SUDIPTO BHATTACHARYA
ARNOUD W.A. BOOT
ANJAN V. THAKOR

The Economics of Bank Regulation

We review the economics of bank regulation as developed in the contemporary literature. We begin with an examination of the central aspects of modern banking theories in explaining the asset transformation function of intermediaries, optimal bank liability contracts, coordination problems leading to bank failures and their empirical significance, and the regulatory interventions suggested by these considerations. In particular, we focus on regulations aimed primarily at ameliorating deposit-insurance-related moral hazards, such as cash-asset reserve requirements, risk-sensitive capital requirements and deposit insurance premia, and bank closure policy. Moreover, we examine the impact of the competitive environment (bank charter value) and industry structure (scope of banks) on these moral hazards. We also examine the implications of banking theory for alternatives to deposit insurance.

THE PRINCIPAL OBJECTIVE of this paper is to survey the modern literature on bank regulation, with a focus on exploring the implications of banking theory for optimal regulation. The 1980s and the ongoing 1990s have been witness to exciting developments in banking. On the academic front, the contributions of Leland and Pyle (1977), Diamond (1984), and Ramakrishnan and Thakor (1984) on financial intermediary existence, and those of Bryant (1980) and Diamond and Dybvig (1983) on bank runs and deposit insurance, generated new interest in the microeconomics of financial intermediaries. The new economics of asymmetric information and contract design have helped generate insights about how banks function and are regulated. These insights have been augmented by those in the literatures on credit market functioning under asymmetric information (Bernanke and Gertler 1990; Greenwald and Stiglitz 1990; Stiglitz and Weiss 1981, for example), corporate financing and governance (Myers and Majluf 1984 and Stiglitz 1985), and incomplete contracting (Hart 1991).

During this time governmental regulation of banking has also evolved. In the United States, for example, major banking legislations enacted in the 1930s (for example, the Glass-Steagall Act and the Bank Holding Company Act) have seen important

The authors thank two referees for helpful comments and Kathleen Petrie for excellent research assistance. Thakor thanks the Edward J. Frey Chair in Banking and Finance for support.

SUDIPTO BHATTACHARYA *is professor of economics in the London School of Economics.* ARNOUD W.A. BOOT *is professor of corporate finance and financial markets at the University of Amsterdam.* ANJAN V. THAKOR *is the Edward J. Frey Professor of Banking and Finance at the University of Michigan*

Journal of Money, Credit, and Banking, Vol. 30, No. 4 (November 1998)

changes. The market-induced "disintermediation" of the 1970s caused regulators to be concerned about the erosion in the competitiveness of banks, and led to the easing of regulatory constraints on banks in the early 1980s. However, the experience with bank deregulation in the 1980s was not entirely pleasant. Many S&L's failed in the late 1980s and early 1990s with total losses exceeding $200 billion, over $2,000 per U.S. household. This has led to a rethinking of the framework of banking regulation, and contemporaneous new regulatory legislation such as the Federal Deposit Insurance Corporation Improvement Act (1991) and the 1988 BIS agreement on capital requirements. In addition, there has been rapid financial innovation, accompanied by an expanding role of financial markets, and this has led to fundamental reconfiguration of the financial services sector; see Berger, Kashyap and Scalise (1995).

Partly due to these developments, many issues in bank regulation remain unresolved, such as these:

(a) Is *deposit contracting* (the right to demand withdrawal of contractual claims at any point in time from the issuer) important for investor welfare on the scale at which it is present in banking today?
(b) Should *deposit insurance* continue to be provided for such claims? If so, how universal (across intermediaries) and up to what scale should the coverage be?
(c) *How* should financial intermediaries with *insured* liabilities be regulated? What should be the role of bank capital controls and closure rules for troubled institutions?
(d) What role should the government play in managing idiosyncratic and systematic *liquidity shocks* experienced by banks?
(e) What should be regulatory policy toward *banking scope* and *interbank competition* in loan markets? Should *portfolio restrictions* on banks be relaxed or made more stringent? Should commercial firms be allowed to own banks or bank holding companies?

Existing theories of banking [see the Bhattacharya and Thakor (1993) survey] and of corporate governance and capital structure [see Hart (1991) and Dewatripont and Tirole (1994b)] have provided valuable insights into these issues but have not been entirely successful in providing answers. In what follows, we first briefly sketch in section 1 the salient features of recent banking theories in explaining the asset transformation functions of intermediaries, optimal bank liability contracts, and the coordination problems leading to bank failures, and their regulatory implications. Section 2 focusses on the empirical evidence on bank runs. In section 3, we extend the basic framework described in section 1 to consider incentive problems created by governmental deposit insurance, and also examine regulations designed to cope with these problems. In particular, we examine the role of the following in attenuating deposit-insurance-related moral hazards: (i) cash-asset reserve requirements, (ii) risk-sensitive capital requirements and deposit insurance premia, (iii) partial deposit insurance, (iv) bank closure policy, and (v) bank charter value. In section 4, we examine alternatives to deposit insurance, and also explore the implications of the theory for the desirability of universal banking. Section 5 concludes.

1. BANKING THEORIES

To examine the economics of bank regulation, we first need to comprehend the economic role of banks. Modern theories of banking have been concerned with explaining (a) why financial intermediaries exist, focusing in particular on the benefits of delegating monitoring for lending and other resolutions to market imperfections, (b) the nature of optimal bank liability contracts, such as deposits, intended to provide insurance for liquidity needs to investors, and (c) the (coordination) problems of imperfect functioning of these contracts, leading to phenomena such as bank runs, and measures to cope with these. In this section, we briefly review the salient features of these theories and the related empirical evidence. Our focus is on extracting the implications of these theories for bank regulation.

A. Asset and Liability Intermediation Services: Why Do Banks Exist?

There are two dominant paradigms for explaining why we need banks and other intermediaries. One focusses on the asset side of the intermediary's balance sheet and the other on the liability side, and each produces different implications for regulation. On the asset side are the contributions of Allen (1990), Leland and Pyle (1977), Boyd and Prescott (1986), Diamond (1984), and Ramakrishnan and Thakor (1984). In these models, institutions are viewed as monitoring—either in the ex post sense of verifying cash flows or in the interim sense of screening quality—the attributes of investment projects. Without intermediation, such monitoring would be duplicated by the many investors involved in funding such projects. Alternatively, investors would have been forced to take large (undiversified) stakes (Leland and Pyle 1977). Markets for information sellers for such monitored knowledge are assumed to function imperfectly, owing to problems of credibility (Ramakrishnan and Thakor 1984) or the sellers' inability to capture the full returns from monitoring (see Allen 1990). Furthermore, intermediaries that monitor many projects with imperfectly correlated rates of return achieve diversification, which in turn allows them to credibly communicate the attributes of their diversified portfolios to ultimate investors at lower cost than possible with nonintermediated, bilateral contracting (Diamond 1984).

On the liability side are the models developed by Bryant (1980) and Diamond and Dybvig (1983). The Diamond-Dybvig model, which has received the most attention, formalizes the following idea. There are ex ante identical investors (depositors) who are risk averse and uncertain about the timing of their future consumption needs. These investors can invest their date 0 endowments in illiquid technologies that will pay off at date 2. Without an intermediary, *all* investors are locked into illiquid long-term investments that yield high payoffs only to those who consume late (date 2); those who consume early (date 1) get very low payoffs because early consumption requires premature liquidation of long-term investments. Improved risk sharing and thus enhanced ex ante welfare are attained by an intermediary that promises investors a higher payoff for early consumption and a lower payoff for late consumption, relative to the nonintermediated case.

The intermediary achieves this by designing a nontraded demand deposit contract that promises the desired vector of future payoffs. Since this vector contains a higher payoff for early consumption than the nonintermediated case, the contribution of the intermediary is to improve liquidity and risk sharing.

B. An Integrated Model of Banking and Its Implication for Regulation

Because the two theories of banking we have discussed focus on very different financial intermediation services, it is important to paint an *integrated* picture of why banks exist, so that we can extract relevant regulatory implications. We do that in this subsection.

The Model with Risk-Neutral Agents. Consider an economy in which at date $t = 0$ there is an unspecified number of risk-neutral investors/depositors, each with \$1 available to invest. There are also n risk-neutral entrepreneurs, each with a project that needs $\$N$ and will pay off a stochastic cash flow at $t = 2$ that has a probability density function with support $[0, \tilde{R}]$ and mean $\$R$, where $R > N$. The intermediate date $t = 1$ comes into play later. Assume that entrepreneurs will have private information about the ex post cash flows at $t = 2$, and will thus be inclined to misreport these cash flows as being too low in order to minimize their repayment to investors. Suppose that an investor can learn the entrepreneur's cash flow by investing $\$K$ in direct monitoring. Then the monitoring cost per project is NK.

An alternative to direct monitoring is for entrepreneurs to raise their funds through repayment contracts that call for the entrepreneur to make a fixed repayment to investors (independent of the realized cash flow at $t = 2$) as long as the reported cash flow exceeds the fixed repayment and to surrender the reported cash flow plus a predetermined "nonpecuniary penalty" if the reported cash flow is below the fixed repayment. Diamond (1984) shows that such a contract is incentive compatible—induces truthful cash flow reports—with the nonpecuniary penalty chosen appropriately. Moreover, the contract resembles straight debt. Let S represent the entrepreneur's expected cost of using this debt contract, that is, the expected nonpecuniary penalty incurred by an entrepreneur reporting truthfully in equilibrium.

Since the entrepreneur will choose the lowest-cost financing method, the financing cost per entrepreneur (or project) without intermediation will be

$$\text{Min}[NK, S]\,. \tag{1}$$

Let us now introduce a bank to intermediate between investors and entrepreneurs. For the bank, the monitoring cost per project[1] will be $K(n) \leq K$. Moreover, the bank faces the same credibility problem in communicating to investors information about entrepreneurs' realized cash flows that the entrepreneurs themselves faced in the nonbank case. The bank could use debt contracts in raising financing from investors. Let

1. $K(n)$ could be strictly lower than K if cash flows are correlated across the n entrepreneurs who are funded. This would generate cross-sectional informational reusability and produce a monitoring cost per project that is lower than K. Millon and Thakor (1985) develop such a model.

$S(n)$ be the expected cost per project faced by the bank in using the debt contract; note that $S(n)$ is the *average* across the expected costs for the n projects. Thus, the bank's cost per project will be

$$K(n) + S(n)\,. \tag{2}$$

From (1) and (2), we see that intermediation enhances welfare if

$$[K(n) + S(n)] < \text{Min}[NK, S]\,. \tag{3}$$

The intermediary (bank) in this case borrows from $N \times n$ depositors and lends to n entrepreneurs.

If project cash flows are independent and identically distributed (i.i.d), then $S(n) \to 0$ almost surely as $n \to \infty$ since the probability that *aggregate* realized project cash flows will fall below the bank's aggregate debt repayment obligation converges to zero. Moreover, $K(n) < NK$ for any $N \geq 2$. Thus, a sufficiently large bank always improves welfare.[2]

Although we have assumed that the entrepreneur's private information is about his realized cash flow, we could alternatively assume that the entrepreneur is privately informed ex ante; for example, he knows the mean cash flow R but no one else does. An investor can discover R at a cost K, or the entrepreneur can credibly signal this information at a cost S. Once again, the financing cost per entrepreneur will be given by (1). And viewing $S(n)$ as the intermediary's cost of motivating each agent within it to discover R and truthfully communicate it to investors, intermediation enhances welfare if (3) holds.

Ramakrishnan and Thakor (1984) show that the first best can be attained in the limit as $n \to \infty$, that is, as the intermediary tends to be infinitely large, $S(n) \to 0$ almost surely. Thus, this view of intermediation—with ex ante informational asymmetry—is consistent with the model with ex post informational asymmetry.

To recapitulate, the main predictions of this theory—with ex ante or ex post informational asymmetry—are (a) intermediaries will be very large, (b) their portfolios will have almost zero unsystematic risk, and (c) their liabilities will be debt contracts (with repayment levels adjusted for systematic risk factors) that will be honored almost surely.

The main *regulatory implications* of the theory, as developed thus far, are these:

Implication 1: Regulatory restrictions that limit banks to financing themselves primarily with debt will not sacrifice efficiency.

Implication 2: There should not be regulatory restrictions on how large banks can be.

2. Ramakrishnan and Thakor (1984) also show the optimality of an infinitely large intermediary as a certifying agency. This argument holds even if project cash flows are not i.i.d., as long as they are pairwise imperfectly correlated. Puri's (1996) evidence supports the bank's certification role.

Extension of the Model to Risk-Averse Investors. We continue to assume that each entrepreneur is risk neutral and needs $\$N$. However, we now assume risk-averse investors.

Let all investors be ex ante identical at $t = 0$. After investing his \$1 endowment at $t = 0$, each investor finds out at $t = 1$ if he must consume then, with utility $U(C_1)$, or can wait until $t = 2$ to consume, with utility $U(C_2)$. The probabilities of the two events, distributed independently across investors, are P and $1 - P$, respectively. As in Diamond and Dybvig (1983), agents' conditional preferences are extreme: consumption has positive utility either at $t = 1$ or $t = 2$. The bank's investment choices include (i) entrepreneurs' projects paying off $R > N$ at $t = 1$, and (ii) a short-term technology that yields \$1 at $t = 1$ for every \$1 invested at $t = 0$. Moreover, if an entrepreneur's project is liquidated prematurely at $t = 1$, it yields $\hat{R} \leq N < R$. Banks, serving many agents in Bertrand competition over contracts, must choose consumption/investment levels $\{C_1$ or $C_2\}$ and investment patterns per capita $\{L, 1 - L\}$, where fraction L of funds is invested in the short-term technology and $1 - L$ in entrepreneurs' projects.

What should the bank maximize in making the above choices? In Diamond and Dybvig (1983), investors directly own the projects that are owned by entrepreneurs in our model,[3] so the competitive bank maximizes each representative investor's ex ante expected utility, $[PU(C_1) + \{1 - P\}U(C_2)]$, subject to the budget-balancing constraints $PC_1 = L$ and $[1 - P]C_2 = R/N[1 - L]$. If $U(\cdot)$ has relative risk aversion greater than unity, then the resulting contract has the "visible" insurance feature:

$$1 < C_1 < C_2 < \frac{R}{N}. \tag{4}$$

Diamond and Dybvig (1983) interpret $\{C_1, C_2; L\}$ as a demand deposit contract giving investors the unconditional right to withdraw C_1 at $t = 1$ *or* C_2 at $t = 2$.[4] The contribution of the bank to enhancing welfare comes from the improved liquidity and ex ante ($t = 0$) risk sharing embodied in (4).

To focus on the asset *and* liability intermediation services of the bank in an integrated framework, we have assumed that investors and entrepreneurs are distinct. But this makes it more difficult to answer the earlier question about what the bank maximizes. To simplify, assume that each investor's reservation expected utility is $\bar{U}$, which exceeds that which investors can get by lending directly to entrepreneurs.[5]

3. In the context of our model, N investors can be viewed as owning each project.

4. An alternative implementation of the contract $\{C_1, C_2; L\}$ is for each intermediary to pay a dividend stream $\{L \text{ and } R/N(1 - L)\}$, with interim trading of the bond/share by agents who wish to consume earlier. With the ex post corner preferences assumed in the Diamond-Dybvig model, such trading leads to consumption patterns $\{C_1 \text{ or } C_2\}$ for the two types of agents in a Walrasian equilibrium, and an interest rate of I, satisfying $1 < I = C_2/C_1 < R/N$. This mimics the risk sharing arrangement above. These observations were made in Jacklin (1987) and Bhattacharya and Gale (1987). Hellwig (1993) obtains analogous results when there is a stochastic (short-term) investment technology between times 1 and 2, with rate of return $\tilde{R}$. von Thadden (forthcoming) develops a continuous-time version of the Diamond-Dybvig model.

5. If investors deal directly with entrepreneurs, they suffer two costs. One is the monitoring-related cost

Similarly, assume that each entrepreneur has a reservation expected payoff of U_0 from his project. U_0 and $\bar{U}$ are exogenous to the model and determined by nonbank financing alternatives such as the capital market. If we now assume that it is the investors' funds that are in relative scarce supply, then the bank's task is to choose $\{C_1, C_2; L\}$ to maximize $PU(C_1) + [1 - P]U(C_2)$ subject to budget-balancing constraints and the constraint that the *expected net payoff* to each entrepreneur is U_0. The expected net payoff to each entrepreneur is equal to the expected value of the project payoff net of monitoring-related costs ($K(n) + S(n)$) and the expected cost of providing investors with the consumption vector $\{C_1, C_2\}$ per investor. For banks to be viable, the maximized value of each investor's expected utility must (weakly) exceed $\bar{U}$. Let $\{C_1^*, C_2^*; L^*\}$ represent the equilibrium solution.

Banks now provide both asset and liability intermediation services. However, banks are vulnerable *to a coordination failure.* If the fraction of investors withdrawing C_1 at $t = 1$ exceeds P, then the promised C_2 becomes infeasible. For a sufficiently large fraction f withdrawing early, liquidation occurs at $t = 1$ because after the depositors withdrawing at $t = 1$ are paid off, the payoff available per investor at $t = 2$ is less than C_1. Thus, others are induced to withdraw, causing a bank run. This bank run is productively disruptive because it forces premature liquidation of entrepreneurs' projects.

There are thus two pure strategy Nash equilibria in the withdrawal game among depositors. In the first, we have $1 < C_1^* < C_2^*$. In the second, which is the Pareto-inferior bank-run equilibrium, we have $C_1 = C_1^*$ with probability P/L, and zero with probability $([L - P]/L)$, and $C_2 = 0$.

This motivated the Diamond-Dybvig treatment of bank regulation. First, they noted that when preference shocks are i.i.d across many depositors at a bank, so that the realized proportion of "early diers" is almost surely P, a precommitment not to liquidate more than a fraction $L = PC_1$ of the bank's total deposits eliminates the bank-run equilibrium, since the promised C_2 is always feasible. They suggest that this rationalizes the pre-deposit insurance suspension of convertibility in U.S. banking.

When preference shocks at a given bank are correlated across depositors, matters are more problematic. Suspension of withdrawals at a fraction of total deposits $L = PC_1$, when P is the ex ante probability of the partially correlated early withdrawal shocks, will imply that those who *need* to will be unable to withdraw when their realized fraction f exceeds P. Diamond-Dybvig investigate an alternative, and superior, intervention: deposit insurance. Under this mechanism, subject to a sequential service constraint, banks let agents withdraw a predetermined C_1 irrespective of f, but this withdrawal is backed up by governmental funds, with the government taxing withdrawals at rate t, so that $C_1(1 - t)$ is the expected-utility-maximizing quantity of withdrawal, given f proportion of early diers. Anticipating the resulting preservation of banks' long-term investment payoffs, there are no runs either.

---

represented by (1) and the other is the lost risk sharing due to $C_1 = 1$ and $C_2 = R - \text{Min}[NK, S]/N$ in the nonintermediated case; the idea here is that investing in the short-term technology for consumption at $t = 1$ involves no monitoring. Thus, the analog of (4) in our model with monitoring-related costs would be: $1 < C_1 < C_2 < \{R - \text{Min}[NK, S]\}/N$.

One criticism of the Diamond and Dybvig bank run is that it lacks a trigger mechanism. Bank runs are pure "sunspot" phenomena, uncorrelated with other economic variables. Gorton (1988), Calomiris and Gorton (1991), and Calomiris and Schweikart (1991) argue persuasively that such was *not* the case in U.S. banking history prior to deposit insurance. Business activity measures, such as small business failure rates, were very useful in predicting bank runs, suggesting a role for adverse information about banks' returns as a trigger mechanism. In the context of our model, at $t = 1$ some "late consumers" have private information about the random project payoff. If the value of their expected future contractual (with potential default) return is lower than their current withdrawal rights, they precipitate a run.

Models of such information-based runs were developed by Bryant (1980), Jacklin and Bhattacharya (1988), Chari and Jagannathan (1988), and Gorton (1988). The basic point is that adverse information about bank returns that is in the possession of a subset of depositors (either real or imagined by the uninformed depositors) could trigger runs, so that a bank run could arise as a unique Nash equilibrium. Moreover, if the shocks to bank returns are correlated across banks, the run may spread to other banks, causing a panic. Suspension of convertibility can eliminate panic runs, but only at a cost: early consumers do not all get their withdrawal when there are many of them, even though there is no adverse asset return shock, and hence no information-based run by informed agents. Deposit insurance does better. It not only prevents information-based runs, thereby eliminating randomization of consumption across liquidity seekers and adversely informed depositors, but also precludes any panic-run equilibrium.

Note, however, that such deposit insurance is *not* socially costless. When asset returns are low or when there are large early withdrawals calling for governmental provision of liquidity at $t = 1$, other sectors must be taxed to make up the shortfall, with possible deadweight costs. Thus, a more rigorous public finance analysis of the optimal governmental intervention would be of interest.

What is insured in our framework, of course, is the demand deposit contract $\{C_1^*, C_2^*\}$ which is *not* traded. Jacklin (1987) points out that even if the contract were traded, it would lead to the same allocations as the optimal nontraded contract as long as investor preferences are extreme as in Diamond and Dybvig (1983). However, if preference shocks are less extreme, then the traded and nontraded deposit contracts are *not* welfare equivalent. Jacklin shows that the resulting traded-debt allocation will be strictly worse in ex ante welfare compared to the nontraded demand deposit contract. The reason is that trading leads to ex post efficient allocations, whereas the nontraded contract permits ex ante efficient allocations that need not honor ex post efficiency constraints. We thus have the following additional implications.

Implication 3: Banks should be permitted to finance with nontraded demand deposit contracts, without restrictions on the associated interest rate patterns.[6]

Implication 4: Governmental insurance of the nontraded demand deposit contract

6. That is, banks should be free to choose the (depositors') utility-maximizing $\{C_1^*, C_2^*\}$ vector.

used by banks should be preferred to suspension of convertibility if the deadweight costs of taxation are exceeded by the costs of randomization in meeting investors' liquidity needs.

To summarize, contemporary banking theories suggest that no regulatory restrictions should be placed on bank size, and banks should be allowed to fund themselves with nontraded demand deposit contracts, and that these contracts should be federally insured if the banking-related liquidity needs of investors are sufficiently high. Next we briefly review the related empirical evidence.

C. *The Empirical Evidence*

There is considerable evidence supporting the prediction that banks provide borrowers with valuable asset intermediation services. Examples are James (1987), Lummer and McConnell (1989), Hoshi, Kashyap, and Scharfstein (1990a, 1990b), Billett, Flannery, and Garfinkel (1995), and Puri (1996).

There is also empirical evidence that illuminates why illiquid loans are financed with demand deposits. Allen, Saunders, and Udell (1991) and Nakamura (1993) provide evidence that shows how information generated by deposit accounts may be used in evaluating loan quality and designing future credit contracts; thus, the existence of a bank-borrower relationship increases firm value. Berger and Udell (1995), in their examination of the role of relationship lending in small business finance, discuss asset-backed lending in which banks take accounts receivable as collateral; borrowers deposit these receipts into a transactions account, and these are automatically credited to the loan account.

In the next section we examine the historical facts related to the theories of bank runs discussed in this section, as a prelude to further extensions of the basic theoretical framework.

2. SOME HISTORICAL FACTS AND RELATIVE IMPORTANCE OF PHENOMENA

How important have bank runs and panics been in the *pre-deposit insurance era*? What have been the resulting losses for depositors, given their much lower equity capital to assets ratios (average of 7–8 percent compares to 35–40 percent for nonfinancial corporations in the United States)? What has been the impact of bank failures on credit allocation and economic activity? We briefly review the empirical evidence on these issues, as discussed in Kaufman (1992), Bernanke (1983), Calomiris (1993), Calomiris and Gorton (1991), and Romer (1993). This should illuminate any discussion of the desirable extent of safety-net provisions for bank liabilities and the related regulation of their capital structure and business activities. In particular, we wish to empirically discriminate between the sunspots and the adverse-information theories of bank runs, because this has relevance for evaluating specific regulations implied by extensions of the basic theoretical framework that we will discuss in the next section.

A. Bank Failures: The Short-Term Implications

The short-term consequences of bank failures are summarized below:

i. Although stock returns of banks show greater within-industry correlations than for firms in other industries, abnormal negative returns of *other* banks given failures of a given bank arise only for banks in the same product or market area (Kaufman 1992).

ii. Average depositor losses in failed banks as a percentage of all deposits was only 0.21 over 1865–1933, less than 1 percent even in crisis years. In failed national banks, losses were 10 percent of deposits over 1865–1930. Between 1950–1980, these latter loss rates decreased to 2 percent due to deposit insurance, but have averaged 12 percent over 1981–1990. However, adjusting for unearned interest on assets sold later by the FDIC, losses averaged nearly 30 percent (James 1991). This compares with market-based value loss estimates of 62 percent on average on defaulted bonds of nonbank firms over 1971–91.

iii. Calomiris and Gorton (1991) have documented that American banking panics were *not* sunspots phenomena. They were uniquely predictable based on a "dual threshold" criterion. Prepanic quarters can be identified using stock price declines and business failure increases.

On balance, the evidence suggests the following conclusions. Bank runs do not seem to have been historically important in causing banks in the aggregate to become insolvent because of lack of liquidity. However, *information-based* runs were important, and provided potentially timely discipline of banks in the preinsurance era. Without such market discipline, regulator-induced discipline may not be as effective as one might like; the average losses to the FDIC in the volatile but not crisis-stricken U.S. economy of the 1980s matched the crises for national banks during 1865–1930.

B. Bank Failures: Medium-Term Implications

Turning now to medium-run (1–5 years) effects of bank failures on credit markets and aggregate investments, consider the evidence summarized in Calomiris (1993) for the Great Depression. He argues, as does Bernanke (1983), that the sharp and unanticipated drop of about 40 percent in nominal prices over 1929–32 in the United States caused a "debt deflation" that lowered borrowers' net worths and increased defaults. Both effects, the first through an increase in the lemons premium in credit markets, and the second through banks' lowering their loan-deposit ratios to guard against default or runs, increased the costs of external credit, thus lowering investment and output; by 1932, U.S. industrial production declined by more than 60 percent in real terms relative to the peak of 1929. Moreover, these effects caused bank failures that lowered investment and output even further. Evidence in support of this hypothesis, a modern version of Irving Fisher's classic debt-deflation theory from the 1930s, includes the following. First, the differential between low-(Baa) and high-grade (U.S. government) bond yields increased from less than 200 basis points in 1929 to more than 750 basis points in 1932 (the trough of the depression in the United States), but

declined to less than 300 basis points by 1934, and was back to 1929 levels by 1937. Surviving banks curtailed their loan to deposit ratios from 0.85 in 1929 to 0.58 in 1933, confirming the debt deflation predictions. The changes in these measures correlated well with the sharp output drop in the United States over 1929–32.

However, Canada, with its more concentrated banking system, experienced far fewer bank failures, and a low correlation between its output drop and measures of financial distress, casting doubt on the debt deflation theory. Calomiris (1993) argues that this evidence does not disprove the debt-deflation theory, since leverage was much smaller in Canada to start with; for example, debt service/GNP ratios in the United States increased from 9 percent to 19.8 percent over 1929–33, versus only from 3.9 percent to 6.4 percent in Canada. Moreover: (i) the industrial output drop in Canada over 1929–32 was also over 50 percent, with recovery to 1929 levels only occurring by 1937, and (ii) the share of consumption goods in the drop of production was much greater in the United States than in Canada, suggesting severe drops in Canada's level of durable good production and investments; see Romer (1993). Thus, it is unclear whether the largest debt deflation in memory had *very* different output effects in economies differing significantly in their net worth to external financing ratios.

This is not to deny that unanticipated shocks, particularly large price-level shocks to which nominal debt contracts are not indexed, may cause borrowers short-term liquidity problems and losses in net worth, that in turn may cause banks to be reluctant to lend. However, it would be premature to conclude that this implies an inherent fragility that warrants an all-embracing safety net. For example, such a safety net may be at odds with prudent risk taking by banks,[7] an issue that we turn to next.

3. REGULATORY IMPLICATIONS OF EXTENDING THE THEORY TO INCLUDE INCENTIVE PROBLEMS

In section 1 we concluded that banking theory makes a case for governmental deposit insurance in some circumstances, and in section 2 we examined the relevant empirical evidence. But thus far we have not considered incentive problems between banks and regulators, an issue we address next.

A. Deposit Insurance and Two Types of Moral Hazard

In section 1, we assumed that the bank's investment choice was to determine how much to invest in the short-term liquid asset and how much in entrepreneurs' projects. But the bank took the project cash flow distributions *as given.* We extend the model now to permit the bank to make *unobservable* project choices after deposits are received. Suppose the bank can choose which entrepreneurs to lend to, and entrepre-

7. The modeling of financial fragility, as for example in Bernanke and Gertler (1990), amounts to noting that with low endowment and high external borrowing, entrepreneurs will be induced to choose risky projects with lower expected payoffs than alternative riskless ones, that this problem is greater when borrowers' net worth is low, and if project investigation involves a fixed cost, then investment may stop if net worth is sufficiently low. Short of such a breakdown, however, the effect of borrower net worth changes on aggregate investment activity is *ambiguous.*

neurs' cash flow distributions can be rank-ordered on the basis of second-order stochastic dominance (Rothschild and Stiglitz 1970). That is, suppose θ represents project riskiness, with $\theta \varepsilon [\theta_{min}, \theta_{max}]$ cross-sectionally, and $\theta_1 > \theta_2$ implies that the probability distribution associated with θ_1 is riskier than that associated with θ_2 in the second-order stochastic dominance sense.

Asset-substitution moral hazard may exist in this setting. The bank accepts deposits at competitively determined terms that are set *before* the bank makes its project choices. The bank then has an incentive to invest in the riskiest project—the one with $\theta = \theta_{max}$—if riskier projects carry higher repayment obligations from entrepreneurs to the bank. Doing this results in an ex post wealth transfer from depositors to the bank's shareholders because a deposit is a *debt contract.*

What may deter the bank from choosing excessive risk is *market discipline.* Calomiris and Kahn (1991) and Peters (1994) develop models in which some depositors receive possibly unverifiable private signals about the bank's asset choices at an interim date ($t = 1$) and exercise their early withdrawal rights then if their information is adverse. This helps discipline bank management against the temptation to commit outright fraud or invest in excessively risky projects. Governmental deposit insurance, however, reduces incentives for information acquisition and control by depositors and thus weakens market discipline; see Saunders and Wilson (1996) for empirical evidence on the role of informed depositors.

Similarly, deposit insurance may induce a bank to lower its liquid reserves below the L^* we described in section 1. This leads to the following regulatory implication.[8]

Implication 5: Deposit insurance invites insured banks to seek excessive portfolio risk and keep lower liquid reserves relative to the social optimum. Thus, regulatory restrictions aimed at limiting risk taking may be necessary.

In what follows, we evaluate different regulatory mechanisms to control moral hazard.

B. Regulatory Measures to Cope with Moral Hazard

The regulatory measures we examine in this subsection are (1) cash-asset reserve requirements, (2) risk-sensitive capital requirements and deposit insurance premia, (3) partial deposit insurance and market discipline, (4) bank closure policy, and (5) bank charter value.

1. Cash Asset Reserve Requirements. It appears that, if banks' liquidity reserves are externally verifiable, then a simple solution to the underinvestment problem would be for the regulator to mandate the optimal level of reserves for each bank. In the static model that we have examined thus far, this will ameliorate the liquidity-related moral hazard, and can be viewed as a rationale for legal cash asset reserve requirements.

8. See Bhattacharya and Thakor (1993). Merton (1977, 1978) first recognized the isomorphic correspondence between put options and deposit insurance, highlighting the attendant moral hazards. Cooperstein, Pennacchi and Redburn (1995) develop a multiperiod version.

Banks exist in a dynamic world, however, and face continuous fluctuations in their deposit levels. In such a setting, mandated reserve requirements, in ratio form relative to total deposits, may actually interfere with the task of meeting the bank's liquidity needs because a deposit withdrawal extinguishes reserves, necessitating a *new* deposit inflow to replace the lost reserves needed to support the remaining old deposits as well as provide the reserves needed to support the new deposits.

2. Risk-Based Capital Requirements and Deposit Insurance Premia. One way to deal with the moral hazards considered earlier in this section would be to link the bank's shareholders' capital infusion to the risk of the bank. This imposes a cost on the shareholders for increasing the bank's failure probability through higher asset risk and permits the regulator to control the bank's portfolio choice. Similarly, making the deposit insurance premium risk sensitive facilitates regulatory risk control since the higher expected profits associated with higher risk are now taxed at a higher rate.

The key, of course, is the regulator's ability to observe the bank's risk. The issue of designing risk-based capital requirements and risk-based deposit insurance premia under asymmetric information is formally examined by Chan, Greenbaum and Thakor (1992). They show that incentive compatibility is sacrificed if the regulator wishes to have a *fairly priced* (actuarially neutral) deposit insurance pricing schedule for a competitive banking system. In an informationally rich environment, with either private information and/or moral hazard, it is therefore impossible to implement a fairly priced and completely risk-sensitive deposit insurance pricing scheme unless banks are permitted access to rents, either through explicit regulatory subsidies—that may be manifested in underpriced deposit insurance—or through restricted entry into banking.[9] The authors note the irony in the decision of U.S. bank regulators to move to risk-sensitive schedules for capital requirements and deposit insurance premia at a time when entry restrictions have been substantially eased and regulatory subsidies for banks have been lowered.

To summarize, the theories developed thus far suggest an important role for linking both capital requirements and deposit insurance premia to bank risk; this may resolve moral hazard and private information problems. These theories provide some support for the Basle risk-based capital guidelines and the recent move by the FDIC in the United States to risk-sensitive deposit insurance premia.[10]

3. Partial Deposit Insurance and Market Discipline. The advantage of complete deposit insurance is that runs are eliminated. The disadvantage is that now it is the responsibility of the insuring agency to monitor bank assets, estimate their value relative to the par value of deposits, and to ensure that the bank is adequately capitalized. The purpose of such monitoring is to prevent excessively risky bank investments. In their review paper, Berlin, Saunders and Udell (1991) observe that a key issue in de-

9. In a somewhat different setting, Friexas and Gabillon (1994) reach a conclusion qualitatively similar to that in Chan, Greenbaum, and Thakor (1992).

10. A common theoretical assumption is that increasing capital is costly for the bank. However, Berger's (1995) evidence casts doubt on its validity. Berger finds *positive* Granger causality from capital to earnings for U.S. commercial banks during 1983–89, although this relationship does not hold for 1990–92. It appears then that whether a capital requirement is costly for a bank depends on whether it elevates the bank's capital above its private optimum.

posit insurance reform is the optimal mix of private and public information production for monitoring of banks. A possibility is the issuance of (uninsured) subordinated debt. Alternatively, Peters (1994) argues that the optimal arrangement is *partial* deposit insurance. With no deposit insurance, *excessive* information production by depositors—who monitor and discipline bank management—is likely and ex post inefficient bank runs may arise too often. On the other hand, complete deposit insurance destroys all potentially beneficial information production and monitoring by depositors. Underlying the partial insurance conclusion is the presumption that informed depositors—with their own endowments at risk—will monitor banks better than governmental regulators do. The empirical evidence in section 2 is supportive of this. This is particularly relevant when we consider that there may be an agency problem between taxpayers and regulators that could lead to lax regulatory monitoring of banks. We turn to this issue next.

4. Bank Closure Policy. There have been recent analyses of the incentives of regulators to close banks in a manner that results in socially optimal bank portfolio choices. The advantage of such regulatory monitoring and control, as compared to the market discipline of bank runs, could lie in (i) nonduplication of monitoring costs and (ii) lowering the transactions costs of managing financial crises. The disadvantage might arise from closure decisions that are too lax to discipline the ex ante asset choices of bank management.

Boot and Thakor (1993) formalize this in a model in which regulators' payoffs depend on a reputation for monitoring ability. The underlying problem is the usual one of risky asset choices by levered bank insiders, but now in *two* separate periods. All deposits are insured. The regulator can either (i) monitor the bank's initial risk choice and ask for a change if the choice is suboptimal and/or (ii) choose the level of bank capital at which to close the bank. The ex post socially efficient criterion is to close the bank when its second-period capital is low enough that it would choose a negative NPV investment in that period. However, regulators also care about their monitoring reputation. In the reputational sequential equilibrium of the game between the bank and regulator, the regulator's optimal bank closure policy is more lax than socially optimal, and this also increases the bank's first-period risk choice.

The social efficiency of closure is analyzed in Acharya and Dreyfus (1989), Fries, Mella-Barral, and Perraudin (1997) and Mailath and Mester (1994). Acharya and Dreyfus derive an optimal closure rule and fair premium rate for banks operating in a competitive environment. While Acharya and Dreyfus focus exclusively on the optimal closure rule that *minimizes* the costs to the guarantor, Fries et al. consider a richer environment characterized by recapitalization possibilities and bankruptcy costs. Their research suggests that "forbearance" might sometimes be optimal, while in Acharya and Dreyfus this would always be suboptimal. Mailath and Mester introduce a social opportunity cost of closing a bank, in the form of lost intermediation services for a period. This cost must be traded off against the influence of closure policies on bank risk taking. In general, the subgame-perfect closure decision may lack a simple structure.

More research is needed on regulatory closure decisions versus depositor-induced

runs as disciplinary devices for bank management, the imperfections of secondary markets for liquidation of bank loan portfolios, and the role of private interim liquidity provision versus governmental deposit insurance under informational asymmetries. In particular, we need to learn more about the empirical costs and benefits of prompt closure (for example, Jones and King 1995). However, the empirical evidence on U.S. bank and thrift failures, including that discussed in section 2, provides powerful reasons to seriously consider an expanded role for monitoring by uninsured depositors and a lesser reliance on regulators for disciplining banks.

5. The Role of Bank Charter Value. The banking theory we have reviewed in section 1 examines a competitive bank in a static setting. But suppose we extend our model to permit heterogeneity among banks that creates rent-generation possibilities. For example, banks could differ in their monitoring costs $K(n)$. Then, lower-cost banks could earn rents. In a dynamic setting, the present value of expected future rents would represent the bank's *charter value.* If the deposit insurer threatens to close the bank whenever it fails, then a high bank charter value can deter risk taking. The reason is transparent. The higher is the bank's charter value, the greater is the cost associated with losing it, and hence the higher is the bank's *private cost* of asset portfolio risk. Suares (1994) formalizes this intuition in a somewhat different setting,[11] and Keeley (1990) provides supporting empirical evidence. Many others have viewed bank charter value as a strategic regulatory instrument and examined its interaction with other instruments. Most of the models are partial equilibrium, however, and do not completely endogenize banks' rents. Papers that attempt to endogenize these industrial organization aspects of banking include Matutues and Vives (1994, 1996).

In Boot and Greenbaum (1993), banks differ in their monitoring ability. The borrower chooses among projects where the bank's (costly) monitoring intensity affects the payoff distribution. The bank's expected profit is decreasing in project risk, but this risk can be reduced with greater bank monitoring. If loan payoffs are low despite monitoring, the bank is closed and future rents are lost. By expending monitoring effort, the bank not only improves these expected future rents, but also its monitoring reputation with uninsured depositors, thereby lowering its subsequent funding cost. An increase in either the expected rent or the funding-related reputational benefit induces greater monitoring. However, unless regulators make an analogous calculation in setting deposit insurance premia, the reputational benefit is lost with insured deposits, which is especially undesirable when enhanced competition decreases future rents.[12]

To summarize, there is clearly an interaction between the roles of bank charter value and regulatory closure policy in controlling banks' risk choices. A high charter

11. Suares (1994) emphasizes the importance of a credible closure policy that threatens to eliminate the charter, and also argues that recapitalization by distressed banks could make banks prefer more risk. The analysis in the literature of the implications of bank recapitalization for regulation has been only cursory. Approaches similar to that of Leland (1993) may be useful in future recapitalization analyses.

12. Bensaid, Pages, and Rochet (1993) examine some issues similar to those in Boot and Greenbaum (1993), focussing on a risk-averse bank's optimal choices of capital, scale, reserves, and monitoring effort, given the regulator's optimal contract for the bank. They interpret the optimal contract they derive as including a capital requirement.

value deters risk taking, but only if the regulator implements an appropriate closure policy. Likewise, an appropriate closure policy deters moral hazard, but only if banks have sufficiently high charter value.

C. *Moral Hazard and Bank Regulation: A Recapitulation*

We have seen in this section that deposit insurance, intended to solve a liquidity problem, creates moral hazard of its own. The literature we have reviewed suggests several ways to attack this moral hazard (see also Dewatripont and Tirole 1994a), without reaching definitive conclusions. We know that rents produce risk-control or monitoring incentives, but what is the best means of generating such rents (entry restrictions, expanded banking powers, deposit interest rate controls)? Capital requirements may also improve risk-control incentives, but not necessarily. In particular, they may not improve bankers' monitoring incentives if "outside" equity is involved.[13] Tough closure rules help control risk, but they may face implementation difficulties due to regulators' reputational concerns and may not be consistent with banks' recapitalization incentives. It would be interesting to consider the interaction of these issues in an empirically testable model. To summarize:

Implication 6: Partial governmental deposit insurance encourages market discipline through bank monitoring by informed depositors, and regulatory measures such as limited regulatory forbearance and tough bank closure rules may control bank risk taking.

Implication 7: A high bank charter value may not only directly deter risk taking but also makes feasible the design of incentive-compatible risk-sensitive capital requirements and deposit insurance premia to curb risk taking in environments plagued by private information and moral hazard.

In view of the ambiguities surrounding the potential benefits of deposit insurance vis à vis alternative liquidity-provision mechanisms and the lack of consensus about the optimal set of regulatory responses to the moral hazard generated by deposit insurance, it would be useful to question the role of deposit insurance in the context of recent credit market developments. We do this in the next section.

4. ARE THERE ALTERNATIVES TO DEPOSIT INSURANCE TODAY?

Our discussion in this section is organized in three subsections. In the first, we take the need for bank-provided liquidity as given, and ask whether there are ways *other* than deposit insurance to provide this liquidity. In the second, we explore the basic

13. Besanko and Kanatas (1996) show that when the distinction between outside and inside equity is explicitly recognized, increasing capital requirements may reduce the bank's incentive to monitor its borrowers and hence increase risk. Gennotte and Pyle (1991) and Boot and Greenbaum (1993) also show that there are circumstances in which capital controls do not reduce risk. Thakor (1996) develops a model in which a binding capital requirement exacerbates credit rationing, and also provides supporting empirical evidence.

question of whether banks are unique in providing liquidity, the extent to which the asset and liability intermediation services provided by banks can and are being provided by the capital market, and the regulatory implications of the transformation of banking into an off-balance-sheet business. Finally, we examine recent regulatory initiatives aimed at expanding the scope of U.S. banking, and their implications for the evolution of bank regulation.

A. *Interbank Borrowing-Lending and the Lender of Last Resort as Alternatives to Deposit Insurance*

Since the role of governmental deposit insurance in the contemporary theoretical framework is the provision of interim liquidity to absorb bank-specific liquidity shocks, we would like to know first if private arrangements can serve this function. One such arrangement is interbank borrowing-lending (the federal funds market, for example), and it has been examined by Bhattacharya and Gale (1987). The idea is that if bank-specific liquidity shocks are imperfectly correlated across banks, then banks could co-insure each other through a private market in which banks with surplus interim liquidity can lend to the deficit-liquidity banks. However, Bhattacharya-Gale show that, with unconstrained Walrasian access to an interbank borrowing-lending market, each bank *underinvests* in liquid assets relative to first best.[14]

A complement to interbank borrowing-lending is the Central Bank in its capacity as a lender of last resort (LLR). The Central Bank could make liquidity available to deficit-liquidity banks through the discount window whenever required, and this could eliminate runs. If the Central Bank made this liquidity available *unconditionally,* then there would be *no* difference between the LLR and deposit insurance, and the LLR facility would be vulnerable to all of the problems with deposit insurance discussed thus far. Therefore, we could envision the Central Bank advancing funds to a bank subject to a run *after* verifying that its underlying asset returns prospects do not warrant a run. Such a step would make it unnecessary for the bank to liquidate its investments early at a loss. On the other hand, if the Central Bank's investigation reveals that the bank is insolvent—the expected discounted value of future returns is lower than the level of the current withdrawal rights of its deposit contracts—then the bank would be closed down. Such an arrangement would be potentially superior to deposit insurance if the Central Bank's investigation of the liquidity-seeking bank is not too error-prone. Otherwise, with access to the discount window predicated on noisy Central Bank investigation, the desirability of employing the LLR facility in lieu of deposit insurance will depend on the trade-off between the various costs of deposit insurance and the cost of randomizing liquidity provision through the discount window which depends on the transparency of the bank's assets. Modern financial intermediation theory predicts bank assets that are opaque to outsiders. In Sharpe (1990) and Rajan (1992), for example, this opaqueness arises from intermediaries acquiring pro-

14. This is a static result. Bhattacharya and Padilla (1994) and Fulghieri and Rovelli (1993) have shown that in intergenerational economies, short-term assets and deposit inflows from new generations may suffice to take care of the liquidity needs of each generation's early diers.

prietary information about their loans through time which permits the extraction of monopoly rents but also diminishes the liquidity of the bank's assets. Of course, an active interbank market for liquid reserves would generate additional signals for the Central Bank to use; these signals would be embedded in the borrowing rates of liquidity-seeking banks in the interbank market.

The upshot of this discussion is that interbank borrowing-lending and the LLR facility—particularly the two working in concert—*may* be effective substitutes for governmental deposit insurance.

B. Recent Capital Market Developments and Their Regulatory Implications

There have been two recent capital market developments that are of particular relevance for bank regulation. One is that improvements in information technology and innovations in financial contracts (for example, securitization and asset-backed commercial paper) have improved capital market access for corporate borrowers. In the context of the model in section 1, one can view this as diminished ex ante and ex post informational asymmetries. This has two consequences. First, the greater capital market access raises borrowers' reservation utilities for bank financing, that is, $\bar{U}$ increases. If we extend the basic model to permit borrower heterogeneity in $\bar{U}$, then an increase in $\bar{U}$ for even a subset of borrowers will reduce the demand for bank credit. Thus, less insured deposits will be needed. Second, the enhanced tradeability of assets has improved the interim liquidity of bank assets. This makes it easier for an individual bank to realize the ("true") discounted (at $t = 1$) value of its long-run ($t = 2$) payoff, so that coping with a proportion of early withdrawals greater than P is *not* a problem. This reduces the value of deposit insurance.

The other relevant capital market development is the proliferation of liquid investment opportunities that are close substitutes for insured deposits from the investors' standpoint. Money market mutual funds have grown explosively at the expense of bank deposits. This elevated access to nonbank liquidity connotes an increase in investors' reservation utilities, U_0, in the context of the model in section 1. If investors are heterogenous—for example, they have different P values—then the theory would predict a decline in the supply of deposits.[15] Again, this causes a decrease in the need for deposit insurance.

In the context of the asset and liability intermediation services described in section 1, these market developments mean a smaller role for traditional banking. Not surprisingly then, banking has increasingly become an off-balance-sheet business with an explosion of new products such as loan commitments and standby letters of credit (Berger, Kashyap, and Scalise 1995). The theory underlying these instruments asserts the following: (i) these instruments are put options sold by the bank to its customers

15. Another deposit-insurance-related issue that deserves mention is that there are various "consumer protection" regulations of U.S. banks that are tied to governmental deposit insurance and impose distortionary allocations on banks. One example is the Community Reinvestment Act (CRA). Thakor and Beltz (1994) develop a theoretical model in which there is a coordination failure among banks in equilibrium that results in the banking industry opting for deposit insurance even though the costs of the investment distortions generated by the CRA outweigh the privately beneficial effects of deposit insurance for the majority of banks.

and therefore create contingent liabilities for the bank (Thakor, Hong, and Greenbaum 1981), (ii) some of these instruments, like loan commitments, improve borrowers' asset choice incentives making the bank *less* risky, and may also resolve problems of informational monopolies in bank-borrower relationships [see Houston and Venkataraman (1994) for the theory and Avery and Berger (1991) for empirical evidence], and (iii) loan commitments sold by competitive banks are wealth enhancing for borrowers [see Shockley (1995) for empirical evidence].

The regulatory implications are twofold. First, banks should be required to post capital against these off-balance-sheet claims. Consistent with the theory, under the 1987 Basle capital guidelines, banks must hold capital against loan commitments as well as standby letters of credit. Second, since these claims do *not* involve deposit funding, they permit banks to enhance borrower welfare and generate (fee) revenue without reliance on deposits. This means a smaller role for deposit insurance.

To summarize, recent capital market developments have led to a significant increase in asset liquidity, leading to declines in both investors' desire for insured deposits and borrowers' demand for bank credit financed with insured deposits. This suggests a decrease in the need for deposit insurance and with it a declining rationale for public regulation of banks.

C. *Portfolio Restrictions and Universal Banking*

Recent literature has also begun to focus on *banking scope,* that is, the degree to which banks can engage in different activities. Banking scope clarifies the distinction between *universal banking* and *functionally separated* banking. Universal banks perform both investment and commercial banking functions, while in a functionally separated system, these functions are allocated to different institutions.[16] Most European systems can be characterized as universal. The U.S. system is best described as functionally separated, although it is converging to the European system.

There is an ongoing debate about the desirability of universal banking, particularly in the United States. A considerable research effort is now directed at this issue, but the literature is still in its infancy. In the following discussion we review the arguments for and against universal banking.

The principal argument in favor of universal banking appears to be that (artificial) limitations on bank activities could potentially constrain optimal configurations that would arise endogenously. More specifically, this viewpoint assumes scope economies that are lost by separating commercial and investment banking. So separation

16. We do not focus on equity holdings. While some have emphasized this issue, particularly in the context of German universal banks, it is unclear whether universal banks voluntarily choose *to fund* corporations by buying equity. In many universal-banking countries, we do *not* observe pervasive equity holdings, and those that are observed are *strategically* motivated. A rationale for the absence of joint equity and debt funding by banks is provided by Dewatripoint and Tirole (1994b) and Gorton and Kahn (1993); they show that combined debt and equity holdings may undermine the disciplinary role of debt and exacerbate problems arising from the softness of bank's budget constraint. A related explanation can be found in banks' comparative advantages in delegated monitoring (Diamond 1984), and in the ability of debt contracts to minimize verification costs (Townsend 1979 and Gale and Hellwig 1985). Indeed, the theory in section 1 implies that banks will optimally choose to finance themselves with debt.

could impair the cross-sectional reusability of information between lending and underwriting activities,[17] and the poorer information utilization may exacerbate adverse selection. Separation may also undermine a bank's incentive to produce information and would consequently elevate borrowers' funding costs. Similarly, functionally separated banking may lower *intertemporal* reusability of information when a borrower "matures" from bank borrowing to financial market funding (Rajan 1995), and thus result in lower relationship-specific investments.

While better cross-sectional *and* intertemporal reusability of information represents an advantage of universal banking, there are also arguments against universal banking, three of which are potentially compelling. First, there could be serious conflicts of interest in permitting banks to underwrite their borrowers' capital market issues. Second, universal banking may adversely affect the development of the capital market. And third, universal banking increases the set of risky investments banks can make and is likely to create even larger banks that could escalate taxpayer exposure under the "Too Big to Fail" (TBTF) doctrine. The latter—TBTF argument—could be readily understood. The sheer size of universal banks makes tough closure decisions subject to substantial political backpedaling.

The first argument has been theoretically modeled by Kanatas and Qi (1994) and Rajan (1993). The bank may, for example, abuse its deposit insurance umbrella by extending a loan to an uncreditworthy borrower who might otherwise default on a capital market issue earlier underwritten by the bank. Or the bank may misrepresent the financial condition of a borrower whose capital market issue it is underwriting for the purpose of using the proceeds to pay off the bank's loan. However, the empirical evidence in Kroszner and Rajan (1994) and Gande, Puri, Saunders, and Walter (1995) suggests that market discipline is likely to be quite effective in preventing abuses.

As for the second argument, Boot and Thakor (1997b) develop a model in which the post-lending monitoring incentives of commercial banks and the financial innovation incentives of investment banks are endogenized. They show that a universal bank, which includes as its subsidiaries a commercial and an investment bank, stochastically innovates less than a stand-alone investment bank. This relative retardation of financial innovation leads to a less-developed capital market in a universal banking system than with a functionally separated banking system. On the other hand, a universal banking system is characterized by a better attenuation of borrower-specific asset-substitution moral hazard than a functionally separated banking system. Thus, the theory points to important trade-offs in financial system design.

The third argument, namely that expanded powers for banks could expand federal deposit insurance liability, remains perhaps the single biggest impediment to further expansion of bank powers. For example, it was recently reported that House Banking Committee Chairman Jim Leach was opposed to dismantling all of the Glass-Steagall restrictions on banks because, "the restrictions are necessary to protect federal deposit insurance" (see McConnell 1996). To summarize:

17. Recent evidence provided by Berger, Humphrey, and Pulley (1996) indicates, however, that the so-called one-stop-shopping advantages of universal banking may be overstated.

Implication 8: The growth of off-balance-sheet contingent claims in banking can be welfare enhancing, and with the market-based increase in asset liquidity, it reduces the need for insured deposits.

Implication 9: Governmental insurance of bank deposits may be the single biggest impediment to universal banking, but financial innovation may be impaired by universal banking.

5. CONCLUSION

What does all this tell us about how a banking system should be designed and how it should be regulated? This question is not easy to answer since we lack a sufficiently rich understanding of the relative efficiencies of banks and capital markets in processing and aggregating information; see Allen (1993), Bhattacharya and Chiesa (1995), and Boot and Thakor (1997a), and Yosha (1995) for initial attempts. However, the literature we have reviewed thus far offers the following conclusions (some tentative):

1. There should be no regulatory restrictions on bank size, other than to limit market power. This rationalizes recent regulations, such as the Neal-Riegle Interstate Banking Act of 1994, that have facilitated bank mergers.
2. The use of sequentially service-constrained demandable debt, without interest rate restrictions, in bank financing may provide superior intertemporal risk sharing. Given this, productively disruptive bank runs and panics can arise as Nash equilibrium phenomena.
3. Governmental deposit insurance should be preferred to suspension of convertibility to cope with bank runs if the deadweight costs of taxation are exceeded by the costs of randomization in meeting investors' liquidity needs.
4. Governmental deposit insurance distorts the behavior of insured institutions and engenders forms of moral hazard that elevate bank risk and taxpayer liability.
5. Risk-sensitive capital requirements and risk-calibrated deposit insurance premia are potentially useful regulatory tools in coping with moral hazard.
6. Improving bank closure policy and bringing market discipline to bear could attenuate deposit-insurance-related moral hazard.
7. Increasing (and preserving) banks' charter values can also help to dampen the risk-taking propensities of insured banks.
8. Improved asset liquidity in the capital market and the evolution of banking into an off-balance-sheet business diminish the need for deposit insurance.
9. Imposing portfolio restrictions on banks' investments may limit the liability of the deposit insurance fund. However, these impinge on the optimal (endogenous) configuration of banking and may as a result diminish charter values.
10. Permitting universal banking facilitates cross-sectional and intertemporal reusability of information, stimulating relationship-specific investments. This improves the banks' ability to cope with borrower moral hazard, elevating bank charter values and reducing moral hazard in banks' behavior. The princi-

pal drawbacks of universal banking are that it could retard financial innovation and financial market development, and lead to noncompetitive outcomes.

The basic message of contemporary banking theory is that banks may be inherently fragile in their role as providers of liquidity, and that this creates a role for a public safety net, either through governmental deposit insurance or through other mechanisms like a Central Bank functioning as a lender of last resort. However, the public safety net has numerous costs, including deadweight taxation costs, distorted asset portfolio choices of banks, and artificial restrictions on banking activities that may be efficiency-depleting. Given the explosive growth in market-provided liquidity as well as off-balance-sheet banking, a serious look at the desirability of governmental deposit insurance is called for.

Much remains to be done. The blurring distinctions between financial institutions and markets challenge us to seriously rethink our convenient theoretical distinctions between institutions and markets. How could this seamlessness be incorporated into a model in which the need for bank regulation arises endogenously? Would there be a need for insured deposits in such a framework? A full-blown analysis of financial system architecture seems necessary.

LITERATURE CITED

Acharya, Sankarshan, and Jean-Francois Dreyfus. "Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance." *Journal of Finance* 44 (1989), 1313–33.

Allen, Franklin. "The Market for Information and the Origin of Financial Intermediation." *Journal of Financial Intermediation* 1 (1990), 3–30.

_____. "Stock Markets and Resource Allocation." In *Capital Markets and Financial Intermediation,* edited by Colin Mayer and Xavier Vives. Cambridge University Press, 1993.

Allen, Linda, Anthony Saunders, and Gregory Udell. "The Pricing of Retail Deposits: Concentration and Information." *Journal of Financial Intermediation* 1 (1991), 355–61.

Avery, Robert, and Allen Berger. "Loan Commitments and Bank Risk Exposure." *Journal of Banking and Finance* 15 (1991), 173–92.

Bensaid, Bernard, Henri Pages, and Jean-Claude Rochet. "Efficient Regulation of Banks' Solvency." University of Toulouse, working paper, 1993.

Berger, Allen N. "The Relationship between Capital and Earnings in Banking." *Journal of Money, Credit, and Banking* 37 (1995), 432–56.

Berger, Allen N., David B. Humphrey, and Laurence B. Pulley. "Do Consumers Pay for One-stop Banking? Evidence from an Alternative Revenue Function." *Journal of Banking and Finance* 20 (1996), 1601–21.

Berger, Allen N., Anil K. Kashyap, and Joseph M. Scalise. "The Transformation of the U.S. Banking Industry: What a Long Strange Trip It's Been." *Brookings Papers on Economic Activity* 2 (1995).

Berger, Allen N., and Gregory Udell. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business* 68 (1995), 351–81.

Berlin, Mitchell, Anthony Saunders and Gregory Udell. "Deposit Insurance Reform: What Are the Issues and What Needs to be Fixed." *Journal of Banking and Finance* 15 (1991), 735–53.

Bernanke, Benjamin. "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression." *American Economic Review* 73 (1983), 257–76.

Bernanke, Benjamin, and Mark Gertler. "Financial Fragility and Economic Performance." *Quarterly Journal of Economics* 105 (1990), 87–114.

Besanko, David, and George Kanatas. "Bank Capital Regulations: Time Consistency, Hedging, and Incentive Compatibility." *Journal of Financial Intermediation* 5 (1996), 160–83.

Bhattacharya, Sudipto, and Gabriella Chiesa. "Proprietary Information, Financial Intermediation, and Research Incentives." *Journal of Financial Intermediation* 4 (1995), 328–57.

Bhattacharya, Sudipto, and Douglas Gale. "Preference Shocks, Liquidity, and Central Bank Policy." In *New Approaches to Monetary Economics,* edited by W. A. Barnett and K. J. Singleton. Cambridge University Press, 1987.

Bhattacharya, Sudipto, and Jorge Padilla. "Dynamic Banking: a Reconsideration." CEMFI, Madrid, working paper, 1994.

Bhattacharya, Sudipto, and Anjan Thakor. "Contemporary Banking Theory." *Journal of Financial Intermediation* 3 (1993), 2–50.

Billett, Matthew, Mark Flannery, and Jon Garfinkel. "The Effect of Lender Identity on a Borrowing Firm's Equity Return." *Journal of Finance* 50 (1995), 699–718.

Boot, Arnoud, and Stuart I. Greenbaum. "Bank Regulation, Reputation, and Rents: Theory and Policy Implications." In *Capital Markets and Financial Intermediation,* edited by Colin Mayer and Xavier Vives. Cambridge University Press, 1993.

Boot, Arnaud, and Anjan V. Thakor. "Self-Interested Bank Regulation." *American Economic Review* 83 (1993), 206–12.

_____. "Financial System Architecture." *The Review of Financial Studies* 10–3 (1997a), 697–733. (a)

_____. "Banking Scope and Financial Innovation." *The Review of Financial Studies* 10–4 (1997b), 1099–1131. (b)

Boyd, John, and Edward C. Prescott. "Financial Intermediary Coalitions." *Journal of Economic Theory* 38 (1986), 211–32.

Bryant, John. "A Model of Reserves, Bank Runs, and Deposit Insurance." *Journal of Banking and Finance* 4 (1980), 335–44.

Calomiris, Charles. "Financial Factors in the Great Depression." *Journal of Economic Perspectives* 7 (1993), 61–86.

Calomiris, Charles, and Gary Gorton. "The Origins of Banking Panics: Models, Facts, and Bank Regulation." In *Financial Markets and Financial Crises,* edited by R. Glenn Hubbard. Chicago: University of Chicago Press, 1991.

Calomiris, Charles, and Charles Kahn. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81 (1991), 497–513.

Calomiris, Charles, and Larry Schweikart. "The Panic of 1857: Origins, Transmission, and Containment." *Journal of Economic History* (December 1991), 807–34.

Chan, Yuk-shee, Stuart Greenbaum, and Anjan Thakor. "Is Fairly Priced Deposit Insurance Possible?" *Journal of Finance* 47 (1992), 227–46.

Chari, Varadarajan V., and Ravi Jagannathan. "Banking Panics, Information, and Rational Expectations Equilibrium." *Journal of Finance* 43 (1988), 749–61.

Cooperstein, Richard, George Pennacchi, and F. Stevens Redburn. "The Aggregate Cost of Deposit Insurance: A Multiperiod Analysis." *Journal of Financial Intermediation* 4 (1995), 242–71.

Dewatripont, Mathias, and Jean Tirole. *The Prudential Regulation of Banks.* MIT Press, 1994. (a)

_____. "A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence." *Quarterly Journal of Economics* 109 (1994), 1027–54. (b)

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

_____. "Liquidity, Banks, and Markets: Effects of Financial Development on Banks and the Maturity of Financial Claims." University of Chicago, working paper, 1996.

Diamond, Douglas, and Phil Dybvig. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (1983), 401–19.

Dowd, Kevin. "Competitive Banking, Bankers' Clubs, and Bank Regulation." *Journal of Money, Credit, and Banking* 26 (1994), 289–308.

Freixas, Xavier, and Enrique Gabillon. "Optimal Regulation of a Fully Insured Banking System." Universitat Pompeu-Fabra, working paper, 1994.

Fries, Steven, Pierre Mella-Barral, and William Perraudin. "Optimal Bank Reorganization and the Fair Pricing of Deposit Guarantees." *Journal of Banking and Finance* 21 (1997), 441–68.

Fulghieri, Paolo, and Ricardo Rovelli. "Capital Markets, Financial Intermediation, and Optimal Investment." IGIER working paper, Milan, 1993.

Gale, Douglas, and Martin Hellwig. "Incentive Compatible Debt Contracts: The One-Period Problem." *Review of Economic Studies* 52 (1985), 647–63.

Gande, Amar, Manju Puri, Anthony Saunders, and Ingo Walter. "Bank Underwriting of Debt Securities: Modern Evidence." Working paper, New York University, November 1995.

Gennotte, Gerard, and David Pyle. "Capital Controls and Bank Risk." *Journal of Banking and Finance* 15 (1991), 805–25.

Goodhart, Charles A.E. "Is the Concept of an Equilibrium Demand for Money Meaningful? A Reply to 'Has the Demand for Money Been Mislaid?'" *Journal of Post Keynesian Economics* 14 (1991), 134–36.

Gorton, Gary. "Banking Panics and Business Cycles." *Oxford Economic Papers* 40 (1988), 751–81.

Gorton, Gary, and Charles Kahn. "The Design of Bank Loan Contracts, Collateral, and Renegotiation." University of Pennsylvania, working paper, 1993.

Greenwald, Bruce, and Joseph E. Stiglitz. "Asymmetric Information and the New Theory of the Firm." *American Economic Review* 80 (1990), 160–65.

Greenbaum, Stuart, and Anjan Thakor. *Contemporary Financial Intermediation.* Orlando, Fl.: Dryden Press, 1995.

Hart, Oliver. "Theories of Optimal Capital Structure: A Managerial Discretion Perspective." National Bureau of Economic Research reprint, 1991.

Hellwig, Martin. "Liquidity Provision, Banking, and the Allocation of Interest Rate Risk." University of Basle working paper, 1993.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein. "Bank Monitoring and Investment: Evidence from the Changing Structure of Japanese Corporate Banking Relationships." In *Asymmetric Information, Corporate Finance, and Investment,* edited by R. Glenn Hubbard. Chicago: University of Chicago Press, 1990.

___ . "The Role of Banks in Reducing the Costs of Financial Distress in Japan." *Journal of Financial Economics* 27 (1990), 67–88.

Houston, Joel, and Subaru Venkataraman. "Information Revelation, Lock-In, and Bank Loan Commitments." *Journal of Financial Intermediation* 3 (1994), 355–78.

Jacklin, Charles J. "Demand Deposits, Trading Restrictions, and Risk Sharing." In *Contractual Arrangements for Intertemporal Trade,* edited by E. C. Prescott and N. Wallace. University of Minnesota Press, 1987.

Jacklin, Charles J., and Sudipto Bhattacharya. "Distinguishing Panics and Information-based Bank Runs: Welfare and Policy Implications." *Journal of Political Economy* 96 (1988), 568–92.

James, Christopher. "Some Evidence on the Uniqueness of Bank Loans." *Journal of Financial Economics* 19 (1987), 217–36.

_____. "The Losses Realized in Bank Failures." *Journal of Finance* 46 (1991), 1223–42.

Jones, David S., and Kathleen K. King. "The Implementation of Prompt Corrective Action: An Assessment." *Journal of Banking and Finance* 19 (1995), 491–510.

Kanatas, George, and Jiangping Qi. "Underwriting by Commercial Banks: Conflicts of Interest vs. Economies of Scope." University of South Florida, working paper, 1994.

Kaufman, George. "Bank Contagion: Theory and Evidence." Federal Reserve Bank of Chicago (1992).

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

Kroszner, Randall, and Raghuram Rajan. "Is the Glass-Steagall Act Justified? A Study of the U.S. Experience with Universal Banking." *American Economic Review* 84 (1994), 810–32.

Leland, Hayne. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." University of California at Berkeley, working paper, 1993.

Leland, Hayne, and David Pyle. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32 (1977), 371–87.

Lummer, Scott L., and John J. McConnell. "Further Evidence on the Bank Lending Process and the Capital-Market Response to Bank Loan Agreements." *Journal of Financial Economics* 25 (1989), 99–122.

Mailath, George, and Loretta Mester. "A Positive Analysis of Bank Closure." *Journal of Financial Intermediation* 3 (1994), 272–99.

Matutes, Carmen, and Xavier Vives. "Imperfect Competition, Risk-Taking, and Regulation in Banking." Autonomous University of Barcelona, working paper, 1994.

_____. "Competition for Deposits, Fragility, and Insurance." *Journal of Financial Intermediation* 5 (1996), 184–216.

McConnell, Bill. "Leach Seeks Broadening of Powers for Bank Units." *American Banker,* Friday, December 20, 1996, pp. 2–3.

Merton, Robert. "An Analytic Derivation of the Cost of Deposit Insurance Loan Guarantees." *Journal of Banking and Finance* 1 (1977), 3–11.

_____. "On the Costs of Deposit Insurance When There Are Surveillance Costs." *Journal of Business* 51 (1978), 439–52.

Millon, Marcia, and Anjan Thakor. "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies." *Journal of Finance* 40 (1985), 1403–22.

Myers, Stewart, and Nicholas Majluf. "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have." *Journal of Financial Economics* 13 (1984), 187–221.

Nakamura, Leonard. "Commercial Bank Information: Implications for the Structure of Banking." In *Structural Change in Banking,* edited by Michael Klausner and Laurence J. White. Homewood, Ill.: Irwin, 1993.

Peters, Stephen. "Why Are Demand Deposits Sequential Service Constrained and Insured?" University of Connecticut, working paper, 1994.

Puri, Manju. "Commercial Banks in Investment Banking: Conflicts of Interest or Certification Role?" *Journal of Financial Economics* 40 (1996), 373–401.

Rajan, Raghuram. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *Journal of Finance* 47 (1992), 1367–1400.

_____. "Conflict of Interest and the Separation of Commercial and Investment Banking." University of Chicago, working paper, 1993.

_____. "The Entry of Commercial Banks into the Securities Business: A Selective Survey of Theories and Evidence." Conference on Universal Banking held at New York University, 1995.

Ramakrishnan, Ram, and Anjan Thakor. "Information Reliability and a Theory of Financial Intermediation." *Review of Economic Studies* 51 (1984), 415–32.

Romer, Charles. "The Nation in Depression." *Journal of Economic Perspectives* 7 (1993), 19–40.

Rothschild, Michael, and Joseph Stiglitz. "Increasing Risk: A Definition." *Journal of Economic Theory* 2 (1970), 225–43.

Saunders, Anthony, and Berry Wilson. "Contagious Bank Runs: Evidence from the 1929–1933 Period." *Journal of Financial Intermediation* 5 (1996), 409–23.

Sharpe, Steven. "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships." *Journal of Finance* 45 (1990), 1069–87.

Shockley, Richard. "Bank Loan Commitments and Corporate Leverage." *Journal of Financial Intermediation* 4 (1995), 272–301.

Stiglitz, Joseph. "Credit Markets and the Control of Capital." *Journal of Money, Credit, and Banking* 17 (1985), 133–152.

Stiglitz, Joseph, and Andrew Weiss. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (1981), 393–410.

Suares, J. "Closure Rules, Market Power, and Risk-Taking in a Dynamic Model of Bank Behavior." University Carlos III, Madrid, working paper, 1994.

Thakor, Anjan. "Capital Requirements, Monetary Policy, and Aggregate Bank Lending: Theory and Empirical Evidence." *Journal of Finance* 51 (1996), 279–324.

Thakor, Anjan, and Jess Beltz. "A Barter Theory of Bank Regulation and Credit Allocation." *Journal of Money, Credit, and Banking* 26 (1994), 679–705.

Thakor, Anjan, Hai Hong, and Stuart Greenbaum. "Bank Loan Commitments and Interest Rate Volatility." *Journal of Banking and Finance* 5 (1981), 497–510.

Townsend, Robert. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (1979), 265–93.

von Thadden, Ludwig. "Intermediated versus Direct Investment: Optimal Liquidity Provision and Dynamic Incentive Compatibility." *Journal of Financial Intermediation,* forthcoming.

Yosha, Oved. "Information Disclosure Costs and the Choice of Financing Source." *Journal of Financial Intermediation* 4 (1995), 3–20.