per firm varies across different countries. They find that, for low levels of bank fragility (measured by credit rating), the number of relationships per firm decreases as fragility increases, but the relationship does become positive at higher levels of bank fragility; however, they do not control for firm risk.

¹⁶ Carletti (2000) shows that having multiple banks has the drawback of decreasing each bank's incentive to monitor; however, total monitoring may increase if monitoring costs are sufficiently convex in monitoring intensity.

¹⁷ If the probability of bank liquidity problems is very high, however, firms switching banks face *less* of an adverse selection problem: outside banks know that the firm is more likely to switch because its old bank had problems rather than because the firm is a "lemon." Thus, for high bank fragility, single banking may again dominate.

A final piece of evidence comes from Houston and James (1996), who examine the mix of bank debt and public (i.e., arm's-length) debt for a sample of publicly-traded U.S. firms. To the extent the hold-up problem of Sharpe (1990) and Rajan (1992) is significant, it is likely to be most costly for firms that have many growth opportunities that need funding, and so these firms should use more public debt. Conversely, if the hold-up problem is not an issue, the advantages of relationship lending should make bank debt more attractive for firms with greater growth opportunities. Houston and James find that firms with a single bank relationship tend to rely less on bank debt as growth opportunities are higher, but the opposite is true for firms with multiple bank relationships. This is consistent with the notion that having multiple bank relationships mitigates the hold-up problem.¹⁸

In sum, theory and evidence both suggest pros and cons to bank-borrower relationships. Benefits include increased flexibility and access to funding; drawbacks include hold-up problems and negative spillovers from bank fragility. Which of these is dominant depends critically on both the nature of the borrowing firm and the nature of the banks that the firm has access to. Increased competition among banks tends to undermine relationships, but too much monopoly power may have the same effect.

B. Monitoring and Loan Structure

The papers just discussed generally assume that a firm is funded either by one bank or by many dispersed lenders with loans of the same size, that a single bank lender costlessly monitors whereas dispersed lenders do not, and that any debt has a simple structure. We now discuss research that has focused on these missing details: how bank monitoring interacts with loan features and a borrower's overall financial structure.

As noted at the beginning of this section, bank loans are often quite complex, and the terms are often renegotiated over time. Moreover, if one compares bank loans and other privately-held debt with publicly-held bonds and notes, privately-held debt typically has more covenants and other terms and is much more likely to be renegotiated than publicly-held debt. These general facts make sense if banks are to be delegated monitors; after all, complex covenants are only useful if the lender observes whether these have been violated, and dispersed lenders will lead to duplication of effort and free-riding problems in monitoring. Similarly, renegotiation is likely to be inefficient if lenders are poorly-informed or dispersed ("arm's-length" lending); again, see Bolton and Scharfstein (1996).

These ideas have led to a more dynamic view of loan features. In the old view, best represented by Smith and Warner (1979), covenants and repayment schedules simply prevent borrowers from exploiting lenders. For example, a borrower may shift into a riskier line of business, capturing any

¹⁸ For further references on the pros and cons of bank relationships, see the survey by Boot (2000).

increased upside over and above the promised payments on the debt while using limited liability to share any increased downside with the lender (“risk-shifting” or “asset-substitution”); a covenant forbidding change of business focus prevents this. In the new view, covenants and repayment schedules are tripwires which give an informed lender the right to threaten the borrower with default, after which renegotiation can occur. In our example, a lender faced with a borrower’s proposal to change to a riskier line of business can make several choices. If the change is a reaction to deteriorating conditions, and the firm’s assets are best used elsewhere, the lender can force default and liquidation. If the change is truly the best option, the bank can simply increase the interest rate to reflect increased risk, and perhaps tighten other terms to prevent any increased chance of future exploitation.

Berlin and Mester (1992) are among the first to model these issues. They show that the ability to renegotiate covenants can substantially improve their usefulness, because renegotiation allows the use of unverifiable information that the borrower and lender may have. If the borrower’s actual situation is poor, the borrower can comply with the covenant, eschewing exploitative behavior and preventing default. If the borrower’s actual situation is good, so that violating the covenant is actually good for overall firm value, the borrower can violate the covenant knowing that it will be in the lender’s interest to renegotiate rather than allow default and associated costs to occur. As a result, in settings where lenders are concentrated (so that the lenders have more incentive to be informed and renegotiation is easier), covenants can be set more stringently, improving overall firm value on an ex ante basis. Since the value of allowing renegotiation increases with the ex ante risk of the borrower, private debt contracts should be more prevalent among riskier borrowers.

Berlin and Mester’s work has some weaknesses. First, renegotiation is possible only when the covenant is tripped, so it is possible that the firm may actually be in trouble based on unverifiable information and yet its covenants are not violated. In such a case, a loan with short-term maturity would be better, since the bank could use all information (verifiable or not) in deciding whether to renew the loan or instead call for immediate repayment, triggering default. Second, the bank observes the borrower’s condition at no cost, abstracting from the need to provide the bank with incentives to monitor. Subsequent research has addressed both issues.¹⁹

One problem with a short-term loan is the hold-up problem of Sharpe (1990) and Rajan (1992); indeed, Rajan shows that a severe hold-up problem may make long-term bank debt attractive. One can then think of long-term debt with covenants as a happy medium: unlike pure long-term debt, this gives

¹⁹ A more technical issue is that Berlin and Mester focus on the case where the verifiable information in covenants is almost perfectly correlated with the firm’s actual (unverifiable) situation, ruling out comparative statics on the relative precision of verifiable information.

the bank *some* power to force renegotiation, but this power is limited to cases where verifiable information suggests that the firm is likely to be in trouble.

Rajan and Winton (1995) highlight another possible disadvantage of short-term loans: paradoxically, by restricting the bank's power to call for repayment, long-term debt with covenants may increase the bank's incentive to monitor in the first place. Suppose that a firm has a number of claimants besides the bank (trade creditors, shareholders, etc.). Two signals of the firm's condition are available: a free but noisy public signal, and a costly but more precise private signal that is only partially verifiable. If the public signal is not too noisy, a bank that holds short-term debt may prefer to rely on this signal; the bank would bear the entire cost of the private signal, but the benefits of more efficient continuation/liquidation decisions would accrue to all claimants. By allowing the bank to call the loan only when costly verifiable information suggests that the firm is in trouble, long-term debt with covenants forces the bank to observe the private signal. Of course, if the public signal is very imprecise (as for very young firms), even a bank with short-term debt will monitor; conversely, if the public signal is very precise, costly monitoring is unnecessary, and the firm can rely on publicly-held arm's-length debt. These results are consistent with the fact that firms that rely on long-term privately-placed debt tend to be larger and older than firms that rely on relatively short-term bank debt, whereas both types of firms tend to be smaller than firms that issue public debt.

Gorton and Kahn (2000) explore a different aspect of the dynamic nature of bank loans. Suppose that an entrepreneur who borrows money may engage in two forms of moral hazard: as in Rajan (1992), he may continue the firm when liquidation is more efficient, and he may also choose to increase the firm's risk in continuation even though this increase in risk is costly. There are situations where it is better to forgive some of the debt so as to deter additional risk-shifting rather than forcing liquidation, but this requires that the debt be held by a single lender so as to allow renegotiation. Even with a single bank lender, short-term debt may lead to excessive liquidation, and so it may be better to give the bank long-term debt with limited ability to call the loan early. The upshot is that, for bank debt, initial terms are not set to price default risk but rather are set to efficiently balance bargaining power in later renegotiation, and renegotiated interest rates may not be monotonic in firm risk.

Yet another strand of research has focused on how the need for bank monitoring affects the mix of bank debt and public debt and the allocation of maturity or seniority between these two groups of creditors. One early paper on this topic is Besanko and Kanatas (1993). As in Rajan (1992), entrepreneurs require outside financing, which then reduces their incentive to exert effort. A bank can force the entrepreneur to exert effort ("monitor") at a cost that is increasing in the effort level desired, but such monitoring is not contractible. Since the bank chooses monitoring levels without considering the

entrepreneur's cost of effort, it is optimal to use some public debt as a way of reducing the bank's claim on the firm and thus its incentive to monitor excessively.

Besanko and Kanatas say little about the relative priority or maturity of bank and public creditors. A starting point for this topic can be found in the work of Diamond (1993a,b) and Berglöf and von Thadden (1994). In these papers, splitting a firm's financing into a short-term senior component and a long-term junior component creates a credible threat of liquidation: the firm's value in liquidation may be enough to satisfy short-term senior debtholders even if junior claimants would prefer to let the firm continue. An appropriate mix of short-term senior debt and long-term junior debt balances credible liquidation *threats* against *actual* inefficient liquidation in a way that either forces the firm to avoid risk-shifting (Berglöf and von Thadden) or minimizes the firm's total cost of capital (Diamond).

Although Diamond (1993b) argues that an active monitor should hold short-term senior debt and focus on deciding when to liquidate the firm, none of these papers explicitly analyze monitoring incentives. It is not immediately clear that active monitors should be senior; if a senior secured creditor's claim is safe regardless of what happens to the firm, the creditor will have no incentive to monitor. Models of costly state verification such as Winton (1995b) suggest that the investor with greatest monitoring skills or lowest costs should be most junior and thus monitor more often, and indeed, venture capital and other private equity funds hold relatively junior claims and monitor intensively.

On the other hand, bank loans are often secured, hence senior to other debt. If banks are active monitors, it follows that their security cannot make them completely safe. Rajan and Winton (1995) show that giving the bank collateral can improve monitoring incentives if the collateral's value is sensitive to inefficient continuation of the firm's current business strategy. In this case, the bank must monitor the firm's situation so that it can use the threat of calling the loan to force the firm to adopt a more conservative policy or even outright liquidation so as to preserve the value of the bank's collateral. Examples of collateral with this feature include inventory and accounts receivable, both of which may suffer drastic declines in value if the borrowing firm continues to run its operations inefficiently. By contrast, a loan secured by real estate whose value was independent of the firm's operations would give the bank little incentive to monitor.²⁰

Although Rajan and Winton explore a mechanism where the bank's taking of additional collateral signals to public creditors that the borrower is in trouble, they do not examine the implications that this has for the optimal mix of public versus private debt. Repullo and Suarez (1997) focus on this mix in a

²⁰ Along similar lines, Manove, Padilla, and Pagano (2001) show that if entrepreneurs have sufficiently high low-risk collateral, similar considerations undermine banks' incentives to screen good entrepreneurial projects from bad ones.

somewhat simpler setting. An entrepreneur borrows money now and exerts costly effort that increases his firm's chance of eventual success; if it does not succeed, the firm is worth zero. If the firm is liquidated at an interim date, proceeds are less than the initial investment amount, but this may still be better than letting the firm continue if initial effort is too low and the chance of failure is high. For a fixed cost, an investor can commit to monitor the entrepreneur's effort choice at this interim date.

If liquidation values are low or the entrepreneur has sufficient investable wealth, uninformed finance (arm's length debt or equity) is optimal. If liquidation values are high relative to the entrepreneur's wealth, it is optimal to have a single investor (bank) monitor and hold short-term debt, forcing liquidation if the entrepreneur does not exert enough effort. For intermediate levels, however, giving the bank all of the firm's debt does not achieve first-best effort: if the entrepreneur deviates slightly from this effort level, the bank's share of future proceeds in success is high enough that it prefers to let the firm continue rather than forcing liquidation. In this case, giving junior debt to arm's length investors restores the credibility of the bank's liquidation threat by shrinking the bank's share of future proceeds relative to its senior claim on liquidation proceeds. As liquidation value falls, the optimal mix of debt shifts towards public (junior) debt and away from private (senior) debt.

A weakness with this model is that it requires that the bank can commit to monitoring. Without commitment, the bank would not monitor if it knew that the entrepreneur would choose first-best effort for sure, and this lack of monitoring would in turn reduce the entrepreneur's effort incentives. Park (2000) examines monitoring choice when commitment is impossible. He shows that it is optimal to give the bank senior debt that is *not* fully collateralized, so that the bank is somewhat impaired when it forces interim liquidation. This gives the bank an incentive to monitor so that it only liquidates when liquidation is efficient. Even so, subject to this incentive constraint, it is best to use as little senior bank debt as possible; otherwise, as in Repullo and Suarez (1997), the liquidation threat is less credible. Note the contrast with the results of Besanko and Kanatas (1993). In Besanko and Kanatas, monitoring automatically forces the entrepreneur to exert effort, and the problem with too much bank debt is excessive monitoring. In Park, monitoring only forces effort through the liquidation threat, and too much bank debt undermines efficient liquidation.

Longhofer and Santos (2000) provide another motivation for making banks senior to other creditors, one that combines monitoring incentives with active benefits from relationship lending. For relationship lending to have value, the bank must have incentive not only to monitor the borrowing firm's situation but also to help the firm when times are bad but the firm is essentially sound, which is just when uninformed investors are unwilling to lend more money. If the bank had a junior claim, it might not be willing to advance the firm *additional* money in bad times. This is a version of the underinvestment problem of Myers (1977): in bad times, even senior creditors' claims may be somewhat risky, in which

case some of the benefit from the bank's additional investment flows to the senior creditors. If the bank is senior, it internalizes this benefit and is willing to lend in bad times.

These theories of public and private debt are consistent with a number of stylized facts: bank loans tend to be secured and relatively short-term, public debt is more likely to be subordinated or relatively long-term, etc. Nevertheless, there has been relatively little empirical work on the detailed implications of these theories. As noted in Section II, James (1995, 1996) finds that, for firms in financial distress, banks generally do not make concessions unless public debt holders do, and banks typically make fewer concessions than public debt holders. This is consistent with the view that banks hold more senior claims and are tougher negotiators than public, more junior debt holders.

Other empirical work has focused on how the mix of public debt and bank debt varies with firm characteristics. As noted above, Houston and James (1996) find that firms that rely on a single bank use more public debt as growth opportunities are larger, which is consistent with the view that firms with lower tangible assets (and thus liquidation value) should use more public debt. Although Houston and James find that the opposite relationship holds for firms with multiple banks, they note that multiple banks are a step in the direction of public, diffusely-held debt. Johnson (1997) finds that the proportion of firm debt held by banks is greater as the firm has a lower market to book value ratio, which again is consistent with a positive link between public debt and growth opportunities. He also finds that smaller and younger firms use more bank debt, consistent with the notion that banks focus on firms with relatively greater information costs. In a study of initial public debt offerings, Datta, et. al. (2000) find that the negative average stock price response to these offerings is mitigated for firms with higher growth opportunities. Since such offerings reduce the fraction of financing provided by bank debt, this too is consistent with the notion that firms with fewer tangible assets should use more public debt. On the other hand, one could argue that this is simply a case where having well-known positive NPV projects mitigates the "lemons"-type announcement effect of issuing securities.

The upshot is that many of the complexities found in actual loan contracts – seniority, collateral, covenants – can be motivated as mechanisms for fine-tuning the bank's monitoring and control incentives as a function of firm characteristics. The same applies to the mix of (concentrated) bank debt and (diffuse) public debt.

C. Beyond Lending: Equity Stakes, Board Seats, and Monitoring

Thus far, we have assumed that banks only make loans, i.e., they do not hold equity. Although this is generally the case in the U.S., other countries have allowed banks to hold equity, Japan and Germany being well-known examples. Moreover, as James (1995) shows, U.S. banks have been allowed to take equity as a consequence of loan restructuring, and in many cases they hold these stakes for a

considerable time after the restructuring. Banks may also gain power over firms by proxy voting of shares that the banks hold in trust.²¹ Finally, even in the United States, bank officers may sit on the board of directors of firms to which they lend. In this subsection, we discuss research on how a bank's relationship with a borrower is affected by having the bank hold shares, vote proxies, or occupy board seats.

At first glance, allowing banks to hold equity has several potential advantages. In a discussion piece, Stiglitz (1985) argues that, although institutional shareholders might have goals more aligned with value-maximization, their shareholdings are usually too small to give them much direct control. By contrast, banks have a lever of control that an institution holding 1% of a firm's shares does not have – namely, the ability to refuse to renew loans. Since lenders get at most a fixed payment (interest plus principle), they care little about a firm's upside and much more about its downside. To the extent they exercise control over management, they will focus on avoiding bad outcomes; moreover, for large firms with extremely low chances of default, banks may simply do nothing. If banks hold significant equity stakes along with their loans, then they would care more about maximizing overall firm value.

John, John, and Saunders (1994) examine this issue in a simple model of firm risk-taking. If a bank's loan covenants give it effective veto power over the borrower's choice of project risk, then allowing bank holdings of equity does improve the efficiency of the borrower's risk choice; the intuition is precisely that of Stiglitz. The downside is that, because all else equal equity is riskier than debt, and because the firm chooses a higher (albeit more efficient) level of risk, holding equity increases the bank's overall portfolio risk. Thus, allowing banks to hold equity may exacerbate costs linked to bank failure.²²

Mahrt-Smith (2000) shows that bank holdings of equity have another advantage: they can reduce the hold-up problem analyzed in Rajan (1992). In Mahrt-Smith's model, a firm that has already received funds from an informed inside bank needs financing for additional investment. As in Rajan, the Winner's Curse faced by outside banks gives the inside bank informational rents which in turn distort the firm's investment decisions. If the inside bank holds both debt and equity, then any debt that outside banks provide is senior to the inside bank's equity claim. The more the inside bank's claim consists of equity, the more senior and so the less risky are the outside banks' claims. Less risk means less sensitivity to information, diminishing the Winner's Curse problem faced by the outside banks. Thus, having the inside

²¹ Although voting shares held in trust is typically associated with German banks (see Gorton and Schmid, 2000), Haubrich and Santos (2002) show that modern-day U.S. bank trust departments also have significant voting powers.

²² On the other hand, if the bank is a relatively passive investor, John, John, and Saunders show that letting the bank hold equity is more likely to *reduce* the risk of both the firm and the bank. The intuition follows Green (1984). If the bank holds only debt, the firm's controlling shareholders engage in some risk-shifting, so firm risk and thus the bank's loan risk are inefficiently high. Equity lets the bank share in the firm's upside potential, diminishing the controlling shareholders' gains from risk-shifting. Santos (1999) pursues the implications that this has for optimal regulation, and finds that regulation restricting bank equity holdings is either not binding or inefficient.

bank hold equity reduces the informational rent earned by that bank. The caveat is that, as in Jensen and Meckling (1976), greater use of external equity finance reduces the amount of effort that the firm's owner exerts, and so there is a cost to having the firm's inside bank hold equity. Because the ability to reduce the outside bank's risk is tied to the firm's need for additional funds, bank equity holdings are most attractive for financing firms with high growth potential.

Berlin, John, and Saunders (1996) focus on how a bank's shareholdings in a firm affect potential collusion between the bank and the firm against the interests of other nonequity stakeholders. Bank debt is assumed to be senior to the stakeholders' claims, and the firm and the bank observe an interim signal of the firm's situation, whereas the stakeholders do not. If the signal is good and the firm is really healthy, the firm and the bank may collude to convince stakeholders that the firm is distressed and concessions are required from creditors; these unnecessary concessions boost the firm's profits. The bank finds this attractive only if its equity share of the increase in profits offsets the concessions it must make on its debt. This puts a cap on the optimal amount of equity the bank can hold. On the other hand, if the firm is really distressed and concessions are required to prevent costly bankruptcy, the bank and the firm may collude to convince the stakeholders that the firm is healthy. This is a problem when deadweight bankruptcy costs are small, so that the bank prefers its senior claim on net bankruptcy proceeds to making concessions. In this case, having the bank hold some equity subordinates part of its claim to the stakeholders and rules out the bank's incentive to collude.

One caveat to Mahrt-Smith (2000) and Berlin, John, and Saunders (1996) is that subordinating the bank's loans accomplishes the same ends as having the bank hold equity. Indeed, in Mahrt-Smith's model, subordinated debt dominates equity because it does not cause under-exertion of effort by the firm's owner. Berlin, John, and Saunders suggest that the bank may not be able to credibly subordinate its debt claim because it could always take additional collateral when the firm's position begins to deteriorate. This would be most applicable to firms with many collateralizable assets. All else equal, such firms also tend to have lower bankruptcy costs, so this is consistent with their prediction that bank equity holdings are most useful for firms with small bankruptcy costs.

A more far-reaching objection to these models is that, even in countries where banks are permitted to hold equity, bank portfolios overwhelmingly favor loans over shareholdings. Santos (1998) finds that, as of year-end 1995, bank shares and equity participations were less than five percent of bank assets in a number of developed countries, including the G7. Germany, Japan, and Switzerland topped the list with shareholdings at 4.8%, 4.6%, and 4.9% of total bank assets, respectively.²³ Large banks in these

²³ Because these numbers include equity investments in financial firms, some of which may be unconsolidated affiliates, Santos' numbers may overstate the extent of bank shareholdings in unaffiliated firms.

countries had somewhat greater shareholdings, but even in Germany (the highest), shareholdings for large banks were only 6.3% of assets. Moreover, Mahrt-Smith and Berlin, John, and Saunders motivate bank shareholdings as a way to resolve frictions between a well-informed inside bank and uninformed outsiders, yet there is evidence that bank shareholdings tend to focus on the shares of large, publicly-traded firms, where such information asymmetries should be smallest. For example, Saunders and Walter (1994) report that, in 1989, German banks as a group held only .6% of all industrial firms' shares, but roughly 5% of the top 100 firms' shares. Thus, there must be some countervailing friction that weighs against bank shareholdings in general and bank shareholdings in smaller firms in particular.

Winton (2001) motivates these patterns as responses to banks' liquidity considerations. When a bank seeks funding to meet unexpected withdrawals, loan takedowns, and so forth, there is some chance that this funding must be explicitly or implicitly backed by the bank's claims on firms that it monitors. These are precisely the assets on which the bank has more information than its providers of funds have, giving rise to an adverse selection problem. In equilibrium, the bank prefers to let some liquidity needs go unmet, creating liquidity costs. Because, all else equal, equity is more risky, hence more exposed to inside information, than debt, the bank can minimize adverse selection costs by holding debt rather than equity. Of course, the bank can hold *some* equity without having a high chance that it will have to use these assets as backing for funding, but the bulk of its claims on firms that it monitors should be debt. Similarly, because the bank's information advantage is smallest for large, publicly-traded firms, holding equity in these firms leads to fewer adverse selection costs. Finally, to the extent that larger banks are better diversified against individual customer liquidity needs, they are less likely to face very severe liquidity needs, so they are able to hold more equity as a fraction of assets.

Thus far, our discussion has focused on the impact of bank shareholdings in terms of cash flow rights: Stiglitz, Mahrt-Smith, and Berlin, John, and Saunders emphasize shares' junior status and their claim to the firm's potential upside, and Winton focuses on the higher risk and thus information-sensitivity that these features create. Shares' voting rights provide banks with control rights over and above those associated with bank loans, and a bank with a significant equity stake in a firm may be pivotal in proxy or takeover contests. Such control rights can be amplified beyond the bank's actual shareholdings if it uses stock pyramids to concentrate its voting power, or if it can exercise the votes of shares that it holds in trust for investors. Moreover, whether through shareholdings or through their role as lenders, banks may use their influence to win seats on a firm's board of directors. We conclude this segment with a discussion of research on these issues.

There has been relatively little theoretical research on how the additional control rights possible through shares or board seats affect optimal contracting between a bank and its borrower. By contrast, such modeling is very common in work on venture capital, where, in contrast to bank loans, contracts

typically provide far more control rights and far more upside via equity or conversion features.²⁴ Although there has been little work on the circumstances in which venture capital financing dominates traditional bank lending, von Thadden (1995) is a partial exception: he does not explicitly model debt versus equity lending contracts, but he does show that giving delegated monitors significant control rights and claims on a firm's cash flow "upside" is optimal in circumstances that resemble venture capital.

In von Thadden's model, an entrepreneur chooses between short- and long-term projects. Interim returns reveal information about the project's innate quality, and this is the *only* information about project quality that the entrepreneur and arm's-length investors receive. There are cases where long-term projects are ex ante optimal, yet the firm has a high chance of poor interim returns and should be liquidated if these occur. Short-term arm's-length finance implements such liquidation, but it may cause the entrepreneur to myopically prefer the short-term project, which has a higher chance of good interim returns. By monitoring at a cost, a "bank" can get clearer information on the firm's eventual chance of success, avoiding inefficient liquidation based on interim returns alone and achieving first-best results. The contract that implements the first-best outcome is a long-term contract that gives the "bank" total control over project continuation but requires that the bank compensate the entrepreneur whenever the interim return is poor, even if the project is then liquidated.

There are two critical ways in which this model differs from "standard" models of bank control such as Rajan (1992). First, the "bank" can learn information that even the entrepreneur does not have. Second, monitoring is optimal for projects that are very likely to be (optimally) liquidated, yet have a very small chance of producing very high returns. It follows that the "bank" gets most of its return in unlikely but very high return states. Both features resemble venture capital settings: the venture capitalist typically has more financial and general business expertise than the entrepreneur, and most target firms do poorly and are terminated but a few offset this by doing very well. Nevertheless, von Thadden does not model the "bank's" incentives to monitor, nor does he allow for intermediate levels of monitoring that would correspond to bank lenders being less informed than venture capitalists but better informed than arm's-length investors. The choice between bank loans and venture capital remains largely unmodeled.

Empirical work on bank control rights through shareholdings and board seats typically focuses on two competing hypotheses. On the one hand, it is possible that these additional control rights enhance bank's ability to control firm agency problems in a positive way. On the other hand, because a bank's

²⁴ For further discussion of venture capital contracting, see Sahlman (1990) and Kaplan and Strömberg (2001).

shareholdings are typically small relative to its loans to and other dealings with a firm, the bank may emphasize policies that help it at the expense of shareholders as a group.²⁵

Findings depend critically on the economic and institutional environment in which banks operate. Examining German banks in 1975 and 1986, Gorton and Schmid (2000) find that banks use their shareholdings and board representation to improve firm performance. They find little evidence that banks use their voting powers for shares held in trust, whether for good or bad. Gorton and Schmid also find that banks' use of control is more beneficial than that of a nonblank blockholder. This suggests that bank blockholders are less likely to emphasize private benefits that harm shareholder value, perhaps because banks' "private benefits" take the form of loans, which do benefit from improved firm performance. Similar results are found by Kaplan and Minton (1994), who find that Japanese banks are more likely to get seats on a firm's board following poor performance by the firm, and that turnover of incumbent top executives increases in the year of the appointment. On the other hand, Weinstein and Yafeh (1998) and Morck, Nakamura, and Shivdasani (2000) find that Japanese firms with a main bank relationship tend to have lower growth and profitability and pay higher interest rates than firms without such relationships, and the effects tend to rise with the fraction of shares held by the main bank. Although the results on growth might be due to selection bias (banks prefer to hold more shares in more stable firms), the results on interest expense suggest that banks are not entirely innocent of pursuing private benefits of control.

By contrast with these results, Kroszner and Strahan (1999, 2001) find that U.S. banks are more likely to sit on the boards of firms that are large and stable, have high proportions of tangible assets, and rely little on short-term financing. Although these firms are more likely to borrow from a bank that has seats on their board than from other banks, this does not seem to alter loan terms: the terms do not differ significantly from those of loans that the same bank makes to similar firms whose board it does not sit on. Kroszner and Strahan interpret this as evidence that U.S. legal doctrines such as equitable subordination and lender liability discourage banks from sitting on boards of firms that have high expected costs of financial distress, and from sitting on boards of informationally-opaque firms where conflicts of interest might be easier to hide, hence more tempting.²⁶

²⁵ This is a variant of the corporate governance concern that shareholders with more voting power than cash flow rights tend to pursue private benefits of control rather than enhancement of shareholder value. See the discussion and references in "Corporate Governance and Control," by Becht, Bolton, and Roell (this volume).

²⁶ Under equitable subordination, a bank that is found to have exercised managerial control over and above what any arm's length lender might do may have all its claims subordinated to those of other creditors. Lender liability allows creditors to sue the bank if it has exercised such control. Berlin and Mester (2001) show that these features can be an optimal way of getting a large investor to both monitor and try to assure good outcomes for the firm as a whole – if a court finds that the investor misused its private information and control rights, it is subordinated and so penalized.

D. Banking Sector Structure and Lending

We now turn to research on how banking sector structure interacts with banks' role as delegated monitors of borrowing firms. This research largely falls into two broad and overlapping areas: the role of scale economies or diseconomies in monitoring, and the role of economies of scope in monitoring.

Much of the literature on scale economies in monitoring focuses on the role of diversification. In the early work on delegated monitoring discussed in Section II, a better-diversified bank is better able to resolve the agency concerns of its own investors, giving it a funding advantage. Since fixed costs of monitoring or evaluating borrowers are at the heart of these models, it follows that, rather than diversifying by making smaller loans to more borrowers, it is cheaper to diversify by keeping loan size fixed and lending to more borrowers. For this reason, larger banks should find it cheaper to diversify than small banks, giving them an efficiency advantage.

The early papers did not go much beyond this insight. In Diamond (1984), a single bank is examined in isolation; in Boyd and Prescott (1986) or Williamson (1986), infinite numbers of perfectly diversified banks compete perfectly. Also, all of these papers assumed that all loans are stochastically independent of one another, and that the marginal cost of monitoring additional loans is constant, so that infinitely-diversified banks offer depositors risk-free investments. In reality, there are only finite numbers of different borrowers, default risk across loans is partly systematic (defaults rise in recessions), and larger banks may face various diseconomies of scale. All of these issues have been the focus of later research into banking sector structure.

Yanelle (1989, 1997) and Winton (1995a, 1997) focus on how the finiteness of the economy affects equilibrium banking sector structure. As Yanelle (1989) notes, one immediate problem is that when finite numbers of banks compete for finite numbers of depositors and borrowers, the paradigm of perfect Bertrand (price) competition is not reasonable. To see why, suppose that two banks are competing; one offers a lower lending rate than the other, but the bank with the higher rate has offered a higher rate on its deposits and is expected to capture the deposit market. All else equal, a borrower that chooses the bank with lower lending rates might find itself at a bank without any funds to lend. If everyone expects this to be the case, the bank with higher rates may end up capturing the market.

A related feature of these models is the existence of multiple equilibria. This is due to "adoption externalities," where a depositor's or borrower's utility from a given bank's offered rates depends on how many other depositors and borrowers plan to use that bank. To see this, suppose that two banks compete and offer the same deposit and lending rates. If the market splits (the classic Bertrand outcome), both banks will have equal size and diversification. Now suppose instead that a number of depositors "switch" from the first bank to the second, and that several borrowers at the first bank anticipate this and also switch so as to get funding. If this occurs, the second bank is better diversified than the first, and to the

extent that the second bank is now less likely to fail, its depositors are better off. Anticipating this, more depositors (and thus firms) may also want to switch to the second bank. Thus, if agents can coordinate, one bank should dominate. On the other hand, if they cannot coordinate, the market-splitting outcome is also possible.²⁷ These adoption externalities mean that agents' beliefs about how other agents will react to a given set of bank rates have dramatic effects on equilibrium outcomes, leading to multiple equilibria.

Winton (1995a) looks at bank structure when agents can freely become investors, banks, or firms, after which banks compete for deposits and loans simultaneously. There are usually multiple equilibria, so without coordination there is no guarantee that the most efficient outcome will emerge. If a single bank is most efficient and regulators charter one bank to attain this, the monopoly bank might exploit its power so greatly that direct lending without banks is preferable. This is least likely to be a problem when monitoring costs are high; intuitively, this is when direct lending is most likely to cause credit rationing (Williamson, 1986) or even autarky, so that the reduction in monitoring costs through delegation to the monopoly bank is most attractive.²⁸

Both Yanelle (1997) and Winton (1997) address the problem of multiple equilibria in models where numbers of firms and investors are fixed exogenously. Yanelle (1997) assumes that fixed numbers of banks compete for each side of the market sequentially, and applies game-theoretic equilibrium refinements such as coalition-proofness and evolutionary stability. Outcomes depend on which side of the market "moves" first. When banks compete for loans first, multiple banks can coexist and earn positive profits in equilibrium. The coalition-proof equilibrium has the maximum feasible number of banks coexisting and earning zero profits, but this is not evolutionarily stable, so refinements do not yield clear predictions. If instead banks compete first for deposits, the coalition-proof outcome is evolutionarily stable. In it, heavy competition for deposits makes both deposit and lending rates Walrasian levels. At most one bank is active, capturing the market and earning zero profits. Since higher deposit rates increase banks' chance of default and associated costs, it is possible that direct lending is preferable to delegated monitoring, in which case banks compete themselves out of existence.²⁹

Rather than use "standard" single-period equilibrium refinements, Winton (1997) examines which investor beliefs are most plausible at different stages of a banking system's development. Early on,

²⁷ Since a single depositor's wealth cannot finance an entire firm, switching by a single depositor does not increase a bank's effective lending capacity, and market-splitting is a Nash Equilibrium.

²⁸ Yosha (1997) also examines tradeoffs between diversification and competition, but his focus is on Cournot competition among risk-sharing intermediaries that resemble mutual funds more than banks.

²⁹ The difference between the "loans-first" and "deposits-first" results is caused by Yanelle's assumption that deposits are in excess supply. Winton (1997) focuses on competition for deposits when loans are in excess supply, and gets results similar to Yanelle's "loans-first" case.

banks are not yet well established, and investors can only coordinate on banks via the rates that banks offer. Here, free entry tends to the maximal number of small, fragile banks. Although a larger, better-diversified bank that pays somewhat lower deposit rates would be best, the only way for banks to achieve this profitably is to lower their rate from the competitive equilibrium – a move that is unlikely to attract depositors to an untried bank. Over time, some banks fail, and investor beliefs begin to focus on the survivors in the sense that, all else equal, they expect the survivors will maintain their market shares. This creates an endogenous barrier to entry: expecting incumbent banks to be better diversified than new entrants, depositors are willing to accept lower deposit rates from incumbents. Barriers to entry and smaller numbers of surviving banks both promote collusive outcomes.³⁰ Evidence from this history of relatively unregulated banking regimes is consistent with these conclusions.

Up to this point, our discussion has assumed that a larger, better-diversified bank is always potentially more efficient than a smaller, less-diversified bank. There are several reasons why this may not be true. As noted in Section II, Krasa and Villamil (1992a,b) show that the combination of increasing costs for monitoring larger banks and nondiversifiable risk lead to an optimal bank size; Winton (1995a) shows that a banker's own limited capital can reduce expected costs of monitoring, which is most beneficial for smaller banks. Cerasi and Daltung (2000) focus on a third possibility, which is that the marginal cost for monitoring additional loans may be increasing. Again, the motivation stems from an individual banker's limited resources. Because any one banker has limited time and attention, his or her cost of monitoring additional loans increases with the number of loans already being monitored. Even if the bank creates a hierarchy of bankers monitoring other bankers, greater size should lead to more layers of monitoring and thus higher costs per loan. Since the diversification benefit from an additional loan diminishes as bank size grows, this cost structure leads to a finite optimal bank size. Although Cerasi and Daltung focus on the case where loans are stochastically independent, it is clear that any systematic loan risk will reinforce the effect of their scale diseconomy.

Another objection to these models is that, whereas they all assume that diversification is a passive function of bank size, in practice, diversification is to some extent a choice variable. For example, in 1982, Continental Illinois was one of the ten largest banks in the U.S., yet roughly 20% of its loan portfolio was in the energy sector – a fact that proved nearly fatal when oil prices dropped precipitously that year. To the extent that depositors cannot observe loan concentrations in timely fashion, banks may be tempted to choose a more concentrated and risky portfolio than depositors initially expected:

³⁰ Deposit insurance reduces this barrier to entry, since depositors will no longer care as much (if at all) about a bank's chance of failure. Applying deposit insurance early on leads to even greater entry, fragmentation, and fragility, but fairly-priced insurance applied to a system that is already concentrated can be beneficial: by threatening incumbents with entry, deposit insurance can reduce their ability to sustain collusion.

shareholders gain from increased upside, while increased downside is shared with depositors because of limited liability. In a single-period setting where depositors cannot observe bank portfolio choice before risk is resolved, Hellwig (1998) shows that the bank will concentrate risk as much as possible. Even though depositors anticipate this and charge a higher rate as compensation, some diversification benefits are lost; indeed, if all funds can be concentrated on a single (large) borrower, delegated monitoring is completely undermined.

In a more dynamic setting, matters are less stark. As Marcus (1984) first showed, future rents or quasi-rents reduce a bank's incentive to take on more risk now, since higher risk means a greater chance of failure and loss of future value. Besanko and Thakor (1993) embed this in a model that combines active choice of diversification across two loan sectors with future relationship rents along the lines of Petersen and Rajan (1995). If bank competition increases, banks earn lower relationship rents and generally have more incentive to concentrate their portfolios.³¹

Winton (1999) suggests that risk-shifting via loan concentration per se may not be the biggest problem in banking in developed economies. If loans are relatively illiquid and it is difficult to change loan concentrations very quickly, investors are able to get *some* information about loan concentrations and adjust deposit rates accordingly before the loans mature and the outcome of risk shifting is realized. This gives banks some ability to commit to diversification strategies: loan concentrations will be detected before the bank can exploit depositors, so more “virtuous” strategies will be rewarded with lower deposit rates, and vice versa. By contrast, loan monitoring is more difficult to observe in a timely fashion, which may let a bank shift risk simply by not monitoring its loans.

This possibility does not arise in the previous papers on diversification because they model monitoring as ex post costly state verification that can be committed to up front. As we have seen, a more realistic view is that the bank must have incentives to monitor, and monitoring is useful because it lets the bank intervene before matters deteriorate too greatly – it is proactive. In other words, monitoring must be done ex ante and it is mostly of use when a borrower is in bad shape ex post. Similarly, banks' ex ante screening activities (e.g., Boyd and Prescott (1986)) seek to avoid making bad loans. In both cases, the emphasis is on avoiding or ameliorating bad outcomes. This is in contrast to most nonfinancial firms, where firms' actions may also seek to improve good outcomes. Loans have limited upside, and the emphasis is on avoiding downside.

It follows that a bank that does not monitor saves the cost of monitoring but makes bad outcomes worse or more likely. Since credit risk is correlated across loans in a given sector, the ex post gains from

³¹ Strictly speaking, Besanko and Thakor assume that bank deposits have flat-rate government insurance. Nevertheless, as Hellwig's (1998) analysis shows, even with risk-sensitive deposit rates or insurance premiums, banks would have incentive to engage in risk shifting if there were no future rents at stake.

monitoring are greatest when a loan sector is in a downturn. If the chance of troubled loans is very low when the sector is doing well (as for many commercial loans), not monitoring loans is a form of risk shifting. In good times, the bank saves monitoring costs and does not have many more defaults; in bad times, the bank is more likely to fail, leaving its debtholders with much of the worse performance. In this case, diversifying across sectors can improve monitoring incentives; diversification reduces the dispersion of the bank's loan returns (monitored or no), decreasing the gains to risk shifting. On the other hand, if the risk of troubled loans is so high in good times that monitoring pays for itself even then (as might be the case for credit card loans), monitoring incentives are strong even if the bank focuses on a single sector.³² Winton argues that this may be one reason why Continental Illinois' focus on "Rust-Belt" and energy sector commercial loans was accompanied by lack of monitoring and eventual failure, whereas specialized credit card banks such as CapitalOne or MBNA had strong monitoring skills and cultures that withstood repeated sector downturns in the 1990s.

Since diversification is in part a matter of choice as well as a passive function of size, it follows that the decision to diversify rather than specialize will depend on the presence of scope economies as well as scale economies. The role of scope economies in monitoring and bank sector structure began with models of spatial differentiation, where transportation or other distance costs give local banks an advantage over more distant rivals. As Besanko and Thakor (1992) note, "local" advantage in the lending market need not be geographic; a bank with a lending focus on one industry or sector may be more effective in making loans to that sector than rivals whose focus is somewhat different.³³

Both Besanko and Thakor (1992) and Gehrig (1995) use spatial models to examine the impact of increased entry on overall welfare and bank risk. In both papers, banks, firms, and depositors are located equidistantly around a "circular" economy of constant size. In Besanko and Thakor, an exogenous increase in the number of banks increases competition for both deposits and loans, making depositors and firms better off and bank shareholders worse off. The increase in deposit rates also increases the amount of deposits in the system and reduces bank capital ratios, increasing banks' chance of failure. Gehrig improves on this analysis in two ways. First, he endogenizes entry, and shows that entry can be either excessive or insufficient relative to the social optimum. Second, whereas Besanko and Thakor assume

³² If there are deadweight costs to bank failure (such as ex post verification), diversification may still be attractive, but it is also possible that diversification can actually *increase* the bank's chance of failure for plausible levels of leverage. This occurs because loan returns are highly skewed to the left: there is a high chance loans pay off in full, and a low chance that they are troubled and produce losses. If losses in a sector downturn are high but the chance of a downturn is sufficiently low, a diversified bank may fail if any sector it is exposed to has a downturn, whereas a specialized (one-sector) bank only fails if its sector has a downturn.

³³ There is a great deal of evidence that the pricing of small loans and retail deposits does in fact vary regionally within the U.S.; see for example Petersen and Rajan (1995) and Neumark and Sharpe (1992).

that downturns affect all locations at once, Gehrig assumes that downturns hit an economic region of random size and location. Since larger banks are less likely to have their whole portfolio suffer a negative shock, they are less likely to fail. Increased entry tends to increase the risk of bank failure by shrinking bank size, and so increased incidence of regional shocks make entry less attractive, all else equal.

Matutes and Vives (1996) use a simpler spatial model to examine the impact of deposit insurance on bank competition and diversification. Two banks compete for deposits, which they then invest; depositors face distance costs, and larger banks are less likely to fail. As in Yanelle (1989, 1997) and Winton (1995a, 1997), there are multiple equilibria. Deposit insurance eliminates multiple equilibria and increases the supply of deposits, but increases the probability of bank failure and associated costs. The positive effects tend to dominate negative effects when uninsured banks would be local monopolies (some depositors don't use either bank); the opposite is true when uninsured banks would serve the entire market and compete directly.

Another approach to scope economies focuses on explicit information differentials between banks that are competing for the same borrowers. The first paper to address this is Broecker (1990). He shows that when banks get imperfectly-correlated binary signals about a borrower's quality, whichever bank gets the borrower's business knows that other banks were likely to have received more negative signals – a version of the “Winner's Curse.” As there are more banks bidding for a given pool of borrowers, the equilibrium probability that *some* bank will accept a given borrower increases, and so the average quality of actual loans goes down. Riordan (1993) takes this a step further by examining continuous signals. He finds that, as the number of banks increases, banks apply more conservative acceptance standards, but the overall percentage of bad loans granted still increases.

We have already seen another implication of the Winner's Curse in the work of Sharpe (1990), Rajan (1992), and von Thadden (forthcoming): banks that are better informed about a given borrower have a comparative advantage over uninformed banks, enabling them to earn rents. Dell'Ariccia, Friedman, and Marquez (1999), Marquez (2001), and Dell'Ariccia (2000) explore how this affects banking sector structure. In Dell'Ariccia et al. (1999), two banks compete, first offering rates to all firms that they do not have relationships with, then offering rates to their existing customer bases. The Winner's Curse gives each bank an advantage over its existing customers that are in good shape. Because Dell'Ariccia et al. assume that banks cannot distinguish between naturally unattached firms and “lemons” that have left their existing bank, the bank with smaller customer base faces relatively more adverse selection when competing for new business. As a result, the smaller bank breaks even on new business, and the larger bank earns higher overall profits. Moreover, if a third bank with no market share tries to enter the market, it loses money in equilibrium; the Winner's Curse is a barrier to entry. Marquez (2001) shows that entry is easier as borrower turnover is higher (so that the pool of unattached firms has

relatively fewer “lemons”) or as entrants’ ability to screen is higher. Dell’Ariccia (2000) endogenizes entry and market shares by incorporating a spatial setting in which firms face “distance” costs of borrowing. He shows that the equilibrium number of banks under free entry decreases as information asymmetries increase; intuitively, this worsens the Winner’s Curse problem faced by entrants.

Since informational rents can lead to ex ante effort distortions as per Rajan (1992), it might be welfare improving if banks could commit to share their information. In fact, credit bureaus provide just this function and are becoming widespread. Pagano and Jappelli (1993) show that lenders’ incentives to join such an information-sharing arrangement are greater as there is more borrower heterogeneity, as borrowers are more mobile, and as the lending market is larger – all of which tend to increase adverse selection problems. Padilla and Pagano (1997) show that such information sharing intensifies competition by reducing Winner’s-Curse effects, reducing effort distortions as well. Nevertheless, the higher effort makes lenders better off, so it is possible that banks participate despite the loss of potential rents. In addition, when default information is shared, borrowers may further increase their efforts so as to avoid defaulting and being labeled a bad credit risk (Padilla and Pagano, 2000).³⁴

In these papers, a bank’s screening ability is innate and any signals that it receives are free. Some recent work has sought to endogenize banks’ screening or monitoring ability. Thakor (1996) and Ruckes (1998) endogenize the probability with which banks engage in costly screening, and focus on how this interacts with macroeconomic conditions. We will return to their work in the next subsection. Gehrig (1998) examines the impact of integrating two previously separate monopoly banking markets when banks can choose their costly screening technologies’ Type I and Type II error probabilities. Although integration may increase rate competition in a beneficial way, entry may be blocked as in Dell’Ariccia et al. (1999); if so, each bank focuses exclusively on its “home” market. Also, in some cases, integration reduces banks’ investment in screening, reducing overall welfare.

Boot and Thakor (2000) focus on banks’ decision to invest in expertise that differentiates them from their competitors. Firms can get “transaction” loans in which the bank simply lends and takes borrower quality (default chance) as given, or “relationship” loans in which the bank monitors at a cost and has some chance of improving the return of firms that would otherwise fail. Because relationship loans are assumed to be a differentiated product, increases in competition between banks undermine rents

³⁴ Gehrig and Stenbacka (2001) raise an important caveat to these results. Just as in switching-cost models in the industrial organization literature, the prospect of future information rents can make banks compete heavily ex ante for borrowers. By reducing future rents and thus current competition for unattached borrowers, information sharing may actually increase bank collusion and reduce social welfare.

on transaction loans relatively more than rents on relationship loans. As a result, banks invest less in relationship expertise, but more of their loans are relationship loans.³⁵

In Boot and Thakor (2000), both interbank competition and relationship expertise are modeled in reduced form. Hauswald and Marquez (2000) endogenize these features and get a richer set of implications for bank strategy choice. In their model, banks are spatially differentiated and can make unscreened “transaction” loans or make screened “relationship” loans. Screening ability deteriorates as the “distance” between bank and borrower increases. Banks can improve their screening ability by investing in “sector” expertise that is location-specific or in “transferable” expertise that reduces the negative effects of distance. As in Boot and Thakor (2000), as costs of entry decline and more banks enter, transaction lending decreases and relationship lending increases; however, banks invest less in transferable expertise and more in sector-specific expertise. Thus, greater entry makes banks compete less heavily in peripheral lending markets, freeing up resources which are used to bolster their position in their core lending markets.³⁶

Because all of these models of differential information focus exclusively on expected loan returns, they ignore how differential information affects bank risk. Winton (1999) addresses this issue. Just as effective monitoring or screening reduces the potential downside of loans, weaker or monitoring or screening skills exacerbate this downside. As a result, a bank diversifying into sectors where it lacks expertise faces greater downside risk from this sector, offsetting the potential benefits of risk diversification. Moreover, this increased downside risk can undermine the bank’s incentive to monitor not only loans in the new sector but in the bank’s core sector as well. Indeed, a number of well-known large banks that aggressively diversified during the 1980s found themselves with poor loan performance in many sectors, Citicorp being a case in point.

The general thrust of these theoretical papers on banking sector structure and monitoring is that the assumption that perfectly-diversified-and-perfectly-competitive banks form quickly and efficiently is overstated. Even if diversification offers scale economies, competitive forces may lead to high entry and fragmentation in the early stages of a banking system. Diseconomies of scale and informational economies of scope make slower diversification or even a strategy of specialization attractive. To the extent that a bank does want to expand, the results on Winners-Curse-type barriers to entry suggest two

³⁵ Boot and Thakor also allow firms to access the capital market at a cost that decreases in borrower quality. A decrease in this cost has the opposite effect as an increase in interbank competition: bank lending rents and the total number of bank loans fall, reducing entry into banking; this makes relationship expertise *more* attractive.

³⁶ In a related vein, Almazan (2002) examines banks’ specialization decisions in a spatial model, where monitoring is more expensive the further a borrower is from the bank. Because capital improves bank monitoring incentives, capital and “expertise” (distance from a borrower) are substitutes, and high capital banks have greater market share.

strategies. One is to focus on lending types or sectors where the entrant bank has the least disadvantage relative to incumbents; an example is Citicorp's successful expansion of credit card lending in East Asia in the 1990s, which exploited Citicorp's technological advantages vis-à-vis local banks in credit scoring and information systems.³⁷ An alternative strategy is to acquire banks that are already established in unfamiliar sectors, effectively buying lending expertise; drawbacks are that the acquiring bank may find itself buying *banks* that are "lemons," and the acquiring bank may find it harder to monitor its lenders in such unfamiliar sectors.

There are many empirical papers on these issues, mostly focusing on the impact of bank size and diversification on lending and loan portfolio risk. This in turn is somewhat subsumed in the even larger literature on bank efficiency. Rather than survey the efficiency literature in detail, we present highlights here, and refer the interested reader to the surveys by Berger and Mester (1997), and Berger, Demsetz, and Strahan (1999). Essentially, throughout the 1980s, studies tended to find few significant scale economies in banking past banks of moderate size; however, more recent studies (such as Berger and Mester, 1997) *have* found significant scale economies for banks of sizes as high as \$25 billion. One possibility is that advances in technology have led to significant advantages for large banks; a case in point is credit card operations, which benefit from specialized expertise in credit scoring, account servicing, and customer information retrieval. A second possibility, which we explore below, is that larger banks are better diversified and can invest more of their assets in risky loans rather than safer but less profitable cash and government securities.

That larger U.S. banks *do* take on more risk is beyond doubt; see e.g. Boyd and Graham (1991), Chong (1991), Akhavein, Berger, and Humphrey (1997), Demsetz and Strahan (1997), and Hughes, Lang, Mester, and Moon (1999). Whether or not this is wholly due to greater risk-bearing capacity or to exploitation of deposit insurance and "too-big-to-fail" is less clear. Boyd and Graham (1991) present evidence that, in the 1970s and 1980s, a higher percentage of large banks failed than of small banks, suggesting that banks may have gone beyond the exploitation of scale economies of diversification. Examining publicly-traded U.S. banks during 1980-1993, Demsetz and Strahan (1997) find that larger banks' stock returns have less firm-specific risk, and that banks that are more diversified (both by loan sector and by geographic region) have lower firm-specific risk. Nevertheless, up through 1991, larger banks' total stock return variance was no less than that of smaller banks, suggesting that larger banks took on more risk to offset diversification advantages. Hughes et al. (1999) estimate a structural model of the decisions of risk-averse bank managers. They find that geographically-diversified banks gain both in efficiency and in reduced "insolvency risk" (standard deviation of return on equity normalized by

³⁷ Dell'Ariccia and Marquez (2001) provide a theoretical model of such behavior.

expected gross return on equity). Together with Demsetz and Strahan's results, this may indicate a reduction in large bank's risk-taking behavior in the 1990s. On the other hand, Winton (1999) notes that bank loan portfolio risk is highly-skewed to the left, with losses peaking during infrequent downturns. Thus, risk measures such as variance which work well for normal distributions may not perfectly capture bank failure risk, especially if the sample does not include a major downturn.

Another focus of the bank-size literature has been whether larger banks are less likely to lend to small firms. There are several overlapping motivations here. One focuses on diseconomies of monitoring more loans, as per Cerasi and Daltung (2000); if this is true, larger banks may prefer to focus their lending on large firms, since this requires fewer loans per dollar of assets. Since larger firms tend to be better diversified than smaller firms in the same industry, this need not reduce overall diversification very much, and saves on costs. An alternative argument is that large organizations favor the use of "hard" information; this may favor lending to large firms, since more information is publicly-available for them. Finally, large firms may prefer large banks simply because the level and complexity of their financial needs is beyond the capacities of a small bank or small group of small banks.

Again, a full discussion of the literature is beyond the scope of our paper, so we report highlights and refer the reader to Berger, Demsetz, and Strahan (1999) for references not otherwise given. The general finding is that large banks focus more on larger firms, and small banks on small firms. Thus, a number of U.S. studies find that small loans are a smaller fraction of total assets at large banks than at small banks. Berger and Udell (1996) find that small firm loans at large banks have significantly lower rates and collateral requirements than those at small banks. Since a number of studies suggest that higher collateral goes along with more risky borrowers, this is consistent with the idea that large banks focus more on relatively safe, "transparent" small firms, while small banks take on the riskier, "opaque" small firms which require greater monitoring. Most U.S. studies find that mergers among larger banks reduce small business lending; Sapienza (2000) finds similar results following mergers of Italian banks. Conversely, recent studies of *de novo* U.S. banks find that these small banks focus more on small loans than do established banks of similar size, which is consistent with entry aimed at filling the financial needs of small firms that have been abandoned by large-bank mergers.

We now turn to empirical evidence for scope economies linked to geographic or sector focus. Generally, in contrast to the findings for expansion within the U.S., there is evidence of geographic barriers to entry when banks expand internationally. This makes sense, since international cultural, legal, and informational differences are much greater than regional differences within the U.S. Berger, DeYoung, Genay, and Udell (2000) survey this literature, and provide a more comprehensive test by examining the efficiency of banks from a number of countries in five different "home" countries (France, Germany, Spain, U.K., U.S.). Home country banks are generally more efficient than foreign banks, but

when the foreign bank is from the U.S., matters are usually reversed. Berger et al. interpret this as evidence that at least some U.S. banks are simply superior and able to more than overcome geographical barriers to entry. On the other hand, the sample may have been relatively favorable for foreign expansion: the “home” countries are relatively advanced economies, where informational barriers may be relatively less severe, and the sample period is from the mid-1990s, a relatively recession-free period.³⁸

Acharya, Hasan, and Saunders (2001) examine the impact of commercial loan diversification on the profitability and risk of Italian banks. Consistent with models of specialization advantages, they find that diversification across industries reduces average returns and increases levels of doubtful or nonperforming loans. The effect of geographic diversification is mixed. Although such diversification generally hurts returns, for banks whose loans have moderate levels of risk, it actually improves returns slightly and reduces risk. This is consistent with the model of Winton (1999), which suggests that diversification is most likely to improve monitoring incentives when loans have moderate levels of risk.

Indirect evidence for economies of scope and specialization advantages is provided by DeLong (2001), who finds that stock market reaction (bidder plus acquirer) to mergers between U.S. banks is only positive for mergers between banks with similar sector focus or geographic scope. This is consistent with the bulk of research on nonfinancial mergers from the 1980s on, which find that focusing mergers add value whereas diversifying mergers tend to destroy value.

Finally, there is evidence that Winner’s-Curse effects can be significant. Shaffer (1998) finds that, consistent with Broecker (1990) and Riordan (1993), loan loss rates are higher in local U.S. markets with more banks. He also finds that, during every year from three to nine years after founding, *de novo* banks have loan loss rates that are significantly worse than those of the average bank. Since *de novo* banks are required to have experienced bank management and are often started by an existing bank holding company, Shaffer argues that this is due to the Winner’s Curse rather than lack of basic lending skills.

To summarize, research on the industrial organization of the bank lending sector suggests caveats to the initial models of delegated monitoring. Because of inherent adoption externalities, laissez-faire banking need not quickly lead to a competitive, well-diversified banking sector. Informational problems such as the Winner’s Curse can compound this problem by creating endogenous barriers to entry; conversely, scope economies tied to specialized monitoring expertise may make “well-diversified” loan portfolios suboptimal.

³⁸ At a more micro level, Berger, Klapper, and Udell (2001) examine lending in Argentina and find that smaller borrowers are less likely to borrow from foreign banks, especially those headquartered outside Latin America. Since smaller borrowers tend to be more “opaque,” this is consistent with cultural and informational barriers to entry.

E. Credit Cycles and the Effect of Bank Funding on Lending

The papers we have discussed in this section typically focus on microeconomic concerns such as the nature of interactions between bank and borrower or the structure of the banking sector. We now briefly discuss work that focuses on how such microeconomic concerns can interact with macroeconomic business conditions. Given the vast literature on this subject, our approach is selective, aiming at key points and a few illustrative papers. This also leads us to a discussion of papers that examine the interaction between a bank's funding and lending activities, since some have argued that this is a critical source of business cycle effects.

In principle, effects can flow both from business conditions to bank lending decisions and vice versa. Worsening business conditions will clearly increase the risk of many potential borrowers, making banks more conservative ("broad lending channel"). To the extent that bank borrowers tend to be smaller and riskier than firms that access public debt markets, bank-dependent borrowers may be hit harder by higher interest rates or worsening business conditions, and bank lending may fall further than public debt borrowings. Conversely, if banks are an important source of funds for firms and consumers, bank-specific shocks that make banks more conservative will reduce borrowers' ability to invest or consume, and again this will hit bank-dependent borrowers hardest ("bank lending channel"). Thus, the evidence that the ratio of bank loans to public debt drops more when monetary policy is tight (see e.g. Kashyap, Stein, and Wilcox, 1993) is consistent with both models.

Since it seems likely that both channels occur in practice, and that there is feedback between them, we will ignore the perennial debate in the macroeconomic literature on which of these channels is more important than the other. Instead, we will discuss models that illustrate how these two effects come about and empirical evidence on these effects.

The simplest view of how worsening business conditions affect bank lending is that of Bernanke and Gertler (1989, 1990): worsening business conditions hurt borrowers' net worth, increasing agency costs that lenders such as banks face. Since loans are less attractive, fewer loans are made, and rates on any loans that are made are higher to compensate for higher costs of distress. Thus, the presence of agency costs exacerbates business cycles.

Although this focuses on the monitoring view of bank lending, Ruckes (1998) shows that similar results obtain when competing banks screen potential borrowers. Intuitively, screening does two things: it avoids making a bad loan, and (if the borrower is good) the screening bank has an information advantage over rivals that have not screened that borrower. When times are good, the chance of bad borrowers is low and any information rents are small, so banks do not screen intensively. As the proportion of bad borrowers grows, banks at first screen more intensively; eventually, however, there are so many bad

borrowers that banks screen few borrowers and make few loans. These endogenous screening effects make lending contract and expand more than it would in the absence of differential information.

Whereas Bernanke and Gertler and Ruckes assume that banks are out to maximize loan value, Rajan (1994) motivates credit cycles through an agency problem between bank managers and their investors. Suppose bank managers vary in their lending ability: only poorly-run banks have bad loans in expansions, but all banks have bad loans in recessions. In expansions, poorly-skilled managers have incentives to renew bad loans so as to hide their incompetence; in recessions, they liquidate bad loans because good banks can also have bad loans, so liquidation does not send a signal of their ability. The upshot is that banks have overly loose lending standards during expansions. One caveat to this argument is that it assumes that the relative difference in loan quality between poorly-run banks and well-run banks is highest in expansions. It seems more likely that normal recessions create more difficulties for weaker borrowers, which should then hit poorly-run banks harder than well-run banks.

By focusing on problems within banks, Rajan's work is also a step in the direction of "bank lending channel" models, which focus on how problems at banks can then spill over to their borrowers and thus the entire economy. This literature has two major strands: papers that focus on how a bank's net worth (its level of equity capital) affects its lending behavior, and papers that focus on how adverse selection and other funding costs affect a bank's ability to make loans. Since capital constraints would not be an issue if banks could raise additional equity at no cost, and costs related to adverse selection are a significant part of the costs of raising additional equity, the two strands are interrelated. We begin with bank capital effects.

Thakor (1996) examines how changes in risk-based capital requirements affect bank lending decisions when it is costly to screen borrowers. Because equity capital is assumed to be more expensive than deposits, an increase in risk-based capital requirements makes loans less attractive on the margin relative to risk-free securities. As a result, higher capital requirements tend to reduce banks' willingness to screen and thus to lend. Conversely, banks that are more constrained by capital requirements are less likely to lend than are their less-constrained rivals.

Holmström and Tirole (1997) obtain similar results in a model of agency costs between borrowers and lenders. Firms are more likely to engage in risk-shifting as their net worth declines. Banks can prevent this via monitoring, but because banks borrow money from other investors, they too may engage in risk-shifting by not monitoring. As in Thakor (1996), banks with more capital monitor borrowers more

intensively, allowing banks to credibly lend more. Thus, the link between capital shocks and lending depends critically on whether the net worth of nonfinancial firms or of banks is most affected.³⁹

Whereas Thakor and Holmström and Tirole assume that bank managers are fully aligned with their shareholders' interests, Besanko and Kanatas (1996) distinguish between insider shareholders and external shareholders. Raising external equity capital dilutes bank insiders' incentives to monitor loans – another variant of the Jensen and Meckling (1976) agency problem between managers and outside shareholders. In some cases, the reduction in monitoring more than offsets the additional “cushion” against bank failure that the additional capital provides; thus, that higher bank capital requirements can sometimes increase the risk of bank failures. Although these results are striking, they seem most applicable to small banks, where issuing additional equity can substantially dilute top management's stake in the bank. In a large bank such as Citigroup, even a relatively small stock or option-based stake can leave management with significant risk in absolute terms, and it seems less likely that issuing equity will significantly affect top management's incentive to monitor loans effectively.

Diamond and Rajan (2000) also incorporate external equity capital, using their “bank fragility” model (see Section II) as a base. Recall that, in their model, short-term deposits limit the rents a bank manager can extract, but bank failure is costly because the bank manager's expertise is lost. Because it is junior to deposits, external equity cushions the bank against costly failure; however, shareholders have a weaker bargaining position than depositors and allow the bank manager to appropriate rents when the bank does not fail. It follows that, as the probability of good loan returns increases, the optimal level of bank capital falls. Also, because banks that are more capital-constrained face a greater risk of failure, their threat to liquidate borrowers is more credible, and they liquidate cash-constrained borrowers more often. By contrast, capital-constrained banks may find that their ability to extract cash now out of cash-rich borrowers is weakened, depending on the relationship between the firm's current cash flows, its liquidation value now, and its liquidation value in the future.

Diamond and Rajan's results are heavily influenced by their use of a Hart and Moore (1998) “incomplete contracts” setting, where everything is observable but only liquidation values are verifiable. Thus, “good loan returns” are linked to high firm liquidation values rather than high firm cash flows per se. As we have argued, banks may have private information that outside investors do not have. Also, in

³⁹ In a related paper, Covitz and Heitfeld (1999) examine the link between bank market power and bank lending rates. When bank incentives to risk-shift are weak (i.e., banks are well-capitalized or have large future rents at stake) but firm incentives to risk-shift are strong, monopoly banks charge lower rates than competitive banks. The intuition is that lower rates reduce firm risk-taking incentives, but when bank competition increases, borrowers have more say in equilibrium rates, and they prefer high rates so that banks prefer to “go along for the ride” on risk-shifting rather than engage in costly monitoring and liquidation. Conversely, when risk-shifting incentives are reversed, monopoly rates are higher than competitive rates.

developed economies such as the U.S. or Japan, banks may be able to attach borrower cash flows. In this case, a capital-constrained bank may have incentive to squeeze cash out of small borrowers without much access to alternative funding sources; conversely, the bank may let a large, cash-constrained borrower with low liquidation value continue in the hope that its cash flows recover before outside investors realize the extent of the bank's (or the firm's) problems. Thus, Diamond and Rajan's results may be most applicable to economies where legal and institutional safeguards are less advanced.

Note that all four of these papers suggest that higher bank capital tends to increase lending, but whether this is good or bad depends on whether or not bank managers are aligned with bank shareholders. In Thakor (1996) and in Holmström and Tirole (1997), higher capital improves banks' monitoring incentives and thus the quantity of loans banks can credibly make. In Besanko and Kanatas (1996) and Diamond and Rajan (2000), higher capital loosens banks' lending standards, but this has either a bad (Besanko and Kanatas) or mixed (Diamond and Rajan) impact on credit quality.

A large empirical literature has examined the link between bank capital and lending. Much of this work stems from the debate over whether implementation of the 1988 Basel Accord's capital standards caused a "credit crunch" in the U.S.. Sharpe (1995) surveys this literature and finds that, overall, evidence suggests that bank profitability has a positive effect on loan growth, whereas loan losses have a significant negative effect on loan growth.⁴⁰ To the extent that higher profits increase capital and higher losses decrease it, this is consistent with banks cutting back lending when capital is low, but it is also consistent with banks cutting back lending when low profits or high losses suggest that loans will be less profitable going forward. In a more recent study, Beatty and Gron (2001) estimate a structural model of U.S. banks' simultaneous choice of asset growth and capital growth during the period from 1986 to 1995. They find that, for banks whose capital to assets ratio is in the bottom decile, increases in capital go with higher risk-weighted asset growth or higher initial levels of risk-weighted assets, and vice versa. (Risk-weighted assets weight loans most heavily.) For other banks, these relationships are less significant.

Several studies have examined firm-specific effects of bank capital levels. Thakor (1996) finds evidence that announcement of a bank loan commitment causes the borrower's stock price to increase significantly, which is consistent with bank screening activity. For the largest quartile of loan commitments as a fraction of bank capital, the increase is greater when the bank's capital is lower, which is consistent with such banks being more choosy about making large loans. Hubbard, Kuttner, and Palia (1999) find that banks with weaker capital levels charge borrowers higher rates, even controlling for borrower risk characteristics. The effect is significant only for borrowers that are small and unrated, or whose loans are priced over prime, all of which proxy for firms with high informational switching costs.

⁴⁰ Empirical papers on the "credit crunch" are discussed in detail in a later section.

The picture painted by these findings is most consistent with Thakor (1996) and Holmström and Tirole (1997): as capital levels fall, banks become more conservative. The findings of Hubbard et al. (1999) suggest that informational frictions as in Rajan (1992) are also important. The caveat is that these findings all come from recent studies of U.S. banks. As noted above, in other countries, different institutional settings may favor models where bank moral hazard becomes more severe as capital falls. We return to the issue of bank moral hazard in Section IV below.

We now turn to papers that examine how adverse selection costs affect bank lending behavior. In these models, banks' private information about their loan portfolios leads them to face adverse selection costs when they seek funds by selling loans or by issuing uninsured liabilities or equity. Stein (1998) shows that, on the margin, these costs lead banks to prefer to fund loans by either selling off liquid securities that they hold or else issuing insured deposits. Tighter monetary policy reduces bank reserves and thus the amount of insured deposits banks can have. Since uninsured liabilities involve adverse selection costs, banks that have fewer liquid securities to draw on for cash cut back on lending. To the extent that smaller banks are less diversified across loans, their private information about loans matters more and they face greater adverse selection costs; thus, such banks should cut back lending more. Kashyap and Stein (2000) find evidence of such behavior: during monetary contractions, small banks with lower securities holdings do cut back on lending significantly more than their more liquid rivals. Ostergaard (2001) examines how lending at the U.S. state level and finds that loan supply in states with many small banks depends positively on banks' internal cash flows, whereas this is not true for states with few small banks.⁴¹

To sum up, there are a number of models that suggest that banks' monitoring or screening incentives can intensify credit cycles, both through changes in the external lending environment and through changes in banks' internal capital and funding situations. Empirical evidence from the U.S. suggests that negative shocks to bank capital have effects over and above any worsening of borrowers' conditions and that these effects are strongest for low-capital banks. Similarly, costs associated with non-core deposit funding also constrain bank lending, with the effects being strongest for small banks.

IV. Banking Panics and the Stability of Banking Systems

A key question about financial intermediaries is whether they are inherently unstable, that is, prone to banking panics? Some researchers believe that a theoretical model of the existence of financial

⁴¹ Bolton and Freixas (2000) and Gorton and Winton (2000) model another implication of adverse selection costs. In downturns, bank loans are more risky because borrowers' chance of default is higher. This increases adverse-selection costs associated with equity or uninsured liability funding, again making banks more conservative about making additional loans. We return to these issues in Section IV below.

intermediaries must simultaneously be a model of panics; banks and panics are inherently intertwined and models should reflect this. Whether this is the correct view or not clearly is at the root of public policies towards banks. In this section we first review the historical evidence on the experience of banking systems with respect to panics. This experience is quite heterogeneous, even though all these systems have banks offering the same asset and liability contracts. We also review the international experience with private bank coalitions. Private bank coalitions are a widespread feature of banking systems, though their roles vary greatly. In some systems they act as lenders-of-last-resort, while in others they are much less important. A great deal of research has been conducted on the economic history of various banking systems in different countries and different periods. Theories of banking panics should be seen in light of this research. We then review the theoretical models that have been proposed as explanations of banking panics. In the final subsection we briefly review the literature on bank regulation, including deposit insurance and capital requirements.

A. Definitions of Banking Panics and the Relation of Panics to the Business Cycle

A great deal of confusion surrounds the notion of a banking panic. One problem is the definition. There is a fairly broad range of phenomena that some researchers seem to have in mind. These are described by a number of terms, such as “financial crisis,” “contagion,” “banking crisis,” “bank run,” and “banking panic,” that are somewhat unclear. Many of the events being referred to are situations in which banking systems are weakened due to shocks, but nothing happened corresponding to a banking panic, as defined below. It is not that such broader phenomena are uninteresting or unimportant. The issue is first of all what exactly happened in these “crises” and then secondly whether such events are inherently related to the structure of bank contracts and bank capital structures. Another problem is researchers’ narrow focus on the U.S. experience and, in particular the Great Depression in the U.S. Many theorists cite Sprague (1910) as providing a description of the phenomenon they are interested in explaining and then proceed to develop a theory. Indeed, Sprague does describe U.S. panics, but those experiences appear to be a somewhat special, compared to the experiences of most other countries

With respect to a definition of a panic, it is difficult to be precise. It is tempting to define a panic in terms of an increase in the currency/deposit ratio, but because of bank suspensions of convertibility this is not accurate. Also, depending on the period over which the decline is defined, there can be large increases in this ratio without a panic having occurred. For these reasons precise definitions have not been used. Bernanke and James (1991), for example, express skepticism about this approach. Instead, definitions rely on a reading of the historical literature. Calomiris and Gorton (1991) define a banking panic as an event in which bank depositors at all or many banks in the banking system suddenly demand that banks convert their debt claims into cash, to such an extent that banks suspend convertibility. In

other words, if the depositors of a *single* bank suddenly demand cash in exchange for their deposits, this is not system-wide event. It may be called a “run,” but it is not a *banking panic*. In reality, however, panics in the U.S. tended to spread spatially and suspension happened in some cities before other cities, and sometimes not at all. Such considerations make precise definitions hard. A similar definition is given by Wicker (1996):

... we define a banking panic to be an exogenous shock whose origins can be found in any sudden unanticipated revision of expectations of deposit loss accomplished by an attempt to substitute currency for checkable deposits, a situation usually described as a run on the banks. A general loss of depositor confidence distinguishes a banking panic from other episodes of bank failures. A transfer of deposits from weak to strong banks during a bank run without any change in the public’s preference for currency does not qualify. (p. 17)

These definitions have in common the feature that a panic is a systemic event in which consumers want to hold currency in exchange for their demand deposits. The structure of the banking contract allows such withdrawals from banks by consumers, and these withdrawals, or attempted withdrawals, is the precipitating event.⁴²

Applying the definition to Pre-Federal Reserve history in the U.S. is not easy. Calomiris and Gorton (1991) identify six panics in the United States prior to 1865, seven during the National Banking Era, and finally the Great Depression (discussed below). Table 3 shows the business cycle chronology and the dates of the panics in the United States during the National Banking Era. Prior to the National Banking Era, there were panics in 1814, 1819, 1837, 1839, 1857, and 1861 (see Calomiris and Gorton (1991)). After the National Banking Era ends, with the founding of the Federal Reserve System in 1914, there were the panics associated with the Great Depression. Sprague (1910) labels 1873, 1893, and 1907 as major panics. Kemmerer (1910) identifies six major panics and fifteen minor panics between 1890 and 1908. Kemmerer’s major panics include 1873, 1893, and 1907, but he adds 1899, 1901, and 1903. Wicker (2000) agrees on 1873, 1893, and 1907, and calls 1884 and 1890 “incipient” panics. The point is that there is no consensus about the events that should be called “panics” in the U.S. prior to the founding of the Federal Reserve System. While there are common elements, corresponding to the definition, each episode has some unique features. There are important papers on some individual panic episodes, e.g., Wicker (1980), Donaldson (1993), Moen and Tallis (1992), Calomiris and Schweikart (1991), and White (1984). Wicker (2000) details each of the U.S. episodes.

Definitions of other crisis phenomena abound. An older literature links problems with the banking system to broader events and the business cycle, e.g. Fisher (1932), but is not specific about the

⁴² Many authors have discussed definitions of “banking crisis.” See, for example, Grossman (1994) and Bernanke and James (1991).

details. Similarly, Bordo (1986) lists what he describes as “key elements of a financial crisis.” One of the key elements listed is “bank runs precipitated by ... threats to solvency” (p. 191). But, there are many other elements listed as well suggesting a link between panics and the macroeconomy. Grossman (1994) argues that historically “banking crises” included one of three elements: (1) a high proportion of banks failed; (2) an especially large or important bank failed; or finally, (3) that government intervention prevented the failures associated with (1) or (2). There are many other examples of attempts at definitions. For the most part, the same events are identified.

There are some notable features to Table 3. First, the table shows the proximity of the panic to the last business cycle peak. The timing of the panics in the U.S. prior to the National Banking Era is similar; see Calomiris and Gorton (1991). The percentage change in pig iron production is a measure of real economic activity. As might be expected, the currency deposit ratio rises sharply in a panic. Remarkably, the losses on deposits and the fraction of (national) banks failing during panics are very small. This is due to the activity of private bank clearinghouses, discussed below. It is, however, worth emphasizing that the actual historical experience of panics, small losses on deposits and few bank failures, seems at odds with the widely-held view of panics, mostly based on the experience of the U.S. Great Depression.

In the twenty-five year period following World War II banking crises all but disappeared. Bordo and Eichengreen (1999) find only one banking crisis between 1945 and 1971 in their sample of twenty-one industrial and emerging market countries. In the 1980s and 1990s, however, the International Monetary Fund counts fifty-four crises in member countries between 1975 and 1997, and the World Bank lists a larger number (see IMF (1998), and Caprio and Klingbiel (1996) for World Bank estimates). In the case of recent international banking “crises” it seems difficult, at least so far, to determine whether these events are panics or whether the banking systems suffered severe losses due to macroeconomic shocks. Five recent studies, for example, all offer different definitions of a “banking crisis.” See Caprio and Klingbiel (1996), Demirgüç-Kunt and Detraiache (1998), Dziobek and Pazarbasioglu (1997), Kaminsky and Reinhart (1999), and Lindgren, Garcia, and Saal (1996). Caprio and Klingbiel (1996) is the root study for many of the lists of crises. Basically, their definition focuses on loan losses and the extent to which the net worth of the banking industry has eroded. If most or all of the capital in the banking system is gone, then there is a crisis.

It is clear that there are situations in which a banking system faces a common shock of sufficient magnitude to bring the soundness of the banking system into question. For example, the U.S. savings and loan debacle of the 1980’s (see e.g., Brumbaugh (1988), Kane (1989), Barth (1991), and White (1991)), and the current situation of the Japanese banking system, were caused by deregulation (see Hoshi and Kashyap (1999)). These events may be called “crises,” but there were no banking panics involved. A

systemic shock to the banking system, whether it is part of a broader macroeconomic downturn or exchange rate shock, or a shock specific to the financial sector, do not obviously call into question the contractual design of financial intermediation. Nevertheless, these crises raise a number of issues about banks, which are discussed further below.

Despite the large number of “crises” it seems that bank panics and bank runs (on individual banks, but not systemic) have been relatively rare. But, this is due to a rather narrow definition of “panic.” Recent history suggests that “panic” and bank crisis” are rather difficult to distinguish, and certainly more research is needed. Lindgren, Garcia, and Saal (1996) provide the most extensive study; they analyze all IMF member countries from 1980 to 1995. By their definition 133 member countries of the 180 members experienced “crises” or significant problems in the banking sector during the period.⁴³ Their analysis then focuses more specifically on 34 countries (19 developing economies, eight transition economies, and seven developed economies). They single out 36 cases of banking crises. In this sample there were seven cases of panic. The study highlights the difficulties in distinguishing “panics” from other phenomena. Indeed, unlike nineteenth century America, banking “crises” or “panics” may well be more complicated now by depositor beliefs about implicit or explicit deposit insurance. Thus, although it is tempting to say that the definition of a “banking panic” is ultimately an empirical matter, the phenomena of interest are very complicated and seem likely to vary for many reasons, not the least of which is data availability, making any definition problematic. Nevertheless, some important empirical regularity has been found. We turn next to this evidence.

It appears that there is an important business cycle component to the timing of panics. Gorton (1988), studying U.S. panics, provides evidence that panics come at or near business cycle peaks. See Table 3. Mishkin (1991) summarizing the findings of his study states that “with one exception in 1873, financial panics always occurred after the onset of a recession” (p. 96). Also, see Donaldson (1992). In the recent international context Demirgüç-Kunt and Detragiache (1998) study a large cross section of countries during the period 1981-1994 and also find that banking crises are more likely to occur with the onset of recession. Lindgren, Garcia, and Saal (1996) also find that banking crises of the modern era are related to the business cycle.

In terms of U.S. history, a seasonal factor in the timing of panics is noted by Andrew (1907), Kemmerer (1910), Miron (1986), Canova (1991) and Donaldson (1992), among others, though Calomiris and Gorton (1991), and Wicker (2000), among others, dispute the evidence. All these studies focus on

⁴³ Lingren, Garcia, and Saal (1996) define a “crisis” as a situation where a large group of financial intermediaries have liabilities exceeding the market value of their assets, and the economy experiences bank runs or significant withdrawals, some financial firms collapse, and there is government intervention. If the banking system is unsound, but there is no crisis, it is termed a “significant problem.”

the U.S. experience. While the timing of panics is, broadly speaking, clear, it is not clear that seasonal money demand shocks *caused* panics. At root the problem is that there are a small number of panic observations. Miron (1986) and Canova (1991) focus on interest rate movements and the inelastic supply of money in the period before the Federal Reserve System. Miron attributes the decreased occurrence of financial crises after 1914 to Federal Reserve activities, because the Fed essentially smoothed out seasonal interest rate movements. Canova argues that the decline in the seasonal pattern of interest rates to Fed activity has been overemphasized. Miron, Mankiw, and Weil (1987) and Barsky, Miron, Mankiw, and Weil (1988) also compare the pre-Fed period with the post-Fed period with respect to interest rate behavior.

B. Panics and the Industrial Organization of the Banking Industry

Another stylized fact about banking panics is summarized by Calomiris (1993a):

International comparisons of the incidence and costs of banking panics and bank failures, and comparisons across regulatory regimes within the U.S., clearly document differences in banking instability associated with different regulatory regimes. The central lesson of these studies is that instability is associated with some historical examples of banking that had common institutional characteristics; it is not an intrinsic problem with banking per se.... the single most important factor in banking instability has been the organization of the banking industry (p. 21).

That the industrial organization of the banking industry is a critical determinant of the propensity for an economy to experience panics has been confirmed in a large literature on the historical and international experience of banking panics. Bordo (1985, 1986), Calomiris and Gorton (1991), and Calomiris (1993a) survey much of this literature and provide some new evidence on the causes of panics.

Calomiris (1993a) examined Scotland, England, Canada, Australia, and the U.S. Bordo (1986) studies the experiences of six countries (U.S., U.K., Canada, Sweden, Germany, and France) over the period 1870 to 1933. One of his conclusions is that most severe cyclical contractions in all the countries are associated with stock market crises, but not with banking panics, except for the United States. He notes that: "In contrast with the U.S. experience, the five other countries in the same period developed nationwide branch banking systems consolidating into a very few large banks" (p. 230). Bordo (1985) surveys banking and securities market panics in six countries from 1870 to 1933 and concludes that: "the United States experienced panics in a period when they were a historical curiosity in other countries" (p. 73). Grossman (1994) examines the experience of Britain, Canada, and ten other countries during the Great Depression to determine the causes of the "exceptional stability" exhibited by their banking systems. He considers three possible explanations: the structure of the banking system, macroeconomic policy and performance, and the behavior of the lender of last resort. He concludes that banking stability is the product of exchange-rate policy and banking structure.

Cross section variation in the United States is also interesting because some states allowed branch banking and some states did not. In the United States, states that allowed branching experienced lower failure rates in the 1920s and smaller banks were more prone to failure (see Bremer (1935) and White (1983, 1984)). Studying this cross section of state experience, Calomiris (1990) reaches the same conclusion about the importance of branching: “States that allowed branch banking saw much lower failure rates- reflecting the unusually high survivability of branching banks... From 1921 to 1929 only 37 branching banks failed in the United States, almost all of which operated only one or two branches. Branching failures were only 4 percent of branch-banking facilities, almost an order of magnitude less than the failure rate of unit banks for this period” (p. 291). Wheelock (1992a) compares the experiences of banks in different states during the 1920s in the U.S. and finds that states that allowed branch banking had fewer failures. Also, see Wheelock (1995). Calomiris (1993a) reviews more evidence.⁴⁴

The importance of industrial organization of the banking system for the incidence of panics is illustrated by a comparison of the U.S. experience with the Canadian experience, which has been the focus of research by a number of scholars. The economies are similar and close in proximity, but Canada is a system that historically has consisted of a small number of highly branched banks, in contrast to the American system of many banks that are not branched across state lines, and sometimes not even within the state. Haubrich (1990), Bordo, Rockoff, and Redish (1994, 1995), and White (1984), among others, have studied the two systems. The contrast in experience is dramatic, as summarized by Bordo, Rockoff, and Redish (1994): “There is an immediate and important difference between the Canadian and United States banking systems. The Canadian experience has been one of considerable stability. There has been only one major bank failure since World War I, and there were no failures during the Great Depression. In contrast, the American system has been characterized by a number of periods of instability. Rates of bank failures were high in the 1920s, and of course the entire system collapsed during the 1930s” (p. 325). Similarly, White (1984) writes: “In Canada, from 1920 to 1929, only one bank failed. The contraction of the banking industry was carried out by the remaining banks reducing the number of their offices by 13.2 percent. This was very near the 9.8 percent decline in the United States...In spite of the many similarities with the United States, there were no bank failures in Canada during the years 1929-1933. The number of bank offices fell by another 10.4 percent, reflecting the shocked state of the economy; yet this was far fewer than the 34.5 percent of all bank offices permanently closed in the United States” (p. 132).⁴⁵

⁴⁴ Carlson (no date) empirically examines U.S. banks during the 1920s and argues that branched banks were less likely to survive because they held riskier portfolios.

⁴⁵ An alternative point of view about the Canadian experience during the Great Depression is that of Kryzanowski and Roberts (1993) who argue that all of the major banks in Canada were insolvent during the Great Depression, but

C. Private Bank Coalitions

Bank coalitions, formal and informal, are an essential part of the industrial organization of the banking system. The existence or nonexistence of coalitions, the extent of their activity, and their interaction with the government are related to the likelihood of panic and to the resolution of panics if they do occur.

Banks are involved with each other because of the clearing of their liabilities. Banks mutually hold claims on each other because of their depositors writing checks and the banks need mechanisms for efficiently netting these claims. Historically, this led to the establishment of clearinghouses, joint associations of banks that had the purposes of organizing the netting of interbank claims. But these coalitions developed into institutions with many more functions. And, more generally, other types of coalitions, sometimes less formal, sometimes organized around a single large bank or even the government's central bank, seem to characterize the banking systems of many countries in many different historical periods. The extent to which these private bank coalitions exist, or existed historically, seems related to the industrial organization of the banking system and the incidence of bank panics.

The U.S. experience with banking panics appears to be an outlier in that it experienced fairly regular panics during the 19th century when few other economies did, as discussed above. Until the last few years, there have been a very large number of rather small, undiversified banks in the U.S. The research cited above suggests that these two facts are linked. U.S. banking history has also been intertwined with the development of the private clearinghouse system. Clearinghouses are private associations of banks that formed in major cities, spreading out across the country during the 19th century. On the U.S. clearinghouse system see Andrew (1908b), Cannon (1910), Gorton (1984, 1985), Gorton and Mullineaux (1987), Timberlake (1984), Sprague (1910), Moen and Tallman (2000), and Wicker (2000), among others.

The U.S. clearinghouse system developed over the course of the 19th century. In particular, the clearinghouses developed methods for coping with banking panics. At first the clearinghouse organized a method of pooling or equalizing reserves. Wicker (2000) argues that such action prevented panics in 1860 and 1861. By the end of the century they had invented a method of turning illiquid loan portfolios into private hand-to-hand money that could be handed out to depositors in exchange for their demand deposits during times of panic. This money, called "clearinghouse loan certificates," originated in the interbank clearing system as a way to economize on cash during a panic. During a banking panic member banks were allowed to apply to a clearinghouse committee, submitting assets as collateral in exchange for

that there was no banking crisis because of implicit support from the government. This viewpoint is disputed by Carr, Mathewson, and Quigley (1995). Also, see Kryzanowski and Roberts' (1999) rejoinder.

certificates. If the committee approved the assets, then certificates would be issued only up to a percentage of the face value of the assets. The bank borrowing against its illiquid assets would have to pay interest on the certificates to the clearinghouse. The certificates could then be used to honor interbank obligations where they replaced cash, which instead could be used to pay out to depositors. The clearinghouse loan certificate process is the origin of the discount window (and is described in detail in the above cited sources), and serves the same function. Notably, the loan certificates were the joint obligations of the clearinghouse member banks; the risk of member banks defaulting was shared by allocating member liabilities in proportion to member bank capital. Thus, the certificates implemented a risk-sharing device, where the members jointly assumed the risk that individual member banks would fail. In this way, a depositor who was fearful that his particular bank might fail was able to insure against this event by trading his claim on the individual bank for a claim on the portfolio of banks in the clearinghouse. This was the origin of deposit insurance. In order for this to work, the clearinghouses in the U.S. developed bank examination and supervision methods, as well as reporting systems for bank information to be made public on a regular basis.

During the Panics of 1873, 1893, and 1907 the clearinghouse loan certificate process was extended, in increasingly sophisticated ways, in a radical innovation. In particular, the clearinghouse loan certificates were issued directly to the banks' depositors, in exchange for demand deposits, in denominations corresponding to currency. The amount of private money issued during times of panic was substantial. During the Panic of 1893 about \$100 million of clearinghouse hand-to-hand money was issued (2.5 percent of the money stock). During the Panic of 1907, about \$500 million was issued (4.5 percent of the money stock). See Gorton (1985). If the depositors would accept the certificates as money, then the banks' illiquid loan portfolios would be directly monetized.

The U.S. clearinghouse system was not the only private central bank-like institution. Before the U.S. Civil War, coincident with the beginnings of the clearinghouse system, the Suffolk Bank of Massachusetts was the focal point of a clearing system and acted as a lender-of-last-resort during the Panic of 1837. The Suffolk banking system operated in New England from 1825 to 1858 and was the first region-wide clearing system in the U.S. The Suffolk system was unique in using a net clearing system (see Rolnick, Smith, and Weber (1998a,b)). Rolnick, Smith, and Weber (1998b) argue that during the Panic of 1837 the Suffolk Bank it essentially acted as a lender-of-last-resort. Also, see Mullineaux (1987), Calomiris and Kahn (1996), Rolnick, Smith, and Weber (1998a, 1998b), and Bodenhorn (1998).

Bank coalitions are also not unique to the United States, though the extent of their activities varies enormously across countries. Most countries did not experience banking panics as frequently as the U.S., but there are many examples of bank coalitions forming on occasion in other countries as well. For example, the Clearing House of Montreal was maintained by the Canadian Bankers' Association and,

according to Watts (1972), was officially recognized in 1901 ‘as an agency for the supervision and control of certain activities of the banks’ (p. 18). According to Bordo and Redish (1987) “the Bank of Montreal (founded in 1817) emerged very early as the government’s bank performing many central bank functions. However, the Bank of Montreal never evolved into a full-fledged central bank as did the Bank of England (or the government’s bank in other countries) perhaps because of the rivalry of other large Canadian banks (for example the Royal Bank).” See Watts (1972), Haubrich (1990), and Breckenridge (1910).

The pattern of the Bank of Montreal (and earlier precursors like the Suffolk Bank) in which the bank coalition is centered on one large bank, is quite common. Another common feature is the cooperation of a (perhaps, informal) coalition of banks with the government to rescue a bank in trouble or stem a panic. For example, major Canadian banks joined with the Canadian government to attempt a rescue of the Canadian Commercial Bank in March 1985. See Jayanti, Whyte, and Do (1993). Similarly, in Germany the Bankhaus Herstatt was closed June 26, 1974. There was no statutory deposit insurance scheme in Germany, but the West German Federal Association of banks used \$7.8 million in insurance to cover the losses.

D. Are Banks Inherently Flawed Institutions?

Diamond and Dybvig (1983), reviewed above, is the most important paper on banking panics. Recall that this model combines preference shocks for early or late consumption with investment opportunities that are long-term. The model becomes one of banking panics with the additional assumption that depositors face a first-come-first-served rule, that is, a sequential service constraint. The assumption of this rule, combined with the irreversibility of long-term investment, means that if all agents decide to withdraw at the interim date, then those in front of the line will receive more than those at the end of the line. Consequently, a panic corresponds to an equilibrium in which agents believe, for whatever reasons, that other agents are intending to withdraw their deposits at the interim date. Such beliefs are self-fulfilling because the best response to the belief that other agents are intending to withdraw is to withdraw oneself. As Wallace (1988) points out, without the assumption of sequential service the model does not generate panics.

The theory of banking panics in Diamond and Dybvig intuitively corresponds to what many view as an irrational element of banking panics. The theory articulates the view that banks are inherently unstable arrangements. The theory is silent on the issue of what kinds of events would cause agents to have beliefs that other agents are going to withdraw their deposits; the events are “sunspots.”⁴⁶ So, it is

⁴⁶ Postlewaite and Vives (1987) modify the Diamond and Dybvig model so that runs are an equilibrium phenomena, though see the comments of Jacklin (1989).

not testable; see Gorton (1988). Moreover, the “sunspots” have to concern all banks in a banking system, in order to generate a system-wide panic, rather than a run on a single bank. The theory is consistent with the cross-country variation in panic incidence only in the trivial sense that some countries have had system-wide sunspots, while other have not.

As an explanation of panics, the theory amounts to the assertion that the sequential service constraint is an inherent feature of reality. This is clearly unsatisfying in the sense that the underlying reality that would give rise to the sequential service constraint is not modeled. Recognizing this several researchers have tried to address this shortcoming.

Wallace (1988) presents a model that rationalizes the existence of the sequential service constraint. He assumes the basic Diamond and Dybvig set-up where consumers’ preferences are such that they need to have assets that can be “cashed” at optional times and where long-term investments are irreversible. The key new assumption is that consumers are isolated in the sense that they cannot coordinate their withdrawals or the amounts the bank will give each of them upon withdrawal. Consumers know where the bank is located and so they can go withdraw. But, their isolation means that at random times they will withdraw and there is no possibility for coordination. Sequential service is an outcome of the assumption that people are isolated from each other at the interim date, the date at which they learn their preferences for early or late consumption. As Wallace explains the assumption “...is consistent with the notion that people hold liquid assets because they may find themselves impatient to spend when they do not have access to asset markets, in which they can sell any asset at its usual price.” Wallace shows that the details of the model have implications for Diamond and Dybvig’s arguments about deposit insurance.

In Wallace’s formulation, following Diamond and Dybvig, bank liabilities do not circulate as a medium of exchange. Instead, when a consumer learns that he has preferences for early consumption, he withdraws from the bank to satisfy those needs. There is no purchase of consumption goods using bank liabilities as money. In the model, the bank is, in effect, also the store. But, in cash-in-advance type models or search-theoretic models, consumers buy goods with bank liabilities without any need to return to the bank to withdraw. This is the essence of a medium of exchange. And that is how bank notes and bank deposits work. While consumption smoothing, and the demand for consumption insurance, are likely important features of reality, it is not clear that consumption smoothing is really a meaningful sense in which bank liabilities are a medium of exchange.

Calomiris and Kahn (1991) rationalize sequential service as an optimal contractual response to depositors being uninformed about the value of their bank’s assets. This information can be produced,

but at a cost. As discussed above, Calomiris and Kahn (1991) assume that information about the banker's decisions must be produced at a cost. Individual depositors who expend resources to produce the information will get into line to withdraw at the bank first. The sequential service constraint, i.e., a first-come-first-served rule, rewards those depositors in line first, so the information-producing depositors will recover more than other depositors. (As noted above, Jean-Baptiste (1999) is also relevant here.)

Green and Lin (1999, 2000) critique the Diamond and Dybvig model. They argue that the Diamond and Dybvig deposit contract is one of the feasible arrangements in the environment of their model. They argue that there are other arrangements that implement an efficient allocation without bank runs. In particular, agents are allowed to send messages of their type, i.e., their consumption timing preferences, to the bank. It turns out that truth-telling is a strictly dominant strategy. Green and Lin do not argue that their contracts are necessarily realistic, but that "Our results imply that environmental features from which Diamond and Dybvig's model abstracts are crucial to a full understanding of banking instability."

Reflecting its importance in the literature, there have been many extensions of the Diamond and Dybvig model. Goldstein and Pauzner (1999) is an important one. They address some of the more fundamental problems with the multiplicity of equilibria in Diamond and Dybvig's model. Agents, for example, do not consider the possibility of a run at the initial date when they deposit in the bank, so the subsequent date is not part of a rational expectations equilibrium. As a result, it is not clear that the deposit contract is optimal. Since there is no theory of selection from the multiplicity of equilibria, the theory is empirically vacuous. Goldstein and Pauzner modify the Diamond and Dybvig model by assuming that consumers do not have common knowledge about the fundamentals; rather they only obtain private signals. A private signal provides information about the final payoff on the bank's portfolio. But, the signal, though private, allows an agent to draw inferences about what signals other agents received. If an agent receives a high signal, the agent believes that other agents are likely to have obtained high signals as well. In this environment, there is a unique equilibrium and the probability of a bank panic is related to news about fundamentals and to the promised payoff on the demand deposit. Morris and Shin (2000)'s set-up is similar.

E. Information-Based Theories of Panics

Another view of panics sees them as rational events where depositors are essentially reacting to new information that is not bank-specific. The basic idea is that depositors learn some information that is relevant for assessing the risk of their bank, but is not specific to any particular bank; it is macroeconomic information. Nevertheless, the macroeconomic information is negative, i.e., a recession is looming, and risk averse depositors, revising their assessment of bank risk, may rationally decide to withdraw their

deposits. In other words, there is consumption smoothing because the consumers realize that a recession is coming and consequently will need draw down their saving. They withdraw from their bank because they want to avoid losing their savings during the recession. Gorton (1987), Chari and Jagannathan (1988), Jacklin and Bhattacharya (1988), Allen and Gale (1998), and Gorton and Huang (2001) all have versions of this basic story.

In Gorton's (1987) model bank portfolio are subject to both idiosyncratic shocks and economy-wide shocks, but only the latter are observable by depositors.⁴⁷ Depositors update their beliefs about the state of bank portfolio based on the economy-wide shocks. Sometimes they seek to withdraw their deposits, a panic, because of fears that the banking system has a low quality portfolio, although they do not know whether their individual bank is in such a situation or not. Banks suspend convertibility to communicate information to depositors. In the model of Jacklin and Bhattacharya (1988) the bank cannot observe the true liquidity needs of depositors (i.e., depositor type) while depositors do not observe the quality of bank assets. A unique (i.e., there are not multiple equilibria) bank run occurs when some of the depositors receive bad news about the realized state of the bank assets. In Chari and Jagannathan (1988) the proportion of depositors wanting to consume early and the state of bank assets are also stochastic. Depositors can observe the initial size of the line of depositors at the bank and act conditional on this observation. The line may be especially long because some depositors received news that the bank's asset portfolio is in a bad state. But, this can be confused with a high proportion of early consumers. Thus, sometimes runs result in liquidating banks that do not have low quality asset portfolios.

Allen and Gale (1998) modify Diamond and Dybvig's model so that panics are related to the business cycle, rather than unexplainable events caused by "sunspots." The two important modifications are, first, that there is aggregate uncertainty about the value of the long-term assets held by banks. This assumption is introduced to link panics to business cycles, modeled as this aggregate risk. Second, the assumption of sequential service is dropped, as unrealistic. Consumers face consumption risk, as in Diamond and Dybvig. They can, however, observe a signal, a "leading indicator," that perfectly predicts the realization of the payoff on the long-term asset, but is not contractible. First best risk sharing can be achieved if contracts could be written on this signal. With noncontingent deposit contracts, but with the signal observable, panics can implement first best risk sharing when there is no cost to early withdrawal. Roughly speaking this is because when the long-term asset is worth zero, the bank's remaining investment is shared equally among the depositors because there is no sequential service. If there is a cost to early withdrawal, then the panic is inefficient and there is a role for the government.

⁴⁷ Gorton assumes that banks exist and he assumes the structure of the contracts.

An important difficulty with the information-based view of panics is that it views the problem as inherent in the banking system, like Diamond and Dybvig. Gorton and Huang (2001) present an information-based explanation, with same basic source of confusion between aggregate and idiosyncratic shocks as the above theories, but in the context of the industrial organization of the banking system. In addition to the asymmetric information setting, bankers may engage in moral hazard if their bank is in a low state. So, depositors must monitor banks. In their model, a panic is a manifestation of depositors monitoring their banks by withdrawing. But, this is not necessarily a panic, in the sense that only in systems of many small banks does it lead to banks being liquidated. Systems of large banks are monitored, and there may be withdrawals, but not panics. This is consistent with, for example, the comparison between the U.S. and Canadian experiences. Gorton and Huang's main result is to show how a coalition resembling a clearinghouse endogenously arises.

F. Other Panic Theories

There are a number of papers that study banking panics in the context of the entire banking system. These are models that focus on interdependencies between banks, either through interbank lending or through interbank clearing systems. The first of these was Bhattacharya and Gale (1987) who extend the Diamond and Dybvig model to examine the effects of preference shocks when there are many banks. Each individual bank faces uncertain liquidity demands, but there are many banks and there is no aggregate uncertainty. Bhattacharya and Gale show that when banks meet unanticipated demands for liquidity by borrowing in the interbank market, there is a free rider problem so that banks under invest in liquid assets. The basic result is that an unregulated interbank market for resources can be improved upon by a central bank that offers restricted opportunities to borrow and lend. Allen and Gale (2000) focus on the transmission of a shock in one location to other locations, suggesting that "contagion" is an important feature of financial crises. Other papers that examine crises and interbank links include Smith (1984, 1991), Donaldson (1992), Champ, Smith, and Williamson (1996), and Rochet and Tirole (1996), among others.

Williamson (1988) is a quite different model of panics. It is a multi-period extension of Boyd and Prescott (1986). His agents are risk neutral so there is no demand for consumption-smoothing insurance, but they do have random preferences and there are both a long-term and a short-term investment opportunity. With decentralized trade there is a possible lemons problem in that agents selling high quality capital cannot distinguish themselves from those selling low quality capital. Banks are large coalition of agents that overcome this information problem. However, the bank must allow for early withdrawals due to the random preferences of the depositors, so it issues demand deposits. The bank can achieve an allocation that is strictly preferred to the decentralized capital markets allocation by all agents

in some states of the world. But, in other states of the world agents are indifferent between the two allocations. In states where agents are indifferent, the bank may dissolve. Williamson interprets this as a bank failure or collapse of the banking system.

G. Tests of Panic Theories

On the basis of the stylized facts about cross-country banking history, reviewed above, it would seem straightforward to observe that banks are not fundamentally flawed institutions. In fact, it does not seem to be an exaggeration to say that most of the theoretical work on panics has been motivated by the U.S. experience, which has then been incorrectly generalized. Panics simply are not a feature of most economies that have banks. The world is more complicated; industrial organization seems to be at the center of the incidence of panics. Not surprisingly, therefore, almost all the empirical work on panics has been on the U.S. experience. Until bank “crises” around the world in the last ten years, there simply has not been much else to study. Clearly, from the point of view of public policy and the design of bank regulation and central bank lender-of-last-resort activity it is important to distinguish between the two views of banking panics outlined above, if only because policies should be in place that are workable in economies where the banking system is susceptible to panics.

With regard to testing, a major difficulty is that Diamond and Dybvig (1983) is not a testable theory, since any observed phenomenon is consistent with “sunspots.” Instead, empirical investigations of panics have focused on the timing of panics in the U.S., checking for patterns that would be consistent with the information-based theories of panics. Importantly then, there are no formal tests that have been conducted that test one hypothesis against any particular alternative. Rather, there has been a variety of empirical work studying the times series behavior of the deposit-currency ratio, interest rates, and other variables, as well as studies of individual panic episodes. Empirical investigations include Gorton (1988), Donaldson (1992), Mishkin (1991), Park (1991), Calomiris and Gorton (1991), and Calomiris and Mason (1997). Wicker (1980, 1996, 2000), Donaldson (1993), Moen and Tallis (1992, 2000), Calomiris and Schweikart (1991), and White (1984) are also relevant.

Gorton (1988) argues that demand deposits are risky, like other securities, and that depositor behavior should correspond to consumption smoothing behavior based on the aggregate information available to them at the time. The basic idea is that when depositors receive information forecasting a recession they know that they will be dissaving, drawing down their bank accounts. But, their banks are more likely to fail during recessions, so they withdraw in advance to avoid such losses. Empirically Gorton analyzes the period 1863-1914 (and also the Great Depression) and shows that the post Civil War period behavior of the deposit-currency ratio displays the hypothesized timing. In fact, on every single occasion that a leading indicator of recession crosses a threshold, there is a recession. The basic

conclusion is that there is nothing special about panic dates compared to nonpanic dates in terms of the behavior of the deposit-currency ratio. While the “sunspots” theory cannot be rejected, the conclusion is that if there are “sunspots” they must be consistent with estimated reduced for description of the deposit-currency ratio.

Donaldson (1992) revisits the issues raised by Gorton (1988) using weekly data, compared to Gorton who used data from the Call Reports, reported five times a year. Donaldson confirms that there are periods that predictably (from the point of view of an econometrician) correspond to instances when panics are more likely to occur, but that the exact starting dates during such periods are unpredictable. One interpretation of his results is that, although panics do tend to occur at business cycle peaks, there is some unknown triggering event that is not predictable, perhaps a “sunspot,” but the data are not fine enough to say anything further.

Calomiris and Gorton (1991) first examine whether pre-panic periods were unusual. That is, do measures of seasonal flows of reserves and deposits show any evidence of tightness or shocks? There is no such evidence. The onset of panics is after the money flows associated with planting and harvesting. However, measures of real economic activity, in particular, the liabilities of failed businesses do decline. Also, stock prices declines did precede panics. Calomiris and Gorton write: “if one posits that the simultaneous violations of thresholds for percentages of real stock price decline and commercial failure increase are sufficient for panic, one can predict panics perfectly” (p. 144). Second, Calomiris and Gorton analyze bank liquidations and deposits losses during and after panics. Basically, there is no evidence of banks failing due to the panic. Rather, weak banks, by pre-panic measures, fail. Finally, Calomiris and Gorton look at sufficient condition for panics to end. The basic point here is that availability of liquidity to satisfy depositor demands does not seem to end panics, with the availability of the discount window during the Great Depression being the outstanding example. Rather, panics end when information becomes available, information typically produced by clearinghouses or the government about which banks are weak.

Mishkin (1991) also studies the National Banking Era in the U.S., as well as the Great Depression. He focuses on the timing of events and financial variables to distinguish between the monetarist and asymmetric information-based views of bank panics.⁴⁸ For example, an observation that interest rate spreads widen and stock market prices decline just prior to the panic, rather than a disruption in the financial markets following the panic, is viewed as evidence in favor of the information theory.

⁴⁸ Some tests of theories of panics have focused, in part, on discriminating between the monetarist views of crises, associated with Friedman and Schwartz (1963) and the information-based theories, discussed above. Monetarists do not proposed a theory of panics, but note that panics reduce the money supply since withdrawals decrease the money multiplier. Thus, monetarists propose a central bank that acts as a lender-of-last-resort. However, if panics are due to asymmetric information, then monetary policy alone cannot eliminate panics or mitigate their effects.

Mishkin analyses each panic episode in U.S. history starting with the Panic of 1857 and concludes that “the asymmetric information approach to financial crises explains the timing of patterns in the data and many feature of these crises which are otherwise hard to explain” (p. 104). Mishkin’s evidence is consistent with that of Gorton and Donaldson.

If asymmetric information is at the root of panics, then panics should end when depositors receive credible information about individual bank shocks. Park (1991) argues that empirically the evidence suggest that panics during in the U.S. did end when information about banks was provided to the public. He focuses on the actions of private bank clearinghouses and the government in providing credible information and concludes: “this empirical finding confirms the crucial link between bank-specific information and bank panics” (p. 285). Calomiris and Mason (1997) study the June 1932 bank panic in Chicago. They compare the attributes of banks that failed during that event to those that did not fail. They conclude that: “the failures during the panic reflected the relative weakness of failing banks in the face of a common asset value shock rather than contagion” (p. 881). “Private cooperation by the Chicago clearing house banks appears to have been instrumental in preventing the failure of at least one solvent bank during the panic” (p. 864). Other papers on the Great Depression are discussed below.

Demirgüç-Kunt and Detragiache (1998) study banking crises in a large sample of countries internationally during the period 1981-1994.⁴⁹ Their basic results “reveal strong evidence that the emergence of banking crises is associated with a deteriorated macroeconomic environment. Particularly, low GDP growth, high real interest rates, and high inflation significantly increase the likelihood of systemic problems in our sample; thus crises do not appear to be solely driven by self-fulfilling expectations as in Diamond and Dybvig (1983). This is consistent with the evidence presented in Gorton (1988) on determinants of bank runs in the U.S. during the eighteenth century” (p. 3-4). While this study is the only study of an international cross section of countries, and therefore is unique, it did not include any variables that might capture cross section variation in the industrial organization of the banking system, which the studies reviewed above suggest would be important. However, the study does include a dummy variable for the presence of explicit deposit insurance and an index of the quality of law enforcement. The presence of explicit deposit insurance significantly increases the likelihood of a banking crisis, while the “law and order” index shows that more “lawful” countries are less likely to have a crisis.

It should be emphasized that none of the above work constitutes a test in a statistical sense. While the evidence is suggestive, the basic finding that panics are associated with business cycle

⁴⁹ The sample ranges from sixty five to forty five countries in different regression, depending on data availability.

downturns does not rule out any theory of panics. In fact, in the Goldstein and Pauzner (1999) and Morris and Shin (2000) extensions of Diamond and Dybvig, the business cycle timing is completely consistent with the self-fulfilling nature of a panic. Morris and Shin write of their extension that: “The theory suggests that depositors will indeed withdraw their money when the perceived riskiness of deposits crosses a threshold value. But, nevertheless, the banking panic is self-fulfilling in the sense that individual investors only withdraw because they expect others to do so” (p. 14-15).

In fascinating recent research Kelley and Ó Gráda (2000) and Ó Gráda and White (2001) study the patterns of withdrawals from a single bank, the Emigrant Savings Industrial Bank, during bank runs in 1854 and 1857. Study at this level of detail can address questions concerning whether depositors respond to a signal that causes them to all crowd at once at the bank, or whether the run builds up slowly. Do rich or poor, less sophisticated or uninformed, line up first? And so on. Kelley and Ó Gráda (2000) find that in 1854 the bank panic followed ethnic patterns, particularly within the Irish community. Ó Gráda and White (2001) document time patterns in withdrawals (or account closings). There are responses to bad news, but there are elements of contagion as well. Moreover, the patterns are different in 1854 and 1857. In 1857, unlike '854, the run was led by business leaders and apparently sophisticated agents, followed by less informed depositors.

What is more important, however, is to keep in mind that, while to date it has not been possible to discriminate between panic theories with data, it is clear that the *prima facie* evidence is against theories that inherently intertwine banks and panics. The previous evidence about the industrial organization of the banking system strongly suggests that, at least historically, there is no necessary link between banks and panics.

H. The Banking Crises During the Great Depression

The Great Depression was a momentous event, resulting in vast institutional change in the U.S., and casting a shadow over the discipline of economics. In the case of U.S. banking, the Great Depression led to enormous change. Deposit insurance was enacted, and the Glass-Steagall Act, separating commercial banking from investment banking, was also passed in response to this event. Much has been written on the Great Depression (e.g., Kindleberger (1973), Temin (1989), Eichengreen (1992), Bordo, Goldin, and White (1998), and James (2001)) and we do not survey this vast literature here. Even the literature more narrowly focused on banking and financial factors during the Great Depression is large. Our focus is only on issues concerning the experience of banks, and banking systems, during the Great Depression, to the extent that these can be separated from other issues. Understanding the experience of banks during the Great Depression is important because much of bank regulatory policy emanates from

this experience, rather than from the earlier panic experience. But, the Great Depression was a very different banking crisis than the earlier episodes in the U.S.

The panics during the Great Depression in the U.S. were certainly different from the previous episodes in terms of the extent of bank failures and losses on deposits. In the U.S. more than nine thousand banks failed during the Great Depression, between 1930 and 1933. That amounts to about one third of the total number of banks in existence at the end of 1929. In previous panics, the numbers of banks failing were miniscule, as shown in Table 3. Internationally, there was a variety of experience with regard to bank failure and system collapse. While the experience in much of Europe was similar to that in the U.S., in that banking systems did collapse (e.g., see Beyen (1949), James (1986), and Kindleberger (1973)), many countries experienced the Great Depression without banking crises (e.g., the United Kingdom, Canada, Czechoslovakia, Denmark, Lithuania, Holland, and Sweden). The international cross section variation with regard to banking crises during the Great Depression, and the magnitude of the failures in the U.S. are puzzles. The experience of Canada, discussed above, is an important example of an economy that had a dramatically different experience than the U.S. We begin with the international experience and then turn to the U.S.

Bernanke and James (1991) use annual data on twenty-four countries to study banking crises during the 1920s and 1930s. They construct a chronology of banking crises during the interwar period and focus on the links between the gold standard, banking crises, and real economic activity. They cite the industrial organization of the banking system as a significant factor in explaining which countries experienced banking panics during this period. In terms of the shock causing panics in those countries whose banking systems were prone to panics, they observe that there “were virtually no serious banking panics in any country after abandonment of the gold standard...” (p. 53), suggesting that deflation was the important shock. Another important point made by Bernanke and James (1991) concerns the real effects of severe banking problems. They argue that countries that experienced panics had deeper depressions than countries that did not experience panics. Bordo (1986), Calomiris (1993a), and Grossman (1994), all mentioned above, also focus on the cross-section variation of banking experiences internationally. Grossman (1994), like Bernanke and James, finds that a combination of macroeconomic policy and banking structure can explain much of the cross section experience in banking crisis. He rules out, as an explanation, lender-of-last-resort behavior of central banks.

Indeed, central banks were relatively new and inexperienced at dealing with bank crises, with the exception of the Bank of England. The Great Depression is a turning point in the history of central banking. According to Capie (1997), there were only eighteen central banks at the beginning of the 20th century. By 1950 there were 59 central banks and by 1990 there were 161. At the beginning of the 20th century, the U.S. Federal Reserve System was not yet established; this would occur in 1914. The Bank of

Canada came into being *after* the Great Depression, in 1934. Prior to the 20th century central banks were established as institutions with monopoly rights over money issuance. But, if a critical element of central banking is the function of lender-of-last-resort, then these institutions generally did not become central banks until later, typically during the 20th century—after the Great Depression.

Although the Federal Reserve System came into existence in 1914, and so there was a central bank with a discount window in existence during the Great Depression, there was no deposit insurance. Prior to the Federal Reserve being enacted there was the private system of clearinghouses that did provide a form of deposit insurance. And after 1934 there is explicit government-provided deposit insurance. However, during the period 1914-1934, there is no deposit insurance in the U.S., either private or public. Although during the Great Depression there were various points at which the government together with the clearinghouses attempted to act, nothing came of this.⁵⁰ Clearinghouses acquiesced to the Federal Reserve, but the Federal Reserve did not play the role that clearinghouses had played in earlier episodes. This accounts for much of how the panics of the Great Depression in the U.S. differed from earlier panics.

First of all, unlike earlier episodes in the U.S. where there was not a single panic near the business cycle peak. Gorton (1988) constructs a counterfactual, based on estimated structural equations and argues that if the private clearinghouse system in existence prior to 1914 had been in existence during the Great Depression (and there had been no Federal Reserve system), then there would have been a single panic in December 1929 (and also in June 1920).⁵¹ Instead of one quarter to one third of the banks failing, Gorton estimates that less than one percent would have failed had the private clearinghouse system been in place. Instead, there was not a single panic during the Great Depression. In fact, it is a matter of dispute which episodes really constituted panics. The dates in question are periods in which there were numerous bank failures; November 1930 to January 1931, April to August 1931, September and October 1931, and February and March 1933. Friedman and Schwartz (1963) were the first to argue that these were four separate national banking panics during the Great Depression. This has been disputed, as we discuss below. The difficulty is not just the matter of the definition of what is a “panic,” but also that these episodes were different than previous U.S. Panics in another way, emphasized by Wicker (1996), namely, that the center of the panic was not the money market in New York City. Rather, the initial banking problems were region specific. Wicker (1996): “There is no discernible pattern in the diffusion of the crisis from certain regional centers to the periphery. One reason for the absence of such a

⁵⁰ In 1930 the Federal Reserve Bank of New York and the New York Clearinghouse Association attempted to arrange a rescue of the Bank of the United States, but the plan failed (see Wicker (1996)). Later, in 1933 there was a proposal to issue clearinghouse loan certificates to the public, but this also failed (see Wicker (1996)).

⁵¹ The fact that there was no panic in 1920 was also significant because the 1920s saw significant numbers of banks fail in the U.S. See Alston, Grove, and Wheelock (1994).

pattern is the fact that the panic did not at any time engulf any of the largest banks of Philadelphia, Pittsburgh, and Chicago. Exactly how the loss of confidence spread across the twelve Federal Reserve Districts is still a matter requiring explanation” (p. 98).

The fact that the panics were more regional, at least they originated outside New York City, has led to disputes among researchers about which of the four events identified by Friedman and Schwartz really were national panics. It has also complicated efforts to test hypotheses about the causes of panics during the Great Depression. The debate over the origins of the panics in the Great Depression echoes the debates about whether panics are irrational contagion or information-based. The fact that rural areas play a role has led to consideration of the fall in agricultural income as an important factor. In a famous quotation, Friedman and Schwartz (1963) put it this way: “In November 1930 ... a crop of bank failures, particularly in Missouri, Indiana, Illinois, Iowa, Arkansas, and North Carolina, led to widespread attempts to convert demand and time deposits into currency ... a contagion of fear spread among depositors starting from agricultural areas, which had experienced the heaviest impact of bank failures in the twenties. But such contagion knows no geographical limits” (p. 308). In contrast to this view of contagion emanating from problems in agricultural, Temin (1976) argued that sharp declines in the value of bank asset portfolios caused bank failures. He constructed proxies for the quality of bank portfolios using traded bond prices and performed annual cross section regressions, attempting to explain the pattern of bank failures. Wicker (1980) presents a third point of view, arguing that the collapse in November 1930 of Caldwell and Company of Nashville, Tennessee was the shock setting off the panic, rather than declines in bank asset values or agricultural incomes.

A number of other authors have contributed to the subsequent debate, attempting to shed light on the three interpretations. The main innovation has been more detailed bank-level data. White (1984) U.S. national banks during four years, 1929-1932. For each of these years, the failed banks are matched with a stratified random sample of non-failing banks based on similar assets and geographical location. White then uses financial ratios to try to discriminate between failed and nonfailed banks, using logit regression. He argues that his results show that Temin and Friedman and Schwartz are not really in conflict. Bank failures are explained by a shocks causing agricultural distress, leaving banks with poorly performing loans. Thies and Gerlowski (1993) revisit White’s analysis a bit differently and confirm his findings. Calomiris and Mason (2000) construct an even more detailed data set to analyze the causes of bank failure during the Great Depression. Their measures of fundamentals include attributes of individual banks, as well as proxies for local, regional, and national economic shocks. In the first three of the Friedman and Schwartz panics, they find no evidence of contagion-like effects. But, for the last episode this does not appear to be the case. Hamilton (1985) also examines the Panic of 1930 and the interpretations of Friedman and Schwartz, Temin, Wicker and White. Hamilton presents a fairly nuanced

view, concluding that “The banking panic of 1930...had no single cause, and none of the various interpretations of the panic’s causes—poor loans and investments made in the twenties, the Caldwell failure, or falling cotton prices—can fully account for the rise in the number of failures and for the shift in the states and regions afflicted with banking difficulties. The wave of failures can be explained by the combined effect of the overextended condition of the failed banks, the Caldwell shock, and the deteriorating agricultural conditions” (p. 607).

There is also controversy concerning the effects of the collapse of the banking system during the Great Depression. Friedman and Schwartz (1963) argue that the collapse of the banking system was only important because it meant a major decline in the supply of money, via the money multiplier when depositors withdrew currency. Temin (1976) sees the collapse of the banking system as a result of real shocks and that even without the banking crisis “the overall story of the Great Depression would not have been much different” (p. 9-10). Bernanke (1983) initiated a revision of the debate when he introduced another interpretation of the events. He argues that “the financial crises of 1930-33 affected the macroeconomy by reducing the quality of certain financial services, primarily credit intermediation” (p. 263). In other words, banks perform a real allocative role, which is important for the functioning of the economy. Without banks, due a collapse of the banking system, output will decline because banks can no longer allocate capital to firms. Bernanke works in the general econometric framework of unanticipated money causing changes in real output (the rate of growth of industrial production), introduced by Barro (1978). His first finding is that declines in money are not quantitatively large enough to explain the output declines of 1930-33. He then includes proxies for the effects of declines in intermediation services, for example, the real deposits of failing banks and the liabilities of failing businesses. These and other proxies improve the explanatory power of the output equation, strongly suggesting his interpretation.

Bernanke’s paper has been very influential and generated a number of responses. Temin (1989) argued that firms which were more reliant on bank loans, namely smaller firms, should suffer the most when the banking system collapses. His analysis is based on dividing industries into bank-reliant ones and non-bank-reliant ones. He finds no pattern, arguing that Bernanke is wrong. Hunter (1982), however, provides more detailed analysis based on firm level characteristics and does find that small firms were affected differently. Also see the discussion in Calomiris (1993b). Another critique of Bernanke is that of Rockoff (1993) who argues that Bernanke’s results are not robust to how money is defined. He constructs a measure of money that takes into account the fact that deposits in banks that have suspended convertibility are not “money” in the same sense as deposits in other banks. When this measure of money is used, Rockoff finds that the nonmonetary proxy variables are not important in the specification. Essentially, Rockoff argues that any times series variable with a spike during 1929-1933

will have statistical significance. Calomiris and Mason (2001), in preliminary work, attempt to examine the issues at a much disaggregated level, asking whether indicators of local banks' conditions can explain cross section variation in state income.

Bernanke's original paper discusses the Canadian experience during the Great Depression, since it provides such a contrast to that of the U.S. suggesting that a more careful study of Canada would be valuable. Haubrich (1990) provides such a study. In Canada, there were no panics and no banking crises, though the number of branches declines from 4049 to 3640 between 1929 and 1933.⁵² Essentially, Haubrich follows Bernanke's path, but studies Canada. He also looks at cross industry comparisons. He finds that measures of financial distress have no economic or statistical significance. His interpretation is that the real effects are due to crisis or panic. The banking system can contract because the demand for loans declines, but real effects only occur when the supply of loans contracts due to crisis. This interpretation is disputed by Calomiris (1993b).

I. Contagion

"Contagion" is the idea that some event can cause a chain reaction or domino effect among banks. For example, when one bank (or possibly a nonfinancial firm) fails, this, it is argued, can cause depositors at other banks to withdraw their deposits. Alternatively, when one bank becomes insolvent this can cause other banks to become insolvent because of a chain of illiquidity stretching through the interbank market. So, one view of "contagion" is that it refers to "interdependence" among banks, more so than nonfinancial firms. In Bhattacharya and Gale (1987) or Allen and Gale (2000), as well as the other papers mentioned above, banks are interdependent so shocks to one or a few banks may have an impact on other banks. This type of shock transmission mechanism is also prominent in discussions of the risks in the payments system. For example, Flannery (1996) develops a model in which banks become wary of lending to other banks, although most banks are in fact solvent. At root, the interbank loan market creates an interdependence that can propagate shocks through the banking system. Another view of "contagion" is informational. Banks are opaque institutions, so that information about a single institution might rationally or irrationally lead to a revision of beliefs about the value of other institutions.

There are a variety of ways of testing contagion hypotheses. Calomiris and Mason (1997) look for informational contagion effects. They adopt the empirical strategy of comparing the ex ante attributes of banks that failed during the Chicago panic of June 1932 with those that did not fail. If banks that failed were just as strong as those that survived, then this would be evidence in favor of confusion on the part of

⁵² As noted above, Kryzanowski and Roberts (1993, 1999) claim that all of the large Canadian banks were insolvent during the 1930s and that the only reason that there was not a banking crisis was that there was implicitly complete deposit insurance provided by the government. This view is disputed by Carr, Matheson and Quigley (1995).

depositors. If banks that failed were weaker, then runs on individual banks were not purely random; weaker banks were run and then failed. They find that weaker banks did fail, and interpret their evidence as being inconsistent with contagion. Rather, there is evidence that while depositors were somewhat confused about the states of individual banks, only the weakest banks were forced into insolvency.⁵³ Though Calomiris and Mason also point to Chicago clearinghouse as the institution that facilitated this. Their study shows that weak banks failed in the panics while the strong survived. So, the panic did not cause *ex ante* stronger banks to fail. But, is it not clear that this is a statement about the causes of the panic. The panic may have been caused by some version of “contagion,” but in the end only weak banks failed.

Other studies of contagion have been event studies that examined the abnormal return on bank stocks when a bank fails, in the post WW II U.S. economy. For example, Aharony and Swary (1983) look at the stock reactions to three large bank failures in the 1970s. They found that other banks’ stock prices did not respond. Similar studies include Aharony and Swary (1996), Swary (1986), and Wall and Peterson (1990). Basically, the empirical results support the idea that the stock market prices respond to new information, rather than to contagion caused by interbank linkages or irrationality. Kaufman (1994) reviews more of these studies.⁵⁴ Event studies test a number of joint hypotheses which makes them difficult to interpret. Chief among these problems is the fact that large U.S. banks are viewed as being “too-big-to-fail.” This is the implicit government policy of rescuing large banks, possibly preventing their failure to the benefit of shareholders. O’Hara, Maureen and Wayne Shaw (1990) find positive announcement effects to encouraging government announcements concerning too-big-to-fail, suggesting that big bank shareholders benefit from this policy. Consequently, there may be detectable contagion effects to a large bank failure were it not for the too-big-to-fail doctrine.

Furfine (2001) analyzes the interbank market by analyzing all individual U.S. federal funds transactions during 1998, a year during which Russia defaulted on its sovereign debt and the hedge fund Long-Term Capital Management (LTCM) was rescued by the private sector. These transactions allow Furfine to trace any chain reaction or domino effect and allow him to identify whether banks, as a group, became fearful of transacting with other banks. Furfine (2001) finds that interest rates did not move from their level intended by the Fed, and that interest rate variability was not really affected by the crises with Russia and LTCM. Aggregate volume in the fed funds market rose in the second half of 1998, during the

⁵³ Esbitt (1986) examined Chicago banks that failed in 1931 and shows that they were plagued by poor management.

⁵⁴ Saunders and Wilson (1996) study deposit flows in a sample of failed and healthy banks over the period 1929-1933 in the U.S. They find evidence of contagion for 1930-1932, but not in 1929 or 1933. However, during 1930-1932, failing-bank deposit outflows exceeded those at a matched control sample of nonfailing banks suggesting that there were informed depositors who distinguished among *ex ante* failing and nonfailing banks.

crises. Credit spreads in the interbank market did not increase, but were often narrower. Finally, individual banks borrowed at least as much during the crises as before. Furfine's results are the strongest results against contagion effects in the interbank market, but these results too are possibly a function of the too-big-to-fail policy of the U.S. government.

There are a large number of studies that examine banking crises in emerging markets in the 1980s and 1990s, arguing that some of these events seem to have elements of contagion. While definitions of "contagion" vary considerably, one view of "contagion" views it as the transmission of real shocks from country to country due to trade links, financial links, or "fear." Some studies of recent crises provide evidence that banks were an important transmission mechanism of shocks, possibly accounting for phenomena labeled "contagion." Peek and Rosengren (1997, 2000) see banks as a transmission mechanism, but do not think of it as contagion. Kaminsky and Reinhart (1998, 1999) identify three channels that may transmit shocks from one country to another: bank lending, liquidity, and trade. Their empirical work is based on forming clusters of countries based on measures of these three channels. They then show that these clusters are regional, a possible source of transmission of shocks. In the case of bank lending they distinguish a cluster of countries that borrows from Japanese banks and one which borrows from U.S. banks. Their main result is that the probability of crisis, conditional on crisis having happened in a certain banking cluster, tends to be higher than the unconditional probability of crisis. However, the clusters associated with each channel overlap so much that it is hard to argue that the common bank lender channel has really been isolated from the other channels. In a similar vein, Caramazza, Ricci, and Salgado (1999) use BIS data define a "common bank lender" for each crisis as the country that lent the most to the first country in crisis in each of the major crises, using a sample of 41 emerging markets. For example, in the Mexican crisis the common bank lender is the United States. Their main result is that countries that experienced crises were more reliant on a common lender than other countries. Van Rijckeghem and Weder (1999, 2000) also investigate the idea that international banks are a major channel for the transmission of shocks. Studies in this area are relatively new, but seem promising.

V. Bank Regulation, Deposit Insurance, Capital Requirements

Government provision of deposit insurance and government intervention into banking markets, including bank supervision and examination, limitations on bank activities, capital requirements, charter requirements and entry restrictions, closure rules, and other rules for banks, are now widespread around the globe. The rationale for deposit insurance and bank regulation is the argument that banks are inherently flawed institutions, being prone to harmful banking panics. Consequently, the government should provide deposit insurance and regulate bank risk taking. Moreover, once deposit insurance has been adopted, there is a further need for government intervention via bank regulation because of the

incentive of banks to take additional risks once they have (underpriced) government deposit insurance (see Buser, Chen, and Kane (1981)).

Most of the vast literature on bank regulation is within this paradigm of panics, deposit insurance, and moral hazard. In general, the literature on bank regulation, and related issues, assumes the need for deposit insurance and government regulation and focuses on the implications of moral hazard problems for the design of bank regulation. There is an associated empirical literature that has attempted to uncover evidence that moral hazard is a problem in banking systems with insured deposits. The empirical literature, while covering a variety of topics, has not been particularly successful at finding evidence of moral hazard problems, despite its dominance as a theory of bank behavior and bank regulation. We start, however, with a discussion of the origins of bank regulation and deposit insurance. This is an important topic because the government provision of deposit insurance, and the associated bank regulation, is a quite recent phenomenon.

A. The Origins of Government Bank Regulation and Government Deposit Insurance

If banks are inherently unstable institutions, prone to panics, then government regulation is perhaps justified, in the form of government deposit insurance, capital requirements, and bank supervision and examination. However, as discussed above, most countries did not have banking panics, or, if they did, panics were infrequent. Why then are government deposit insurance schemes and bank regulation so widespread? Part of the answer is that they were not widespread until recently, reflecting policy advice based on the paradigm of panics, deposit insurance, and moral hazard.

Bank regulation and deposit insurance have their origins in the private arrangements among banks, as described above in the discussion of bank clearinghouses and other private bank coalitions, and theoretically by Gorton and Huang (2001). Governments took over these insurance schemes and regulations fairly recently, although in the U.S. there were various earlier deposit insurance arrangements sponsored by state governments (see White (1983), Calomiris (1990), and Wheelock (1992b)). The first formal nationwide government deposit insurance system in the world was established in the United States in 1934. Other countries did not follow the U.S. lead, even those that had experienced the depression of the 1930s. It was not until after World War II that countries around the globe began to adopt deposit insurance. For example, Canada did not adopt deposit insurance until 1967. Figure 2 shows the number of explicit national deposit insurance programs in countries around the world.⁵⁵ In 1980 only sixteen countries had explicit deposit insurance programs; by 1999, sixty-eight countries had such programs (see

⁵⁵ Demirgüç-Kunt and Detragiache (no date) and Garcia (1999) detail the variation in the schemes adopted around the world.

Kyei (1995), Garcia (1999) and Demirgüç-Kunt and Sobaci (2000)). Two thirds of the deposit insurance programs in the world have been established in the last fifteen years. Widespread banking crises during the 1980s and 1990s were the proximate cause of the spread of government deposit insurance.

Not only is deposit insurance recent, it has been hard to explain why it was adopted in the first place. In the U.S., the federal deposit insurance legislation was originally supported by all but the largest banks, but was widely viewed by others as special interest legislation, a subsidy for banks. Even with the collapse of the banking system during the Great Depression, the Roosevelt administration, the bank regulatory agencies, and large banks opposed the legislation. See Flood (1992) and White (1998). In fact, looking at U.S. history, White (1998) concludes that: “There is no ready model to explain the growth and spread of federal insurance of intermediaries” (p. 87-89). This challenge may have been answered by Kane and Wilson (1998).

That deposit insurance and capital requirements are recent developments is consistent with the above observations that instability is not inherent in banking, that most banking systems do not have problems with banking panics; the U.S. experience is an outlier. Since 1934, when deposit insurance came into being, most of the U.S. banking experience has been quiescent. Figure 1 shows the number of bank failures in the U.S. since 1934. The figure strongly suggests that deposit insurance per se is not subject to moral hazard. For fifty years the banking industry in the U.S. was a rather quiet industry, with few failures and little academic attention. It appears that there was a regime switch in the mid-1980s. We return to this issue below.

B. Deposit Insurance and Moral Hazard

Moral hazard is the idea that bank shareholders have an incentive to take advantage of under-priced deposit insurance by engaging in riskier actions than they would otherwise. The idea that equity holders want to increase risk, at the expense of bondholders, applies to all situations where there is limited liability (and under the standard Black-Scholes assumptions). This is the observation that, viewing equity as a call option on the value of the firm, option values are increasing in volatility. In other words, the equity holders do not care if the payoffs are low because of limited liability, but they will benefit if the payoffs are high. This idea is commonly described as “gambling for resurrection” or “playing the lottery.” Moral hazard has been viewed as being of particular relevance to banking because government deposit insurance premia are not (explicitly) based on the riskiness of the bank. Moreover, the argument is that banks are regulated, in part, to prevent them from engaging in moral hazard. The usual view, however, is that the equity holders will engage in risk-increasing strategies only when bank capital is low or nonexistent. This view implicitly reflects the idea that there are some institutional and contractual

constraints on equity holders, but that these constraints lose their force when equity value is low. It has never been clear what these constraints actually are, or why they lose their force at low equity values.

A starting point for considering the moral hazard issue concerns whether deposit insurance premia are (implicitly) set to reflect the risk of individual banks. If insurance premia are fairly priced, then the incentives to engage in moral hazard are the same as in nonfinancial firms and, presumably it is prevented in the same way. Using stock price data and an option-based approach, Marcus and Shaked (1984) found that the vast majority of large banks are overcharged for deposit insurance. Pennacchi (1987), also using an option-based approach, finds that nearly all the banks in his sample were overcharged. These results suggest that the bank regulators or corporate governance mechanisms exert control over banks to limit their risk-taking so that they are effectively being overcharged even with flat rate premia. Also, see Ronn and Verma (1986). Federal Deposit Insurance Corporation (2000) discusses a variety of risk-based pricing methods.

Buser, Chen, and Kane (1981) and Marcus (1984) raise an important issue with the moral hazard argument, that is that in order to enter the banking industry, a charter from the government is required. The charter is the license to take deposits and make loans. The charter is not a transferable asset and it is lost if the bank fails. Charters are in limited supply because the government, in many countries, does not allow free entry into banking. This valuable charter alters the bank's risk-taking behavior, particularly as posited by those who argue that deposit insurance is under-priced and so bank equity holders have an incentive to engage in moral hazard. As Marcus (1984) put it "...the traditional view of bank finance in the presence of FDIC insurance is overly simplistic in that it ignores the effects of potential bankruptcy costs" (p. 565). Charter value causes the bank equity holders to decrease risk as default becomes more likely.

The pattern of bank failures shown in Figure 1 may be understandable based on the value of commercial bank charters. What changed? Essentially, the story is that prior to the 1980's banks were partially protected from competition. They often had local monopolies in deposit markets (see Hannan and Berger (1991) and Neumark and Sharpe (1992); there was no competition from money market mutual funds and there are interest rate ceilings. In other words, the charter values of banks were high. Keeley (1990) was the first to link Marcus' theoretical insight to the empirical world of banking in the mid-1980s. Keeley (1990) uses Tobin's q as a measure of market power or charters in banking. (Also, see Saunders and Wilson (2001).) For example, banks that have a local monopoly on deposits can issue deposits at below-market rates, and this will be reflected in the bank's stock price. He finds "that [Tobin's] q appears to be useful proxy for market power and that banks with greater market power hold more capital and pay lower rates on CD's" (p. 1186). Keeley's interpretation of banking in the 1980s is that increased competition in banking reduced charter values, causing banks to increase risk in response.

Demsetz, Saidenberg, and Strahan (1996) also conclude that banks with high charter values operate more safely than other banks. Keeley's argument is consistent with moral hazard being operative when commercial bank charter values are low. We return to Keeley's argument below.

The bulk of the empirical work aimed at testing the moral hazard hypothesis as applied to financial intermediaries analyzes the behavior of insolvent or poorly capitalized U.S. savings and loan institutions, "thrifts," during the 1980s. The S&L crisis would appear to be a good testing ground for the moral hazard hypothesis. A series of exogenous interest rate shocks in late 1979 and early 1980, and the deregulation of deposit rates, caused large numbers of thrifts to lose significant amounts of equity. Essentially, deregulation reduced the value of thrift charters while interest rate shocks almost simultaneously reduced their equity value. Between January 1980 and December 1988, nearly 1,200 thrifts failed, though not all were actually closed, later described as a policy of regulatory "forbearance," which makes for an even more interesting testing ground for moral hazard. Brumbaugh (1988), Kane (1989), Kormendi et. al. (1989), Barth (1991), and White (1991), among others, provide background on the thrift crisis.

The first type of tests for the presence of moral hazard in the thrift industry focused on comparing the behavior of insolvent thrifts with those of solvent counterparts. For example, Barth, Bartholomew, and Bradley (1990) find that failed thrifts had disproportionately high concentrations of commercial mortgages, real estate loans, and direct equity investments, compared to the average thrift. DeGennaro, Land, and Thomson (1993) study the investment strategies of the 300 largest thrifts to post capital deficiencies in 1979. The institutions of this group that subsequently failed followed "higher-growth investment strategies" than did those that returned to health. Benston (1985) analyzed a matched sample of solvent and insolvent thrifts. Between January 1, 1980 and August 31, 1985, 202 thrifts failed. Each of these thrifts was matched to two nonfailing thrifts, two just smaller and two just larger. Barth and Bradley (1989) also pursue this empirical strategy. Rudolph and Hamdan (1988) use financial ratios to try to discriminate between failed and solvent thrifts in the post-deregulation period (i.e., after the Depository Institutions Deregulation and Monetary Control Act of 1980 and the Garn-St. Germain Act of 1982). Brewer (1995) look at changes in thrifts' stock prices in response to changes in the mix of asset investments.

A related approach is to look at changes in thrift behavior following major legislation to see whether solvent and insolvent thrifts responded differently. The Depository Institutions Deregulation and monetary Control Act of 1980 and the Garn-St. Germain Depository Institutions Act of 1982 allowed thrifts to invest in previously forbidden assets. McKenzie, Cole, and Brown (1992) estimate the average returns on various types of thrift investments for the years ending June 30, 1987 and June 30, 1988. In particular, they estimate the returns on traditional thrift assets and on the new, nontraditional, investments.

Returns on nontraditional assets are estimated to be lower than on traditional assets, but in particular, the results are more pronounced at capital deficient thrifts. This can be interpreted as evidence that thrifts with low capital were engaging in moral hazard. However, as the authors note, there are a number of other explanations consistent with the finding. First, thrifts that were about to fail may have already sold the more liquid traditional assets, biasing the estimates of returns. Second, nontraditional assets may be easier to use to engage in fraud or “looting,” a hypothesis distinct from moral hazard, as discussed below. Third, regulators were more inclined not to close insolvent thrifts with traditional portfolios.

These approaches to testing for moral hazard are fraught with difficulties. There is no question that failed thrifts are different than solvent thrifts; this is true by definition, since they failed and the others did not. And it is not surprising that the failed thrifts have many common characteristics; they engaged in similar types of investments and those investments did not do well. But, it is not clear that these observations have anything to do with moral hazard. For example, if a thrift is not successful investing in traditional asset categories, for whatever reasons, it may invest in new asset classes allowed by deregulation. If there is a negative exogenous shock to this new asset class, and the thrift subsequently is closed, it may have nothing to do with moral hazard. But, this outcome must be distinguished from the case where the thrift, seeing that its net worth is negative invests in the new asset class because the new assets are viewed *ex ante* as being very risky.

In addition to the problems mentioned above, another problem involves ensuring that the risk-taking behavior is caused by insolvency, rather than the other way around. This issue is related to the use of accounting data, which almost all of the studies rely on. Benston, Carhill, and Olasov (1991) discuss the accounting issues, and then go on to base their analysis on estimates of market values. These authors “conclude that insolvent thrifts did not expand more rapidly than did solvent thrifts and, in general, did not take greater risks” (p. 379). Brickley and James (1986) also avoid the accounting issues by looking at the response of stock prices to changes in Federal Savings and Loan Insurance Corporation closure policy. They find that the response is as if access to underpriced federal deposit insurance is a valuable option, but it is not clear that there is moral hazard. According to standard option theory, call options are more valuable if the maturity is extended, *ceteris paribus*.

Another interesting experiment concerns the U.S. Comptroller of the Currency’s announcement, in 1984, that the eleven largest U.S. banking firms were “too big to fail” (TBTF), implying they would receive *de facto* 100 percent deposit insurance. Did this encourage risk-taking? O’Hara and Shaw (1990) investigate the effect on bank equity of the Comptroller of the Currency’s announcement using event study methodology. They find positive wealth effects accruing to TBTF banks, with corresponding negative effects accruing to non-TBTF banks. Boyd and Gertler (1994b) study the poor performance of banks in the 1980s in a statistical study controlling for location, asset size remains a significant factor in

poor performance of large banks. They find that the poor performance of the U.S. banking industry in the 1980s was due mainly to the risk-taking of the largest banks and interpret this as risk-taking that was encouraged by the U.S. government's too-big-to-fail policy. Also, see Black, Collins, Robinson and Schweitzer (1997). De Nicolo (2001) shows that the link between bank size and risk extends beyond the U.S. He examines banks in 21 industrialized countries during 1988-1998 and finds that larger banks have lower charter values (as measured by Tobin's q ratio) and higher risk of insolvency.

Some authors have attempted to address the shortcomings of studies based on the thrift crisis by examining historical situations where some institutions are covered by insurance, while other similar institutions are not. Wheelock (1992b), Wheelock and Kumbhakar (1995), and Wheelock and Wilson (1995) looked at individual banks in Kansas that participated in a state deposit insurance program in the 1920s, on a voluntary basis. Banks that chose insurance coverage took additional risks. Calomiris (1990) finds that deposit insurance in the early 1900s increased bank risk-taking, more so for states with mandatory insurance than for states with voluntary insurance. Grossman (1992), examining thrifts in the 1930s, found that thrifts entering the voluntary federal insurance program did take on more risks than uninsured thrifts after several years.

Demirgüç-Kunt and Detragiache (2000) analyze international evidence. They estimate the probability of a systemic banking crisis in a panel of 61 countries over the period 1980-1997. This is the period when most countries adopted deposit insurance; see Figure 2.⁵⁶ About 40 banking crises are identified in the panel and for about half the observations a deposit insurance system was present. The main result is that a dummy variable for the presence of deposit insurance is positive and significant. Refining this by distinguishing different types of deposit systems shows that explicit deposit insurance is associated with higher likelihood of crisis. As the authors point out, one possibility for the association between deposit insurance and bank crises is that economies with fragile banking systems tend to adopt deposit insurance. An instrumental variable approach to this issue shows deposit insurance to still be associated with crisis. The authors conclude that "explicit deposit insurance tends to be detrimental to bank stability" (p. 22).

The international evidence of Demirgüç-Kunt and Detragiache raises another issue. The moral hazard argument is the idea that equity holders are motivated to take on risk inefficiently in the hopes that there is a state of the world that could be realized in the future in which their equity would be positively valued. But, Akerlof and Romer (1993) point out that much of what is claimed to be evidence of moral hazard, in fact, appears to be behavior in which there is NO state of the world in which equity would be positively valued. Rather, it is "looting," that is, the equity holders are simply illegally stealing from the

⁵⁶ Years in which banking crises were occurring were excluded.

institution. Akerlof and Romer attempt to estimate the extent of looting during the U.S. S&L crisis and present a range of estimates that would account for a large fraction of the government clean-up costs. The argument is also applied to the banking systems of emerging markets. As Akerlof and Romer write: “..it is a safe bet that many developing countries that have far less sophisticated and honest regulatory mechanisms than those that exist in the United States will be victimized by financial market fraud as their financial markets develop” (p. 59). An important question concerns whether this type of fraud is increased when deposit insurance is adopted.

Returning to the U.S. S&L crisis, there is little dispute that if the government had closed thrifts faster, then the costs of resolving the insolvencies would have been lower. But, closure policies are somewhat complicated by the constraints the government may face. The government may optimally not want to close banks or thrifts and instead engage in a kind of moral hazard itself. See Gorton and Winton (2000). Acharya and Dreyfus (1989) and Mailith and Mester (1994) also analyze government closure policies for financial intermediaries. If there had been no deposit insurance for U.S. savings and loan institutions, then there would likely have been a generalized panic at some point, rather than forbearance by the government. This raises the question of the costs of deposit insurance. White (1998) attempts to address this issue with regard to U.S. commercial banks. While it is a difficult problem, and White admits the tenuousness of his estimates, it certainly raises the point that the social value of deposit insurance is uncertain. White concludes that ‘...it is hard to escape the conclusion that deposit insurance did not substantially reduce aggregate losses from bank failures and may have raised them’ (p. 119).

C. Corporate Governance in Banks and the Moral Hazard Argument

Despite the appealing simplicity of the moral hazard argument, it has been difficult to find compelling evidence of its existence. One issue concerns whether there are counterbalancing forces, arrangements that provide incentives for shareholders to not engage in risky strategies. Another issue concerns whether equity holders are even in a position to make decisions in large institutions anyway. As a practical matter, equity holders are not in a position to dictate policy in large banks, where management may be “entrenched.” The idea that equity holders control the firm, the basis of the moral hazard argument, appears to contradict the standard corporate finance view of large firms, namely that managers are effectively in control, not equity holders. This is the essence of the agency view of corporate governance. This raises issues of corporate governance in banking. Are the corporate governance issues in banking similar to those in unregulated nonfinancial firms? Are there agency conflicts between bank owners and their managers? Is executive compensation for bankers different?

Not surprisingly there is evidence of agency conflicts between managers and shareholders in banking. James (1984) and Brickley and James (1987) demonstrate the importance of the market for

corporate control in banking. James (1984) finds that salary expenses, occupancy expense, and total employment are higher for banks in states that prohibit acquisitions by other banks. James and Brickley (1987) find that banks in states that allow acquisitions have more outside directors on their boards. Schranz (1993) finds that banks in states that allow a more active takeover among banks are more profitable. Houston and James (1993) find that the frequency of management turnover among poorly run commercial banks is about the same as management turnover in poorly run nonbanks. They also find that the costs to managers, in terms of lost income and future job opportunities, of their departure are significant. Cannella, Fraser and Lee (1995) find that the bank managerial labor market discriminates between cases where the bank manager was arguably at fault for the bank failure and those cases where failure was due to exogenous events. Hubbard and Palia (1995) find that the pay-for-performance link in bank management compensation is stronger in states that permit interstate banking, suggesting that the market for corporate control is important. Also, see Barro and Barro (1990) and Crawford, Ezzell, and Miles (1995).

Corporate governance in banks may be similar to nonbanks in some ways, such as frequency of management turnover, but this may be due to regulatory intervention. Prowse (1995) studies corporate governance in U.S. commercial banking, and considers this issue. He borrows the method employed by Morck, Shleifer and Vishny (1989) to a sample of manufacturing firms and studies a sample of U.S. bank holding companies. The results are given in Table 4 (which is Table 1 of Prowse, p. 28). The frequency of control changes, in terms of percentage of the sample that experienced a change, is about the same for banks and nonbanks. But, the composition of the changes is very different. Clearly, regulatory intervention is important for banks, while hostile takeovers are not frequent. In Prowse's sample, management turnover is less important in banks than in nonbanks.

If there is a separation of ownership and control in larger banks, then moral hazard would be a problem only if the interests of equity holders and their managers were in alignment. Houston and James (1995) consider this question, that is, they ask whether executive compensation in banking is structured to promote risk-taking? They find that (on average) CEOs of banks are less likely to participate in a stock option plan; they hold fewer stock options, and they receive a smaller fraction of their total pay in the form of equity-linked remuneration compared to CEOs in other industries. They also find that the relation between equity-linked compensation and the value of the bank charter is positive and significant. Finally, they find that weakly capitalized banks (as identified by the regulators) are no more likely to use equity-linked compensation than other banks. Overall, it does not seem to be the case that equity-holders attempt to induce or entice managers to engage in moral hazard.

We are left with a conundrum. If there are corporate governance problems with large financial institutions, and these appear to be similar in magnitude to the problems in nonfinancial firms, how are

bank and thrift managers being enticed to engage in moral hazard? Keeley's identification of the increase in commercial bank failures and the decline in bank capital ratios in the mid-1980s with increased competition seems compelling. But, it is not clear that the explanation is "moral hazard," that is, an increase in risk by equity holders seeking to maximize the value of their equity (given insured deposits). In the case of U.S. commercial banking, there was no exogenous shock that had the same detrimental effect on equity capital that occurred with thrifts. Competition increased in banking, and over a protracted period of time commercial banks failed. The question is: Who made the decision to engage risky investments, equity holders or bank managers?

Gorton and Rosen (1995) empirically analyze whether the risk-taking identified by Keeley is explained by shareholders making the decisions, i.e., moral hazard, or by managers. They look at the relationship between risk-taking and the ownership structure in banks. The banking environment of the late 1980s was not the usual environment. It was a competitive environment where opportunities were shrinking. In such an environment, managers may resist shrinking their banks through merger, acquisition, or directly by making fewer loans. But, equity value is not yet low. As charter values decline, shareholders prefer to exit the industry and redeploy their resources in investments with higher returns. But, to the extent that managers are entrenched, this could mean that their careers as bankers are over in many cases, as the industry is shrinking. In a shrinking industry, managers and equity holders have different incentives, and they are different than they would be in a growing industry.

Which group, managers or equity holders, make decisions depends sensitively on the relative sizes of the equity holdings of the managers versus the outside equity concentrations. At one extreme, the manager is the owner if he owns all the equity. At the other extreme, the manager may have no equity and face a large outside block holder who is in control. But, more likely combinations involve a manager with a small amount of equity in a bank where all the other shareholders are dispersed or there is a group of outside block holders. In these cases, a manager may not have enough equity to care as much about his pecuniary returns as he does about his private benefits of control, but his equity holding may be large enough for him to exert control. Because the other shareholders are dispersed, he is effectively entrenched and can attempt keep the size of his bank from decreasing in the shrinking industry. These issues mean that the sought after relationship is potentially nonlinear. This turns out to be very important empirically. Gorton and Rosen find that the managerial hypothesis is empirically more important than the moral hazard hypothesis, which can be rejected, in explaining risk-taking by banks.

The relations between managers' equity holdings and the ownership structure of the outside shareholders, i.e., the extent to which there are blocks, makes the relationship between control and risk taking potentially highly nonlinear. Gorton and Rosen use semiparametric methods to tests for the nonlinearities. Saunders, Strock, and Travlos (1990) estimate a linear relation between managerial stock

holdings and the stock price volatility of their bank holding company over the period 1978-1985 and find a positive relation. Demsetz, Saidenberg, and Strahan (1997) use piecewise linear regression. They find that risk increases with the shareholdings of managers, but only for banks with relatively low charter value. The sample period is 1991-1995.

The structure of ownership and risk taking has also been investigated in the thrift industry. Cebenoyan, Cooperman, and Register (1995) analyze balance sheet measures of risk for samples of thrifts in 1988 and 1991. Independent variables include the holdings of insiders, and this variable squared, and a measure of the percentage of equity held by institutional investors. They find that “S&Ls with a high concentration of managerial stock ownership exhibit greater risk-taking behavior than other S&Ls in 1988, a period of regulatory leniency and forbearance on S&L closures, but lower risk-taking behavior in 1991, a period of regulatory stringency and nonforbearance” (p. 63). These results are confirmed in Cebenoyan, Cooperman, and Register (1999) over the period 1986 to 1995 with the same quadratic term included for insider holdings.

D. Bank Capital Requirements

Like deposit insurance, bank capital requirements, or at least, explicit capital requirements, are also a recent development. In the U.S., the Banking Act of 1933 required that bank regulators consider the “adequacy of the capital structure.” In the 1950s, the Federal Reserve mandated capital levels on the basis of a formula based on amounts to be held against different asset categories. This was dropped in the 1970s. But it seems that no attempt was made to enforce explicit capital requirements until 1982 (see Morgan (1992), and Baer and McElravey (1992)). Risk-based capital requirements became effective in the U.S. in 1991 with the passage of Federal Deposit Insurance Corporation Improvement Act, which explicitly required regulators to enforce capital requirements. The Basle Accord, reached on July 12, 1988, was a regulatory agreement among the G-10 countries, together with Switzerland and Luxembourg, which specified risk-based capital requirements.

Intuitively, bank capital reduces the likelihood of bank failure. So, if banks are riskier than socially desirable, why not require that they hold more capital? What is the cost of such a policy? For nonfinancial firms equity finance is costly because of asymmetric information, the explanation developed by Myers and Majluf (1984). Wansley and Dhillon (1989) and Polonchek, Slovin, and Sushka (1989) find that the announcement of common stock issuances by commercial banks in the U.S. results in a significant negative stock price reaction, though the reaction is smaller, on average, than for industrial firms. This is somewhat surprising because banks are generally viewed as more opaque institutions than nonfinancial firms. Cornett and Tehranian (1994) may provide some of the answer. They document that the stock price reaction to a voluntary equity issuance is significantly more negative than those associated

with an involuntary issuance taken to satisfy capital requirements.⁵⁷ But, these results do not address other issues concerning bank capital, to which we now turn.

The recent attention to bank capital requirements seems ironic, as Benveniste, Boyd, and Greenbaum (1991) point out. They observe that in the U.S. there has been a century-long secular decline in bank capital ratios, until about forty years ago. Prior to 1850 about 50% of bank assets were financed with capital. By the turn of the century the fraction financed by capital had shrunk to less than 20 percent. It was about 14 percent in 1929 and just over six percent at the end of World War II. Since then bank capital rose a small amount from the mid-1940s through the mid-1960s, then declined through the mid-1970s, and rose thereafter. In 1986, total capital was 6.8 percent, the same as it was in 1950. When first implemented, the Basle Accord called for a minimum capital of 7.25 percent of total assets by the end of 1990 (of which at least half was to be “core capital”) and at least 8.0 percent of assets by the end of 1992 (against with at least half in core capital).⁵⁸ During the late 1980 and through the 1990s U.S. bank holding companies increased their capital ratios to the highest levels in 50 years. See Saunders and Wilson (1999) and Flannery and Rangan (2001). However, the nation’s largest banks lost about one fourth of their market capitalization in the third quarter of 1998, coinciding with the Russian crisis and the Long Term Capital Management debacle. Hovakimian and Kane (2000) argue empirically that neither the market nor regulators prevented large banks from shifting risk to the government deposit insurance safety net.

These trends raise a number of questions. Why was there such a long downward trend? Why has it been reversed? Why did regulators (and academics) start focusing on capital requirements? What are the trends in other countries? Most of these questions have not been studied. Benveniste, Boyd, and Greenbaum (1991) interpret bank charter value (monopoly rents) as a hidden source of capital that began to dissipate in the late 1970s and 1980s, as discussed above. Flannery and Rangan (2001) argue that the increase in capital ratios in the last fifteen years is due to market forces disciplining banks. This argument contrasts with earlier views that banks always held the minimum amount of required capital, and that increases in capital ratios imposed costs on banks, and possibly on borrowers through a “credit crunch.”

If there are significant costs to raising bank capital, government-imposed capital requirements can have real effects. Banks may choose to exit the industry rather than satisfy the requirements. Such “exit”

⁵⁷ Wagster (1996) studies the wealth effects to shareholders of stocks of banks in Canada, Germany, Japan, the Netherlands, Switzerland, the United Kingdom, and the United States due to announcements of eighteen important events leading up the Basle Accord, starting on September 11, 1985 and ending on March 28, 1990. The main result is that only shareholders of Japanese banks experienced significant wealth effects for all eighteen events. Moreover, the cumulative gain was a positive 31.63 percent.

⁵⁸ “Core capital” includes common equity and minority interests in consolidated subsidiaries, but excludes loan loss reserves.

may occur through a reduction in bank loans rather than a reduction in bank assets per se. In such an event, otherwise worthy borrowers would not obtain bank loans. There would be a “credit crunch” due to capital requirements. Shortly after the Basle Accord in 1988, U.S. banks reduced their investments in commercial loans and increased the investments in government securities. More specifically, the share of total bank assets composed of commercial and industrial loans fell from about 22.5 percent in 1989 to less than 16 percent in 1994. At the same time, the share of assets invested in government securities increased from just over 15 percent to almost 15 percent. See Keeton (1994) and Furfine (2001). This period has been identified as a “credit crunch,” which refers to the possibility that banks were reluctant to lend to worthy loan applicants because of capital requirements.

In fact, the credit crunch, broadly outlined above, could be due to a number of nonmutually exclusive factors. Banks may have voluntarily and autonomously reduced their risk appetites. Regulators may have become tougher, causing the loan contraction. The Basle Accord, because of its risk-sensitive measures of capital, may have encouraged banks to reallocate their assets towards government securities. Finally, aside from the risk-sensitive nature of the capital requirements, the level of the required capital (to assets ratio) may have caused the reallocation. These factors are all supply-related. But, since there was a recession during the early 1990s, the decline in lending may have been demand-related. As noted above in Section III, a large empirical literature has searched for evidence of a supply-related credit crunch. The basic approach is to regress bank loan growth on measures of bank capital and control variables. Versions of this regression have been studied by Bernanke and Lown (1991), Hall (1993), Berger and Udell (1994), Haubrich and Wachtel (1993), Hancock and Wilcox (1994), Brinkman and Horvitz (1995), Peek and Rosengren (1995), and Beatty and Gron (2001). With different sample periods, slightly different econometric specifications, different definitions of capital adequacy, and so on, it is not surprising that the results are mixed. Other researchers have tested whether bank supervisors or regulators become tougher during recessions, causing or contributing to a credit crunch. See, for example, Bizer (1993), Wagster (1999), Furfine (2001), and Berger, Kyle, and Scalise (2001). Although there is some evidence that bank capital does affect bank lending, the studies have a difficult time distinguishing loan demand shifts from loan supply shifts, leaving the question of the relative importance of different effects unanswered.

While we discussed bank capital models generally above, there are other papers focusing on the effects of increasing regulatory capital on bank risk taking. Flannery (1989), Furlong and Keeley (1989, 1990), and Kim and Santomero (1988) are examples of partial equilibrium models that examine the relationship between bank leverage and risk-taking. Because these papers are partial equilibrium, the issue of whether there is anything peculiar about banks raising capital is not addressed. Gennotte and Pyle (1991) is more general equilibrium. Gorton and Winton (2000) present a general equilibrium of

bank capital in which there is a social cost to increasing bank capital requirements, as well as a benefit. The benefits are clear; nontransferable charter value, which is also socially beneficial, is protected more with higher capital ratios. But, there is a unique cost to forcing banks to raise capital, namely, they can supply fewer deposits in general equilibrium. Deposits are in demand for the reasons put forth by Gorton and Pennacchi (1990); they provide a way for uniformed agents to transact. In this context, the government never imposes binding capital requirements because it is not socially optimal for bank equity holders to exit the industry, resulting in a smaller banking industry.

E. Other Issues

There are many other regulatory issues that have been studied. In this subsection, we mention a few of the more important or more interesting issues.

The Glass-Steagall Act, passed in the United States in the aftermath of the Great Depression, separated commercial banking from investment banking, ostensibly because of conflicts-of-interest in undertaking lending activity and underwriting activity. While this is no longer the case under current U.S. law (due to changes that culminated in the passage of the Financial Services Modernization Act of 1999), the question whether such separation is socially desirable remains of interest. If banks produce private information in their lending activity, might they have an incentive to only underwrite their poor-quality borrowers? Such questions have been addressed by Ang and Richardson (1994), Gande Puri, Saunders and Walter (1997), Krozner and Rajan (1994), Puri (1994, 1996), Gande, Puri, and Saunders (1999), Yasuda (1999), and Schenone (2001), among others. An open issue is the way in which these joint activities affect bank relationships with borrowers and borrower corporate governance.

Market discipline of banks refers to the extent to which market participants can determine when and which banks are riskier, and impound this information into asset prices. To the extent that market participants are better able to perform this role, there is less need for government oversight. A large literature has investigated this question. For example, Flannery and Sorescu (1996) consider the extent to which bank subordinated debt prices reflect the riskiness of banks. Berger, Davies, and Flannery (2000) compare the information in market security prices to the information produced by bank supervisors. See Flannery (2001) for a discussion.

Another interesting regulatory issue concerns the behavior of the regulators. Kane (1990) focuses on the incentives of regulators during the resolution of insolvent thrifts in the U.S. According to Kane, the Resolution Trust Corporation, the government agency charged with liquidating insolvent thrifts, was inefficient due to the agency problem of the government wanting to avoid recognizing losses that would be borne by the government insurance fund. Dell'Aricia and Marquez (2001) analyze competition among

regulators when there are externalities across national markets. Competition can result in lower regulatory standards.

VI. Conclusion

We have surveyed the major themes and major developments in research in financial intermediation over the last two decades, or so. There are many related topics that we have not covered. For example, the roles of banks in economic growth, and the role of banks in the transmission of monetary policy, are large subjects that we have not touched on at all. Other topics have been only briefly examined. Nevertheless we hope the reader is convinced of the progress made in understanding financial intermediation. Despite this progress there are still major questions. Some of these are:

1. Why do banking crises and banking panics persist? Most theories of banking panics seem inconsistent with the facts put forth by economic historians. The view that banks are inherently unstable seems based on a misreading of U.S. financial history. Overall, it seems that the question of whether banking is inherently unstable or not remains unresolved.
2. Recent experience in emerging markets again raises the issue stability of banking systems. Why are emerging markets experiencing banking panics and crises? Are these crises fundamentally different than earlier historic episodes? Is the basic problem corruption, bad regulations, moral hazard, or some combination?
3. What features of the industrial organization of banking systems make the banks in the system more prone to panics? Are these features consistent with the usual model of competition? In other words, if a few large banks are not panic prone, are they prone to being monopolists?
4. On a related note, how does the industrial organization of the bank lending sector interact with banks' funding and capital structures? Empirically, how does bank funding structure affect the structure of the bank's loans and loan portfolio?
5. The basic paradigm of bank regulation, namely, moral hazard emanating from mispriced deposit insurance, may have outlived its usefulness, if it was ever relevant. It is clear that in the U.S. this has not been a problem since the inception of deposit insurance, nor has clear-cut evidence been produced for the existence of this problem. Corporate governance issues in intermediaries and the intersection of governance and alleged incentives for moral hazard have yet to be fully explored. What is the rationale for government intervention into banking markets?
6. Is corporate governance in banks fundamentally different than nonbanks?
7. Why are deposit insurance and capital requirements such recent developments, especially if their efficacy is as claimed?

8. The differences between loans and bonds seem clear, but questions remain. The existence of loan sales complicates the distinctions that have been made theoretically. The existence of vulture investors, who act like state-contingent banks, buying up blocks of bonds in distressed firms and then actively participating in restructuring, also complicates the distinction. Finally, despite the large literature on potential conflicts of interest, there has been little work on how bank lending and bank underwriting activities jointly affect the corporate governance of bank borrowers.
9. Our survey has focused on the traditional “bank-like” model of monitored finance. There are other forms of monitored finance, such as venture capital. What determines the choice between these different structures? Are these differences driven by the type of firm that seeks financing, or by financing structure of the intermediary?
10. Most models of banking assume that banks and borrowers are perfectly rational. How can insights from the growing field of behavioral finance change our understanding of banks? This question is only just starting to receive attention.⁵⁹
11. Why are bank liabilities used as media of exchange, but not the liabilities of nonbanks? Why can’t demand deposits be traded without being cleared through the banking system? These questions remain largely unexplored.
12. How have loan sales, credit derivatives, commercial paper conduits, collateralized loan obligations, and other recent financial innovations, affected banking?
13. Are banks and stock markets substitutes? Is there a meaningful distinction between of “bank-based” systems and “stock market-based” systems? What are the welfare implications of “bank-based” systems versus “stock market-based” systems?

⁵⁹ Manove and Padilla (1999) examine how the presence of overly-optimistic entrepreneurs complicates bank lending decisions; under competition, banks may be insufficiently conservative, reducing investment efficiency. Coval and Thakor (2001) show how rational agents may endogenously choose to become intermediaries between overly-optimistic agents, who become entrepreneurs, and overly-pessimistic agents, who become depositors.

References

- Acharya, Sankarshan and Jean-Paul Dreyfus (1989), "Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance," Journal of Finance 64, 1313-1333.
- Acharya, Viral, Iftekhar Hasan, and Anthony Saunders (2001), "The Effects of Focus and Diversification: Evidence from Individual Bank Loan Portfolios." Working paper, London Business School.
- Aharony, Joseph and Itzhak Swary (1983), "Contagion Effects of Bank Failures: Evidence from Capital Markets," Journal of Business 56, 302-22.
- Aharony, Joseph and Itzhak Swary (1996), "Additional Evidence on the Information-based Contagion Effects of Bank Failures," Journal of Banking and Finance 20, 57-69.
- Akerlof, George and Paul Romer (1993), "Looting: The Economic Underworld of Bankruptcy for Profit," Brookings Papers on Economic Activity 2, 1-73.
- Akhavain, Jalal, Allen Berger, David Humphrey (1997), "The Effects of Megamergers on Efficiency and Prices: Evidence from a Bank Profit Function." Review of Industrial Organization 12:1, 95-139.
- Allen, Franklin (1990), "The Market for Information and the Origin of Financial Intermediation," Journal of Financial Intermediation 1, 3-30.
- Allen, Franklin (1993), "Stock Markets and Resource Allocation," in Capital Markets and Financial Intermediation, edited by Colin Mayer and Xavier Vives (Cambridge University Press), 81-104.
- Allen, Franklin (2001), "Presidential Address: Do Financial Institutions Matter,?" Journal of Finance, forthcoming.
- Allen, Franklin and Douglas Gale (1997), "Financial Markets, Intermediaries, and Intertemporal Smoothing," Journal of Political Economy 105(3), 523-546.
- Allen, Franklin and Douglas Gale (1998), "Optimal Financial Crises," Journal of Finance 53, 1245-1284.
- Allen, Franklin and Douglas Gale (1999), "Diversity of Opinion and Financing New Technologies," Journal of Financial Intermediation 8, 68-89.
- Allen, Franklin and Douglas Gale (2000), "Financial Contagion," Journal of Political Economy 108, 1-33.
- Almazan, Andres (2002), "A Model of Competition in Banking: Bank Capital vs. Expertise." Journal of Financial Intermediation 11:1, 87-121.
- Alonso, Irasema (1996), "On Avoiding Bank Runs," Journal of Monetary Economics 37, 73-87.

- Alston, Lee J., Wayne A. Grove, and David Wheelock (1994), "Why Do Banks Fail? Evidence from the 1920s." Explorations in Economic History 31, 409-431.
- Andrew, A. Piatt (1907), "The Influence of Crops Upon Business in America," Quarterly Journal of Economics 20, 323-353.
- Ang, James S. and Terry Richardson (1994), "The Underwriting Experience of Commercial Bank Affiliates Prior to the Glass-Steagall Act: A Re-examination of Evidence for Passage of the Act," Journal of Banking and Finance 18, 351-95.
- Arping, Stefan (2001), "Banking, Commerce, and Antitrust." Working paper, University of Lausanne.
- Asquith, Paul, Robert Gertner, and David Scharfstein (1994), "Anatomy of Financial Distress: An Examination of Junk Bond Issuers," Quarterly Journal of Economics 109, 625-634.
- Baer, Herbert, and John McElravey (1992), "The Changing Impact of Capital Requirements on Bank Growth: 1975 to 1991," Federal Reserve Bank of Chicago, working paper.
- Baliga, Sandeep and Ben Polak (2001), "The Emergence and Persistence of the Anglo-Saxon and German Financial Systems," Yale University, Department of Economics, working paper.
- Baltensperger, E. (1980), "Alternative Approaches to the Theory of the Banking Firm," Journal of Monetary Economics 6, 1-37.
- Barro, Robert (1978), "Unanticipated Money, Output, and the Price Level in the United States," Journal of Political Economy 86, 549-580.
- Barro, Jason and Robert Barro (1990), "Pay, Performance and Turnover of Bank CEOs," Journal of Labor Economics 8, 448-481.
- Barsky, R., G. Mankiw, J. Miron, and D. Weil (1988), "The World Wide Change in the Behavior of Interest Rates and Prices in 1914," European Economic Review 32, 1136-1154.
- Barth, James R. (1991), The Great Savings and Loan Debacle (American Enterprise Institute; Washington D.C.).
- Barth, James and Michael Bradley (1989), "Thrift Deregulation and Federal Deposit Insurance," Journal of Financial Services Research 2, 231-259.
- Barth, James, Philip F. Bartholomew, and Michael G. Bradley (1990), "Determinants of Thrift Resolution Costs," Journal of Finance 45, 731-754.
- Beatty, Anne, and Anne Gron (2001), "Capital, Portfolio, and Growth: Bank Behavior Under Risk-Based Capital Guidelines." Journal of Financial Services Research 20:1, 5-31.
- Beck, Thorsten (no date), "Deposit Insurance as a Private Club: Is Germany a Model?," World Bank mimeo.
- Benston George (1976), "A Transaction Cost Approach to the Theory of Financial Intermediation," Journal of Finance 31, 215-231.

- Benston George (1985), An Analysis of the Causes of Savings and Loan Association Failures, Monograph Series in Finance and Economics; New York University Graduate School of Business, Salomon Brothers Center for the Study of Financial Institutions, Monograph 1985-4/5.
- Benston, George, Mike Carhill, and Brian Olovson (1991), "The Failure and Survival of Thrifts: Evidence from the Southeast," in Glenn Hubbard, editor, Financial Markets and Financial Crises (University of Chicago Press; Chicago), 305-384.
- Benveniste, Larry, John Boyd, and Stuart Greenbaum (1991), "Bank Capital Regulation," Osaka Economic Papers 40, 210-226.
- Berger, Allen, Sally Davies, and Mark Flannery (2000), "Comparing Market and Regulatory Assessments of Bank Performance: Who Knows What When?," Journal of Money, Credit and Banking 32, 641-667.
- Berger, Allen, Rebecca Demsetz, and Philip Strahan (1999), "The Consolidation of the Financial Services Industry: Causes, Consequences, and Implications for the Future." Journal of Banking and Finance 23:2-4 (February), 135-194.
- Berger, Allen, Robert DeYoung, Hesna Genay, and Gregory Udell (2000), "Globalization of Financial Institutions: Evidence from Cross-Border Banking Performance." Brookings-Wharton Papers on Financial Services 3, 23-125.
- Berger, Allen, Leora Klapper, and Gregory Udell (2001), "The Ability of Banks to Lend to Informationally Opaque Small Businesses," Working paper, Federal Reserve Board of Governors.
- Berger, Allen, Margaret Kyle, and Joseph Scalise (2001), "Did U.S. Bank Supervisors Get Tougher During the Credit Crunch? Did They Get Easier During the Banking Boom? Did It Matter for Bank Lending?," in Prudential Supervision: What Works and What Doesn't, edited by Frederic Mishkin (University of Chicago Press).
- Berger, Allen, and Loretta Mester (1997), "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?" Journal of Banking and Finance 21:7 (July), 895-947.
- Berger Allen N. and Gregory Udell (1994), "Did Risk-Based Capital Allocate Bank Credit and Cause a 'Credit Crunch' in the U.S.?" Journal of Money, Credit and Banking 26, 585-628.
- Berger Allen N. and Gregory Udell (1995), "Relationship Lending and Lines of Credit in Small Firm Finance," Journal of Business 68, 351-382.
- Berger Allen N. and Gregory Udell (1998), "The Economics of Small Business Finance: The Roles of Private Equity and Debt Markets in the Financial Growth Cycle," Journal of Banking and Finance 22, 613-673.
- Berlin, Mitchell, Kose John, and Anthony Saunders (1996), "Bank Equity Stakes in Borrowing Firms and Financial Distress." Review of Financial Studies 9:3 (Fall), 889-919.

- Berlin, Mitchell and Jan Loeys (1988), "Bond Covenants and Delegated Monitoring," *Journal of Finance* 43, 397-412.
- Berlin, Mitchell and Loretta Mester (1992), "Debt Covenants and Renegotiation," *Journal of Financial Intermediation* 2, 95-133.
- Berlin, Mitchell, and Loretta Mester (1999), "Deposits and Relationship Lending," *Review of Financial Studies* 12:3 (Fall), 579-607.
- Berlin, Mitchell, and Loretta Mester (2001), "Lender Liability and Large Investors," *Journal of Financial Intermediation* 10:2 (April), 108-137.
- Berglof, Erik, and Ernst-Ludwig von Thadden (1994), "Short-Term versus Long-Term Interests: Capital Structure with Multiple Investors," *Quarterly Journal of Economics* 109:4 (November), 1055-1084.
- Bernanke, Ben (1983), "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review* 73, 257-276.
- Bernanke, Ben and Alan Blinder (1988), "Credit, Money and Aggregate Demand," *American Economic Review*, Papers and Proceedings, May.
- Bernanke, Ben and Alan Blinder (1992), "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review* 82, 901-921.
- Bernanke, Ben, and Mark Gertler (1989), "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review* 79:1 (March), 14-31.
- Bernanke, Ben, and Mark Gertler (1990), "Financial Fragility and Economic Performance," *Quarterly Journal of Economics* 105:1 (February), 87-114.
- Bernanke, Ben and Harold James (1991), "The Gold Standard, Deflation, and Financial Crisis in the Great Depression: An International Comparison," in R. Glenn Hubbard, ed., *Financial Markets and Financial Crises* (Chicago), 33-68.
- Bernanke, Ben and Cara Lown (1991), "The Credit Crunch," *Brookings Papers on Economic Activity* 2, 205-247.
- Besanko, David, and George Kanatas (1993), "Credit Market Equilibrium with Bank Monitoring and Moral Hazard," *Review of Financial Studies* 6:1, 213-232.
- Besanko, David, and George Kanatas (1996), "The Regulation of Bank Capital: Do Capital Standards Promote Bank Safety?" *Journal of Financial Intermediation* 5:2 (April), 160-183.
- Besanko, David, and Anjan Thakor (1992), "Banking Deregulation: Allocational Consequences of Relaxing Entry Barriers," *Journal of Banking and Finance* 16:5 (September), 909-932.

- Best, Ronald and Hang Zhang (1993), "Alternative Information Sources and the Information Content of Bank Loans," Journal of Finance 48, 1507-1522.
- Beyen, J. W. (1949), Money in a Maelstrom (New York).
- Bhattacharya, Sudipto, and Gabriella Chiesa (1995), "Proprietary Information, Financial Intermediation, and Research Incentives," Journal of Financial Intermediation 4:4 (October), 328-357.
- Bhattacharya, Sudipto and Douglas Gale (1987), "Preference Shocks, Liquidity and Central Bank Policy," in W.A. Barnett and K. Singleton, editors, New Approaches to Monetary Economics (Cambridge University Press, Cambridge), 69-88.
- Bhattacharya, Sudipto and Paul Pfleiderer (1985), "Delegated Portfolio Management," Journal of Economic Theory 36, 1-25.
- Bhattacharya, Sudipto and Anjan Thakor (1993), "Contemporary Banking Theory," Journal of Financial Intermediation 3, 2-50.
- Billet, Matthew, Mark Flannery, and Jon Garfinkel (1995), "The Effect of Lender Identity on a Firm's Equity Return," Journal of Finance 50, 699-718.
- Bizer, David (1993), "Regulatory Discretion and the Credit Crunch," U.S. Securities and Exchange Commission, working paper.
- Black, Fisher (1975), "Bank Fund Management in an Efficient Market," Journal of Financial Economics 2, 323-339.
- Black, Harold A., M. Cary Collins, Breck L. Robinson, and Robert L. Schweitzer (1997), "Changes in Market Perception of Riskiness: The Case of Too-Big-to-Fail," Journal of Financial Research 20, 389-406.
- Bodenhorn, Howard (1998), "Making the Little Guy Pay: Payments System Networks, Cross-Subsidization, and the Collapse of the Suffolk System," Journal of Economic History, forthcoming.
- Bolton, Patrick and Xavier Freixas (2000), "Equity, Bonds, and Bank Debt: Capital Structure and Financial Market Equilibrium Under Asymmetric Information," Journal of Political Economy 108, 324-351.
- Bolton, Patrick, and Xavier Freixas (2000), "Equity, Bonds, and Bank Debt: Capital Structure and Financial Market Equilibrium under Asymmetric Information," Journal of Political Economy 108:2 (April), 324-351.
- Bolton, Patrick, and David S. Scharfstein (1996), "Optimal Debt Structure and the Number of Creditors," Journal of Political Economy 104:1 (February), 1-25.
- Bonaccorsi di Patti, Emilia, and Giovanni Dell'Ariccia (2001), "Bank Competition and Firm Creation." Working paper, Bank of Italy Research Department.
- Boot, Arnoud (2000), "Relationship Banking: What Do We Know?" Journal of Financial Intermediation 9:1 (January), 7-25.

- Boot, Arnoud, Stuart Greenbaum, and Anjan Thakor (1993), "Reputation and Discretion in Financial Contracting," American Economic Review 83(5), 1165-1183.
- Boot, Arnoud, and Anjan Thakor (1997), "Financial System Architecture." Review of Financial Studies 10:3, 693-733.
- Boot, Arnoud, and Anjan Thakor (2000), "Can Relationship Banking Survive Competition?" Journal of Finance 55:2, 679-713.
- Bordo, Michael (1985), "The Impact and International Transmission of Financial Crises: Some Historical Evidence, 1870-1933," Revista di Storia Economica 2, 41-78.
- Bordo, Michael (1986), "Financial Crises, Banking Crises, Stock Market Crashes and the Money Supply: Some International Evidence, 1870-1933," in Financial Crises and the World Banking System (The Macmillan Press: London), edited by Forrest Capie and Geoffrey Wood, 190-248.
- Bordo, Michael and Barry Eichengreen (1999), "Is Our International Financial Environment Unusually Crisis Prone?" in David Gruen and Luke Gower, editors, Capital Flows and the International Financial System (Sydney: Reserve Bank of Australia), 18-75.
- Bordo, Michael, Claudia Goldin, and Eugene White, editors, (1998), The Defining Moment: The Great Depression and the American Economy in the Twentieth Century (Chicago: University of Chicago Press).
- Bordo, Michael, Hugh Rockoff, and Angela Redish (1995), "A Comparison of the Stability and Efficiency of the Canadian and American Banking Systems, 1870-1925," NBER Historical Working Paper #67.
- Bordo, Michael, Hugh Rockoff, and Angela Redish (1994), "The U.S. Banking System from a Northern Exposure: Stability versus Efficiency," Journal of Economic History 54, 325-341.
- Boyd, John and Mark Gertler (1994a), "Are Banks Dead? Or Are the Reports Greatly Exaggerated?," Federal Reserve Bank of Minneapolis Quarterly Review 18.
- Boyd, John H., and Mark Gertler, (1994b), "The Role of Large Banks in the Recent U.S. Banking Crisis," Federal Reserve Bank of Minneapolis Quarterly Review 18, 2-21.
- Boyd, John, and Stanley Graham (1991), "Investigating the Banking Consolidation Trend." Federal Reserve Bank of Minneapolis Quarterly Review 15:2 (Spring), 3-15.
- Boyd, John and Edward Prescott (1986), "Financial Intermediary Coalitions," Journal of Economic Theory 38, 211-232.
- Boyd, John and Bruce Smith (1994), "How Good are Standard Debt Contracts? Stochastic versus Nonstochastic Monitoring in a Costly State Verification Environment?" Journal of Business 67, 539-561.
- Bremer, C.D. (1935), American Bank Failures (New York: Columbia University Press).

- Brewer, Elijah (1995), "The Impact of Deposit Insurance on S&L Shareholders' Risk/Return Trade-Offs," Journal of Financial Services Research 9, 65-89.
- Brickley, James and Christopher James (1986), "Access to Deposit Insurance, Insolvency Rules and the Stock Returns of Financial Institutions," Journal of Financial Economics 16, 345-371.
- Brickley, James and Christopher James (1987), "The Takeover Market, Corporate Board Composition, and Ownership Structure: The Case of Banking," Journal of Law and Economics 35, 161-180.
- Brinkman, Emile J. and Paul M. Horvitz (1995), "Risk-Based Capital Standards and the Credit Crunch," Journal of Money, Credit and Banking 27, 848-863.
- Bryant, John (1980), "A Model of Reserves, Bank Runs, and Deposit Insurance," Journal of Banking and Finance 4, 335-44.
- Broecker, Thorsten (1990), "Credit-Worthiness Tests and Interbank Competition," Econometrica 58, 429-452.
- Brumbaugh, R. Dan (1988), Thriffs Under Siege (Ballinger Publishing Company; Cambridge, Massachusetts).
- Buser, Steven, Andrew Chen, and Edward Kane (1981), "Federal Deposit Insurance, Regulatory policy, and Optimal Bank Capital," Journal of Finance 35, 51-60.
- Calomiris, Charles (1989), "Deposit Insurance: Lessons from the Record," Federal Reserve Bank of Chicago Economic Perspectives (May/June), 10-30.
- Calomiris, Charles (1990), "Is Deposit Insurance Necessary? A Historical Perspective," Journal of Economic History 50, 283-95.
- Calomiris, Charles (1993a), "Regulation, Industrial Structure, and Instability in U.S. Banking: An Historical Perspective," in Structural Change in Banking (Business One Irwin: Homewood, Illinois), edited by Michael Klausner and Lawrence White, 19-116.
- Calomiris, Charles (1993b), "Financial Factors in the Great Depression," Journal of Economic Perspectives 7, 61-85.
- Calomiris, Charles and Gary Gorton (1991), "The Origins of Banking Panics: Models, Facts, and Bank Regulation," in Glenn Hubbard, editor, Financial Markets and Financial Crises (University of Chicago Press; Chicago), pp. 109-173.
- Calomiris, Charles and Charles Kahn (1991), "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," American Economic Review 81, 497-513.
- Calomiris, Charles and Joseph Mason (2000), "Causes of U.S. Bank Distress During the Depression," National Bureau of Economics, Working Paper no. 7919.
- Calomiris, Charles and Joseph Mason (2001), "Consequences of Bank Distress During the Great Depression," Columbia University, working paper.

- Calomiris, Charles and Larry Schweikart (1991), "The Panic of 1857: Origins, Transmission, and Containment," Journal of Economic History 51, ?.
- Cannella, Albert, Donald Fraser, and D. Scott Lee (1995), "Firm Failure and Managerial Labor Markets: Evidence from Texas Banking," Journal of Financial Economics 38, 185-210.
- Canova, Fabio (1991), "The Sources of Financial Crises: Pre and Post Fed Evidence," International Economic Review 32, 689-713.
- Cantillo Simon, Miguel (1998), "The Rise and Fall of Bank Control in the United States: 1890-1939," American Economic Review 88:5 (December), 1077-1093.
- Cantillo Simon, Miguel and Julian Wright (2000), "How Do Firms Choose Their Lender?," Review of Financial Studies 13, 155-189.
- Capie, Forrest (1997), "The Evolution of Central Banking," chapter in Reforming the Financial System: Some Lessons From History (Cambridge, England: Cambridge University Press).
- Caprio, Gerald and Daniela Klingebiel (1996), "Bank Insolvencies: Cross Country Experience," World Bank Working Paper No. 1620.
- Caramazza, Francesco, Luca Ricci, and Ranil Salgado (1999), "International Financial Contagion," unpublished background paper to the World Economic Outlook, May 1999 (IMF).
- Carletti, Elena (2000), "The Structure of Bank Relationships, Endogenous Monitoring and Loan Rates." Working paper, University of Mannheim.
- Carr, Jack, Frank Mathewson, and Neil Quigley (1995), "Stability in the Absence of Deposit Insurance: The Canadian Banking System, 1890-1966," Journal of Money, Credit and Banking 27, 1137-1158.
- Cavalcanti, R. and Neil Wallace (1999), "Inside and Outside Money as Alternative Media of Exchange," Journal of Money, Credit and Banking , 443-57.
- Cebenoyan, A. Sinan, Elizabeth Cooperman, and Charles Register (1995), "Deregulation, Reregulation, Equity Ownership, and S&L Risk-Taking," Financial Management 24, 63-75.
- Cebenoyan, A. Sinan, Elizabeth Cooperman, and Charles Register (1999), "Ownership Structure, Charter Value, and Risk-Taking Behavior for Thrifts," Financial Management 28, 43-60.
- Cerasi, Vittoria, and Sonja Daltung (2000), "The Optimal Size of a Bank: Costs and Benefits of Diversification," European Economic Review 44:9 (October): 1701-1726.
- Champ, B., Bruce Smith, and Steve Williamson (1996), "Currency Elasticity and Banking Panics: Theory and Evidence," Canadian Journal of Economics 29, 828-864.
- Campbell, Tim and William Kracaw (1980), "Information Production, Market Signaling and the Theory of Financial Intermediation," Journal of Finance 35, 863-881.

- Canova, Fabio (1991), "The Sources of Financial Crises: Pre and Post Fed Evidence," International Economic Review 32, 689-713.
- Cantillo, Miguel and J. Wright (2000). "How do Firms Choose their Lenders? An Empirical Investigation," Review of Financial Studies 13, 155-189.
- Caprio, Gerard and Daniela Klingebiel (1996), "Bank Insolvencies: Cross-Country Experience," World Bank Policy Research Working Paper No. 1620.
- Carey, Mark, Mitchell Post and Steven Sharpe (1998), "Does Corporate Lending by Banks and Financing Companies Differ? Evidence in Private Debt Contracting," Journal of Finance 53, p. ??.
- Carlson, Mark (no date), "Are Branch Banks Better Survivors? Evidence From the Depression Era," mimeo.
- Carr, Jack, Frank Mathewson, and Neil Quigley (1995), "Stability in the Absence of Deposit Insurance: The Canadian Banking System, 1890-1966," Journal of Money, Credit and Banking 27, 1137-1158.
- Cavalcanti, Ricardo de, and Neil Wallace (1999), "A Model of Private Bank-Note Issue," Review of Economic Dynamics 2, 104-36.
- Cetorelli, Nicola, and Michele Gambera (2001), "Banking Market Structure, Financial Dependence and Growth: International Evidence from Industry Data." Journal of Finance 56:2 (April), 617-648.
- Champ, Bruce, Bruce Smith and Stephen Williamson (1996), "Currency Elasticity and Banking Panics: Theory and Evidence," Canadian Journal of Economics 29, 828-864.
- Chan, Y. (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.
- Chari, V.V. and Ravi Jagannathan (1988), "Banking Panics, Information, and Rational Expectations Equilibrium," Journal of Finance 43, 749-763.
- Chong, Beng Soon (1991), "The Effects of Interstate Banking on Commercial Banks' Risk and Profitability." Review of Economics and Statistics 73:1 (February), 78-84.
- Cooper, Russell and Thomas W. Ross (1998), "Bank Runs: Liquidity Costs and Investment Distortion," Journal of Monetary Economics 41, 27-38.
- Coval, Joshua, and Anjan Thakor (2001), "Financial Intermediation as a Beliefs-Bridge Between Optimists and Pessimists." Working paper, University of Michigan Business School.
- Covitz, Daniel, and Erik Heitfeld, "Monitoring, Moral Hazard, and Market Power: A Model of Bank Lending." Working paper, Federal Reserve Board of Governors.
- Cornett, Marcia and Hassan Tehranian (1994), "An Examination of Voluntary versus Involuntary Security Issuances by Commercial Banks," Journal of Financial Economics 35, 99-122.

- Crawford, Anthony, John Ezzell and James Miles (1995), "Bank CEO Pay-Performance Relations and the Effects of Deregulation," Journal of Business 68, 231-256.
- Da Rin, Marco, and Thomas Hellmann (1996), "Banks as Catalysts for Industrialization." Graduate School of Business Research Paper No. 1398, Stanford University.
- Datta, Sudip, Mai Iskandar-Datta, and Ajay Patel (2000), "Some Evidence on the Uniqueness of Initial Public Debt Offerings." Journal of Finance 55:2 (April), 715-743.
- DeGennaro, R. P., L. Lang, and J.B. Thomson (1993), "Troubled Savings and Loan Institutions: Turnaround Strategies Under Insolvency," Financial Management (Autumn), 163-175.
- Dell'Ariccia, Giovanni (2000), "Asymmetric Information and the Structure of the Banking Industry (2000)," European Economic Review, forthcoming.
- Dell'Ariccia, Giovanni, Ezra Friedman, and Robert Marquez (1999), "Adverse Selection as a Barrier to Entry in the Banking Industry." RAND Journal of Economics 30:3 (Autumn), 515-534.
- Dell'Ariccia, Giovanni and Robert Marquez (2001a), "Competition Among Regulators and Credit Market Integration," University of Maryland, working paper.
- Dell'Ariccia, Giovanni, and Robert Marquez (2001b), "Flight to Quality of Captivity? Information and Credit Allocation." Working paper, University of Maryland.
- DeLong, Gayle (2001), "Stockholder Gains from Focusing versus Diversifying Bank Mergers." Journal of Financial Economics 59:2 (February), 221-252.
- Demarzo, Peter and Darrell Duffie (1999), "A Liquidity-Based model of Security Design," Econometrica 67, 65-99.
- Demirgüç-Kunt, Asli, and Enrica Detragiache (1998), "The Determinants of Banking Crises: Evidence from Developed and Developing Countries," IMF Staff Papers 45(1).
- Demirgüç-Kunt, Asli, and Enrica Detragiache (2000), "Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation." Working paper, International Monetary Fund.
- Demirgüç-Kunt, Asli, Enrica Detragiache, and Poonam Gupta (2000), "Inside the Crisis: An Empirical Analysis of Banking Systems in Distress," World Bank, working paper.
- Demirgüç-Kunt, Asli and Vojislav Maksimovic (1990), "Law, Finance and Firm Growth," Journal of Finance 53, 2107-2137.
- Demirgüç-Kunt, Asli and Tolga Sobaci (2000), "Deposit Insurance Around the World: A Data Base," World Bank, mimeo.
- Demsetz, Rebecca, Marc Saidenberg, and Philip Strahan (1996), "Banks with Something to Lose: The Disciplinary Role of Franchise Value," Economic Policy Review 2, 1-14.

- Demsetz, Rebecca, Marc Saidenberg, and Philip Strahan (1997), "Agency Problems and Risk Taking in Banks," Federal Reserve Bank of New York, working paper.
- Demsetz, Rebecca, and Philip Strahan (1997), "Diversification, Size, and Risk at Bank Holding Companies." Journal of Money, Credit, and Banking 29:3 (August), 300-313.
- De Nicro, Gianni (2001), "Size, Charter Value and Risk in Banking: An International Perspective," Working paper, International Monetary Fund.
- Detragiache, Enrica (1994), "Public versus Private Borrowing: A Theory with Implications for Bankruptcy Reform," Journal of Financial Intermediation 3, 327-354.
- Detragiache, Enrica, Paolo Garella, and Luigi Guiso (2000), "Multiple versus Single Banking Relationships: Theory and Evidence." Journal of Finance 55:3 (June), 1133-1161.
- Diamond, Douglas (1989), "Asset Services and Financial Intermediation," in Financial Markets and Incomplete Information, edited by Sudipto Bhattacharya and George Constantinides (Rowman and Littlefield; Maryland), 272-283.
- Diamond, Douglas (1984), "Financial Intermediation and Delegated Monitoring," Review of Economic Studies 51, 393-414.
- Diamond, Douglas (1989), "Reputation Acquisition in Debt Markets," Journal of Political Economy 97, 828-862.
- Diamond, Douglas (1991), "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," Journal of Political Economy 99, 689-721.
- Diamond, Douglas W. (1993a), "Seniority and Maturity of Debt Contracts." Journal of Financial Economics 33:3 (June), 341-368.
- Diamond, Douglas W. (1993b), "Bank Loan Maturity and Priority When Borrowers Can Refinance." In Mayer, Colin, and Xavier Vives, eds., Capital Markets and Financial Intermediation, 46-68. Cambridge University Press, Cambridge, New York, and Melbourne, 1993.
- Diamond, Douglas (1997), "Liquidity, Banks, and Markets," Journal of Political Economy 105, 928-956.
- Diamond, Douglas and Philip Dybvig (1983), "Bank Runs, Deposit Insurance, and Liquidity," Journal of Political Economy 91, 401-419.
- Diamond, Douglas and Raghuram Rajan (2000), "A Theory of Bank Capital," Journal of Finance 55, 2431-2465.
- Diamond, Douglas and Raghuram Rajan (2001a), "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," Journal of Political Economy, forthcoming.
- Diamond, Douglas and Raghuram Rajan (2001b), "Banks and Liquidity," American Economic Review, forthcoming.

- Dinc, Serdar (2000), "Bank Reputation, Bank Commitment, and the Effects of Competition in Credit Markets." Review of Financial Studies 13:3 (Fall), 781-812.
- Donaldson, R. Glen (1993), "Financing Banking Crises: Lessons from the Panic of 1907," Journal of Monetary Economics 31, 69-95.
- Donaldson, R. Glen (1992), "Sources of Panics: Evidence From the Weekly Data," Journal of Monetary Economics 30, 277-305.
- Donaldson, R. Glen (1992), "Costly Liquidation, Interbank Trade, Bank Runs, and Panics," Journal of Financial Intermediation 2, 59-85.
- Dow, James and Gary Gorton (1997), "Stock Market Efficiency and Economic Efficiency: IS There a Connection?," Journal of Finance 52, 1087-1129.
- Dziobek, Claudia and Ceyla Pazarbasioglu (1997), "Lessons from Systemic Bank Restructuring: A Survey of 24 Countries," IMF Working Paper 97-161.
- Economopoulos, A.J. (1988) "Illinois Free Banking Experience," Journal of Money, Credit, and Banking 20 (2), 249-64.
- Economopoulos, A.J. (1990) "Free Banking Failures in New York and Wisconsin: A Portfolio Analysis," Exploration in Economic History 27, 421-41.
- Eichengreen, Barry (1992), Golden Fetters: The Gold Standard and the Great Depression, 1919-1939 (New York: Oxford University Press).
- Elsas, Ralf, and Jan Pieter Krahen (1998), "Is Relationship Lending Special? Evidence from Credit-File Data in Germany." Journal of Banking and Finance 22:10-11 (October), 1283-1316.
- Esbitt, M. (1986), "Bank Portfolios and Bank Failures During the Great Depression: Chicago," Journal of European History 46, 455-62.
- Gehrig, Thomas (1995), "Excessive Risks and Banking Regulation." Working paper, University of Basel.
- Fama, Eugene (1980), "Banking and the Theory of Finance," Journal of Monetary Economics 10, 10-19.
- Fama, Eugene (1985), "What's Different About Banks?," Journal of Monetary Economics 17, 239-249.
- Farinha, Luísa, and João Santos (1999), "Switching from Single to Multiple Bank Lending Relationships: Determinants and Implications." Journal of Financial Intermediation (forthcoming).
- Federal Deposit Insurance Corporation (2000), "Options Paper," (August).
- Fischer, Karl-Hermann (2000) "Acquisition of Information in Loan Markets and Bank Market Power – An Empirical Investigation." Working paper, J. W. Goethe University.

- Fisher, Irving (1932), Booms and Depressions (New York: Adelphi).
- Flannery, Mark (1989), "Capital Requirements and Insured Banks' Choice of Individual Loan Default Risk," Journal of Monetary Economics 24, 235-258.
- Flannery, Mark (1994), "Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing Banking Firms," American Economic Review 84, 320-331.
- Flannery, Mark (1996), "Financial Crises, Payments System Problems, and Discount Window Lending," Journal of Money, Credit and Banking 28, 804-24.
- Flannery, Mark (2001), "The Faces of Market Discipline," Journal of Financial Services Research 20, 107-119.
- Flannery, Mark and Kasturi Rangan (2001), "Market Forces at Work in the Banking Industry: Evidence from the Capital Buildup of the 1990s," University of Florida, working paper.
- Flannery, Mark and Sorin Sorescu (1996), "Evidence of Bank Market Discipline in Subordinated Debenture Yields: 1983-1991," Journal of Finance 51, 1347-1377.
- Flood, Mark D. (1992), "The Great Deposit Insurance Debate," Federal Reserve Bank of St. Louis Review 74 (July/August), 51-77.
- Foglia, A., S. Laviola, and P. Reedtz (1998), "Multiple Banking Relationships and the Fragility of Corporate Borrowers," Journal of Banking and Finance 22:10-11 (October), 1441-1456.
- Fohlin, Caroline (1998), "Relationship Banking, Liquidity, and Investment in the German Industrialization," Journal of Finance 53, 1737-1758.
- Fohlin, Caroline (1999), "Universal Banking in Pre-World War I Germany: Model or Myth?," Explorations in Economic History 36, 305-343.
- Franks, Julian and Walter Torus (1994), "A Comparison of Financial Recontracting in Distressed Exchanges and Chapter 11 Reorganizations," Journal of Financial Economics 35, 349-370.
- Freeman, Scott (1996a), "The Payments System, Liquidity, and Rediscounting," American Economic Review 86, 1126-38.
- Freeman, Scott (1996b), "Clearinghouse Banks and Banknote Over-Issue," Journal of Monetary Economics 38, 101-15.
- Freeman, Scott (1988), "Banks as the Provision of Liquidity," Journal of Business 61, 45-64.
- Friedman, Milton and Anna J. Schwartz (1963), A Monetary History of the United States, 1867-1960 (New York: National Bureau of Economic Research).
- Furfine, Craig (2001), "Bank Portfolio Allocation: The Impact of Capital Requirements, Regulatory Monitoring, and Economic Conditions," Journal of Financial Services Research 20, 33-56.

- Furlong, Frederick and Michael Keeley (1989), Capital Regulation and Bank Risk-Taking: A Note,” Journal of Banking and Finance 13, 883-891.
- Furlong, Frederick and Michael Keeley (1990), “A reexamination of Mean-Variance Analysis of Bank Capital Regulation,” Journal of Banking and Finance 14, 69-84.
- Gale, Douglas, and Martin Hellwig (1985), “Incentive-Compatible Debt Contracts: The One-Period Problem.” Review of Economic Studies 52:4, 647-663.
- Gande, Amar, Manju Puri, and Anthony Saunders (1999), “Bank Entry, Competition and the Market for Corporate Securities Underwriting,” Journal of Financial Economics 54, 165-195.
- Gande, Amar, Manju Puri, Anthony Saunders, and Ingo Walter (1997), “Bank Underwriting of Debt Securities: Modern Evidence,” Review of Financial Studies 10, 1175-1202.
- Garcia, Gillian (1999), “Deposit Insurance: A Survey of Actual and Best Practices,” IMF Working Paper No. 96/83.
- Gehrig, Thomas (1998), “Screening, Cross-Border Banking, and the Allocation of Credit.” Research in Economics 52:4 (December), 387-407.
- Gehrig, Thomas, and Rune Stenbacka (2001), “Information Sharing in Banking: A Collusive Device?” CEPR Discussion Paper No. 2911.
- Genotte, Gerard and David Pyle (1991), “Capital Controls and Bank Risk,” Journal of Banking and Finance 15, 805-824.
- Gibson, Michael (1995), “Can Bank Health Affect Investment? Evidence from Japan,” Journal of Business 68, 281-308.
- Gilson, Stuart (1990), “Bankruptcy, Boards, Banks, and Blockholders: Evidence on Changes in Corporate Ownership and Control When Firms Default,” Journal of Financial Economics 27, 355-388.
- Gilson, Stuart, Kose John, and Larry Lang (1990), “Troubled Debt Restructurings: An Empirical Study of Private Reorganization of Firms in Default,” Journal of Financial Economics 27, 315-353.
- Goldstein, Itay and Ady Pauzner (1999), “Demand Deposit Contracts and the Probability of Runs,” Eitan Berglas School of Economics, Tel Aviv University, mimeo.
- Gorton, Gary (1988), “Banking Panics and Business Cycles,” Oxford Economic Papers 40, 751-781.
- Gorton, Gary (1985), “Clearinghouses and the Origin of Central Banking in the U.S.,” Journal of Economic History 45, 277-283.
- Gorton, Gary (1996), “Reputation Formation in Early Bank Note Markets,” Journal of Political Economy 104, 346-397.

- Gorton, Gary (1999), "Pricing Free Bank Notes," Journal of Monetary Economics 44, 33-64.
- Gorton, Gary and Joseph Haubrich (1987), "Bank Deregulation, Credit Markets, and the Control of Capital," Carnegie-Rochester Conference Series on Public Policy 26, 289-334.
- Gorton, Gary and Lixin Huang (2001a), "Bank Panics and the Endogeneity of Central Banking," Wharton School, University of Pennsylvania, mimeo.
- Gorton, Gary and Lixin Huang (2001b), "Banking Panics and the Origin of Central Banking," in The Origin and Evolution of Central Banking (Cambridge University Press), forthcoming.
- Gorton, Gary and Donald Mullineaux (1987), "The Joint Production of Confidence: Endogenous Regulation and the 19th Century Commercial Bank Clearinghouse," Journal of Money, Credit, and Banking 19, 457-468.
- Gorton, Gary and George Pennacchi (1990), "Financial Intermediaries and Liquidity Creation," Journal of Finance 45, 49-72.
- Gorton, Gary and George Pennacchi (1995), "Bank and Loan Sales: Marketing Non-Marketable Assets," Journal of Monetary Economics 35, 389-411.
- Gorton, Gary, and Frank Schmid (2000), "Universal Banking and the Performance of German Firms," Journal of Financial Economics, 58:1-2 (October/November), 29-80.
- Gorton, Gary, and Andrew Winton (2000), "Liquidity Provision, Bank Capital, and the Macroeconomy." Working paper, University of Minnesota.
- Green, Edward (1997), "Money and Debt in the Structure of Payments," Bank of Japan Monetary and Economic Studies 15, 63-87.
- Green, Edward and Ping Lin (2000), "Diamond and Dybvig's Classic Theory of Financial Intermediation: What's Missing?," Federal Reserve Bank of Minneapolis Quarterly Review 24, 3-13.
- Green, Edward and Ping Lin (1999), "Implementing Efficient Allocations in a Model of Financial Intermediation," Centre for Public Policy Studies, Lingnan University, Hong Kong, Working Paper #92.
- Green, Richard C. (1984), "Investment Incentives, Debt, and Warrants," Journal of Financial Economics 13:1 (March), 115-136.
- Greenbaum, Stuart, George Kanatas, and Itzhak Venezia (1989), "Equilibrium Loan Pricing Under the Bank Client Relationship," Journal of Banking and Finance 13, 221-235.
- Grossman, Richard (1992), "Deposit Insurance, Regulation, and Moral Hazard in the Thrift Industry: Evidence from the 1930's," American Economic Review 82, 800-22.
- Grossman, Richard (1994), "The Shoe That Didn't Drop: Explaining Banking Stability During the Great Depression," Journal of Economic History 54, 654-682.

- Grossman, Sanford and Joseph Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets," American Economic Review 61, 393-408.
- Gupta, Atul and Lalatendu Misra (1999), "Failure and Failure Resolution in the U.A. Thrift and Banking Industries," Financial Management 28, 87-105.
- Hadlock, Charles and Christopher James (2000), "Bank Lending and the Menu of Financing Options," University of Florida, mimeo.
- Hall, Brian (1993), "How Has the Basel Accord Affected Bank Portfolios?" Journal of the Japanese and International Economies 7, 408-440.
- Hamilton, David E. (1985), "The Causes of the Banking Panic of 1930: Another View," Journal of Southern History LI, 581-608.
- Hancock, Diana and James A. Wilcox (1994), "Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulation," Journal of the American Real Estate And Urban Economics Association 22, 59-94.
- Hannan, Timothy and Alan Berger (1991), "The Rigidity of Prices: Evidence form the Banking Sector," American Economic Review 81, 938-945.
- Hart, Oliver, and John Moore (1998), "Default and Renegotiation: A Dynamic Model of Debt." Quarterly Journal of Economics 113:1 (February), 1-41.
- Haubrich, Joseph (1989), "Financial Intermediation, Delegated Monitoring, and Long-Term Relationships," Journal of Banking and Finance 13, 9-20.
- Haubrich, Joseph (1990), "Nonmonetary Effects of Financial Crises: Lessons From the Great Depression in Canada," Journal of Monetary Economics 25, 223-252.
- Haubrich, Joseph and Robert King (1990), "Banking and Insurance," Journal of Monetary Economics 26, 361-386.
- Haubrich, Joseph, and João Santos (2002), "Alternative Forms of Mixing Banking and Commerce: Evidence from American History." Working paper, Federal Reserve Bank of New York.
- Haubrich, Joseph and Paul Wachtel (1993), "Capital Requirements and Shifts in Commercial Bank Portfolios," Federal Reserve Bank of Cleveland Economic Review 29, 2-15.
- Hauswald, Robert, and Robert Marquez (2000), "Competition and Strategic Focus in Lending Relationships." Working paper, University of Maryland.
- Hellwig, Martin (1994), "Liquidity Provision, Banking, and the Allocation of Interest Rate Risk." European Economic Review 38:7, 1363-1389.
- Hellwig, Martin (1998), "Allowing for Risk Choices in Diamond's 'Financial Intermediation as Delegated Monitoring'." Discussion Paper No. 98-04, Sonderforschungsbereich 504, University of Mannheim.

- Himmelberg, Charles and Donald Morgan (1995), "Is Bank Lending Special?," in Is Bank Lending Important for the Transmission of Monetary Policy, Conference Series No. 39, edited by J. Peek and E. Rosengren, Federal Reserve Bank of Boston.
- Holmstrom, Bengt and Jean Tirole (1998), "Private and Public Supply of Liquidity," Journal of Political Economy 106, 1-40.
- Holmstrom, Bengt and Jean Tirole (1997), "Financial Intermediation, Loanable Funds, and the Real Sector," Quarterly Journal of Economics 112, 663-691.
- Hoshi, Takeo, Anil Kashyap (1999), "The Japanese Banking Crisis: Where Did It Come From and How Will It End?," NBER Macroeconomics Annual, ???.
- Hoshi, Takeo, Anil Kashyap, and David Scharfstein (1990a), "Bank Monitoring and Investment: Evidence from the Changing Structure of Japanese Corporate Banking Relationships," in R. Glenn Hubbard, editor, Asymmetric Information, Corporate Finance and Investment (University of Chicago Press, Chicago, Illinois).
- Hoshi, Takeo, Anil Kashyap, and David Scharfstein (1990b), "The Role of Banks in Reducing the Costs of Financial Distress in Japan," Journal of Financial Economics 27, 67-88.
- Hoshi, Takeo, Anil Kashyap, and David Scharfstein (1991), "Corporate Structure, Liquidity and Investment: Evidence from Japanese Industrial Groups," Quarterly Journal of Economics 106, 33-60.
- Houston, Joel and Christopher James (1995), "CEO Compensation and Bank Risk: Is Compensation in Banking Structured to Promote Risk Taking?," Journal of Monetary Economics 36, 405-431,
- Houston, Joel, and Christopher James (1996), "Bank Information Monopolies and the Mix of Private and Public Debt Claims." Journal of Finance 51:5 (December), 1863-1889.
- Hovakimian, Armen and Edward Kane (2000), "Effectiveness of Capital Regulation at U>S> Commerical Banks, 1985-1994," Journal of Finance 55, 451-468.
- Hubbard, R. Glenn, Kenneth Kuttner, and Darius Palia (1999), "Are There 'Bank Effects' in Borrower's Costs of Funds? Evidence from a Matched Sample of Borrowers and Banks." Journal of Business (forthcoming).
- Hubbard, Glenn and Darius Palia (1995), "Executive Pay and Performance: Evidence from the U.S. Banking Industry," Journal of Financial Economics 39, 105-130.
- Hughes, Joseph, William Lang, Loretta Mester, and Choon-Geol Moon (1999), "The Dollars and Sense of Bank Consolidation." Journal of Banking and Finance 23:2-4 (February), 291-324.
- Hunter, Helen M. (1982), "The Role of Business Liquidity During the Great Depression and Afterwards: Differences Between Large and Small Firms," Journal of Economic History 42, 883-902.

- International Monetary Fund (1998), "Financial Crises: Characteristics and Indicators of Vulnerability," World Economic Outlook (May), 74-97.
- Jacklin, Charles (1987), "Demand Deposits, Trading Restrictions, and Risk-Sharing," in Contractual Arrangements for Intertemporal Trade, edited by Edward Prescott and Neil Wallace (University of Minnesota Press; Minneapolis, MN), p. 26-47.
- Jacklin, Charles (1989), "Banks and Risk-Sharing: Instabilities and Coordination," in Financial Markets and Incomplete Information, edited by Sudipto Bhattacharya and George Constantinides, Rowan and Littlefield, Maryland, 317-327.
- Jacklin, Charles and Sudipto Bhattacharya (1988), "Distinguishing panics and Information-Based Bank Runs: Welfare and Policy Implications," Journal of Political Economy 96, 568-592.
- James, Christopher (1996), "Bank Debt Restructurings and the Composition of Exchange Offers in Financial Distress," Journal of Finance 51, 711-727.
- James, Christopher (1995), "When Do Banks Take Equity? An Analysis of Bank Loan Restructurings and the Role of Public Debt," Review of Financial Studies 8, 1209-1234.
- James, Christopher (1987), "Some Evidence on the Uniqueness of Bank Loans," Journal of Financial Economics 19, 217-233.
- James, Christopher (1984), "An Analysis of the Effect of State Acquisition Laws on Managerial Efficiency: The Case of the Bank Holding Company Acquisitions," Journal of Law and Economics 27, 211-226.
- James, Harold (1986), The German Slump: Politics and Economics, 1924-1936 (Oxford).
- James, Harold (2001), The End of Globalization: Lessons from the Great Depression (Harvard University Press).
- Jean-Baptiste, Eslyn (1999), "Demand Deposits as an Incentive Mechanism," Wharton School, University of Pennsylvania, mimeo.
- Jean-Baptiste, Eslyn (2001), "Information Monopoly and Commitment in Intermediary-Firm Relationships," Graduate School of Business, Columbia University, working paper.
- Jensen, Michael, and William Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," Journal of Financial Economics 3:4 (October), 305-360.
- John, Kose, Theresa John, and Anthony Saunders (1994), "Universal Banking and Firm Risk-Taking," Journal of Banking and Finance 18, 307-323.
- Johnson, Shane (1997), "An Empirical Analysis of the Determinants of Corporate Debt Ownership Structure," Journal of Financial and Quantitative Analysis 32:1 (march), 47-69.
- Kaminsky, Graciela and Carmen Reinhart (1999), "The Twin Crises: The Causes of Banking and Balance of Payments Crises," American Economic Review 89, 473-500.

- Kaminsky, Graciela and Carmen Reinhart (1998), "On Crises, Contagion, and Confusion," Journal of International Economics, forthcoming.
- Kanatas, George, and Jianping Qi (2001), "Imperfect Competition, Agency, and Financing Decisions." Journal of Business 74:2 (April), 307-338.
- Kane, Edward J. (1989), The S & L Mess: How Did it Happen? (Urban Institute Press; Washington D.C.).
- Kane, Edward (1990), "Principal Agent Problems in S&L Salvage," Journal of Finance 45, 755-764.
- Kane, Edward and Berry Wilson (1998), "A Contracting Theory Interpretation of the Origins of Federal Deposit Insurance," Journal of Money, Credit and Banking 30, 573-595.
- Kang, Jun-Koo, and Rene Stulz (2000), "Do Banking Shocks Affect Borrowing Firm Performance? An Analysis of the Japanese Experience." Journal of Business 73:1 (January), 1-23.
- Kaplan, Steven, and Bernadette Minton (1994), "Appointments of Outsiders to Japanese Boards: Determinants and Implications for Managers." Journal of Financial Economics 36:2 (October), 225-258.
- Kaplan, Steven, and Per Strömberg (2001), "Venture Capitalists As Principals: Contracting, Screening, and Monitoring." American Economic Review 91:2, 426-430.
- Kashyap, Anil, Raghuram Rajan, and Jeremy Stein (2001), "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking," Journal of Finance, forthcoming.
- Kashyap, Anil and Jeremy Stein (1994), "Monetary Policy and Bank Lending," in Gregory N. Mankiw, editor, Monetary Policy, Studies in Business Cycles, Vol. 29, University of Chicago Press; Chicago, 221-56.
- Kashyap, Anil, and Jeremy Stein (2000), "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" American Economic Review 90:3 (June), 407-428.
- Kashyap, Anil, Jeremy Stein, and David Wilcox (1993), "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance," American Economic Review 83, 78-98.
- Kaufman, George (1994), "Bank Contagion: A Review of the Theory and Evidence," Journal of Financial Services Research 8, 123-150.
- Keeley, Michael C. (1990), "Deposit Insurance, Risk, and Market Power in Banking," American Economic Review 80, 1183-1200.

- Keeton, William (1994), "Causes of the Recent Increase in Bank Security Holdings," Federal Reserve Bank of Kansas City Economic Review (Second Quarter), 45-57.
- Kelley, Morgan and Cormac Ó Gráda (2000), "Market Contagion: Evidence from the Panics of 1854 and 1857," American Economic Review 90(5), 1110-1124.
- Kemmerer, E. W. (1910), Seasonal Variations in the Relative Demand for Money and Capital in the United States (National Monetary Commission. 61st Cong., 2nd sess., Senate Doc. 588, Washington D.C.; Government Printing Office).
- Kihlstrom, Richard and Steven Mathews (1990), "Managerial Incentives in an Entrepreneurial Stock Market Model," Journal of Financial Intermediation 1, 57-79.
- Kim, Daesik and Anthony Santomero (1988), "Risk in Banking and Capital Regulation," Journal of Finance 43, 1219-1233.
- Kindleberger, Charles (1973), The World in Depression (Berkeley: University of California Press).
- King, Robert and Ross Levine (1993), "Finance and Growth: Schumpeter Might Be Right," Quarterly Journal of Economics 108, 717-38.
- Kormendi, Roger, Victor Bernard, S. Craig Pirrong, and Edward Snyder (189), Crisis Resolution in the Thrift Industry (Kluwer Academic Publishers; Boston).
- Krasa, Stefan and Ann Villamil (1992), "Monitoring the Monitor: An Incentive Structure for a Financial Intermediary," Journal of Economic Theory 57, 197-221.
- Krasa, Stefan and Ann Villamil (1993), "A Theory of Optimal Bank Size," Oxford Economic Papers 44, 725-749.
- Kroszner, Randall and Philip Strahan (1999), "Bankers on Boards: Monitoring, Conflicts of Interest, and Lender Liability," Journal of Financial Economics, forthcoming.
- Kroszner, Randall, and Philip Strahan (2001), "Throwing Good Money after Bad? Board Connections and Conflicts in Bank Lending." Working paper, University of Chicago.
- Kroszner, Randall and Raghuram Rajan (1994), "Is the Glass-Steagall Act Justified? A Study of the United States Experience with Universal Banking Before 1933," American Economic Review 84, 810-32.
- Kroszner, Randall and Raghuram Rajan (1997), "Organizational Structure and Credibility: Evidence from Commercial Bank Activities Before the Glass-Steagall Act," Journal of Monetary Economics 39, 475-516.
- Krysanowski, Lawrence and Gordon S. Roberts (1993), "Canadian Banking Solvency, 1922-1940," Journal of Money, Credit and Banking 25, 361-376.
- Krysanowski, Lawrence and Gordon S. Roberts (1999), "Perspectives on Canadian Banking Insolvency During the 1930s," Journal of Money, Credit and Banking 31, 130-136.

- Kyei, Alexander (1995), "Deposit Protection Arrangements: A Survey," IMF Working Paper 95/134 (Washington: International Monetary Fund).
- Kyle, Albert S. (1985), "Continuous Auctions and Insider Trading," Econometrica 53, 1335-1355.
- Levine, Ross (1997), "Financial Development and Growth: Views and Agenda," Journal of Economic Literature 35, 688-726.
- Levine, Ross (1998), "The Legal Environment, Banks, and Long-Run Economic Growth," Journal of Money, Credit, and Banking 30, 596-613.
- Levine, Ross (2000), "Bank-based or Market-based Financial Systems: Which is Better?," University of Minnesota, mimeo.
- Levine, Ross and Sara Zervos (1998), "Stock Markets, Banks and Economic Growth," American Economic Review 88, 537-58.
- Lindgren, Carl-Johan, Gillian Garcia, and Matthew Saal (1996), Bank Soundness and Macroeconomic Policy (Washington: International Monetary Fund).
- Lizzeri, Alessandro (1999), "Information Revelation and Certification Intermediaries," Rand Journal of Economics 30, 214-231.
- Longhofer, Stanley, and Joao Santos (2000), "The Importance of Bank Seniority for Relationship Lending," Journal of Financial Intermediation 9:1 (January), 57-89.
- Lummer, Scott and John McConnell (1989), "Further Evidence on the Bank Lending Process and the Capital Market Responses to Bank Loan Agreements," Journal of Financial Economics 25, 99-122.
- Mahrt-Smith, Jan (2000), "Should Banks Own Equity? A Corporate Finance Perspective." Working paper, London Business School.
- Mailath, George and Loretta Mester (1994), "A Positive Analysis of Bank Closure," Journal of Financial Intermediation 3, 272-299.
- Maksimovic, Vojislav (1990), "Product Market Imperfections and Loan Commitments," Journal of Finance 45:5 (December), 1641-1653.
- Manove, Michael, and Jorge Padilla (1999), "Banking (Conservatively) with Optimists," RAND Journal of Economics 30:2 (Summer), 324-350.
- Manove, Michael, Jorge Padilla, and Marco Pagano (2001), "Collateral versus Project Screening: A Model of Lazy Banks," RAND Journal of Economics 32:4 (Winter), 726-744.
- Marcus, Alan (1984), "Deregulation and Bank Financial Policy," Journal of Banking and Finance 8, 557-565.
- Marcus, Alan and Israel Shaked (1984), "The Valuation of FDIC Deposit Insurance Using Option Pricing Estimates," Journal of Money, Credit and Banking 16, 446-460.

- Marquez, Robert (2001), "Competition, Adverse Selection, and Information Dispersion in the Banking Industry." Review of Financial Studies, forthcoming,
- Matutes, Carmen, and Xavier Vives (1996), "Competition for Deposits, Fragility, and Insurance." Journal of Financial Intermediation 5:2 (April), 184-216.
- Mayer, Colin (1990), "Financial Systems, Corporate Finance, and Economic Development," in Asymmetric Information, Corporate Finance, and Investment, edited by Glenn Hubbard (Chicago: University of Chicago Press).
- McAndrews, James and William Roberds (1999), "Payment Intermediation and the Origins of Banking," Federal Reserve Bank of Atlanta, Working Paper # 99-11.
- McKenzie, Joseph, Rebel Cole, and Richard Brown (1992), "Moral Hazard, Portfolio Allocation, and Asset Returns for Thrift Institutions," Journal of Financial Services Research 5, 315-339.
- Merton, Robert (1977), "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees: An Application of Modern Option Pricing Theory," Journal of Banking and Finance 1, 3-11.
- Mester, Loretta and Mitchell Berlin (1992), "Debt Covenants and Renegotiation," Journal of Financial Intermediation 2, 95-133.
- Mester, Loretta, Leonard Nakamura, and Micheline Renault (2001), "Checking Accounts and Bank Monitoring," Federal Reserve Bank of Philadelphia Working paper #01-3.
- Mikkelson, Wayne and Megan Partch (1986), "Valuation Effects of Security Offerings and the Issuance Process," Journal of Financial Economics 15, 31-60.
- Miron, Jeffery A. (1986), "Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed," American Economic Review 76, 125-140.
- Miron, Jeffrey, G. Mankiw, and D. Weil (1987), "The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve," American Economic Review 77, 358-374.
- Mishkin, Frederic (1991), "Asymmetric Information and Financial Crises: A Historical Perspective," in Glenn Hubbard, editor, Financial Markets and Financial Crises (University of Chicago Press; Chicago), 69-108.
- Morck, Randall, Masao Nakamura, and Anil Shivdasani (2000), "Banks, Ownership Structure, and Firm Value in Japan." Journal of Business 73:4 (October), 539-567.
- Moen, Jon and Ellis Tallman (2000), "Clearinghouse Membership and Deposit Contraction during the Panic of 1907," Journal of Economic History 60, 145-163.
- Moen, Jon and Ellis Tallman (1992), "The Bank Panic of 1907: The Role of Trust Companies," Journal of Economic History 52, 611-630.

- Morgan, George (1992), "Capital Adequacy," in The New Palgrave Dictionary of Money and Finance, edited by Peter Newman, Murray Milgate, and John Eatwell (Macmillan Press).
- Morgan, Donald (1998), "The Credit Effects of Monetary Policy: Evidence Using Loan Commitments," Journal of Money, Credit, and Banking 30, 102-18.
- Morgan, Donald (2000), "Rating Banks: Risk and Uncertainty in an Opaque Industry," Federal Reserve Bank of New York, Staff Report No. 105.
- Morris, Stephen and Hyun Song Shin (2000), "Rethinking Multiple Equilibria in Macroeconomic Modeling," NBER Macroeconomics Annual, forthcoming.
- Mullineaux, Donald (1987), "Competitive Monies and the Suffolk System: A Contractual Perspective," Southern Economic Journal 53, 884-898.
- Mullineaux, Donald and Dianna Preece (1996), "Monitoring, Loan Renegotiability, and Firm Value: The Role of Lending Syndicates," Journal of Banking and Finance 20, 577-593.
- Myers, Stewart and Nicholas Majluf (1984), "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," Journal of Financial Economics 13, 187-221.
- Neumark, David, and Steven A. Sharpe, "Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits." Quarterly Journal of Economics 107:2 (May), 657-680.
- Ó Gráda, Cormac and Eugene N. White (2001), "Who Panics During Panics? Evidence from a Nineteenth Century Savings Bank," Rutgers University, Working Paper.
- O'Hara, Maureen and Wayne Shaw (1990), "Deposit Insurance and Wealth Effects: The Value of Being "Too Big To Fail," Journal of Finance 45, 1587-1600.
- Ongena, Steven, and David C. Smith (2000), "What Determines the Number of Bank Relationships? Cross-Country Evidence." Journal of Financial Intermediation 9:1 (January), 26-56.
- Ongena, Steven and David C. Smith (2000), "The Duration of Bank Relationships," Journal of Financial Economics, forthcoming.
- Ostergaard, Charlotte (2001), "External Financing Costs and Banks' Loan Supply: Does the Structure of the Bank Sector Matter?" Working paper, Norwegian School of Management.
- Padilla, A. Jorge, and Marco Pagano (1997), "Endogenous Communication among Lenders and Entrepreneurial Incentives." Review of Financial Studies 10:1, 205-236.
- Padilla, A. Jorge, and Marco Pagano (2000), "Sharing Default Information as a Borrower Discipline Device." European Economic Review 44, 1951-1980.
- Pagano, Marco, and T. Japelli (1993), "Information Sharing in Credit markets." Journal of Finance 48, 1693-1718.

- Park, Sangkyun (1991). "Bank Failure Contagion in Historical Perspective," Journal of Monetary Economics 28, 271-286.
- Park, Cheol (2000), "Monitoring and Structure of Debt Contracts." Journal of Finance v55, n5 (October 2000): 2157-2195.
- Peek, Joe and Eric Rosengren (1995), "The Capital Crunch: Neither a Borrower Nor a Lender Be," Journal of Money, Credit and Banking 27, 625-638.
- Peek, Joe and Eric Rosengren (1997), "The International Transmission of Financial Shocks: The Case of Japan," American Economic Review 87(4), 495-505.
- Peek, Joe and Eric Rosengren (2000), "Collateral Damage: Effects of the Japanese Bank Crisis on Real Economic Activity in the United States," American Economic Review 87(4), 30-45.
- Pennacchi, George (1987), "A Reexamination of the Over- (or Under-)Pricing of Deposit Insurance," Journal of Money, Credit and Banking 19, 340-360.
- Petersen, Mitchell and Raghuram Rajan (1994), "The Benefits of Lending Relationships: Evidence From Small Business Data," Journal of Finance 49, 3-37.
- Petersen, Mitchell and Raghuram Rajan (1995), "The Effect of Credit Market Competition on Lending Relationships," Quarterly Journal of Economics 110, 407-443.
- Polonchek, John, Myron Slovin, and Marie Sushka (1989), "Valuation Effects of Commercial Bank Securities Offerings: A Test of the Information Hypothesis," Journal of Banking and Finance 13, 443-461.
- Postlewaite, Andrew and Xavier Vives (1987), "Bank Runs as an Equilibrium Phenomenon," Journal of Political Economy 95, 485-491.
- Preece, D. and Donald Mullineaux (1989), "The Nonuniqueness of Bank Loans," University of Kentucky, mimeo.
- Prowse, Stephen (1995), "Alternative Methods of Corporate Control of Commercial Banks," Federal Reserve Bank of Dallas Economic Review, Third Quarter, 24-36.
- Puri, Manju (1994), "The Long-term Default Performance of Bank Underwritten Security Issues," Journal of Banking and Finance 18, 397-418.
- Puri, Manju (1996), "Commercial Banks in Investment Banking: Conflict of Interest or Certification Role?," Journal of Financial Economics 40, 373-401.
- Qi, Jianping (1994), "Bank Liquidity and Stability in an Overlapping Generations Model," Review of Financial Studies 7, 389-417.
- Rajan, Raghuram (1992), "Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt," Journal of Finance 47, 1367-1400.

- Rajan, Raghuram (1994), "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence." Quarterly Journal of Economics 109:2 (May), 399-441.
- Rajan, Raghuram, and Andrew Winton (1995), "Covenants and Collateral as Incentives to Monitor." Journal of Finance 50:4 (September), 1113-1146.
- Rajan, Raghuram and Luigi Zingales (1998), "Financial Dependence and Growth," American Economic Review 88, 559-86.
- Ramakrishnan, R.T.S. and Anjan Thakor (1984), "Information Reliability and a Theory of Financial Intermediation," Review of Economic Studies 51, 415-432.
- Repullo, Rafael , and Javier Suarez (1998), "Monitoring, Liquidation, and Security Design." Review of Financial Studies 11:1 (Spring), 163-187.
- Riordan, Michael (1993), "Competition and Bank Performance: A Theoretical Perspective." In Mayer, Colin, and Xavier Vives, eds., Capital Markets and Financial Intermediation, 328-343. Cambridge University Press, Cambridge, New York, and Melbourne, 1993.
- Ritchken, Peter, James Thomson, Ramon DeGennaro, and Anlong Li (1993), "On Flexibility, Capital Structure and Investment Decisions for the Insured Bank," Journal of Banking and Finance 17, 1133-46.
- Rochet, Jean-Charles and Jean Tirole (1996), "Interbank Lending and Systemic Risk," Journal of Money, Credit and Banking 28, 733-62.
- Rockoff, Hugh (1974), "The Free Banking Era: A Reinterpretation," Journal of Money, Credit and Banking 6, 141-167.
- Rockoff, Hugh (1975), The Free Banking Era: A Reconsideration (Arno Press; New York).
- Rockoff, Hugh (1993), "The Meaning of Money in the Great Depression," National Bureau of Economics, Historical Research Paper No. 52.
- Rolnick, Arthur and Warren Weber (1982), "Free Banking, Wildcat Banking and Shinplasters," Quarterly Review, Federal Reserve Bank of Minneapolis (Fall).
- Rolnick, Arthur and Warren Weber (1983), "New Evidence on the Free Banking Era," American Economic Review 73(5), 1080-91.
- Rolnick, Arthur and Warren Weber (1984), "The Causes of Free Bank Failures: A Detailed Examination," Journal of Monetary Economics 14, 267-292.
- Rolnick, Arthur and Warren Weber (1988), "Explaining the Demand for Free Bank Notes," Journal of Monetary Economics 21, 47-72.
- Rolnick, Arthur, and Warren Weber (1998a), "The Suffolk Banking System Reconsidered," Federal Reserve Bank of Minnesota, Working Paper 587D.

- Rolnick, Arthur, Bruce Smith and Warren Weber (1998a), "The Suffolk Bank and the Panic of 1837: How a Private Bank Acted as a Lender-of-Last-Resort," Federal Reserve Bank of Minnesota, Working Paper 592.
- Ronn, Ehud and Avinash Verma (1986), "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model," Journal of Finance 41, 871-895.
- Ruckes, Martin (1998), "Competing Banks, Credit Standards, and Corporate Conservatism." Working paper, University of Mannheim.
- Sahlman, William (1990), "The Structure and Governance of Venture-Capital Organizations." Journal of Financial Economics 27, 473-521.
- Santos, João (1998), "Banking and Commerce: How Does the United States Compare to Other Countries?" Economic Review, Federal Reserve Bank of Cleveland, 34:4, 14-26.
- Santos, João (1999), "Bank Capital and Equity Investment Regulations." Journal of Banking and Finance 23, 1095-1120.
- Sapienza, Paola (2000), "The Effects of Banking Mergers on Loan Contracts," forthcoming, Journal of Finance.
- Saunders, Anthony, E. Strock, and N.G. Travlos (1990), "Ownership Structure, Deregulation, and Bank Risk Taking," Journal of Finance Vol ? , 643-654.
- Saunders, Anthony and Berry Wilson (1996), "Contagious Bank Runs: Evidence from the 1929-1933 Period," Journal of Financial Intermediation 5, 409-423.
- Saunders, Anthony and Berry Wilson (1999), "The Impact of Consolidation and Safety-net Support on Canadian, US, and UK Banks: 1893-1992," Journal of Banking and Finance 23, 537-571.
- Saunders, Anthony and Berry Wilson (2001), "An Analysis of Bank Charter Value and Its Risk-Constraining Incentives," Journal of Financial Services Research 19, 185-195.
- Schenone, Carola (2001), "The Effect of Lending Relations on the Firm's Cost of Equity Capital in Its IPO." Working paper, Northwestern University.
- Schranz, Mary (1993), "Takeovers Improve Firm Performance: Evidence from the Banking Industry," Journal of Political Economy 101, 299-326.
- Seward, James (1990), "Corporate Financial policy and the Theory of Financial Intermediation," Journal of Finance 45, 351-375.
- Shaffer, Sherrill (1998), "The Winner's Curse in Banking," Journal of Financial Intermediation 7:4 (October), 359-392.
- Sharpe, S. (1990), "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships," Journal of Finance 45, 1069-1087.

- Sharpe, Steven (1995), "Bank Capitalization, Regulation, and the Credit Crunch: A Critical Review of the Research Findings." Finance and Economics Discussion Series Paper No. 95/20, Board of Governors of the Federal Reserve System.
- Sheard, Paul (1994), "Main Banks and the Governance of Financial Distress," in Masahiko Aoki and Hugh Patrick, editors, The Japanese Main Bank System (Oxford University Press; Oxford), 188-230.
- Sheard, Paul (1989), "The Main Bank System of Corporate Monitoring and Control in Japan," Journal of Economic Behavior and Organization 11, 399-422.
- Slovin, Myron, Shane Johnson, and John Glascock (1992), "Firm Size and the Information Content of Bank Loan Announcements," Journal of Banking and Finance 16, 1057-1071.
- Slovin, M., M. Sushka, and C. Hudson (1988), "Corporate Commercial Paper, Note Issuance Facilities, and Shareholder Wealth," Journal of International Money and Finance 7, 289-302.
- Slovin, Myron, Marie Sushka and John Polonchek (1992), "Informational Externalities of Seasoned Equity Issues: Differences Between Banks and Industrial Firms," Journal of Financial Economics 32, 87-101.
- Slovin, Myron, Marie Sushka and John Polonchek (1993), "The Value of Bank Durability: Borrowers as Bank Stakeholders." Journal of Finance 48:1, 247-266.
- Smith, Bruce (1984), Private Information, Deposit Insurance Rules and the Stability of the Banking System," Journal of Monetary Economics 14, 293-317.
- Smith, Bruce (1991), "Bank Panics, Suspensions and Geography: Some Notes on the 'Contagion of Fear' in Banking," Economic Inquiry 24, 230-248.
- Smith, Clifford and Jerold Warner (1979), "On Financial Contracting: An Analysis of Bond Covenants," Journal of Financial Economics 7(2), 117-61.
- Sprague, O.M.W. (1910), History of Crises Under the National Banking System (Government Printing Office: Washington D.C.).
- Stein, Jeremy (1998), "An Adverse-Selection Model of Bank Asset and Liability Management with Implications for the Transmission of Monetary Policy." RAND Journal of Economics 29:3 (Autumn), 466-486.
- Stiglitz, Joseph (1985), "Credit Markets and the Control of Capital." Journal of Money, Credit, and Banking 17:2 (May), 133-152.
- Stiglitz, Joseph and Andrew Weiss (1981), "Credit Rationing in Markets with Imperfect Information," American Economic Review 71, 393-409.
- Stiglitz, Joseph and Andrew Weiss (1983), "Incentive Effects of Terminations: Applications to the Credit and Labor Markets," American Economic Review 73, 919-927.

- Stiglitz, Joseph and Andrew Weiss (1987), "Credit Rationing With Many Borrowers," American Economic Review 77, 228-231.
- Stomper, Alex (2001), "A Theory of Banks' Industry Expertise, Market Power, and Credit Risk." Working paper, University of Vienna.
- Swary, Itzhak (1986), "Stock Market Reaction to Regulatory Action in the Continental Illinois Crisis," Journal of Business 59, 451-73.
- Temin, Peter (1976), Did Monetary Forces Cause the Great Depression? (New York: W.W. Norton).
- Temin, Peter (1989), Lessons From the Great Depression (Cambridge: M.I.T. Press).
- Thakor, Anjan (1996), "Capital Requirements, Monetary Policy, and Aggregate Bank Lending: Theory and Empirical Evidence," Journal of Finance 51:1 (March), 279-324.
- Thies, Clifford and Daniel Gerlowski (1993), "Bank Capital and Bank Failure, 1921-1932: Testing the White Hypothesis," Journal of Economic History 53, 908-914.
- Timberlake, Richard H. (1985), "The Central Banking Role of Clearinghouse Associations," Journal of Money, Credit, and Banking 16, 1-15.
- Townsend, Robert (1979), "Optimal Contracts and Competitive Markets with Costly State Verification," Journal of Economic Theory 21, 265-293.
- Van Rijckeghem, Caroline and Beatrice Weder (1999), "Sources of Contagion: Finance or Trade?," International Monetary Fund Working Paper 99/146.
- Van Rijckeghem, Caroline and Beatrice Weder (2000), "Spillovers Through Banking Centers: A Panel Data Analysis," International Monetary Fund Working Paper 00/88.
- Von Thadden, Ernst-Ludwig (1995), "
- Von Thadden, Ernst-Ludwig (1998), "Intermediated versus Direct Investment: Optimal Liquidity Provision and Dynamic Incentive Compatibility." Journal of Financial Intermediation 7:2, 177-197
- Von Thadden, Ernst-Ludwig (forthcoming), "Asymmetric Information, Bank Lending and Implicit Contracts: The Winner's Curse." Journal of Finance (forthcoming).
- Wagster, John (1996), "Impact of the 1988 Basle Accord on International Banks," Journal of Finance 51, 1321-1346.
- Wagster, John (1999), "The Basle Accord of 1988 and the International Credit Crunch of 1989-1992," Journal of Financial Services Research 15, 123-143.
- Waldo, Douglas G. (1985), "Bank Runs, the Deposit-Currency Ratio and the Interest Rate," Journal of Monetary Economics 15, 269-277.

- Wall, Larry D. and David R. Peterson (1990), "The Effect of Continental Illinois' Failure on the Financial Performance of Other Banks," Journal of Monetary Economics 26, 77-99.
- Wallace, Neil (1988), "Another Attempt to Model an Illiquid Banking System: The Diamond-Dybvig Model with Sequential Service Taken Seriously," Federal Reserve Bank of Minneapolis Quarterly Review, 3-15.
- Wansley, James and Upinder Dhillon (1989), "Determinants of Valuation Effects for Security Offerings of Commercial Bank Holding Companies," Journal of Financial Research 12, 217-234.
- Weinstein, David and Yishay Yafeh (1998), "On the Costs of a Bank-Centered Financial System: Evidence from the Changing Main Bank Relations in Japan," Journal of Finance 53, 635-672.
- Weiss, Andrew (1989), "Imperfect Information, Credit Markets, and Macroeconomic Fluctuations," in Financial Markets and Incomplete Information, edited by Sudipto Bhattacharya and George Constantinides (Rowman and Littlefield; Maryland), 304-316.
- Wheelock, David (1992a), "Regulation and Bank Failures: New Evidence from the Agricultural Collapse of the 1920's," Journal of Economic History 52, 806-25
- Wheelock, David (1992b), "Deposit Insurance and Bank Failures: New Evidence from the 1920s," Economic Inquiry 30, 530-43.
- Wheelock, David (1995), "Regulation, Market Structure, and the Bank Failures of the Great Depression," Federal Reserve Bank of St. Louis Review (March/April).
- Wheelock, David and Subal Kumbhakar (1995), "Which Banks Choose Deposit Insurance? Evidence of Adverse Selection and Moral Hazard in a Voluntary Insurance System," Journal of Money, Credit and Banking 27, 186-201.
- Wheelock, David and Paul Wilson (1995), "Explaining Bank Failures: Deposit Insurance, Regulation, and Efficiency," Review of Economics and Statistics 77, 689-700.
- White, Eugene N. (1983), The Regulation and Reform of the American Banking System, 1900-1929 (Princeton: Princeton University Press).
- White, Eugene N. (1984), "A Reinterpretation of the Banking Crisis of 1930," Journal of Economic History 44, 119-138.
- White, Eugene N. (1986), "Before the Glass-Steagal Act: An Analysis of the Investment Banking Activities of National Banks," Explorations in Economic History Vol/ ?, 33-55.
- White, Eugene N. (1998), "The Legacy of Deposit Insurance: The Growth, Spread, and Cost of Insuring Financial Intermediaries," chapter 3 of The Defining Moment: The Great Depression and the American Economy in the Twentieth Century, edited by Michael Bordo, Claudia Goldin, and Eugene N. White (University of Chicago Press; Chicago and London).
- White, Lawrence (1991), The S & L Debacle (Oxford University Press; new York and Oxford).

- Wicker, Elmus (1980), "A Reconsideration of the Causes of the Banking Panic of 1930," Journal of Economic History 40, 571-583.
- Wicker, Elmus (1996), The Banking Panics of the Great Depression (Cambridge University Press).
- Wicker, Elmus (2000), Banking Panics of the Gilded Age (Cambridge University Press).
- Williamson, Stephen (1992), "Laissez-Faire Banking and Circulating Media of Exchange," Journal of Financial Intermediation 2, 134-67.
- Williamson, Stephen (1988), "Liquidity Banking, and Bank Failures," International Economic Review 29, 25-43.
- Williamson, Stephen (1986), "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing," Journal of Monetary Economics 18, 159-179.
- Winton, Andrew (1995a), "Delegated Monitoring and Bank Structure in a Finite Economy," Journal of Financial Intermediation 4:2 (April), 158-187.
- Winton, Andrew (1995b), "Costly State Verification and Multiple Investors: The Role of Seniority," Review of Financial Studies 8:1 (Spring), 91-123.
- Winton, Andrew (1997), "Competition among Financial Intermediaries When Diversification Matters," Journal of Financial Intermediation 6:4 (October), 307-346.
- Winton, Andrew (2001), "Institutional Liquidity Needs and the Structure of Monitored Finance." Working paper, University of Minnesota.
- Yanelle, Marie-Odile (1989), "The Strategic Analysis of Intermediation," European Economic Review 33:2-3 (March): 294-301.
- Yanelle, Marie-Odile (1997), "Banking Competition and Market Efficiency," Review of Economic Studies 64:2 (April), 215-239.
- Yasuda, Ayako (1999), "Relationship Capital and Competition in the Corporate Securities Underwriting Market," Wharton School, working paper.
- Yosha, Oved (1995), "Information Disclosure Costs and the Choice of Financing Source," Journal of Financial Intermediation 4:1 (January), 3-20.

Table 1
Net Financing of Nonfinancial Enterprises, 1970-1985 (%)

	Canada	Finland	France	Germany	Italy	Japan	U.K.	U.S.
Retentions	76.4	64.4	61.4	70.9	51.9	57.9	102.4	85.9
Capital transfers	0.0	0.2	2.0	8.6	7.7	0.0	4.1	0.0
Short-term Securities	-0.8	3.7	-0.1	-0.1	-1.3	N.A.	1.7	0.4
Loans	15.2	28.1	37.3	12.1	27.7	50.4	7.6	24.4
Trade Credit	-4.4	-1.4	-0.6	-2.1	0.0	-11.2	-1.1	-1.4
Bonds	8.5	2.8	1.6	-1.0	1.6	2.1	-1.1	11.6
Shares	2.5	-0.1	6.3	0.6	8.2	4.6	-3.3	1.1
Other	1.3	7.4	-1.4	10.9	1.0	-3.8	3.2	-16.9
Statistical Adjustment	1.2	-5.0	-6.4	0.0	3.2	N.A.	-13.4	-5.1
Total	99.9	100.1	100.1	99.9	100.0	100.0	100.1	100.0

Source: Mayer (1990) based on OECD Financial Statistics. See Mayer (1990) for details.

Table 2

Stock Price Response to Announcements of Corporate Security Offerings

Type of Security Offering	Two Day Abnormal Return⁺
Common Stock	-3.14% (155)
Preferred Stock	-0.19% (28)
Convertible Preferred Stock	-1.44% * (53)
Straight Bonds	-0.26% (248)
Convertible Bonds	-2.07% * (73)
Private Placement of Debt	-91.0% (37)
Bank Loans	1.93% * (80)

+ Sample size in parentheses.

* Indicates significantly different from zero.

Source: Smith (1986) and James (1987).

Table 3
Banking Panics During the U.S. National Banking Era

NBER Cycle Peak-Trough	Panic Date	% Change in the Currency-to- Deposits Ratio+	% Change in Pig Iron Production++	Loss Per Deposit \$	% and # Nat'l Bank Failures
Oct. 1873 - Mar. 1879	Sept. 1873	14.53	-51.0	0.021	2.8 (56)
Mar. 1882 - May 1885	Jun. 1884	8.8	-14.0	0.008	0.9 (19)
Mar. 1887 - Apr. 1888	No Panic	3.0	-9.0	0.005	0.4 (12)
Jul. 1890 - May 1891	Nov. 1890	9.0	-34.0	0.001	0.4 (14)
Jan. 1893 - Jun. 1894	May 1893	16.0	-29.0	0.017	1.9 (74)
Dec. 1895 - Jun. 1897	Oct. 1896	14.3	-4.0	0.012	1.6 (60)
Jun. 1899 - Dec. 1900	No Panic	2.78	-6.7	0.001	0.3 (12)
Se. 1902 - Aug. 1904	No Panic	-4.13	-8.7	0.001	0.6 (28)
May 1907 - Jun. 1908	Oct. 1907	11.45	-46.5	0.001	0.3 (20)
Jan. 1910 - Jan. 1912	No Panic	-2.64	-21.7	0.0002	0.1 (10)
Jan. 1913 - Dec. 1914	Aug. 1914	10.39	-47.1	0.001	0.4 (28)

+ Percentage change in the ratio at the panic date from the previous year's average.

++ Measured from peak to trough.

Source: Gorton (1988).

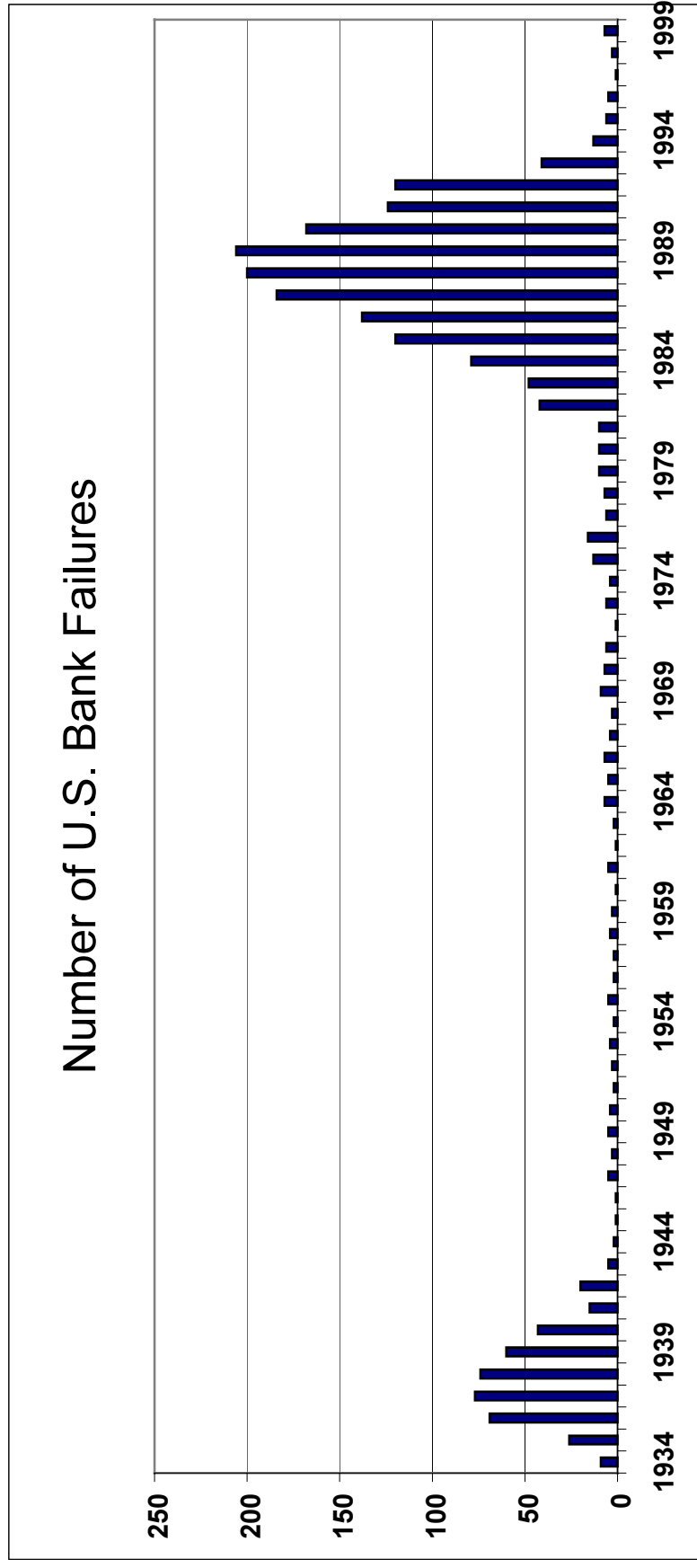
Table 4

**Corporate Control Changes in Banks and Nonbanks
(Percent of total sample)**

	Morck, Shleifer and Vishny Sample of 454 Manufacturing Firms	Prowse Sample of 234 Bank Holding Companies
Hostile Takeover	8.8	1.7
Management Turnover	20.5	10.2
Friendly Merger	7.5	10.7
Market-Based Control Changes	36.8	22.6
Regulatory Intervention	0	14.1
Total Control Changes	36.8	36.7

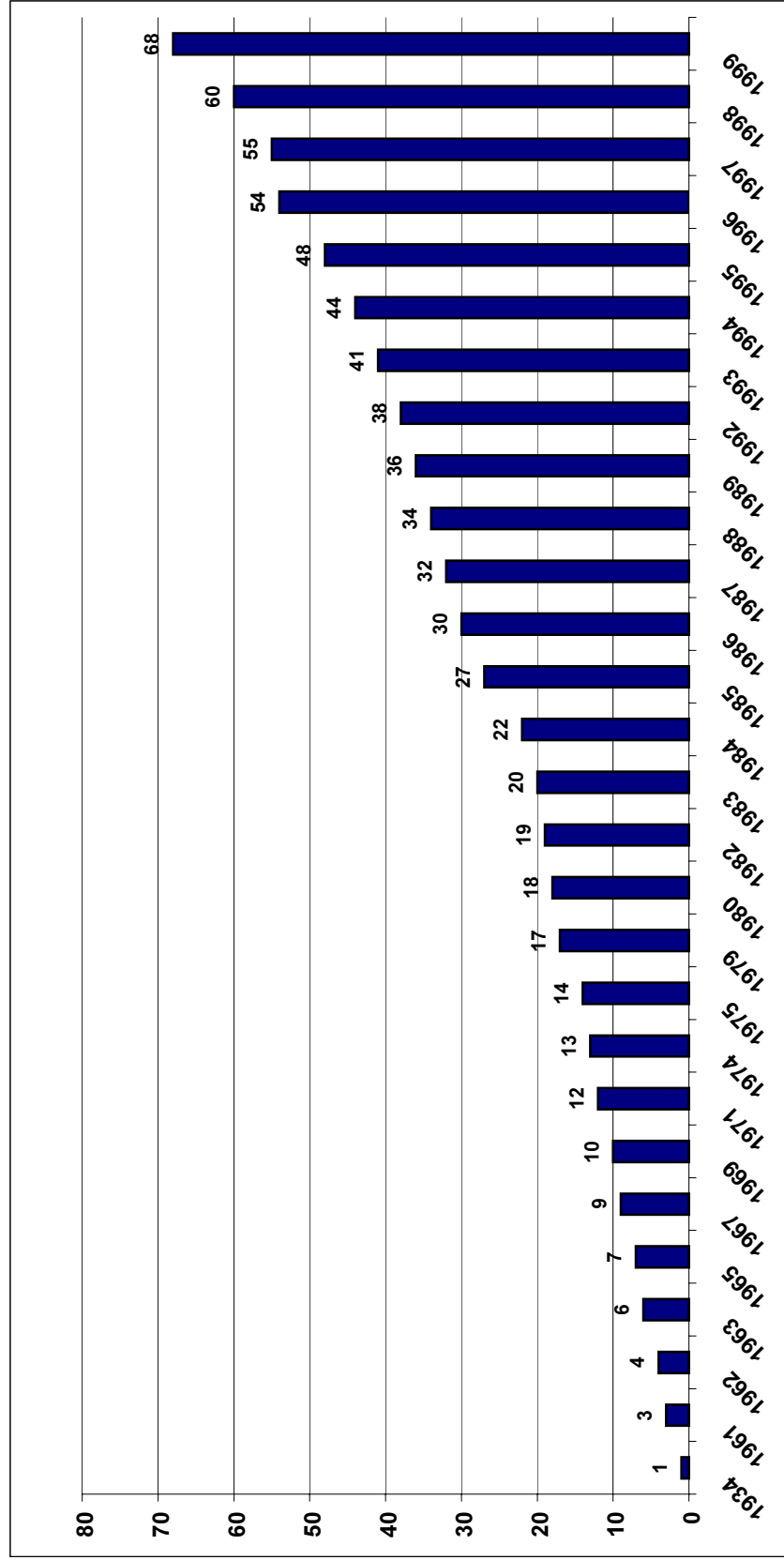
Source: Prowse (1995).

Figure 1



Source: FDIC.

Figure 2
Number of Explicit Deposit Insurance Schemes in the World



Source: Demirgüç-Kunt and Sobaci (2000).