Too Big to Fail: Origins, Consequences, and Outlook

Robert L. Hetzel

It should be emphasized that the views expressed in this article are the author’s and not necessarily those of the Bank or the Federal Reserve System.

I. INTRODUCTION

The General Accounting Office estimates the cost of the thrift industry bailout to be around $150 billion. George Benston, professor at Emory University, helps his students grasp the immensity of this number by asking them to imagine how hard it is to become a millionaire. He then asks them to imagine 150,000 millionaires being made paupers (New York Times, 6/10/90). The most important lesson of the thrift debacle is the need to close insolvent and nearly insolvent financial institutions promptly [Kane (1989); Bartholomew (1991)].

“Too big to fail” refers to the practice followed by bank regulators of protecting creditors (uninsured as well as insured depositors and debt holders) of large banks from loss in the event of failure. In this article, attention is focused on the resulting transfer of the decision to close a troubled bank from its creditors to bank regulators. The paper argues that the policy of too big to fail created in banking the same kinds of problems of timely closure that existed in the thrift industry.

The policy of too big to fail resulted from a fundamental deficiency in bankruptcy arrangements for banks. In banking there is no arrangement analogous to that existing for nonfinancial corporations unable to meet their debts, where the troubled corporation continues to operate while its creditors determine whether it is viable. It is usually undesirable to liquidate a large corporation immediately following a failure to pay its debts. The corporation may be viable if restructured. Also, an orderly, rather than an immediate, liquidation can increase the salvageable value of its assets. For nonfinancial corporations, therefore, Chapter 11 of the bankruptcy law provides a procedure under which a bankruptcy judge supervises the operation of a corporation that cannot pay its debts. Creditors and existing management then negotiate whether to liquidate or restructure.

Banks are not subject to bankruptcy law. For banks, there is no Chapter 11 administered by the courts. In its absence, the Federal Deposit Insurance Corporation (FDIC) and the Federal Reserve have come to run an informal Chapter 11 for banks. Particularly since the early 1980s, deposit insurance and the discount window have been used to keep in operation banks that otherwise would have been closed by the market.¹ This arrangement has been useful in that it prevents the abrupt closing of large banks. In contrast to corporate bankruptcy proceedings, however, the decision to allow a bank to fail is made by public officials rather than the bank’s creditors. This arrangement has delayed the resolution of insolvencies.

Section II relates the genesis of the policy of too big to fail. Section III examines how it encourages risk taking. Section IV advances a reform that would provide deposit insurance while leaving the decision to close a troubled bank to bank creditors. Using this reform as a benchmark, Section V asks whether the changes in bank regulation mandated by the Comprehensive Deposit Insurance Reform and Taxpayer Protection Act of 1991 will ensure that banks are closed neither too soon nor too late and that banks take neither too much nor too little risk.

II. THE DEVELOPMENT OF TOO BIG TO FAIL

A. A Brief History of Restrictions on Bank Competition

Free entry, the sine qua non of competition, has never had the constitutional protection in banking that it has in other industries. The Constitution

¹ The FDIC has argued that the policy of too big to fail has been imposed on it by statutory restrictions that require the least costly method of resolving a bank failure. Specifically, the FDIC can economically liquidate a small bank, but not a large one. Large bank failures, consequently, are handled by purchase and assumption arrangements, which avoid liquidation, but require the FDIC to inject enough funds into the sale of closed banks to restore their solvency and to avoid depositor losses. This situation reflects the absence of a Chapter 11 arrangement for banks that would allow either the market or regulators to close banks without forcing immediate liquidations.
prohibits states from interfering with trade across their boundaries. The commerce clause (Article I, Sec. 9) states, "No tax or duty shall be laid on articles exported from any state." In 1869, in Paul v. Virginia, the Supreme Court extended the protection of the interstate commerce clause to corporations by ruling that a state could not exclude an out-of-state corporation from doing business in it. [See Butler (1982), especially Ch. IV.] Banks are engaged in commerce in the sense that they are middlemen. They make money on the difference between the rates at which they borrow and lend. It has never been acceptable politically, however, to extend to banking the constitutional protection of free entry across state boundaries accorded other corporations.2

The ability of states to exclude out-of-state banks allowed state legislatures to organize their intrastate banking industries into a large number of small banks, each of which enjoyed some local monopoly power. Around the turn of the century, when the demand for banking services grew in rural areas, state legislatures passed laws prohibiting banks from meeting this demand through branching. By 1929, almost all states had laws either prohibiting or restricting intrastate branching. Along with low capital requirements for establishing a bank, the result was a banking industry consisting of large numbers of unit banks. The National Banking Act included ambiguous language that was often interpreted as forbidding interstate branching by national banks. In 1927, the McFadden Act eliminated any ambiguity by specifically prohibiting national banks from branching across state lines. As a result, banks could not diversify geographically their loans and the sources of their deposits. [For a history of prohibitions on branching, see Mengle (1990)].

Competition in banking was further restricted during the Depression. At the time, many blamed the Depression on excessive competition. It appeared plausible that individual insolvent firms could be made solvent by restricting competition. In many industries, consequently, government regulation attempted to raise prices by restricting competition. The government divided financial intermediation into separate industries that could not compete with each other. Glass-Steagall separated fund-raising for corporations into banking and securities industries. Insurance companies, savings and loans, and credit unions were assigned their own regulators and spheres of influence. The Banking Act of 1933 prohibited the payment of interest on demand deposits, and Regulation Q limited the interest banks could pay on time and savings deposits. Later, the 1956 Bank Holding Company Act restricted the ability of banks to operate nationally through multibank holding companies. Justice Department antitrust guidelines restricted competition by limiting the ability of banks to acquire other banks.

B. Increased Competition for Banks

Beginning in the late 1960s, innovations in communications and computer technology eroded restrictions on competition in banking by making it possible for nonbank institutions like money market mutual funds to offer bank-like services to bank customers. By lowering the cost of bookkeeping and disseminating information, this technology lessened the advantage that banks had possessed formerly in gathering deposits and monitoring the credit risk of borrowers. The emergence of new competitors to banks has reduced the viable size of the traditional banking industry.

The dramatic increase in competition facing banks is apparent in the minimal extent to which businesses relied on domestic banks in 1990 for additional credit. Baer and Brewer (1991) report the following statistics. In 1990, business credit provided by domestic banks, finance companies, U.S. branches of foreign banks, offshore sources, and by nonfinancial commercial paper grew 7.3 percent. Domestic banks provided only a small fraction, 7 percent, of this growth in funding. The year 1990 was unusual in that many banks were attempting to increase their capital-to-asset ratios by restricting asset growth. The figures reflect, however, a longer-run decline in bank business lending. Over the decade of the 1980s, banks' share of short- and medium-term lending to businesses declined at an annual rate of 1.5 percent. Furthermore, as Baer and Brewer point out, the market is continuing to develop new substitutes for bank lending like asset-backed commercial paper (backed by trade receivables) and prime rate funds (backed by commercial loans). These sources, which are not included in the above sources, added about two percentage points in 1990 to growth in short-term business credit.

---

2 Ironically, the European Common Market is using the U.S. federal model to promote open competition in banking. Starting in 1993, banks will be free to branch across national boundaries. Furthermore, in general, host-country regulators cannot place any restrictions on the activities of branches of foreign banks not imposed by the home-country regulator. [See Coleman and Hart (7/29/91).] The combination of guaranteed free entry and host-country regulation sharply curtails the ability of host-country regulators to limit competition in the banking industry.
Beginning in the last half of the 1960s, a series of financial innovations eroded regulatory restrictions on competition in deposit gathering. High market rates of interest produced by high rates of inflation generated incentives to escape the low ceilings on deposit rates fixed by Reg Q and the prohibition of payment of interest on demand deposits. In the 1960s, the Eurodollar market developed as a way of allowing large depositors to put their deposits on banks' books in Europe, where Reg Q did not apply. In the 1970s, automatic teller machines allowed banks to evade some geographical restrictions on banking. Money market funds and NOW accounts allowed depositors with small accounts to avoid ceilings on deposit rates. Money market funds permit savers with small amounts of savings to bypass financial institutions and invest indirectly in commercial paper, which is issued only in large denominations. NOW accounts are checkable deposits that evade the prohibition of payment of interest on demand deposits through the technicality of being labeled a savings account.

Banks have lost their dominant role not only as collectors of relatively cheap funds from small savers but also as suppliers of loans to low-risk businesses. Many large corporations issue commercial paper directly to financial intermediaries like pension funds, rather than borrow from banks. By the end of 1989, money market funds, with more than 20 million accounts, held about $455 billion in assets, much of it commercial paper. Moreover, the Glass-Steagall Act, which forbids banks from underwriting corporate securities, has prevented banks from meeting the changing needs of their corporate clients by underwriting their debt issues. In addition, a variety of firms not regulated as banks lend to corporations and consumers. Subsidiaries owned by AT&T, Ford, General Electric, and Sears make commercial loans. Automobile companies finance car loans through financial subsidiaries. American Express, AT&T, and Sears finance consumer loans through their credit card divisions.

Regional banks retain a comparative advantage in making loans to companies too small to enter the money market. They are, however, losing their comparative advantage in gathering deposits. In particular, they rely on deposit insurance as an aid in competing with money market funds for the relatively large accounts of older depositors.

Securitization, the packaging of illiquid assets like mortgages and car loans into a security that can be sold, has eroded the former natural monopoly possessed by banks to transform a portfolio of illiquid assets into liquid liabilities. Securitization has increased the range of financial intermediaries that can hold the illiquid assets formerly held only by banks. Pension funds, insurance companies, mutual funds and individuals now hold assets that formerly were held chiefly by banks. In the decade of the 1970s, banks held 34.8 percent of the financial assets of financial intermediaries, which include, in addition to banks, other depository institutions, government-sponsored enterprises, insurance companies, pension and retirement funds, and money market and other mutual funds. By 1989, this figure had fallen to 26.6 percent [U.S. Treasury (1991), Ch. 1, Table 7].

Government policies have also created competitors for banks. The sharp rise of market rates in 1972 in combination with fixed Reg Q ceilings on time and savings deposits produced an outflow of funds from thrifts and banks. In order to maintain the flow of funds to housing without raising Reg Q ceilings, the government expanded the financing activities of the Federal National Mortgage Association, or “Fannie Mac,” and the Federal Home Loan Mortgage Corporation, or “Freddie Mac.” Fannie Mae and Freddie Mac purchase mortgages and then either hold them for their own account or package them so they can be held by institutional investors. These federally sponsored credit agencies compete with banks and thrifts for the financing and warehousing of mortgages.

Tax laws have placed banks at a disadvantage in competing for the long-term savings of individuals. For example, certain laws make some portion of an employee’s wages tax-exempt if placed in a long-term savings plan (such as a 401-K plan that allows the employee to defer taxes on interest income from investments). These laws encourage corporations and state and local governments to replace banks as deposit gatherers. Corporations and state and local governments then negotiate directly with pension and thrift funds. Because these funds typically are large enough to evaluate the risk of their assets, they can hold commercial paper and bonds issued by corporations. This financial intermediation completely bypasses banks.

C. The Extension of Deposit Insurance

Had the market forces described above been left unopposed, they would have forced a contraction of the banking industry in the 1980s. Contraction was postponed through the extension of deposit insurance
coverage in the form of the policy of too big to fail. This extension provided a subsidy to banks by lowering their costs of funding relative to the costs they would have incurred if some holders of their liabilities had not been protected from loss. Too big to fail arose from pressures created by the lack of satisfactory institutional arrangements for closing banks rather than from a conscious decision on the part of policymakers.

The FDIC was created in 1933 to protect depositors holding small accounts. It was not created to keep insolvent banks in operation. That task was assigned to the Reconstruction Finance Corporation. [See Todd (1988), App. C, and Todd (1991).] The emergence of too big to fail is recounted in Bailout by Irvine Sprague, a former director of the FDIC. Sprague recounts the transformation of the FDIC from an agency charged with covering losses of insured depositors of already failed banks into a modern day Reconstruction Finance Corporation that prevents failures by protecting all creditors of large banks from loss.3

**Beginnings** From 1950 until 1982, Section 13(c) of the Federal Deposit Insurance Act allowed the FDIC to prevent a bank from failing if the bank were judged “essential to provide adequate banking service in its community.” Sprague points out that legislative history is an unclear guide to the intended use of the “essentiality” clause, but that “this authority was not intended for widespread use” [Sprague (1986), p. 28]. Probably, it was included to prevent the failure of banks in rural areas served by a single bank. In any event, as Sprague points out, the language effectively gave the FDIC complete discretion because “the courts have always upheld an agency’s discretionary authority. . . . No challenge has been successful. So there you have it. A bank can be bailed out if two of three FDIC board members determine it should be” (Sprague, p. 28). Since the passage of the Garn-St. Germain Act in 1982, the FDIC has had the additional authority to prevent failures by arranging purchase and assumption transactions if it determines that liquidation of the bank is a costlier alternative.

In 1971, Unity Bank in Boston became the first bank bailed out under the essentiality doctrine. One of the FDIC directors opposed the bailout on the grounds “that bailouts were bad public policy and doing the first one would lead to many more, possibly an uncontrollable flood” (Sprague, p. 46). That director was persuaded not to vote while the other two directors, Sprague and Wille, voted to keep Unity afloat, even though it was mismanaged. Unity was saved because of a fear that the failure of a bank considered to be a black institution would set off riots in black neighborhoods. Sprague (p. 48) notes, “[M]y vote to make the ‘essentiality’ finding and thus save the little bank was probably foreordained, an inevitable legacy of [the riots in] Watts.” At the time, Sprague reports he believed Unity could not set a precedent because it was “a unique case, one of a kind” (Sprague, p. 49). In fact, shortly thereafter, the FDIC bailed out Bank of the Commonwealth, a large, mismanaged bank in Detroit for the same reason.

In retrospect, Sprague identifies the bailout of Unity as the first step in establishing too big to fail as public policy. “[T]he important precedent was, of course, the irreversible turn we had taken with Unity, away from our historic narrow role of acting only after the bank had failed. . . . Now we were in the bailout business, how deeply no one could then tell” (Sprague, p. 49). Sprague then goes on to describe how over time too big to fail became embedded in banking regulation through the precedent of saving one troubled bank at a time, rather than as a result of a conscious decision. By Chapter 15 of Bailout, Sprague says: “Of the fifty largest bank failures in history, forty-six—including the top twenty—were handled either through a pure bailout or an FDIC-assisted transaction where no depositor, insured or uninsured, lost a penny. In effect, the forty-six enjoyed 100 percent insurance protection. The four lonely exceptions . . . were the result of unusual circumstances” (Sprague, p. 242).4

**Continental Illinois** Because of its size, the bailout of Continental Illinois exemplified most clearly the transformation of the FDIC into a modern Reconstruction Finance Corporation. Although only 10 percent of Continental’s deposits were insured, the FDIC protected all of its depositors. [The

---

3 Even when the too big to fail doctrine is applied, banks can fail in that their charters are revoked and their stockholders lose their investments. Too big to fail means that the FDIC and the Federal Reserve prevent a lack of funding from closing large, troubled banks while the FDIC arranges a takeover by another bank (a purchase and assumption transaction). The FDIC subsidizes the takeover so that bank depositors (uninsured as well as insured) do not incur any losses.

4 The major “lonely exception” was Penn Square, which was liquidated because the FDIC believed that the bank’s negligent and possibly fraudulent loan practices might create such extensive litigation as to render impossible a purchase and assumption transaction.
following draws on Sprague (1986), Part 4. See also Thomson and Todd (1990). In the late 1970s and early 1980s, Continental grew rapidly by taking on high-risk loans. In particular, Continental purchased $1 billion in oil and gas loans from Penn Square, collateralized by drilling rigs and other assets made worthless when the oil drilling business collapsed. Continental’s downfall began with the bankruptcy of Penn Square Bank in Oklahoma City in July 1982. On May 9, 1984, foreign depositors began to withdraw deposits from Continental. Initially, the Federal Reserve Bank of Chicago assisted it with discount window loans, which eventually totaled $7.6 billion.

Because no buyers could be found for Continental, the FDIC decided to keep it in operation through “open bank assistance.” It purchased $1 billion in preferred stock from Continental’s holding company, which lent the funds to Continental. This arrangement amounted to capital forbearance in that Continental was allowed to remain in operation with an amount of private capital below regulatory standards. (It also protected the creditors of the holding company, as well as the bank.)

Recent Developments Two recent bank failures illustrate the blanket coverage extended to uninsured depositors under the policy of too big to fail. On August 10, 1990, the Comptroller of the Currency closed National Bank of Washington and named the FDIC as receiver. Because of the lapse of time between the dissemination of information about the bank’s problems and the closing of the bank, depositors at the Nassau branch had begun to withdraw funds beginning early in 1990. Discount window lending by the Fed might have permitted a sizable portion of these deposit withdrawals. The foreign deposits remaining when the bank was seized, although not formally insured by the FDIC, were protected. According to newspaper accounts, the FDIC protected Washington National’s foreign deposits in order to provide assurance to foreign depositors at money-center banks that their deposits were protected (American Banker, 9/27/90). The policy of too big to fail has effectively extended insurance to deposits in banks’ overseas branches.5

On January 6, 1991, the FDIC took control of the Bank of New England Corporation’s banks—Bank of New England, Connecticut Bank and Trust, and Maine National Bank. The FDIC’s initial estimate of the loss was $2.3 billion (Financial Times, 1/8/91). Newspaper commentary accompanying the rescue makes clear how broad the criteria have become for bailing out a bank’s uninsured creditors (Wall Street Journal, 1/7/91):

The arrangement will protect from loss all depositors, even those with accounts exceeding the $100,000 insurance ceiling. Mr. Seidman made it clear in an interview that the urgency of the rescue transcended the bank’s difficulties. “We’re looking at an eroding economy, particularly in New England,” he said. . . . Over the weekend, government officials stressed the need to improve credit conditions in New England to slow economic deterioration there.

At present, “too big to fail” appears to be a misnomer as even small banks are usually not allowed to fail with a loss to uninsured depositors. As William Seidman, former chairman of the FDIC, noted (Wall Street Journal, 6/5/91):

Some people mistakenly believe that small-bank failures usually are resolved through a payout of insured deposits—a liquidation, where uninsured depositors and creditors suffer some loss. The reality is that, currently, about nine out of ten small-bank failures are resolved through “purchase and assumption” transactions. In a P&A [purchase and assumption transaction], all the deposits (including those over the $100,000 insurance limit) generally are assumed by a healthy bank. Of the 169 banks that failed in 1990, only 20 were resolved through a payout of insured deposits. The rest were resolved through P&As.

While deposit insurance expanded principally through growth of the policy of too big to fail, explicit increases in the size of covered deposits and regulatory actions also expanded its coverage. Regulators used increases in coverage per account to lower the cost of funds and to increase their availability to banks and thrifts when increases in market rates above Reg Q prompted disintermediation. Increases in coverage per account coincided with peaks in market rates: 1966, 1969, 1974, and 1980. FDIC-insured deposits as a fraction of total deposits increased from about 55 percent in 1965 to more than 70 percent in the 1980s (U.S. Treasury (1991), Conclusions and Recommendations, Figures 6 and 7).

During the 1980s, the FDIC also increased the kinds of bank liabilities protected from loss. It extended insurance to the deposits of pension plans by “passing through” the $100,000 insurance limit to the individual participants of the plans. It insured brokered deposits.6 The FDIC also insured private

5 By the end of 1990, these deposits totaled about $300 billion and amounted to 51 percent of the deposits of the nine largest U.S. banks.

6 In 1984, the FDIC and FSLIC adopted regulations to deny deposit insurance to certificates of deposit purchased from a broker. The regulations, however, were overturned by a federal court, and Congress was unwilling to pass legislation allowing the regulatory agencies to prohibit brokered certificates of deposit.
The current low rate of return to banking is consistent with the argument that the policy of too big to fail has kept the banking industry from contracting in response to increased competition. In the 1970s, the return on assets for insured commercial banks was .77 percent. From 1985 to 1989, this figure fell to .55 percent [U.S. Treasury (1991), Ch. 1, Table 6]. The decline in the value of bank stocks relative to the S&P 500, which began in the late 1970s and became more pronounced beginning in 1986, reveals investor skepticism about the future profitability of the banking industry given its current size. Even with the rally in bank stocks in early 1991, the P/E ratio of money-center and large regional banks is only half that of firms in the S&P 500 (American Banker, 3/18/91).

D. Pressures to Postpone Closing Insolvent Banks

Regulators incur a variety of pressures to postpone closing a troubled bank. Closing a bank produces active disapproval from those affected adversely. In contrast, beneficiaries are unaware of the costs indirectly imposed on them by the relaxation of market discipline involved in keeping a troubled bank afloat. Beneficiaries include consumers who benefit from being treated fairly, and they want to be perceived as fair in the media. Given the ambiguity of measures of capital, they naturally tend to close a bank only when it is clearly insolvent. Their desire for fairness presents them with another dilemma over troubled banks: given a chance, some of these banks will recover. If regulators were to close a bank that is not obviously insolvent, they would inevitably be criticized for the "premature and unnecessary" closure of a bank that "if given a chance, would make it." They would receive especially heated criticism from small borrowers with special, ongoing relationships with the troubled bank.

In the case of the Bank of New England, the Comptroller of the Currency only closed the bank when depositors actually began to run it. In an American Banker (1/9/91) article, Comptroller Clarke and FDIC Chairman Seidman were reported as having said that "they did not act until last weekend because only then was it certain that the Bank of New England Corp. had no chance to survive." In the same article, Karen Shaw, president of the Institute for Strategy Development, noted: "There was the hope, even if it was an errant hope, that the bank would survive. It had made it through several crises."

Rather than liquidate a large bank for failure to meet a capital standard, regulators are more likely to swap transactions with banks in receivership.

Significantly, the committee found nothing intrinsically wrong with the intervention by these five senators with federal regulators in 1987 in behalf of Charles H. Keating. . . Each had ample information to justify contacting regulators about the fairness of the regulatory treatment Lincoln was receiving. The case has hung on the nexus between the five senators' intervention and the enormous political contributions that they collected from Keating.

Congressional Quarterly [Cranford (1991), p. 519] also reports Sen. DeConcini's reply to the report of the Ethics Committee:

As the committee effectively acknowledges, in early 1987 I had strong reason to believe that a major Arizona company was being treated unfairly by the federal government. I further had reason to believe that 2,000 Arizona jobs were unfairly at stake.

Pressures on regulators to keep troubled banks afloat do not have to be political. Regulators may temporize because they are genuinely uncertain whether a bank is insolvent. It is often difficult to measure the market value of a bank's assets and, consequently, the market value of its capital. Regulators want to be fair, and they want to be perceived as fair in the media. Given the ambiguity of measures of capital, they naturally tend to close a bank only when it is clearly insolvent. Their desire for fairness presents them with another dilemma over troubled banks: given a chance, some of these banks will recover. If regulators were to close a bank that is not obviously insolvent, they would inevitably be criticized for the "premature and unnecessary" closure of a bank that "if given a chance, would make it." They would receive especially heated criticism from small borrowers with special, ongoing relationships with the troubled bank.

7 "Regulators, attuned to the necessities of congressional relations, and no more willing than other human beings to put other people out of work, will forbear until the death rattle is clearly audible" [Wallison (1991), p. 12].

8 Such criticism is unlikely to acknowledge the moral hazard problems of allowing a troubled bank to remain open. "History and what little we know of human nature suggest that as weak bank managements struggle to survive they will reach for more risky investments to pay for their more costly funds, obscure from examiners the dangerous condition of their enterprise, appeal to their elected representatives as needy constituents and surrender only when all hope is gone" [Wallison (1991), p. 12].
to try to find a merger partner for the bank while allowing it to remain in business. Such partners, however, are hard to find. Possible merger partners have an incentive to wait until the condition of a troubled bank deteriorates to the point where it is actually taken over by the FDIC. They can then negotiate with the FDIC. In this way, banks not only avoid a possible costly court fight with the troubled bank's stockholders, but also open up the possibility that the FDIC will add incentives to make the acquisition more attractive.

An additional incentive to procrastination in closing a bank is the multiplicity of regulators. Understandably, each regulator would like for the other regulator to receive any criticism for closing a bank. For example, the Comptroller of the Currency has the responsibility for declaring national banks insolvent. The Comptroller, however, does not use its own resources to keep a troubled bank afloat. It has an incentive, therefore, to wait and hope that the Fed will effectively close the bank by pulling its discount window loan or that the FDIC will close the bank by refusing to grant a waiver for the bank to attract insured brokered deposits. Also, state regulators may have been reluctant to revoke the charters of insolvent state-chartered banks. Closing down a bank that they examine may appear as an admission of failure.

### III. THE PROBLEMS WITH TOO BIG TO FAIL

Restrictions on the ability of market forces to close banks embodied in the policy of too big to fail started to cause problems when banks began to experience significant external competition in the 1970s and 1980s. The extension of deposit insurance in the form of the policy of too big to fail provided a subsidy to banks that kept the banking industry from contracting. It also encouraged risk taking. Kept from shrinking by an extension of the implicit subsidies of deposit insurance, banks responded to the loss of low-risk corporate customers by turning to riskier investments. Many banks also increased the subsidy from deposit insurance by holding riskier asset portfolios without increasing their capital.

#### A. FDIC Insurance as a Subsidy to Risk Taking

This section first explains why government-sponsored insurance subsidizes risk taking by failing to price risk. The remainder of the section documents the increase in the riskiness of the banking system since the 1960s. Banks have acquired increasingly risky asset portfolios, with no increase over this period in capital ratios.

Insurance offered by a private company pools individual risks by establishing a fund into which premiums are paid and from which losses are met. The insurance company is the residual claimant on the fund. It makes money when the fund grows and loses money when it declines. For this reason, the insurance company places restrictions on its policyholders that limit the risks they take. Private insurance cannot subsidize risk taking. It must price risk accurately or go out of business.

Although the FDIC uses the term “insurance fund,” its fund is fundamentally different from the kind of fund maintained by a private insurance company. With the FDIC fund, there are no residual claimants whose own money is at stake. The Treasury keeps a tally on the cumulative difference between incoming FDIC deposit premia and outgoing FDIC expenditures and includes interest on the positive balance. This tally is the FDIC “insurance fund.” The FDIC fund can be depleted, but it cannot become insolvent. If the current receipts from the premia paid by banks are insufficient to cover current FDIC expenditures, then FDIC deposit insurance commits the taxpayer to pay the difference. It is this commitment that allows deposit insurance to be used to subsidize risk taking.

In contrast to a private insurance arrangement that limits risk taking, FDIC insurance encourages risk taking. The subsidy banks receive from the guarantee of their deposits by the government increases with the riskiness of their asset portfolio and decreases with the amount of capital they hold. FDIC deposit insurance does not lower the cost of funds appreciably for a conservatively managed, highly capitalized bank, but it does for a risk-taking, poorly capitalized bank. Both kinds of banks pay the same flat rate on their insured deposits.\(^9\)

The encouragement given by deposit insurance to risk taking was kept in check as long as restrictions

\(^9\) In principle, risk-based capital guidelines can offset the incentives deposit insurance creates for risk taking. In practice, however, such guidelines are hard to implement. They assign risk on the basis of broad categories, which do not differentiate between riskiness of assets within categories. Also, it is often hard to defend the relative assessment of risk across categories. For example, the 1988 Basle Agreements on risk-based capital guidelines stipulate that mortgage-backed securities, which are often subject to significant risk from interest rate fluctuations, require only one-fifth the capital of a commercial loan. The most important drawback to risk-based capital guidelines is their failure to reward risk reduction through asset diversification.
on competition gave banks a high franchise value. A high franchise value acts like a large amount of capital. It limits risk taking because stockholders bear significant losses if the bank fails. Increased competition in banking, however, has eroded the franchise value of banks, especially since the 1970s [Keeley (1990)].

B. Evidence of Increased Risk in the Banking System

Deposit insurance has allowed the banking industry to pursue riskier investment strategies in the 1980s without increasing its capital. In the early 1960s, insured commercial banks had a ratio of capital to total assets of about 8 percent. In the 1980s, this ratio fell to about 6 percent [U.S. Department of Commerce (1965), Table 607, and U.S. Department of Commerce (1985), Table 826]. The 6 percent aggregate figure conceals significant variation among banks. At year-end 1989, the largest 25 banks had a ratio of capital to assets of only 4.8 percent [U.S. Treasury (1991), Ch. II, Table 2]. Barth, Brumbaugh and Litan (1990, Table 3) report that, as of June 1990, 183 banks with assets of $11.2 billion were operating with capital-to-assets ratios of less than 3 percent. An additional 67 banks with combined assets of $80 billion had capital-to-assets ratios between 3 and 3.5 percent.

There is also evidence that, for troubled banks, capital based on the book value of assets overstates capital based on market values. Using cross section data to study the relationship between book and market values, Mengle (1991, p. 21) finds that, among banks that failed, capital based on book values significantly overstated capital based on market values. Mengle (1991, p. 19) also notes:

Failures are far more common than book value insolvency numbers would suggest; in any given year, the number of failures far exceeds the number of book value insolvencies in either the current or the previous year. . . . only 6 percent of the banks that failed in 1985 had reported themselves to be book value insolvent in 1985.

Moreover, the stock market values many banks less highly than the book value of their capital. Barth, Brumbaugh, and Litan (1991, Table 6) show that for 20 of the 25 largest U.S. banks, the market value of their equity is less than the book value.

While the ratio of book capital to assets has changed very little for the banking system over the last two decades, banks’ asset portfolios have become riskier. The loss of blue chip corporate customers to the commercial paper market left remaining bank loan portfolios riskier. Particularly since the mid 1980s, bank lending has been concentrated in high-risk categories. From 1985 through 1990, 65 percent of the increase in bank loans was in real estate. Of the increase in real estate lending during this period, 43.5 percent was in commercial real estate [Board of Governors (March 1991), Table 1.25, “Assets and Liabilities of Commercial Banks” and Board of Governors (December 1991), Table 1.54, “Mortgage Debt Outstanding”]. Banks purchased large amounts of mortgage-backed securities, which present considerable interest rate risk, and became heavily involved in off-balance sheet activities such as loan commitments and standby letters of credit. In dollar terms, these latter activities grew from 58 percent of assets in 1982 to 116 percent in 1989 [U.S. Treasury (1991), Ch. I, p. 27].

The increase in the riskiness of bank asset portfolios could not have occurred without the subsidy to risk-taking provided by deposit insurance. Consider the contrasting management of the banks described in the following quotations:

Long dependent on a regional economy that rises and falls with the auto industry’s swings, banks in the region [Midwest] rarely stray far from the basics of opening checking accounts, backing small businesses and managing trust funds. Economic uncertainty has generally kept them from seeking quick gains through risky loans to commercial developers, junk bond artists or Third World nations (Wall Street Journal, 4/11/91).

Each week, Walter Connolly [head of the Bank of New England Corp.] would survey the numbers from his banking empire, 450 branches that stretched from the top to the bottom of New England. If the numbers didn’t look right—if a bank had lost even a bit of market share—Connolly would get the bank’s president on the phone and demand to know what had happened. No matter what he was told, he had the same answer: “Grow it, grow it.” “The whole culture was one of growth,” said Donald J. Kauth. . . . “Size was success to Walter,” said one colleague. “He wanted to retire as the chairman of the biggest bank in Boston” (Washington Post, 1/9/91).

Deposit insurance subsidizes the risk-taking bank, not the conservatively managed bank. In particular, deposit insurance makes it possible for a bank to finance rapid growth with cheap deposits even if that growth is achieved by acquiring low quality and risky assets.

C. Systemic Risk

The policy of too big to fail is often defended as a response to inherent instability in the banking sector.
system. This section makes the opposite argument, namely, that deposit insurance and the policy of too big to fail have created instability in the banking system.

Because too big to fail entails closing banks without imposing losses on depositors, there is an ongoing possibility that the FDIC will have to ask Congress for funds to close insolvent banks. The regressive character of the wealth transfers involved in bailing out the creditors of large banks, however, makes it difficult for Congress to appropriate funds for the FDIC. The ongoing possibility of a need for additional funding to operate the FDIC, combined with the uncertainties surrounding the congressional appropriations process, creates a potential for systemic instability. In a situation where many banks are poorly capitalized, a run on a large bank at a time when the deposit insurance fund is depleted could cause bank runs to spread uncontrollably. Under current institutional arrangements, therefore, regulators must maintain control over the timing of the closing of insolvent banks. This control can only be achieved by protecting all creditors of banks (including uninsured depositors) from loss. Current institutional arrangements make too big to fail an imperative.

Too big to fail, however, is part of a vicious circle. Protecting all creditors from loss limits incentives for creditors to monitor the riskiness of bank asset portfolios. Banks can then hold only minimal amounts of capital while making risky investments without increasing the rate they pay on deposits. This behavior, however, creates precisely the weakness in the banking system that makes too big to fail appear to be an imperative. Ironically, deposit insurance has produced the systemic instability it was supposed to prevent.

IV. THE WORLD WITHOUT TOO BIG TO FAIL

A. The New Legislation

The policy of too big to fail resulted in part from a lack of satisfactory institutional arrangements for closing insolvent banks in a timely way. The 1991 Deposit Insurance Reform Act addresses the problem of timely closure by requiring regulators to close a bank when its capital falls below a specified level. [The Shadow Financial Regulatory Committee (1988) has been the primary proponent of this approach.]

The new act requires bank regulators to establish five capital categories: well capitalized, adequately capitalized, undercapitalized, significantly undercapitalized, and critically undercapitalized. It requires regulators to classify as critically undercapitalized any bank with a capital-to-assets ratio of 2 percent or less. Regulators must close such a bank within 90 days. Furthermore, the act legislates a list of strictures that bank regulators must impose on banks in the undercapitalized and significantly undercapitalized categories.

The mandatory early intervention written into the new law probably derived from the political difficulty in introducing market discipline by explicitly limiting deposit insurance. The law intends that regulators close troubled banks before insured deposits are put at risk. Consequently, the FDIC would not have to absorb a loss when a bank fails. It follows that the new law should remove the former subsidy to risk taking created by deposit insurance. Its intent is to make deposit insurance superfluous without explicitly limiting or repealing it.

Will the Deposit Insurance Reform Act end the policy of too big to fail through the prompt closing of troubled banks? Will banks begin to assume an optimal amount of risk? Answering these questions requires a benchmark against which the legislation can be judged. This benchmark is taken to be a banking system in which market discipline controls both the closure decision and the riskiness of bank asset portfolios. How would such a system work?

B. Market Closure

Such a system could apply the arrangements for corporate bankruptcy to the banking industry. Under

---

11 At present, the FDIC possesses only a limited ability to generate additional revenue through increases in the premium it levies on deposits. Additional premium increases will reduce the ability of banks to compete with other financial intermediaries. Banks are already subject to the special tax imposed by noninterest-bearing reserve requirements. In the past, required reserves and FDIC deposit premia have collected similar amounts of revenue. For example, in December 1990, depository institutions held $37.5 billion in required reserves, and the three-month Treasury bill rate was 6.8 percent. At an annual rate, the reserve requirement tax was collecting $3.9 billion in revenue (0.068 x $37.5). In 1989, FDIC assessment income came to $3.5 billion.

12 Direct attempts to limit deposit insurance, such as reducing the $100,000 deposit limit, limiting the number of allowable accounts per household, and providing for a depositor deductible, proved too controversial to include in the new law.
corporate bankruptcy law, the timing of the closure decision of troubled firms is approximately right. Although much of corporate bankruptcy law would not be applicable, it would be possible to recreate its major features for banks. First, bankruptcy law provides a market mechanism for placing corporations in receivership. That is, the decision to place a corporation in receivership is made by a firm's creditors or management rather than by public officials. Second, corporations in receivership are not allowed to fail catastrophically. Chapter 11 allows a bankrupt corporation to continue in operation while it is either reorganized or liquidated in an orderly fashion. Third, bankruptcy law provides a rule for apportioning losses among a firm's creditors. Fourth, although insolvent corporations are not liquidated abruptly, they are not kept in operation with government money. The individuals involved have the necessary incentives to close institutions that are not viable.

These features of corporate bankruptcy law could be approximated for banks. First, the policy of too big to fail could be eliminated by rescinding the ability of any government agency to offer guarantees to bank creditors, including depositors. Bank creditors then would have their own money at risk in a bank failure. Analogous to corporate bankruptcy, bankruptcy of a bank would be triggered by failure to honor an obligation for payment. Second, a petition for bankruptcy would put a bank into the kind of receivership provided for by Chapter 11 in corporate bankruptcy. Third, when a bank went into receivership its depositors would incur an immediate loss (haircut) on the amount of their deposits at the close of the day preceding the bankruptcy decision. Regulators would set the size of the loss so that the failed bank would again possess positive net worth. Apart from the haircut, depositors would have complete access to their funds, unlike creditors in a corporate bankruptcy. Fourth, if necessary, banks in receivership could obtain new funds from debtor-in-possession financing, that is, financing in which new lenders receive priority in repayment over existing holders of subordinated debt.

13 Dotsey and Kuprianov (1990) discuss relevant issues in the context of the problems with the S&L industry in the 1980s.
14 A detailed proposal is available from the author.
15 Although creditors could petition a bankruptcy court to close a bank that did not honor its debts, most petitions for bankruptcy would probably be voluntary petitions by bank management made to stop an incipient run.
16 The American Banker's Association (1990) and Boyd and Rolnick (1988) have advocated depositor haircuts.

C. A Proposal to Restructure Deposit Insurance

The provision of a Chapter 11 arrangement for banks, which would ensure that banks never failed abruptly, would protect the payments system. The issue remains, however, of whether allowing bank depositors to close a bank through a run could precipitate a system-wide run. In particular, economists differ over whether the ability of the Federal Reserve to undertake large-scale open market purchases of securities to supply reserves to the banking system would be sufficient to prevent a general run of the banking system. This issue is likely to remain contentious. A system of deposit insurance, however, could be designed that would keep the closure decision in the hands of bank creditors while still protecting against bank panics. Deposit insurance, which currently is an entitlement granted whenever a bank creates a deposit, could be fixed in quantity and priced by the market.

Specifically, the Treasury would auction deposit insurance certificates to banks, which would sell them to depositors desiring to insure their deposits. The amount of these certificates, however, would be kept less than the total deposits in the banking system. In addition, individual banks would be able to offer insurance certificates only up to a fraction of their total deposits. In this way, all banks would have some depositors genuinely at risk in the event of a failure.

An advantage of this proposal is that it is agnostic on the issue of whether instability is an inherent feature of banking. On the one hand, if policymakers believe that the banking system is prone to instability, they can maintain permanently a relatively high ratio of deposit insurance certificates relative to the total deposits of the banking system. On the other hand,

17 Anna Schwartz argues that it would. She observes (personal communication to the author): "There were runs before the public could confidently count on a lender of last resort to nip them in the bud. I keep asking why Britain had its last run on banks in 1866, but the U.S. continued to experience them until 1933. In the first case, the lender of last resort had learned what to do to prevent a run, and the public knew it." Diamond and Dybvig (1983), in contrast, construct a model in which in principle it is possible to have bank runs that would not be offset by open market purchases.
18 For example, assume a depositor has $20,000 in deposits at the time his bank enters bankruptcy court and $15,000 in deposit insurance certificates. Assume also that depositors receive a haircut of 5 percent. The depositor, then, receives a haircut of 5 percent on the $5,000 of deposits not covered by his certificates.

When a bank is placed into bankruptcy and its depositors incur haircuts, the bank submits the deposit insurance certificates registered with it to the Treasury. The Treasury becomes a claimant on the failed bank for an amount equal to the percentage haircut applied to the bank's certificates.
they may believe that the best way to ensure a stable banking system is through market discipline that gives individual banks an incentive to hold high levels of capital and limit risk taking. They could then gradually eliminate the certificates.

V. BANK CLOSURE AND RISK TAKING WITH THE NEW LEGISLATION

Under the new law, how well will regulators approximate the results that would be achieved by competitive market forces? Will troubled banks be closed neither prematurely nor tardily? Will banks take the right amount of risk, neither too much nor too little?

A. Optimal Closure

The new law continues to rely heavily on the discretion of government regulators. Regulators still decide when to write down the book value of a bank's assets and, as a consequence, when to lower the credit category into which a bank falls. Fear of triggering a run could still make regulators reluctant to downgrade a large bank's capital category.

Alternatively, the new law could create political pressure to close troubled banks prematurely. It gives regulators a whole arsenal of weapons for limiting risk and restricting the activities of troubled banks. If a bank does fail with a significant loss to the FDIC, bank regulators could easily be accused of negligent supervision. In order to avoid the criticism of negligence, regulators might write down the book value of a bank's assets at any sign of trouble. In the spirit of the new law, they might take strong action against any bank whose capital falls below even the most conservative capital categories.

B. Optimal Risk Taking

Even if banks are closed promptly when the book value of their capital falls below 2 percent, losses to the deposit insurance fund may remain large. Limiting losses to the insurance fund will also require limiting the riskiness of bank asset portfolios. The new legislation leaves uncertain whether limiting bank risk will occur through imposition of market discipline or through regulator intervention.

Consider, in this regard, the failure on August 10, 1990, of National Bank of Washington (discussed earlier). National Bank of Washington was the second bank closed under the powers granted by the Financial Institution Reform, Recovery, and Enforcement Act of 1989. These powers allow regulators to seize a bank before it becomes insolvent. On March 31, 1990, the bank had a capital-to-assets ratio of 5 percent (Consolidated Reports of Condition and Income). When seized four months later, it still had a capital-to-assets ratio of 1.4 percent. Measured by the book value of its assets, it was solvent. When National Bank of Washington was closed and sold to Riggs National Bank, however, the FDIC had to retain $539 million in its assets, one-third of the bank's $1.6 billion in assets. The FDIC's initial estimate of its loss was $500 million (American Banker, 9/27/90). National Bank of Washington was in fact deeply insolvent when it was closed.

The new law requires regulators to close within 90 days a bank with a capital-to-assets ratio of 2 percent or less. The experience with National Bank of Washington, which was closed when it had a book value capital ratio of 1.4 percent, suggests that the new law will not necessarily prevent failures that impose heavy losses on the FDIC.

A study by Alton Gilbert (1992, Table 4) of 1,000 banks that failed since 1985 puts the average loss ratio at 27 percent (the loss to the FDIC divided by the book value of the failed bank's assets). Under the system of bank regulation that existed prior to the Deposit Insurance Reform Act, the standard practice was for regulators to close banks when their book capital-to-assets ratio reached 0 percent. Under the new act, they will close banks when this ratio reaches 2 percent. Given the magnitude of the historic loss ratio, 27 percent, the change from 0 to 2 percent should not be expected to have a significant effect in reducing FDIC losses.

Two other changes in the regulatory regime could, however, keep FDIC losses small in the future. First, the public could come to believe that federal regulators will allow large banks to fail with losses to uninsured depositors. Uninsured depositors would then begin to exert a discipline on bank risk taking by demanding a return on their deposits commensurate with the riskiness of their portfolios. Market discipline would force banks to price risk correctly by imposing higher interest rates on the uninsured deposits of banks with risky asset portfolios. Correct pricing by banks would limit the riskiness of their asset portfolios.

\[\text{Pratt's Letter (December 20, 1991) commented:} \quad \text{"When Senate Banking Committee Chairman Riegle prematurely (and unjustly) dispatched Comptroller Clarke to regulatory boot hill, he delivered a powerful message of his own: Mess up and I'll have your head. And now that Congress has legislatively signed what it perceives to be a $95 billion check for the FDIC and RTC, its probity for the Riegle-style witch hunt will increase."}\]
Second, bank regulators might be drawn heavily into limiting bank risk. Because of the difficulties in evaluating risk, this involvement would probably take the form of relatively crude quantitative limits on different kinds of investments. Attempts by regulators to restrict the riskiness of bank asset portfolios, however, could prevent banks from allocating capital efficiently by balancing risk and return. Reducing the riskiness of bank assets is not an end in itself. After all, banks exist because of the need to make risky loans. They are financial intermediaries that specialize in pricing risk and monitoring risky lending. If regulators are drawn into this second alternative, they could greatly harm the banking industry’s ability to extend credit efficiently.

VI. CONCLUDING COMMENT

The new legislation may remove most of the subsidy offered by deposit insurance. Increases in the deposit insurance premium levied by the FDIC could even turn deposit insurance into a net tax. If implicit subsidies in deposit insurance kept banking from contracting in the 1980s, then the new legislation, by removing these subsidies, will precipitate a contraction of banking. The contraction may be particularly severe among large banks that lost corporate customers to the commercial paper market in the 1980s. The difficulties in achieving optimal closure and risk taking in banking discussed above are likely to be exacerbated by the need for the banking industry to contract. This contraction could be impeded by the tendency of failed banks to be merged rather than closed and by legal restraints that prevent banks from diversifying freely into other financial and commercial activities.

Banking is still different from other industries. There is no active market for corporate control in which raiders can acquire a bank and shrink it. There is no market-driven bankruptcy procedure. Consequently, there is no market mechanism to assure efficient shrinkage of the banking industry. It is important that this problem be understood and publicly debated.

REFERENCES


