

# 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 non-financial 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.<sup>1</sup> 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

---

<sup>1</sup> 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.<sup>2</sup>

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 Mingle (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.

---

<sup>2</sup> 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 home-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. I, 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.<sup>3</sup>

**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.

<sup>3</sup> 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.

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).<sup>4</sup>

**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

<sup>4</sup> 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.<sup>5</sup>

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).

<sup>5</sup> 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.

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.<sup>6</sup> The FDIC also insured private

<sup>6</sup> 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.

parties to swap transactions with banks in receivership.

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. I, 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 competition and potential entrants in the banking industry.

The ability of regulatory agencies to keep open a troubled bank inevitably invites political pressures. Congressmen pass on constituent discontent over the job losses and personal disruptions that accompany the closing of a bank. The case of the Keating Five is instructive. Beginning in 1987, five senators apparently intervened with the Federal Home Loan Bank Board in order to keep Lincoln Savings and Loan in operation. The Senate Ethics Committee found the intervention itself appropriate. The only issue was whether it was appropriate to accept money from Mr. Keating while intervening on his behalf with federal regulators. In summarizing the report of the Ethics Committee, *Congressional Quarterly* [Cranford (1991), p. 518] noted:

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.<sup>7</sup> 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.<sup>8</sup>

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

<sup>7</sup> "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].

<sup>8</sup> 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.<sup>9</sup>

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

---

<sup>9</sup> 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.<sup>10</sup> 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 outstrips 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

<sup>10</sup> At the time, this capital ratio was adequate because restrictions on competition made banking relatively safe. The absence of any change in the bank prime rate from August 1960 through December 1965 used to be cited as evidence of the cartelization of banking.

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.<sup>11</sup> 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.

---

<sup>11</sup> 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 \$57.5 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 (.068 x \$57.5). In 1989, FDIC assessment income came to \$3.5 billion.

## 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.<sup>12</sup> 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

---

<sup>12</sup> 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.<sup>13</sup> 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.<sup>14</sup> 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.<sup>15</sup> 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.<sup>16</sup> 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.

<sup>13</sup> Dotsey and Kuprianov (1990) discuss relevant issues in the context of the problems with the S&L industry in the 1980s.

<sup>14</sup> A detailed proposal is available from the author.

<sup>15</sup> 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.

<sup>16</sup> 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.<sup>17</sup> 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.<sup>18</sup>

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,

<sup>17</sup> 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.

<sup>18</sup> 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.<sup>19</sup> 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.

<sup>19</sup> *Pratt's Letter* (December 20, 1991) commented: "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 proclivity for the Riegle-style witch hunt will increase."

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. With the passage of 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 banks' 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.

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

- American Banker*. "Banks Halt Their Binge in Mortgage Securities." May 8, 1990, p. 1.
- . "FDIC Vetoed Plan to Curb Its Coverage in DC Failure." September 27, 1990, p. 1.
- . "Regulators Defend Timing in Seizure of Boston Bank." January 9, 1991, p. 1.
- . "Even After a Five-Month Surge, Bank Stocks Remain Dirt Cheap." March 18, 1991, p. 1.
- American Bankers Association. "Federal Deposit Insurance: A Program for Reform." March 1990.
- Baer, Herbert and Elijah Brewer III. "The Credit Crunch and the Diminishing Role of Banks." Federal Reserve Bank of Chicago, July 1991, processed.
- Barth, James R., Philip F. Bartholomew, and Michael G. Bradley. "Reforming Federal Deposit Insurance: What Can Be Learned from Private Insurance Practices?" Research Paper #161, Federal Home Loan Bank Board, June 1989.
- Barth, James R., R. Dan Brumbaugh, Jr., and Robert E. Litan. "The Banking Industry in Turmoil: A Report on the Condition of the U.S. Banking Industry and the Bank Insurance Fund." Report to the Financial Institutions Subcommittee of the House Committee on Banking, Housing and Urban Affairs, 101st Cong., 2nd sess., December 1990.
- Bartholomew, Philip F. *Estimating the Cost of Regulatory Forbearance During the Thrift Crisis*. Congressional Budget Office, 1991, processed.
- Board of Governors of the Federal Reserve System. *Annual Statistical Digest*. Washington, D.C.: Board of Governors, March 1991.
- . *Federal Reserve Bulletin*. Washington, D.C.: Board of Governors, December 1991.
- Boyd, John H. and Arthur J. Rolnick. "A Case for Reforming Federal Deposit Insurance." Federal Reserve Bank of Minneapolis *Annual Report*, 1988.
- Butler, Henry Nolde. *Legal Change in an Interest-Group Perspective: The Demise of Special Corporate Charters*. Ph.D. Dissertation, Virginia Polytechnic Institute and State University, May 1982.
- Coleman, Martin and Susan Hart. "Europe Can Teach U.S. a Few Lessons on Reform." *American Banker*, July 29, 1991, p. 4.
- Cranford, John R. "Decision in Keating Five Case Settles Little for Senate." *Congressional Quarterly*, March 2, 1991, pp. 517-23.
- Diamond, Douglas W. and Philip Dybvig. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (June 1983): 401-19.
- Dotsey, Michael and Tony Kuprianov. "Reforming Deposit Insurance: Lessons from the Savings and Loan Crisis." Federal Reserve Bank of Richmond *Economic Review* 76 (March/April 1990): 3-28.
- Financial Times*. "Lifeboat Rushes Into New England Storm." January 8, 1991, p. 17.
- Gilbert, R. Alton. "Effects of Legislating Prompt Corrective Action on the Losses of the Bank Insurance Fund." January 1992, photocopy.
- Kane, Edward J. *The S&L Insurance Mess: How Did It Happen?* Washington, D.C.: Urban Institute Press, 1989.

- Kaufman, George G. "Banking Risk in Historical Perspective." Federal Reserve Bank of Chicago *Proceedings of a Conference on Bank Structure and Competition*, 1986, pp. 231-49.
- Keeley, Michael. "Deposit Insurance, Risk and Market Power in Banking." *American Economic Review* 80 (December 1990): 1183-1200.
- Mengle, David L. "The Case for Interstate Banking." Federal Reserve Bank of Richmond *Economic Review* 76 (November/December 1990): 3-17.
- \_\_\_\_\_. "Market Value Accounting and Bank Failures." Federal Reserve Bank of Richmond, June 1991, processed.
- Mengle, David L. and John R. Walter. "How Market Value Accounting Would Affect Banks." Federal Reserve Bank of Chicago *Proceedings of a Conference on Bank Structure and Competition*, forthcoming 1991.
- New York Times*. "Numbers Crunch." June 10, 1990, p. E1.
- Rolnick, Arthur and Warren E. Weber. "The Causes of Free Bank Failures: A Detailed Examination." *Journal of Monetary Economics* 14 (November 1984): 267-91.
- Shadow Financial Regulatory Committee. "An Outline of a Program for Deposit Insurance Reform." Statement No. 38. Mid America Institute for Public Policy Research, Chicago, December 5, 1988.
- Sprague, Irvine H. *Bailout*. New York: Basic Books, 1986.
- Thomson, James B. and Walker F. Todd. "An Insider's View of the Politics of the Too Big to Fail Doctrine." Federal Reserve Bank of Cleveland Working Paper #9017, December 1990.
- Todd, Walker F. "Lessons of the Past and Prospects for the Future in Lender of Last Resort Theory." Federal Reserve Bank of Chicago *Proceedings of a Conference on Bank Structure and Competition*, May 11-13, 1988, pp. 533-77.
- \_\_\_\_\_. "Banks Are Not Special: The Federal Safety Net and Banking Powers" in Catherine England (ed.), *Governing Banking's Future: Markets vs. Regulation*. Boston: Kluwer Academic Publishers, 1991.
- U.S. Department of Commerce. *Statistical Abstract of the United States*. Washington, D.C.: 1965 and 1985.
- U.S. Treasury. "Modernizing the Financial System." February 1991.
- Wall Street Journal*. "U.S. Recession Claims Bank of New England As First Big Victim." January 7, 1991, p. 1.
- \_\_\_\_\_. "Bruised by the S&L Fiasco, Lawmakers Now Try to Show They Are Born Again Bank Guardians." March 19, 1991, p. A26.
- \_\_\_\_\_. "In Great Lakes Region Many Bankers Prove Caution Can Pay Off." April 11, 1991, p. 1.
- \_\_\_\_\_. "The Facts About the FDIC." June 5, 1991, p. A12.
- Wallison, Peter J. "Treasury Bank Plan: Strangle the Weak." *Wall Street Journal*, February 13, 1991, p. 12.
- Washington Post*. "Bank of New England Out-Paced Rivals." January 9, 1991, p. F1.

# Financial Evolution and the Long-Run Behavior of Velocity: New Evidence from U.S. Regional Data

*Peter N. Ireland\**

## I. INTRODUCTION

Monetary economists have devoted considerable effort to establishing a link between the financial innovations of the past two decades and the coincident instability of conventional econometric money demand specifications.<sup>1</sup> They have paid little attention, in contrast, to the more general question of how financial developments may have influenced the demand for money over longer periods of U.S. monetary history. Thus, one survey of the literature notes, new hypotheses about the effects of financial innovation "have for the most part been tested on the same body of data that suggested them in the first place" [Judd and Scadding (1982), p.1001]. It is unclear whether these hypotheses can be useful in understanding the effects of earlier innovations or in predicting the effects of future innovations.

The utility of a stable econometric money demand function, however, lies precisely in its ability to forecast out-of-sample so as to indicate, for instance, what rate of nominal money growth will be consistent with a desired rate of inflation. A satisfactory theory attributing changes in money demand to innovations in the financial sector must therefore account for the effects of a long history of past innovations and be able to predict the effects of future innovations. Such a theory has recently been developed and tested by Michael Bordo and Lars

Jonung (1987, 1990) as part of an extensive research project on the long-run behavior of the income velocity of money.<sup>2</sup>

Bordo and Jonung suggest that the institutional and financial factors that systematically influence the demand for money in an economy over the entire course of its development are of two types. On the one hand, the process of monetization—meaning the growth of the commercial banking system in addition to the expansion of formal market activity at the expense of barter and production for own use—ought to increase the demand for money as an economy grows. On the other hand, the emergence of a variety of nonbank financial intermediaries offering assets that potentially substitute for money and the invention of cash management techniques used to economize on real balances ought to have the opposite effect of lowering money demand. Bordo and Jonung's hypothesis is that the first set of effects will dominate early in the course of economic development but will be eclipsed by the second set in later stages of growth; velocity will therefore tend to trace out a U-shaped pattern over time. In recently published work [Bordo and Jonung (1987, 1990)], they provide evidence that this pattern can indeed be found in both U.S. and international data.

This paper shows how Bordo and Jonung's hypothesis derives from traditional theories of velocity's long-run behavior. It then discusses some objections that have been raised in reviews of their empirical work. In response to these objections, it examines a new data set containing figures for demand deposit velocity by region in the United States since 1929. Regression equations estimated with the new data support Bordo and Jonung's theory. The

---

\* Thanks go to Marvin Goodfriend, Robert Hetzel, Tom Humphrey, Jeff Lacker, Barbara Mace, and Richard Manning for making helpful suggestions, and to Andy Atkeson and Rachel van Elkan for providing unpublished worksheets containing regional demand deposit data. The opinions expressed herein are those of the author and do not necessarily reflect those of the above-mentioned individuals, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

<sup>1</sup> The first among recent empirical studies to attribute money demand instability to financial innovation include Enzler, Johnson, and Paulus (1976) and Goldfeld (1976). Judd and Scadding (1982) and Goldfeld and Sichel (1990) survey the subsequent literature.

<sup>2</sup> The income velocity of money is defined as the ratio of national income (in nominal terms) to the nominal money supply. It is therefore a convenient measure of how money demand compares to income, with lower money demand relative to income translating into higher velocity and vice versa.

regional figures are also found to be consistent with an explanation for the recent weakness in M1 velocity. These results suggest that the new data set represents a valuable untapped source of evidence with which a variety of hypotheses about the behavior of velocity can be tested.

## II. THEORIES OF VELOCITY'S LONG-RUN BEHAVIOR

Figure 1 displays the long-run behavior of the income velocity of the U.S. monetary aggregate M1, using gross national product as the measure of income.<sup>3</sup> It shows that M1 velocity declined secularly from 1869 until the end of World War II and has risen secularly since then.

The downward trend in velocity prior to 1945, as well as the short-run movements that accompanied it, is documented in great detail by Friedman and Schwartz (1963).<sup>4</sup> They propose that real money balances be viewed as a luxury good, having an income elasticity in excess of unity, and therefore

<sup>3</sup> All data sources are listed in the appendix.

<sup>4</sup> Friedman and Schwartz use M2 as their empirical definition of money. The long-run behavior of M2, however, does not differ substantially from that of M1 until after World War II, when M2 velocity levels off and M1 velocity rises sharply. Thus, Friedman and Schwartz's explanation of M2 velocity's initial downward trend works equally well in explaining the prewar behavior of M1.

attribute the secular decline in velocity to the concurrent secular rise in real income. To explain the trend's subsequent reversal, which at the time their volume was written had only just begun, Friedman and Schwartz point to postwar expectations of greater economic stability that worked, they said, to decrease the demand for money as a safe and highly liquid asset.

Friedman and Schwartz's luxury good hypothesis became increasingly difficult to apply in explaining the postwar behavior of M1 as its velocity continued to rise. Thus, a number of researchers, including Latané (1960), Meltzer (1963), and Lucas (1988) argue instead for a unitary income elasticity and a significantly negative interest rate elasticity of money demand, thereby implying that movements in velocity are directly attributable to movements in interest rates. In their later work, Friedman and Schwartz (1982) also include an interest rate variable in regression equations used to describe the demand for money in the United States and the United Kingdom from 1869 to 1975. Figure 2, which shows the behavior of the commercial paper rate from 1869 to 1989, does suggest the existence of a close velocity-interest rate relationship, as both variables trace out U-shaped patterns over time.

After studying the long-run demand for money in two countries, however, Friedman and Schwartz (1982) conclude that movements in velocity cannot

Figure 1  
M1 VELOCITY  
1869-1989

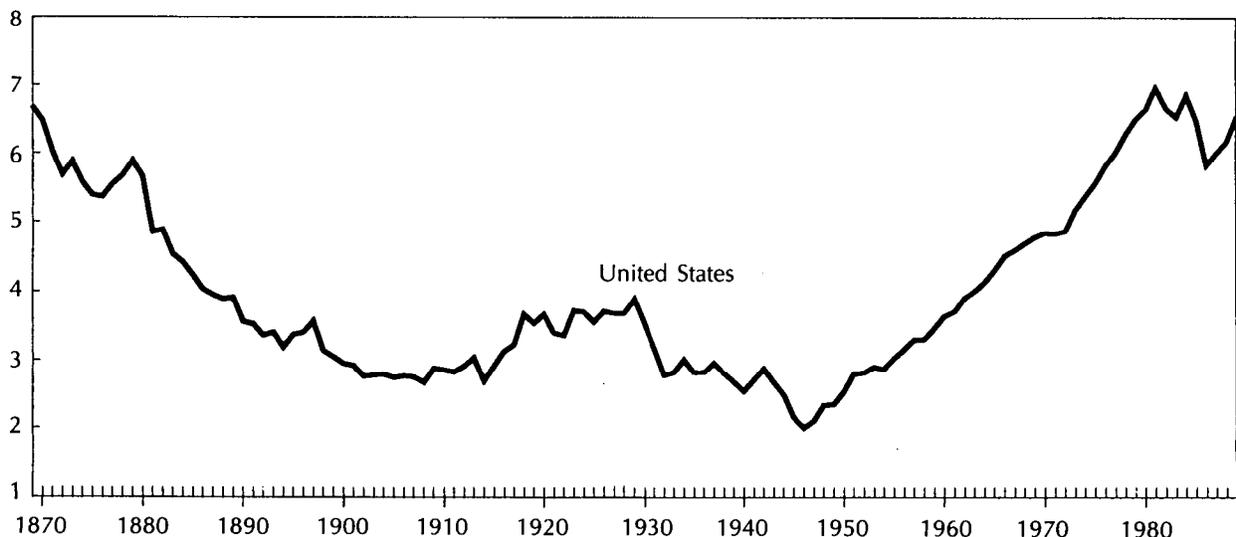
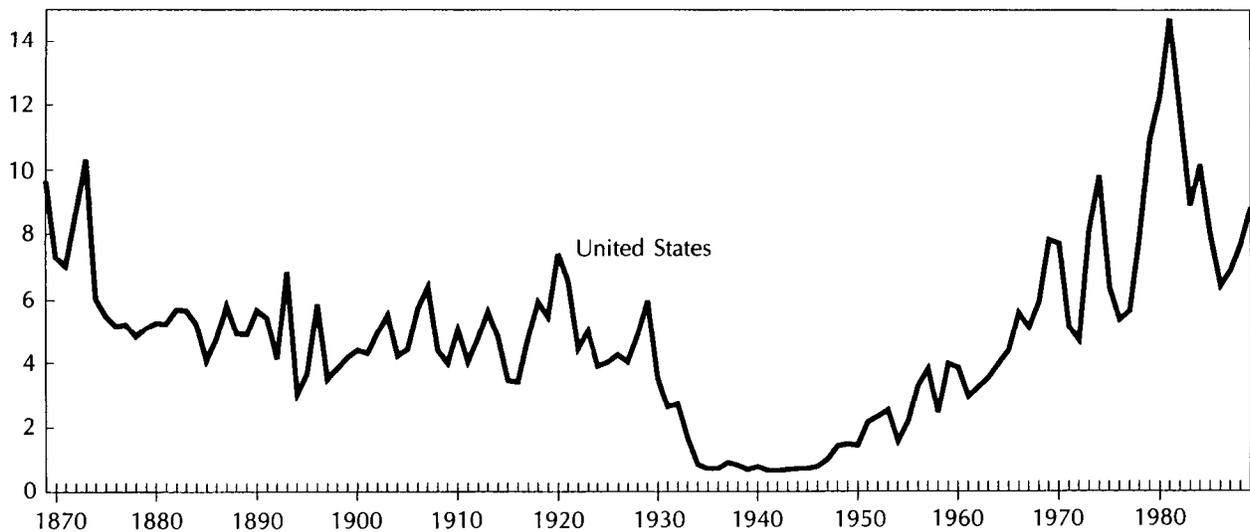


Figure 2  
**COMMERCIAL PAPER RATE**  
 1869-1989



be attributed exclusively to movements in income and interest rates. Figure 3 displays the Friedman-Schwartz M2 velocity and interest rate data for the United States and the United Kingdom from 1880 to 1910.<sup>5</sup> Although interest rates in both countries moved within a narrow range, velocity fell sharply in the United States while remaining remarkably stable in Britain. Friedman and Schwartz (pp. 146-47) explain the divergence in the two velocity series by noting that while in 1880 the United Kingdom's economy was far more financially sophisticated than that of the United States, by 1910 this gap had narrowed:

From 1880 to 1910, the United States population nearly doubled, but the number of banks multiplied more than sevenfold. The fraction of the population residing in rural areas had declined from over two-thirds to only a bit over one-half; the fraction of the work force in agriculture had declined from one-half to less than one-third. . . .

. . . the change in relative financial sophistication of the United Kingdom and the United States . . . was probably by all odds the single most important factor accounting for the divergent trends in real balances.

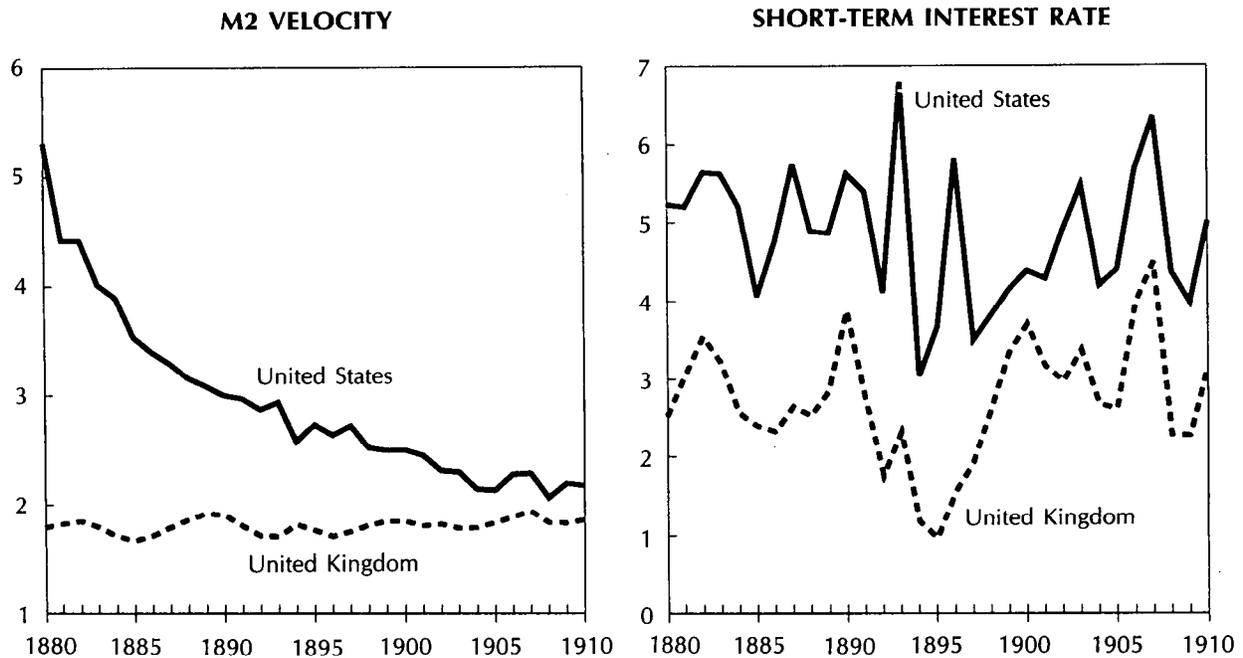
By using a dummy variable in their two-country model to capture the effects of changing levels of financial sophistication in the United States relative to Britain during the late nineteenth and early

<sup>5</sup> Friedman and Schwartz use net national product as the measure of income in computing velocity for both countries. Their interest rates are the six-month commercial paper rate for the United States and the three-month bill rate for the United Kingdom.

twentieth centuries, Friedman and Schwartz take an approach to modifying a money demand equation that mirrors the approach taken by many others to repair conventional money demand specifications for the most recent two decades: they acknowledge that in one instance financial innovation has apparently shifted the relationship between income, interest rates, and the demand for money, without considering the possibility that other episodes of instability in this relationship may have occurred and may yet occur. Friedman and Schwartz's approach is, in fact, exactly the same as that of Hafer and Hein (1982), who use period-specific dummy variables to restore stability to a money demand equation for the years following 1973.

Bordo and Jonung's hypothesis, in contrast, views all observed episodes of money demand instability as symptomatic of an ongoing process of financial evolution; indeed, their hypothesis suggests that it is not appropriate to regard money demand instability as episodic at all, but rather as a predictable and regular phenomenon. That is, the Bordo-Jonung hypothesis implies that when the demand for money equation is properly specified to include proxies for their two types of ongoing financial innovation, the equation will be stable for the 1890s, the 1980s, and all decades in between without needing period-specific variables. Thus, Bordo and Jonung's work both acknowledges and extends that of Friedman and Schwartz and those who have studied the effects of more recent financial innovations.

Figure 3



The idea that financial innovations may systematically influence velocity over long periods of time is not solely Bordo and Jonung's. In fact, they give credit to Knut Wicksell, who argues (1936, Ch. 6, Sec. C) that the velocity of currency is likely to increase as an economy's banking system develops, for inspiring their work. Irving Fisher (1963, Ch. V, Sec. 3) lists among the determinants of velocity "habits as to the use of book credit and to the use of checks," both of which have varied considerably over time. Warburton (1949), who finds that the initial downward trend in velocity seen in Figure 1 extends back to 1799, attributes the trend to changes such as the increase in the share of national output sold in organized markets, the increase in the fraction of the population working for wages instead of producing for their own consumption, and the increase due to specialization in the number of intermediate payments required in production, all of which Bordo and Jonung would classify as aspects of the monetization process. Among more recent studies, both Townsend (1987) and Goodfriend (1991) describe how improvements in communications and information-gathering technologies facilitate the substitution of privately issued securities for currency as means of payment; their analyses suggest that technological progress may simultaneously drive the process of real economic growth and allow the payments system to evolve over time. Bordo and Jonung's work is unique, however, in the extent to

which it attempts to find quantitative evidence, drawn from a variety of data sources, in support of their hypothesis.

### III. EVIDENCE OF VELOCITY'S LONG-RUN BEHAVIOR

Bordo and Jonung take four distinct approaches to document that the financial evolution that accompanies the process of real economic growth exerts an ongoing systematic influence on velocity's long-run behavior. First, the authors show that, as predicted by their theory, the U-shaped velocity pattern found in Figure 1 for the United States can be found in data from a number of other countries as well. Second, they modify the traditional regression equation expressing velocity as a function of income and interest rates by adding proxies for the two types of institutional changes identified by their hypothesis. The ratio of M2 to currency and the fraction of the labor force employed outside of the agricultural sector should both increase as part of the monetization process and therefore have a negative effect on velocity. Meanwhile, the ratio of nonbank financial assets to total financial assets should be a proxy for the rise of nonbank financial intermediaries and the development of money substitutes and therefore have a positive effect on velocity. When estimated using data extending back into the nineteenth century for Canada, Norway, Sweden, the

United Kingdom, and the United States, the regression coefficients on these three proxies all have the expected signs.

Next, Bordo and Jonung focus in detail on the monetization process as it occurred in Sweden from 1871 to 1913. They show that a number of other proxies for institutional change, including the number of commercial bank accounts per capita and the share of agricultural wages paid in cash, enter significantly into regression equations for velocity. Finally, they show that evidence of U-shaped velocity patterns can be found in cross-sectional data from 84 countries; those with low levels of income per capita have tended to experience falling velocity in the postwar years, while those with more developed economies have recently seen velocity rise over time.

In spite of their success in presenting an extensive and diverse body of evidence to support their hypothesis, Bordo and Jonung have not escaped criticism. Reviewers have found problems with their empirical work, with most of the criticism directed at the augmented velocity equations estimated for the five advanced industrialized countries. Raj and Siklos (1988), in commenting on an earlier presentation of Bordo and Jonung's results [Bordo and Jonung (1981)], warn that the significance of the institutional variables in the velocity equations may be the product of a spurious regression rather than a true economic relationship. Granger and Newbold (1986, pp. 205-16) demonstrate that standard test statistics from a regression of one random walk variable on another, independent random walk variable will often incorrectly suggest that the two are correlated. Since Raj and Siklos find that the variables in Bordo and Jonung's regressions behave like random walks, the test statistics from these regressions may be misleading. Bordo and Jonung's proxies continue to be significant when the equations are reestimated in first-differenced form, however. Because differencing serves to remove the random walk component from each variable, Raj and Siklos conclude that the test results are probably not spurious.

In more damaging reviews, Hamilton (1989) and Huizinga (1990) point out that it is difficult to defend the assumption, made implicitly by Bordo and Jonung when they use single-equation econometric methods, that financial variables such as the M2-currency ratio and the ratio of nonbank financial assets to all financial assets, used as independent variables in their model, are truly exogenous in a world in which the supply of as well as the demand for money and other assets responds to

changes in income and interest rates. Both Hamilton and Huizinga note, for example, that the ratio of M2 to currency is approximately equal to the M2 money multiplier. If, through the reserve decisions of banks, the money multiplier depends on the nominal interest rate, Hamilton (pp. 341-43) demonstrates that the M2-currency ratio may appear to be significant in Bordo and Jonung's regressions not because it is standing proxy for the effects of monetization on money demand, but because it is an important variable in determining money supply. Bordo and Jonung's parameter estimates, therefore, potentially suffer from simultaneous-equations bias.<sup>6</sup>

The simultaneity problem is a difficult one to overcome in the study of money demand, and direct attempts to do so have met with only limited success.<sup>7</sup> In fact, by repeating their analysis with a variety of data sets, Bordo and Jonung take what is perhaps the only route toward establishing that their estimates do not suffer from this problem. Indeed, Hamilton (p.343) admits as much:

When one finds, as they document, evidence of a consistent, reproducible pattern that is robust across a large number of specifications, one begins to establish a compelling scientific case that there is a predictable regularity in the correlations warranting a structural interpretation.

Likewise, Friedman and Schwartz (1991) argue that confidence in statistical results on money demand can be established only by repeating the analysis with data from as many time periods and as many countries as possible. The remainder of this paper takes Hamilton's and Friedman and Schwartz's advice: it attempts to answer the critics of Bordo and Jonung's empirical work by looking for evidence to support their financial-innovations hypothesis in a new data set.

#### IV. NEW EVIDENCE FROM U.S. REGIONAL DATA

##### A. A New Data Set

Andy Atkeson and Rachel van Elkan, both working at the University of Chicago, have recently

<sup>6</sup> Of course, the effect of interest rates on reserve decisions is just one of many potential sources of simultaneous-equations bias in Bordo and Jonung's work. For instance, Goodfriend (1991) describes how the spread of the commercial banking system and the coincident development of interbank credit markets allow banks to economize on their holdings of reserves. Thus, the supply of as well as the demand for money is intimately related to the process of monetization.

<sup>7</sup> See Goldfeld and Sichel (1990) for a review of work on the problems of simultaneity and exogeneity as they relate to the estimation of money demand functions.

compiled a data set containing figures for demand deposits by region in the United States from 1929 to 1988. Along with the Commerce Department's state personal income data, these figures may be used to construct series for demand deposit velocity by region over a 60-year period. The Commerce Department's regional definitions, used here, are given in Table 1.

As the public's currency holdings cannot be broken down geographically, demand deposit velocity is the closest analog to M1 velocity available at the regional level. Indeed, Figure 4 reveals that the U-shaped pattern found in the aggregate M1 series is shared by all eight regional demand deposit velocity series, with velocity in each region falling before World War II and rising thereafter.

### B. Empirical Strategy

Since the patterns found by region may simply be reflections of the pattern found in the aggregate, the time-series properties of the data shown in Figure 4 should not necessarily be thought of as providing new and independent evidence in support of the assertion that velocity ought to follow a U-shaped pattern as an economy develops. There is, however, considerable variation in levels of velocity across regions at any given point in time, suggesting the possibility of using the cross-sectional properties of

Table 1  
**Regional Definitions,  
U.S. Department of Commerce**

---

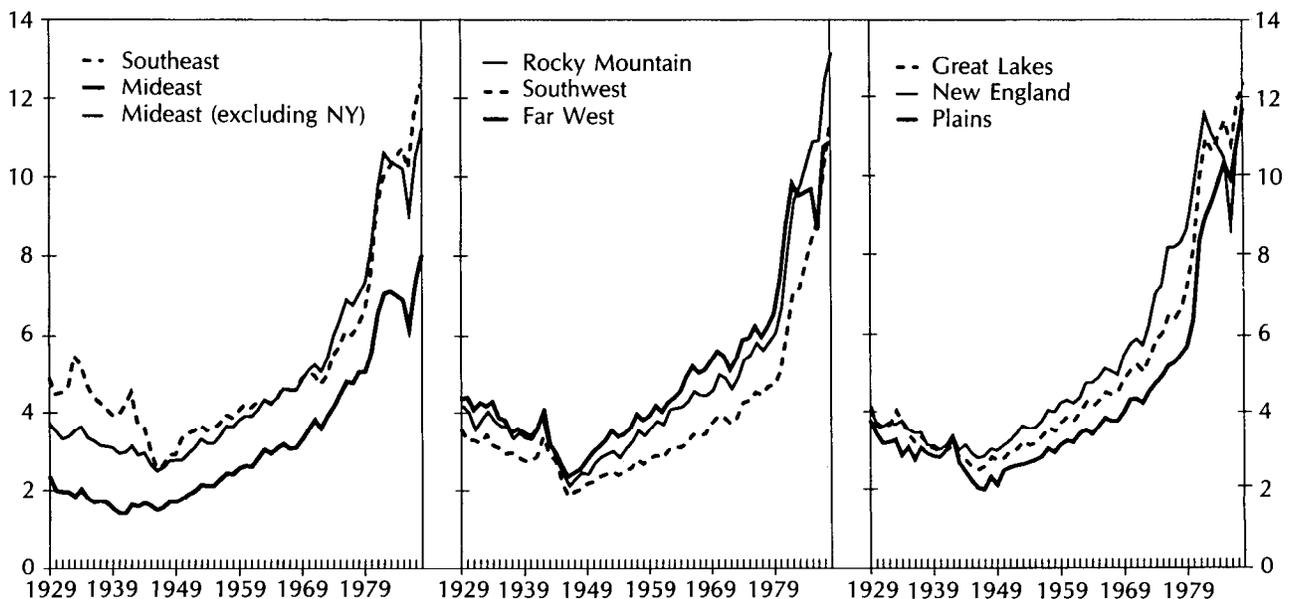
<i>New England:</i> Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
<i>Mideast:</i> Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania
<i>Great Lakes:</i> Illinois, Indiana, Michigan, Ohio, Wisconsin
<i>Plains:</i> Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
<i>Southeast:</i> Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia
<i>Southwest:</i> Arizona, New Mexico, Oklahoma, Texas
<i>Rocky Mountain:</i> Colorado, Idaho, Montana, Utah, Wyoming
<i>Far West:</i> California, Nevada, Oregon, Washington

---

the data to test Bordo and Jonung's hypothesis. Specifically, the theory predicts that for earlier years (during which the commercial banking system was still expanding geographically) a negative correlation should be observed between velocity and indexes of

Figure 4

### DEMAND DEPOSIT VELOCITY



financial sophistication across regions. For later years (with the spread of nonbank intermediaries) this correlation should turn positive.

### C. Cross Section Regression Equations

State personal income data provide two proxies for the level of financial development by region. The first, personal income per capita, is a simple proxy under the assumption that the processes of real economic growth and financial evolution are synchronized. The second, the share of total earnings originating in finance, insurance, and real estate (FIRE), the narrowest industrial class including banks and nonbank financial intermediaries for which data beginning in 1929 are available, is a more direct measure of financial sophistication. Although the share of earnings in FIRE does not distinguish between growth in banking and growth in nonbank finance, such a distinction is not necessary because changes in the proxy should primarily reflect changes in banking early on and changes in nonbank finance later.

Let  $v_{it}$  denote demand deposit velocity in region  $i$  at time  $t$ ,  $PIPC_{it}$  denote personal income per capita in region  $i$  at time  $t$ , and  $FIRE_{it}$  denote the share of total earnings originating in finance, insurance, and real estate in region  $i$  at time  $t$ . Bordo and Jonung's hypothesis predicts that  $v_{it}$  should be negatively correlated with both  $PIPC_{it}$  and  $FIRE_{it}$  for early  $t$  and positively correlated for later  $t$ . Thus, consider the cross section regression equations

$$(1) \quad v_{it} = \alpha_t + \beta_t PIPC_{it} + e_{it}$$

$$(2) \quad v_{it} = \delta_t + \gamma_t FIRE_{it} + u_{it}$$

to be estimated for each  $t$  from 1929 to 1988 (a total of 120 regressions). Bordo and Jonung's hypothesis predicts that  $\beta_t$  and  $\gamma_t$  should be negative for the early years and positive later. More generally, the coefficients should increase as functions of  $t$ .<sup>8</sup>

### D. Results: 1929-1980

Figure 4 reveals that the Mideast region inclusive of New York has considerably lower levels of velocity than the other regions. Comparing the numbers

<sup>8</sup> For the reasons given in Section IV.B, equations (1) and (2) focus only on the cross-sectional patterns appearing in the regional data. An alternative and perhaps equally informative approach would be to pool all the observations in the data set and specify a model that simultaneously tests both the cross-sectional and time-series implications of Bordo and Jonung's hypothesis. This task is left for future research efforts.

for the Mideast as a whole to those for the Mideast excluding New York State indicates that it is the New York data that make this region an outlier. In fact, if figures for New York are included in the data set used to estimate (1) and (2), they dominate the regressions, generating coefficients  $\beta_t$  and  $\gamma_t$  that are negative for all  $t$ . Including New York reveals only that as a financial center for the world, New York City has a high concentration of demand deposits, a high level of income per capita, and a large fraction of its labor force employed in finance. Thus, the data for the Mideast region exclusive of New York State, used to obtain the results discussed below, are more informative in assessing the relevance of the Bordo-Jonung hypothesis.

Equations (1) and (2) are estimated by ordinary least squares. The slope coefficients  $\beta$  and  $\gamma$  are plotted as functions of  $t$  in Figure 5 to see if they increase over time as expected. Since each coefficient is estimated using only eight observations, the standard errors are quite large. The point estimates, however, show the data from 1929 to 1980 to be consistent with Bordo and Jonung's hypothesis. The coefficients follow upward paths over the first 50 years for which data are available; for years prior to World War II they are negative, and after World War II they become positive. In fact, beginning in 1959 for  $\beta$  and in 1969 for  $\gamma$ , the slope estimates are at least one standard deviation greater than zero.

The changes that underlie the switch from negative to positive slope coefficients in the 1940s may be seen in Table 2. Between 1942 and 1946, velocity fell in every region, but the decline was less pronounced in New England, the Mideast, and the Great Lakes than in the other regions. Attributing these relative changes to the geographic expansion of the banking system is consistent with data presented in Goldsmith (1958, Ch. V), which document the spread of commercial banking from the northern United States to the South and West during 1929-1949.

### E. Results: 1981-1988

In contrast to the first 50 years of data, the most recent figures fail to display the pattern predicted by Bordo and Jonung's theory, with the slope coefficients falling over time and even becoming negative again. Driving this reversal, as may be seen in Table 3, is declining velocity in the more financially sophisticated regions: New England, the Mideast, and the Far West. The timing of the breakdown in the expected

Figure 5

**OLS COEFFICIENTS**  
(with  $\pm 1$  standard error band)

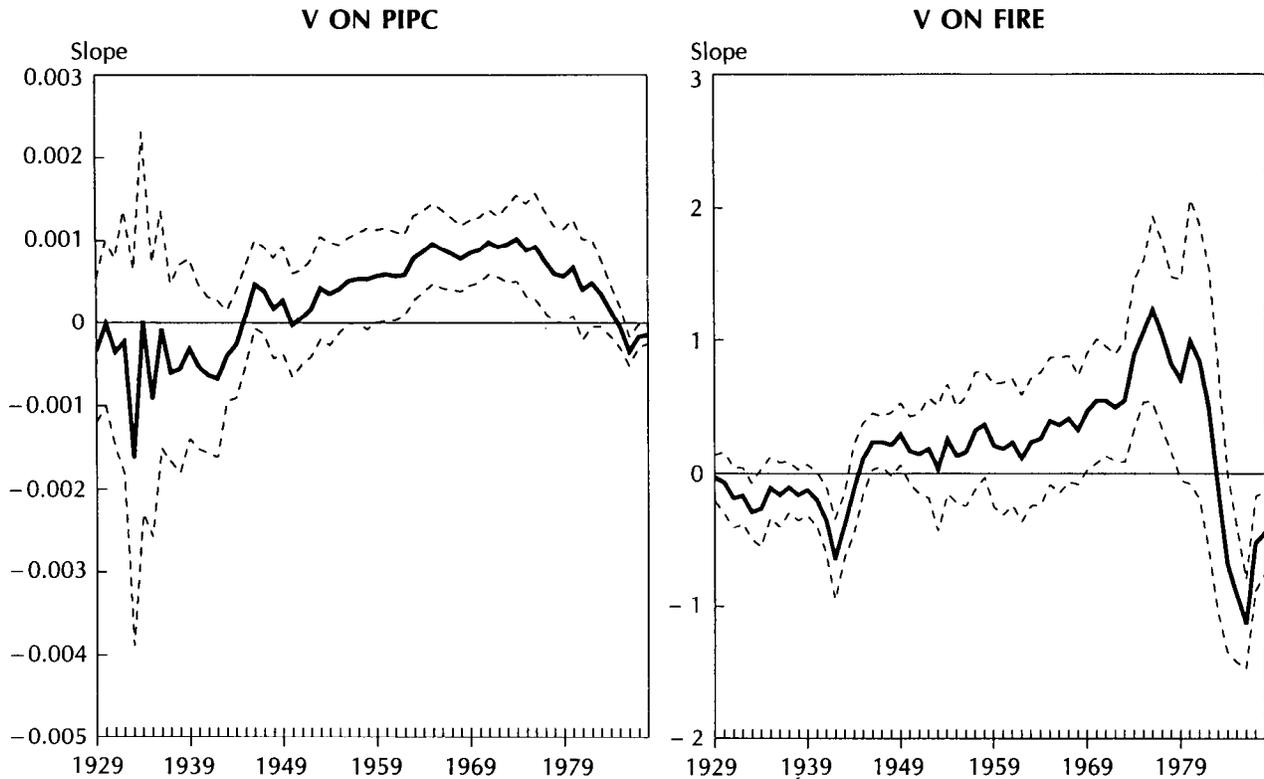


Table 2

**Changes in Demand Deposit Velocity  
1942-1946**

New England	-0.58
Mideast (excluding New York)	-0.65
Great Lakes	-0.63
Plains	-1.26
Southeast	-1.93
Southwest	-1.51
Rocky Mountain	-1.97
Far West	-1.60

Table 3

**Changes in Demand Deposit Velocity  
1981-1986**

New England	-1.90
Mideast (excluding New York)	-0.66
Great Lakes	+0.80
Plains	+1.58
Southeast	+0.97
Southwest	+2.84
Rocky Mountain	+2.73
Far West	-0.16

pattern of coefficients suggests that it may be related to the coincident break in M1 velocity's postwar trend (see Figure 1), which has received considerable attention in the money demand literature.<sup>9</sup>

<sup>9</sup> See, for example, Rasche (1987), Stone and Thornton (1987), Darby, Mascaro, and Marlow (1989), and Hetzel and Mehra (1989).

In fact, one explanation that has been offered for M1 velocity's mysterious behavior is also consistent with the surprising cross-sectional pattern to have emerged in the past decade. Stone and Thornton (1987) argue that the weakness in M1 velocity during the 1980s is most likely the result of two distinct forces. First, the nationwide introduction of NOW accounts in 1981 attracted funds

out of other interest-bearing assets; since M1 includes NOW account balances, this substitution caused its velocity to decrease. Second, if the demand for money is a function of permanent income (or wealth) rather than current income, as suggested by Friedman (1959), then velocity measured using current income will fall as permanent income increases relative to current income. In particular, expectations of improved future income following the recessions of the early 1980s can explain the decrease in measured M1 velocity.

If the availability of NOW accounts has drawn funds out of demand deposits as well as out of non-M1 assets, then demand deposit velocity should have increased even as M1 velocity fell. Table 3 reveals that, in fact, demand deposit velocity did continue to rise throughout the 1980s in regions other than New England, the Mideast, and the Far West. Interest-bearing checkable deposits were available before 1981 in New England and New Jersey, so the nationwide introduction of these accounts would not have put the same upward pressure on velocity in New England and the Mideast region (the two regions with large decreases in velocity) as it did elsewhere. Moreover, patterns in real estate prices suggest that the nationwide increase in wealth during the 1980s was concentrated in New England, the Mideast, and the Far West, putting downward pressure in measured velocity through the permanent income effect in those three regions, but not in the others. Thus, NOW accounts explain why velocity rose in some regions, while changes in permanent income explain why velocity fell in others; together, the two parts of Stone and Thornton's hypothesis explain why the regression coefficients change sign in the 1980s.

If Stone and Thornton's theory is correct, the regional regression results for the 1980s do not contradict Bordo and Jonung's hypothesis. The introduction of NOW accounts was a consequence of regulatory change rather than institutional or technological innovation, for interest-bearing checkable deposits existed prior to their prohibition in 1933. To the extent that regional patterns in velocity during the 1980s are the product of this regulatory change, the patterns say nothing about the

accuracy of Bordo and Jonung's predictions for the effects of financial evolution. Since Bordo and Jonung use permanent income as the scale variable in their velocity equations, they would also predict that velocity measured using current income would fall when permanent income increases; regional patterns induced by changes in permanent income are not evidence against Bordo and Jonung's theory either.

On the other hand, since deviations of permanent income from current income are by definition transitory, Bordo and Jonung's hypothesis implies that barring further significant regulatory change, the slope coefficients in equations (1) and (2) should soon become positive again. Thus, only if the experience of the 1980s is found, in light of future developments, to be an exceptional aberration in an otherwise unbroken pattern will the most recent data help in establishing their hypothesis as a useful guide for predicting the effects of financial innovation.

## V. CONCLUSIONS

The empirical results obtained here show Bordo and Jonung's hypothesis to be consistent with 50 years of regional demand deposit data from 1929 to 1980. Although the figures for 1981-1988 fail to fit the expected pattern, they have an explanation that is not inconsistent with the Bordo-Jonung hypothesis. Overall, therefore, the regional data can be counted as part of the large and diverse body of evidence that Hamilton (1989) and Friedman and Schwartz (1991) argue is necessary to support the claim that the correlations found between velocity and various proxies for financial sophistication reflect the structural relationship implied by Bordo and Jonung's theory.

In addition, the regional figures support Stone and Thornton's (1987) explanation for the weakness in M1 velocity during the 1980s. This finding suggests that regional data are a valuable untapped source of evidence with which competing hypotheses about the recent behavior of M1 can be tested. In particular, any theory that purports to explain the fall in velocity must also explain why this reversal in trend has been confined to the east and west coasts.

## APPENDIX: DATA SOURCES

*M1, United States, 1869-1891:* The M1 series is extended back to 1869 by multiplying M2 figures by 0.962, the ratio of M1 to M2 in 1892, the first year for which M1 data is available. The M2 data is taken from Milton Friedman and Anna J. Schwartz, *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975*. Chicago: University of Chicago Press, 1982. Table 4.8, Column 1. *1892-1914:* U.S. Bureau of the Census, *Historical Statistics of the United States, Colonial Times to 1957*. Washington, D.C., 1960. Series X-267. Adjusted upward to link with later data. *1915-1970:* U.S. Bureau of the Census, *Historical Statistics of the United States, Colonial Times to 1970*. Washington, D.C., 1975. Series X-414. *1971-1989:* *Economic Report of the President*. Washington, D.C., 1991. Table C-67.

*Gross National Product, United States, 1869-1928:* Nathan S. Balke and Robert S. Gordon, "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy* 97 (February 1989): 38-92. Table 10. *1929, 1933, 1939-1989:* *Economic Report of the President* (1991). Table B-1. *1930-1932, 1934-1938:* U.S. Department of Commerce, *The National Income and Product Accounts of the United States, 1929-1982: Statistical Tables*. Washington, D.C., September 1986. Table 1.1.

*Six-Month Commercial Paper Rate, United States, 1869-1975:* Friedman and Schwartz (1982). Table 4.8, Column 6. *1976-1989:* *Economic Report of the President* (1991). Table C-71.

*M2 Velocity, United States, 1880-1910:* Friedman and Schwartz (1982). Table 4.8, Column 1 (M2) divided by Column 2 (Net National Product).

*M2 Velocity, United Kingdom, 1880-1910:* Friedman and Schwartz (1982). Table 4.9, Column 1 (M2) divided by Column 2 (Net National Product).

*Three-Month Bill Rate, United Kingdom, 1880-1910:* Friedman and Schwartz (1982). Table 4.9, Column 6.

*Demand Deposits by Region, United States, 1929-1988:* Unpublished worksheets compiled by Andy Atkeson and Rachel van Elkan. For 1929-1949, their data are figures for total demand deposits less interbank and federal government demand deposits at all banks, taken from Board of Governors of the Federal Reserve System, *All Bank Statistics, United States, 1869-1955*. April 1959. For 1950-1968, the data are figures for business and personal demand deposits at all banks, as reported in various issues of Federal Deposit Insurance Corporation, *Assets and Liabilities and Capital Accounts—Commercial and Mutual Savings Banks*. For 1969-1977, the data are figures for demand deposits at all insured commercial banks, as reported in various issues of *Assets and Liabilities and Capital Accounts—Commercial and Mutual Savings Banks*. For 1978-1988, the data are unpublished figures for demand deposits at all insured commercial banks, obtained directly from the FDIC.

*Personal Income by Region, United States, 1929-1982:* U.S. Department of Commerce, *State Personal Income: 1929-1982*. Washington, D.C., February 1984. *1983-1988:* U.S. Department of Commerce, *Survey of Current Business*. August issues, 1986-1990.

*Personal Income per Capita by Region, United States, 1929-1982:* *State Personal Income, 1929-1982* (1984). *1983-1988:* *Survey of Current Business*. August issues, 1986-1990.

*Earnings by Region and by Industry, United States, 1929-1982:* *State Personal Income, 1929-1982* (1984). *1983-1988:* *Survey of Current Business*. August issues, 1986-1990.

## REFERENCES

- Bordo, Michael D. and Lars Jonung. "The Long-Run Behavior of the Income Velocity of Money in Five Advanced Countries, 1870-1975: An Institutional Approach." *Economic Inquiry* 19 (January 1981): 537-45.
- . *The Long-Run Behavior of the Velocity of Circulation: The International Evidence*. Cambridge: Cambridge University Press, 1987.
- . "The Long-Run Behavior of Velocity: The Institutional Approach Revisited." *Journal of Policy Modeling* 12 (Summer 1990): 165-97.
- Darby, Michael R., Angelo R. Mascaro, and Michael L. Marlow. "The Empirical Reliability of Monetary Aggregates as Indicators: 1983-1987." *Economic Inquiry* 27 (October 1989): 555-85.
- Enzler, Jared, Lewis Johnson, and John Paulus. "Some Problems of Money Demand." *Brookings Papers on Economic Activity* (1976): 261-80.
- Fisher, Irving. *The Purchasing Power of Money*. New York: Augustus M. Kelley, 1963.
- Friedman, Milton. "The Demand for Money: Some Theoretical and Empirical Results." *Journal of Political Economy* 67 (August 1959): 327-51.
- Friedman, Milton and Anna J. Schwartz. *A Monetary History of the United States, 1867-1960*. Princeton: Princeton University Press, 1963.
- . *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975*. Chicago: University of Chicago Press, 1982.
- . "Alternative Approaches to Analyzing Economic Data." *American Economic Review* 81 (March 1991): 39-49.
- Goldfeld, Stephen M. "The Case of the Missing Money." *Brookings Papers on Economic Activity* (1976): 683-730.
- Goldfeld, Stephen M. and Daniel E. Sichel. "The Demand for Money" in Benjamin M. Friedman and Frank H. Hahn (eds.), *Handbook of Monetary Economics*. Amsterdam: North-Holland, 1990.
- Goldsmith, Raymond W. *Financial Intermediaries in the American Economy Since 1900*. Princeton: Princeton University Press, 1958.
- Goodfriend, Marvin. "Money, Credit, Banking, and Payments System Policy." Federal Reserve Bank of Richmond *Economic Review* 77 (January/February 1991): 7-23.
- Granger, C. W. J. and Paul Newbold. *Forecasting Economic Time Series*, 2d ed. Orlando: Academic Press, 1986.
- Hafer, R. W. and Scott E. Hein. "The Shift in Money Demand: What Really Happened?" Federal Reserve Bank of St. Louis *Review* 64 (February 1982): 11-16.
- Hamilton, James D. "The Long-Run Behavior of the Velocity of Circulation: A Review Essay." *Journal of Monetary Economics* 23 (March 1989): 335-44.
- Hetzl, Robert L. and Yash P. Mehra. "The Behavior of Money Demand in the 1980s." *Journal of Money, Credit, and Banking* 21 (November 1989): 455-63.
- Huizinga, John. "Comments on 'The Long-Run Behavior of Velocity: The Institutional Approach Revisited.'" *Journal of Policy Modeling* 12 (Summer 1990): 199-204.
- Judd, John P. and John L. Scadding. "The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature." *Journal of Economic Literature* 20 (September 1982): 993-1023.
- Latané, Henry A. "Income Velocity and Interest Rates: A Pragmatic Approach." *Review of Economics and Statistics* 42 (November 1960): 445-49.
- Lucas, Robert E., Jr. "Money Demand in the United States: A Quantitative Review." *Carnegie-Rochester Conference Series on Public Policy* 29 (Autumn 1988): 137-68.
- Meltzer, Allan H. "The Demand for Money: The Evidence from the Time Series." *Journal of Political Economy* 71 (June 1963): 219-46.
- Rasche, Robert H. "M1-Velocity and Money-Demand Functions: Do Stable Relationships Exist?" *Carnegie-Rochester Conference Series on Public Policy* 27 (Autumn 1987): 9-88.
- Raj, Baldev and Pierre L. Siklos. "Some Qualms About the Test of the Institutional Hypothesis of the Long-Run Behavior of Velocity." *Economic Inquiry* 26 (July 1988): 537-45.
- Stone, Courtenay C. and Daniel L. Thornton. "Solving the 1980s' Velocity Puzzle: A Progress Report." Federal Reserve Bank of St. Louis *Review* 69 (August/September 1987): 5-23.
- Townsend, Robert M. "Economic Organization with Limited Communication." *American Economic Review* 77 (December 1987): 954-70.
- Warburton, Clark. "The Secular Trend in Monetary Velocity." *Quarterly Journal of Economics* 63 (February 1949): 68-91.
- Wicksell, Knut. *Interest and Prices*. London: Macmillan, 1936.