

Eugene N. White is a professor of economics at Rutgers University. The author thanks Michael Bordo, Per Hansen, Naomi Lamoreaux, David Weil, David Wheelock, and the other participants in the Federal Reserve Bank of St. Louis' Twenty-Second Annual Economic Policy Conference for comments.



Were Banks Special Intermediaries in Late Nineteenth Century America?

Eugene N. White

The financial crises and vastly increased competition of the last two decades have radically reshaped the American financial system. One key feature of this transformation has been the declining importance of banks' traditional activities. Weakened by crises and regulatory disadvantages, banks' share of intermediation has shrunk while the shares of other financial intermediaries and markets have expanded. The shrinking banking sector has raised concerns because banks are important "special" lenders to small firms and other borrowers, they operate the payments system and provide liquidity, and monetary policy is carried out by altering their balance sheets. To put in historical perspective the issue of banks' declining role in lending, this article examines the nature of bank lending in the late nineteenth century and why banks remained the dominant intermediaries, even when disadvantaged by regulation and challenged by competitors.

In banking history, the late nineteenth century is termed the National Banking Era. Beginning in 1864 with the passage of the National Banking Act and ending with the founding of the Federal Reserve System in 1913, the National Banking Era was a period of rapid economic growth and price stability. Growth was accompa-

nied by the spread of financial intermediation and innovation. Given the virtual prohibition of branch banking and low capital requirements, the demand for banking services drove up the number of commercial banks that were chartered under the National Banking Act and state laws from 467 in 1864 to 21,478 in 1913. Commercial banks' portfolios were shaped by regulations that prohibited investment in equities, limited mortgages, and encouraged short-term loans. Their liabilities were predominantly demand deposits, and although there had been experiments with insurance of bank liabilities before the Civil War, there was no insurance until very late in the period, when seven states created deposit guarantee funds after the Panic of 1907 (White 1983; Calomiris 1993).

As they do today, commercial banks felt competitive pressures from other regulated financial markets and intermediaries, including trust companies, investment banks, insurance companies, and thrifts. Combining deposit and loan banking with other financial activities, trust companies competed vigorously in the Northeast and Midwest. Commercial banks could not easily meet the demand for longer-term finance by the newly emergent modern corporations. Instead, investment banks created the large bond and equity markets to finance big business. These new financial instruments were absorbed by life insurance companies, often allied with investment banks, which had a steadily rising flow of policy premiums to invest. Banks also faced competition from the money markets. Improvements in transportation and communications enabled commercial paper houses to intrude on banks' territory, offering access to a national market for short-term credit. Mutual savings banks catered to small depositors and the mortgage market, although mortgage companies and savings and loan associations became increasingly important competitors late in the century.

Table 1

Financial Intermediaries' Shares of Assets (percent)

	Commercial Banks	Mutual Savings Banks	Savings and Loan Associations	Life Insurance Companies	All Other Intermediaries
1880	62.6	22.6	1.3	10.5	3.1
1900	62.9	15.1	3.1	10.7	8.2
1922	63.2	8.8	3.7	11.6	12.7
1950	50.8	7.6	5.7	21.1	14.8
1990	27.0	2.1	8.9	11.1	50.9

SOURCES: Goldsmith (1958), U.S. Dept. of Commerce (1975), U.S. Comptroller of the Currency (1975), Snowden (1987), *Federal Reserve Bulletin* (1991).

In spite of these fast-growing challengers, commercial banks retained their preeminent position in the National Banking Era. Table 1 reports the shares of all financial intermediaries' assets. Although it is difficult to reconstruct a complete picture of the financial system before 1900, the table demonstrates that commercial banks retained their dominant position among intermediaries well into the early twentieth century. There was little change between 1880 and 1922, when commercial banks steadily held approximately 63 percent of assets. The twentieth-century decline is evident in 1950; by 1990, commercial banks held only 27 percent of all financial intermediaries' assets.¹ The sources of this recent decline have been studied intensively (Boyd and Gertler 1993; Wheelock 1993; and Berger, Kashyap, and Scalise 1995). Banks' commercial and industrial lending, a "special" function of commercial banks, has been at the center of this contraction. As a share of all short-term debt of nonfinancial corporations, banks' commercial and industrial loans fell from more than 80 percent in 1970 to 60 percent by the early 1990s (Wheelock 1993). Banks have lost ground in lending to both nonbank intermediaries and markets. Finance company loans have supplanted bank loans, while offshore bank loans, not subject to reserve requirements, have competed with domestic banks and even domestic offices of foreign banks to grab a bigger share of commercial and industrial lending (Boyd and Gertler

1993). Instead of commercial and industrial loans, many corporations with good credit histories have found it cheaper to borrow on the commercial paper market.

However, banks' traditional lending operations have declined more than their total operations (Berger, Kashyap, and Scalise 1995). Banks have survived and prospered by moving some traditional business off their balance sheets. They have unbundled traditional functions in intermediation by offering loan commitments and standby letters of credit, and by selling and securitizing loans.² Banks remain on the scene in the commercial paper market by providing borrowers with standby lines of credit. Boyd and Gertler (1993) show that off-balance sheet items, in terms of credit equivalents, are roughly equal to half of banks' commercial and industrial lending. Berger, Kashyap, and Scalise (1995) conclude that while banks have lost considerable business to foreign banks, nonbanks, and markets, their share of intermediation has not shrunk as much as measured by traditional activities on their balance sheets.

Nevertheless, even with these qualifications, banks at the end of the twentieth century are no longer the preeminent financial institutions that they were at the beginning. Recent theoretical work argues that this smaller role for commercial banks and depository institutions, in general, is a predictable development. According to Diamond (1997), banks will be key

¹ Banks' share of intermediation declined even if one takes a broad definition of banking to include mutual savings banks and savings and loan associations. By this measure, banks' share of intermediation falls from 87 percent in 1880 or 81 percent in 1900 to 64 percent by 1950 and 38 percent in 1990.

² Noting that bank income from off-balance-sheet activities rose from 20 percent of total income in 1979 to 33 percent in 1991, Boyd and Gertler (1993) argue banking has not shrunk as much as would be indicated by commercial banks' share of assets. However, rising income from off-balance-sheet activities is not new. In the 1920s, this income rose from 9 percent to 14 percent of bank income (White 1984). Unfortunately, there is not sufficient data to make long-term comparisons of the relative shares of intermediaries by alternative measures.

providers of liquidity and allocators of capital when there is limited participation in markets in the early stages of economic development. More liquidity is created by banks' offerings of demand deposits backed by long-term assets, the price of which is raised by the expansion of the banking sector. The eventual growth of markets increases the use of long-term debt and equity. More participation in markets induces banking sector shrinkage, and bank holdings of long-term assets are reduced relative to short-term assets. Diamond's analysis emphasizes the role of banks as providers of liquidity and intermediation rather than as firms that solve a problem of asymmetric information by specializing in the evaluation and monitoring of high-risk, low-information borrowers. The special informational advantages of banks is, instead, the focus of contemporary theoretical analysis of banking lending, and the more central role of commercial banks in the nineteenth century suggests that these advantages loomed even larger during the National Banking Era. However, banking theory in the nineteenth century was concerned with very different issues and had very strong prescriptions for lending that do not accord with the modern literature.

To begin, this article reviews both contemporary and late nineteenth-century banking theory. The limited available published data, complemented by two case studies, provide some empirical evidence on the special character of banking. The preeminence of banking during the National Banking Era highlights the key attributes of commercial bank lending, while the growing competition from markets and other intermediaries reveals the origins of the change in banking's relative importance in the twentieth century.

BANKS AS SPECIAL INTERMEDIARIES IN MODERN THEORY

The growing similarity among financial intermediaries makes banks, defined by their functions of taking

deposits and making loans, seem less special to the operation of the financial system. In recent theoretical literature, banks are considered special not because of the functions they perform but because they overcome important informational asymmetries.³ According to this literature (Bhattacharya and Thakor 1993), the varied forms of intermediation are responses to various informational problems that prevent markets managed by brokers from efficiently selling borrowers' liabilities to savers. The special character of banks is best understood by comparing them to their closest market competition, money market mutual funds (MMMFs). Both banks and MMMFs provide transaction services and increase divisibility and diversification for portfolios of large-denomination assets. But banks are viewed by theorists as different because they are delegated by depositors to monitor borrowers. By monitoring borrowers through their transaction activities and by covenant enforcement, banks obtain continuous information on their customers' creditworthiness before and after the creation of a new loan.

Banks invest in the acquisition of this information to serve various types of borrowers. Bank loans may be useful to borrowers who are relatively poor credit risks and for whom information is relatively volatile (Berlin and Mester 1992). Relatively new borrowers, without well-established reputations, may also gain more from bank monitoring and choose bank loans instead of the capital market (Diamond 1991). Although small firms may find that banks are some of the few sources of credit, the higher interest they must pay for a bank loan may indicate that the cost of such loans is less than the cost of contracting for marketable debt. For large firms with access to the market, periodic signals from short-term bank loans are a useful signal to the market (Fama 1985).⁴ Although banks do sell loans and securitize bundles of relatively homogeneous collateralized loans (automobile loans and mortgages), loans are often difficult to market because outsiders without access to banks' proprietary information find it difficult to value

³ Fama (1985) first suggested that banks had this special character, examining why they finance loans with both demand deposits and certificates of deposit (CDs). Banks are at a disadvantage vis-a-vis other lenders because reserve requirements are an implicit tax on their liabilities. Banks can compete by using demand deposits, whose transaction services allow them to pay lower interest. CDs do not provide special transaction services and must pay the same interest as commercial paper and bankers' acceptances. The viability of CDs implies that borrowers regard bank's loans as special and willingly pay a higher rate of interest.

⁴ Some empirical evidence for the special character of bank loans has been found by James (1987) and Lummer and McConnell (1989), who discovered abnormal positive returns on the stock of firms signing or renegotiating credit agreements.

loans originated by banks. Many loans are thus not marketed and, instead, are held until maturity, with the banks bearing the residual risk from nonperformance. It may be difficult to sell loans without recourse, because resale leaves the originating banks with no incentive to produce the information needed to monitor the borrower. Banks have a strong incentive to monitor borrowers continuously because bank loans are usually last in line of debt seniority. Thus, renewal of a bank loan credibly signals other, more senior creditors of a firm that they do not need to invest in a costly analysis of the borrower.

Another key feature of banks that is highlighted by this theoretical literature is the fact that loans are financed with shorter-maturity liabilities (Thakor 1992). This maturity transformation requires banks to bear interest-rate risk, for which they are rewarded. A positive term premium in the yield curve gives banks an incentive to engage in a maturity mismatch. The greater the mismatch, the higher the return and volatility on a bank's equity (Deshmukh, Greenbaum, and Kanatas 1983). However, a maturity mismatch also imposes market discipline on the bank, which induces it to screen and monitor loans, as deposits may be withdrawn faster than loans are paid off. This maturity mismatch makes banks prone to panics in a system without deposit insurance.

The threat of a panic arises because, as delegated monitors, the banks themselves need to be monitored. For transaction services, depositors need a very low-risk asset. Demandable deposits, secured by a diversified portfolio of loans, are such an asset. But security is guaranteed only when depositors can discipline the bank managers by quick redemption of their deposits (Calomiris and Kahn 1991). When depositors believe that their bank's risk has increased, they can withdraw their deposits or refuse to roll over their short-maturity CDs. However, given that loans are difficult for outsiders to value, a change in the economic environment may cause depositors to panic. Some banks may indeed be in trouble, but the inability to

value the portfolio of all banks correctly may lead to contagion in which a run on weak banks spreads to strong banks.

In the current literature, banks perform a central role by lending on nonmarketable information they have produced and providing continuous monitoring of borrowers. Funded by liabilities that create a maturity mismatch, banks are, in turn, monitored by their depositors.

THE REAL-BILLS DOCTRINE IN THEORY

Nineteenth-century banking theorists would have found this contemporary description of banks' role in the financial system somewhat puzzling. Banks were regarded as very special intermediaries in the nineteenth century, but for different reasons than we think of today. There was less competition from other intermediaries and markets, and hence there was less concern about the special character of bank loans. Instead, banking theorists were more worried about the safety and liquidity of loans as they affected the banks' ability to pay their depositors on demand.

Virtually all students of banking in the nineteenth century paid homage to the "real bills" doctrine. According to this theory, banks should offer only short-term loans to finance the production or shipment of goods. The sale of goods would then be used to pay off the loans. These loans were considered to be safe because they financed real short-term commercial transactions.⁵ Warning against borrowing to pay off existing debts or make speculative investments, Homans (1857, p. 32) praised the safety of real bills: "When money is to be invested in the purchase of merchandise, cattle, flour, or other property in the regular course of the borrower's business, the investment yields to the borrower a means of repayment; nothing is hazarded by ordinary integrity, and ordinary exemption from disasters." Proponents of the real-bills doctrine believed that if banks followed its prescriptions, the quantity of loans and liabilities

⁵ A bill of exchange to finance the shipment of goods was thus an example of a real bill.

would be limited to the legitimate needs of business, and banks would remain liquid. In this era, liquid loans were simply loans that were paid off at maturity. Liquidity meant that an asset was automatically paid off at maturity, not necessarily that it was easy to market. The more modern idea of liquidity and the idea of holding a diversified portfolio of readily marketable assets for a secondary reserve did not gain wide acceptance until the beginning of the Federal Reserve period (James 1978).

Proponents of the real-bills doctrine offered blunt prescriptions to bankers. In 1876, the Comptroller of the Currency addressed the American Bankers' Association: "As banks are commercial institutions, created for commercial purposes, preferences in discounts should always be given to paper based upon actual commercial transactions. Banks are not loan offices. It is not part of their business to furnish their customers with capital..." (Bolles 1890, p. 70). The Comptroller was emphatic that all paper should be paid off at maturity, enabling banks to meet withdrawals with funds from maturing loans.

As late as 1915, Kniffin (1915, p. 209) wrote in a standard text on banking that "the secret of sound banking is to have a steady stream of money coming in by way of maturing loans, so that the constant stream of obligations falling due daily by reason of the demands of the checking depositors may be met. A demand obligation cannot be met by a time security and only as a bank keeps its funds liquid—that is, flowing in and out—can it meet every demand made on it without hardship."

The strict prescriptions that real-bills advocates proffered to commercial banks raise the question of how such advocates would manage savings banks, which by their very design held much longer-term assets. Although funded by savings deposits, savings banks did experience runs in the late nineteenth century and could not rely on a rapid payoff of loans to meet their customers' demands.⁶ Few writers of the period addressed this

problem. Bolles (1888, pp. 208-13) was an exception, but his analysis was somewhat strained, given real-bills strictures. He admitted that mortgages were "less readily convertible than some other securities" but argued that "no property is more stable in value, and none less likely to depreciate, than real estate." Bolles wrote, "These institutions represent the industry and frugality of the masses, and every effort should be made to put them on the soundest footing." The best way to retain confidence was to invest "savings deposits in mortgages properly secured on the farms, the shops and the homes of the people. If these are not real values, what are?" Recognizing that savings banks' advantage probably lay in local real estate, he warned against lending on real estate out of the region because of the difficulty in ascertaining the mortgages' underlying security. Like other theorists of the period, he opposed permitting large deposits in savings banks. The organizers of mutual savings banks had established rules to keep wealthy individuals from making deposits on the grounds that these had been established primarily to promote thrift among the poorer classes. Bolles, on the other hand, objected to participation by large depositors because he believed that they would be likely to withdraw their funds in a crisis. Limit deposits to small savers and loans to local real estate, said Bolles, and savings banks would be safe institutions.

The real-bills prescription for lending required that loans be short term. The implication was that banks should minimize the maturity mismatch of assets and liabilities. While this would reduce the earnings that a bank might obtain from the term premium as a result of a mismatch, it would supposedly increase a bank's ability to meet a run on the bank, satisfying its customers with the proceeds of loans that were being rapidly paid off. Thus, even though panics were frequent in the late nineteenth century, depositors did not play as large a disciplinary role in monitoring banks as envisioned by contemporary theory.

⁶ Savings banks did have one advantage in a crisis. While commercial banks were required to make payment on demand, savings banks had the right under state law to restrict payment under certain conditions.

THE REAL-BILLS THEORY IN PRACTICE

While banking theorists offered very strict prescriptions to banks in the nineteenth century, banking practice diverged from theory even as bankers believed they held to the theory in spirit. The following paragraphs explore some of these discrepancies.

The Problem of Lending

The real-bills theoreticians favored the use of two-name paper, but bankers began to employ other financial instruments. Before the Civil War, commercial transactions were usually financed by a trade acceptance, a two-name bill of exchange that provided recourse to the acceptor or endorser of the bill in case of default. In *The Banker's Common-Place Book* (1857), Homans advised his readers to accept only notes endorsed by men of wealth and good reputation and stated, "Banks... never regularly lend money without receiving the security of more than one person who is deemed safe for the debt; and a good banker will err on the side of excessive security, rather than accept security whose sufficiency may reasonably be questioned." In his standard text on banking, Bolles (1888, pp. 52-53) explained why two-name paper was essential for making banking "a very safe and easy business":

I should say that the first and most important function of a bank is, by the use of the capital which it controls, to bridge over the periods of credit which necessarily intervene between production and consumption, in such a manner as to give back to each producer, or middleman, as quickly as possible, the capital invested by him in such products, in order that he may use it over again in new production or new purchases. ... Thus defined, banking is not only one of the most useful; but it is also one of the most safe and healthy of business operations. Its safety lies in the fact that every loan of

the character described, is based on property of intrinsic value. ... The several makers of the paper, though debtors in form, are only insurers, or guarantors, in fact. They pledge their respective property to the payment of the loans; but the primary and generally sufficient pledge is the property for which the notes are given. The wealth of the makers is a necessary margin or guaranty, because the property sold may be destroyed or the value may fall.

An ideal real bill, like a trade acceptance, was secured by a real transaction, endorsed by a respectable, wealthy individual, and was short term. Evaluating the quality of this form of lending relied as much on an evaluation of the endorser as it did on the issuer of the note and the safety of the underlying transaction.

The crucial difference between contemporary banking theory's positive description of banking and the real bills' normative description is that they are predicated on banks' specializing in the collection of different types of information. According to contemporary theory, a modern bank collects financial and transaction information from its customers to assess their creditworthiness. In contrast, nineteenth-century banks were told that in addition to verifying that bills represented bona fide transactions, they needed to monitor and collect information on the endorsers of bills. Bolles (1888, pp. 97-99) described the process whereby a bank would decide every day what loanable funds were available and examine the bills offered. Some makers and endorsers were better known, and a bank would select the most desirable offerings and decline the remainder. In effect, the tasks of judging the creditworthiness of the final borrower and monitoring his performance were delegated partly to the endorsers of the bills, to whom the bank had recourse.

In spite of the admonitions of real-bills advocates, markets in the nineteenth century moved away from two-name paper. After the Civil War, the single-name

Table 2

**Composition of Bank Loan Portfolios
(percentage of total loans)**

	Total Loans (\$ millions)	Demand, Unsecured by Collateral	Demand, Secured by Collateral	Time, Two-Name Unsecured	Time, One-Name Unsecured	Time, Secured by Collateral	Mortgages or Secured by Mortgages & Other
National banks 1895	2,042	5.1	13.9	46.9	15.6	15.6	0
National banks 1910	5,455	9.7	17.2	33.2	19.1	20.4	0.5
Nonnational banks 1910	7,066	3.8	13.5	14.3	7.5	15.8	45.2

SOURCE: U.S. Comptroller of the Currency, annual reports, 1895, 1910.

unsecured promissory note—commercial paper—became the leading short-term instrument for farmers and merchants. This instrument was criticized for its lack of adequate security in the form of collateral or personal guarantees from one or more endorsers, and it required a more modern direct evaluation and monitoring of the customer. This increasingly popular instrument could be taken directly by a commercial bank or handled by a commercial paper house. Data on the types of loans and their characteristics during the National Banking Era are very scarce. Greef (1938, p. 68) reports one estimate that single-name paper constituted 75 percent of the market by 1894. Myers (1931, pp. 322-23) calculated that the ratio of two-name paper to total loans and discounts fell from 50 percent in 1886 for New York banks to 20 percent in 1900, with the proportion of single-name paper rising from 10 to 20 percent. For country banks, these ratios fell from 50 percent to 33 percent for two-name paper and rose from 10 percent to 20 percent for single-name paper. According to Myers (1931, p. 136), only 3 percent of all domestic credit transactions were financed by trade acceptances at the end of the century. Some of the limited data from the Comptroller of the Currency's annual reports is presented in Table 2. For national banks, the share

of two-name unsecured paper fell from 47 percent to 33 percent of all loans and discounts between 1895 and 1910, while unsecured time and demand loans both rose. The amount of two-name paper held by nonnational banks—all state banks, savings banks, loan and trust companies, and private banks—was already a low 14 percent by 1910.

Discounts were unsecured loans made on the general credit of the borrower. Loans were usually secured by a pledge of collateral, including stocks, bonds, receivables, merchandise, or real estate. With the decline in two-name paper, collateralized loans were of increasing importance in the nineteenth century, as collateral provided an alternative to the guarantee of an endorser. The Philadelphia National Bank, for example, ventured into granting loans on warehouse receipts, which proved to be a profitable line of business.⁷ Considered poor collateral for commercial banks because it was illiquid, real estate did not fit the real-bills doctrine, and mortgage loans on real estate were prohibited to national banks until 1913. The share of both demand and time loans secured by collateral in national banks' loan portfolios rose between 1895 and 1910 (see Table 2). The largest item of secured lending for national banks was, of course, mortgages. Mortgage lending was dominated by

⁷ Although backed by collateral involved in a real transaction, as prescribed by real-bills theory, these loans were not completely safe. In 1888, the bank found itself the owner of a warehouse full of overvalued prunes. Wainwright (1953), p. 155.

savings banks, and mortgages constituted the largest item in their portfolios.

Given the hold of the real-bills doctrine, with its emphasis on evaluating specific transactions and the quality of endorsers, there was a slow development of alternative methods of evaluating loans and the quality of a bank's portfolio. Bankers had supplemented their own knowledge of business borrowers by subscribing to reports of credit agencies like R.G. Dun and Company. Yet, these reports were usually based on estimates of a firm's worth and reports from lawyers and other business people about the character of its proprietors, not financial statements (Lamoreaux 1994). Some banks required borrowers to maintain compensating balances on deposit to gain additional information.⁸ Contemporary writers also urged banks to discover what constituted a firm's fixed assets and its quick or convertible assets, recommending that loans could be granted if a borrower's liabilities did not exceed 50 percent of his quick assets. The maximum recommended term was six months. The result would be short-term, self-liquidating loans.

Offering advice on how to judge a potential borrower, Moulton (1918, p. 655) wrote the following:

The amount that may be safely loaned... can be ascertained only from an intimate personal acquaintanceship with the borrower and his business or from a study of a balance sheet or financial statement setting forth the condition of the business. The growing impersonality of modern business in the larger centers and the growing size and complexity of business enterprise has more and more necessitated the use of the balance sheet as a basis of credit extension.

He believed that it was in the late 1870s that financial statements were first used in procuring loans, but it was not until the 1890s that their use became common even in large banks. Few small suburban banks or country banks used them.

The intimate and often qualitative knowledge of local clients possessed by a banker was not easily replaced by financial statements in the late nineteenth century. Perhaps the most important reason for the failure to use financial statements was that there were no uniform accounting standards for business. This feature of business practice added to the asymmetry of information that gave banks their special role. For all nineteenth-century firms, accountants had little authority to impose standardized accounting practices on clients, and there were few statutory requirements governing accounting behavior (Brief 1966). Without uniform accounting methods, there was no ready alternative evaluation method to a banker's qualitative judgment of his customer.

The longstanding problem posed by an absence of accounting standards came into sharp focus when the Federal Reserve Board wanted to guarantee the quality of member banks' paper eligible for discount by using an objective analysis of financial condition instead of subjective judgments about the borrower's character (Miranti 1986). To ensure that lending officers analyzed reliable data when granting credit, the Board wanted borrowers' statements to be certified by public accountants, and it issued Circular No. 13 in 1914 to set down the rules. When notes were offered for rediscount, they were to be accompanied by a statement setting forth the condition of the borrowers and stating that the funds were used for the purpose of financing current transactions, not fixed capital or permanent working capital. The member bank offering a note for rediscount was required to have these statements on file and to certify that they were in compliance with the circular (Willis 1923). Member banks quickly realized that, under these rules, they would be able to discount very little of their paper. They protested to the Fed that the average business, and especially the average farmer, could not furnish the appropriate type of statement, the demand for which would be seen as insulting and burdensome.

A special committee was convened by the Board to reconsider Circular No. 13.

⁸ Compensating balances could be used to raise interest rates above usury rates, but this tactic was less important in the late nineteenth century, when market rates had declined well below usury rates in most states. Lamoreaux (1994), pp. 101.

The committee offered the following observation:

We believe that the country banks which constitute the majority of our members are generally without credit files as known to the large city bank. Borrowers are personally known by the officers and directors who are usually their neighbors, and the means, business and character of such borrowers are matters of intimate personal knowledge to the bank officer. To bring about a uniform understanding among country bankers as to what is and what is not eligible paper within a narrow or even technically exact interpretation of the [Federal Reserve] Act will take a long time and a still longer time will be necessary to arrange for the filing of financial statements by borrowing customers of country banks (Willis 1923, p. 914).

The impossibility of imposing financial and accounting standards on member banks and their customers led the Fed to back down and issue a new circular in 1915 that lowered requirements for eligible paper. Most importantly, this circular waived regulation for loans below \$2,500, which exempted most country bank paper. The Federal Reserve would now discount bank paper, but it was not easily marketable. As this episode demonstrates, the absence of generally accepted standard accounting practices in the nineteenth century made independent loan evaluation difficult, thus augmenting the asymmetry of information and ensuring the special position of banks, which could observe lenders firsthand.

For banks, sophisticated borrower evaluation—beyond reliance on personal knowledge of local business—required specialization. However, the banks of this period were predominantly small with very modest staffs and limited management structures. Larger banks employed a cashier who headed daily operations, several tellers and clerks, and perhaps a bookkeeper. Smaller banks might have only a cashier. The largest bank in the

major financial center of Philadelphia, the Philadelphia National Bank, had 34 men on its payroll in 1879, including one assistant cashier to help manage its expanding operations (Wainwright 1953). Management was in the hands of the directors, one of whom was selected as president. The directors verified the cashier's accounts and decided how much to lend and to whom (Lamoreaux 1994). The directors thus had no staff to draw up detailed reports on customers and instead relied on their local knowledge of business and their customers' "character." The growth of business and the shift to lending outside the community created a need to professionalize bank management. Writers advocated that a professional bank staff should evaluate real bills with objective standards, keeping banks safe and sound and avoiding excesses from insider lending. But most banks outside the major urban centers were too small to be able to follow these recommendations.

According to Margaret Myers, the first credit department was established by the Importer's and Trader's National of New York in the 1880s (Myers 1931). At the same time, the Philadelphia National Bank found it necessary to add a credit department because directors no longer intimately knew all borrowers (Wainwright 1953). The idea began to spread slowly after the Panic of 1893 (Westerfield 1924), but by 1899 there were still only 10 banks in New York with credit departments. Credit departments gradually made granting credit more impersonal, examining the financial records, not the character, of prospective borrowers. However, their role in the National Banking Era remained small overall.

The Maturity Mismatch

Another feature of lending practices that contradicted the real-bills doctrine was the length of loan contracts. The real-bills doctrine assumed that banks would operate without a large maturity mismatch in order to reduce a bank's exposure to a run, but the common practice of renewing loans produced a significant mismatch. If

Table 3

**Composition of Bank Deposits
(percentage of total loans)**

	Total Deposits (\$ millions)	Demand Deposits	Demand Certificates of Deposit	Time Certificates of Deposit	Savings Deposits	Other
National banks, 1910	5,287	80.1	7.6	8.2	0.0	4.1
Nonnational banks, 1910	9,996	35.9	2.2	9.7	48.7	3.5

SOURCE: U.S. Comptroller of the Currency, annual report, 1910.

one looked at bank portfolios without inquiring into loan or borrower histories, it appeared that banks did keep their lending short-term. According to James (1978), most bank loans had short-term maturities. He concluded that typical loans were for 30, 60, or 90 days, with one year being an upper bound. An average maturity was about 60 days. One survey by the Comptroller of the Currency in 1913 found that 57 percent of all bank loans had maturities of fewer than 90 days (James 1978, p. 61). In 1913, the Comptroller of the Currency (1913, p. 100) calculated that 57 percent of bank loans had maturities of fewer than 90 days.

Lending practices differed quite sharply from what appeared on banks' books. Many loans were rolled over in accordance with the working-capital needs of firms and farmers. Moulton (1918) saw little evidence that loans were automatically liquidated at maturity. He found that country banks granted repeated renewals, extending a loan for years to finance working capital. In commercial centers, bankers estimated that 40 percent to 50 percent of unsecured loans were typically renewed. The continuous needs for working capital required continuous credit. Unnerved by the Panic of 1907, the Chicago banks asked Mr. Armour to liquidate his loans so they could replenish their reserves. He replied, "What? I who am liquidating the country and taking the cattle, sheep, and hogs that are being daily sent to

market to liquidate bank loans! . . . What would be the condition of your bank loans if I turned these cattle back to the farms?" (Moulton 1918, pp. 719-20).

The maturity mismatch from funding loans of a few months' maturity with demand deposits thus was even greater, given the actual maturities of loans. If commercial banks held substantial time deposits, this mismatch would have been reduced. However, commercial banks, and especially national banks, primarily held demand deposits. The reserve requirements set by the National Banking Act of 1864 made no distinction among demand deposits, savings, or time deposits, and thus yielded no incentive to increase deposits with longer maturities. Table 3 offers some limited data on the composition of bank deposits in 1910. Demand deposits, at 80 percent of all deposits, were of overwhelming importance for national banks. The picture for nonnational banks was more complex. States often set lower reserve requirements for time deposits (White 1983), encouraging the use of these liabilities. One study of Minnesota banks (James 1978) found that two-thirds of deposits in country banks were time deposits, while time deposits were just 10 percent to 25 percent of deposits in city banks. Longer-term loans and a maturity mismatch are what one would expect to see, according to Diamond (1997), when the capital market is not yet well developed.

The design of the Federal Reserve Act had been partly informed by the real-bills doctrine, and the Fed's early regulations reflected the theory. Even in the 1920s, bankers were expected to conduct an annual "clean up" of debt to demonstrate their creditworthiness and to ensure that the bank was not financing any permanent capital. However, Jacoby and Saulnier (1942) reported that while bankers continued to offer short-term loans almost exclusively, there was a full expectation on the part of both borrowers and lenders that these would be renewed. One study of Iowa banks for 1914-24 showed that while notes were dated with six-month maturities, actual maturities ranged from 10 months to 32 months (Jacoby and Saulnier, p. 13). Jacoby and Saulnier observed that many businesses continued to retire their loans for a short period each year by borrowing from other institutions. They commented that this had the limited value of showing that the borrower could get credit from another institution. Only after the crisis of the 1930s did regulators concerned about the absence of long-term credit to business actively encourage longer-term loans. Bank examiners were instructed not to criticize loans because they had maturities in excess of six months, and the Banking Act of 1935 permitted Federal Reserve Banks to lend on security of any sound asset, regardless of maturity. Beginning in the late 1930s, long-term loans, encouraged by federal regulators and the cessation of new issues on the capital markets, finally became acceptable assets in bank portfolios, even though long-term credits had been implicitly given in the nineteenth century.

The maturity mismatch may have widened after the crisis of the Great Depression and New Deal legislation. Table 4 presents two surveys of member bank loans in 1946 and 1955. One third of member bank loans in these years had maturities of more than one year. The stated maturity structure of these banks appears to be much longer than the structure claimed by most contemporaries and historians for the National Banking Era.

Given that 75 percent of commercial bank deposits were demand deposits in 1950 (*Historical Statistics*, Part II, p. 1022), this would imply a greater maturity mismatch. But this mismatch appears to have shrunk in more recent years. The survey of new loans for 1996 also presented in Table 4 shows that only 12 percent had maturities over one year, and more than 72 percent were for one month or less. This shortening of maturities also appears to have reduced the need for collateral, which had fallen. By January 1996 (*Federal Reserve Bulletin*, May 1996), demand deposits accounted for only 15 percent of all deposits. Even if NOW accounts are included, the total is only 24 percent. There is obviously close liability management and a closing of the maturity mismatch. With the advent of highly developed capital markets, commercial banks, as Diamond (1997) argued, had fewer long-term assets.

The other major nineteenth-century depository institutions, the mutual savings banks, had the bulk of their liabilities in the form of savings deposits. In Table 3, the large fraction of savings deposits in nonnational banks reflects the inclusion of mutual savings banks. Savings deposits were assumed to be less volatile than demand deposits and a good match for a portfolio composed primarily of mortgages. Mutual savings banks' exposure to maturity mismatch was less than might have been expected because mortgage contracts were different from today's contracts. A census study in 1895 of a sample of mortgages from 35 counties across the country found that the ratio of a mortgage to the underlying property value was moderate, and the average life of a loan was relatively brief. For farms across regions, the ratio of the mortgage to property value ranged from 32 percent to 44 percent, and the average life of a mortgage from 2.81 years to 6.62 years. For homes, the ratio of values varied between 33 percent and 48 percent, and the loan life from 1.92 years to 5.99 years (Snowden 1987). According to Snowden (1991), these short mortgage contracts

Table 4

Terms of Lending for Commercial and Industrial Loans

Type & Maturity of Loan	Member Banks, November 20, 1946			Member Banks, October 5, 1955			Commercial Banks, February 5-9, 1996		
	Billions of dollars	Percent	Percent collateralized	Billions of dollars	Percent	Percent collateralized	Billions of dollars	Percent	Percent collateralized
Total short-term loans	8.7	65.9	44.8	20.3	65.9	46.3	56.9	87.9	31.7
Demand loans	2.1	15.9	76.2	4.5	14.6	77.8	19.5	30.1	44.7
Overnight loans							13.6	21.0	11.8
1 month or less							13.8	21.3	21.5
Under 6 months	5.6	42.4	33.9	13.1	42.5	35.9			
More than 1 month							9.9	15.3	44.7
6 months to 1 year	1.0	7.6	40.0	2.7	8.8	44.4			
Total long-term loans	4.5	34.1	44.4	10.5	34.1	59.1	7.9	12.2	64.5
Total loans	13.2	100.0	43.9	30.8	100.0	51.0	64.7	100.0	31.7

SOURCES: *Federal Reserve Bulletin*, June 1947, ~~Sept~~ September 1959, and May 1996.

permitted renegotiation to adjust to altered circumstances. The relative brevity of these contracts meant that the exposure to risk from security mismatches was reduced for any financial intermediary making mortgages because a high proportion of these mortgages fell due each year.

WERE LATE NINETEENTH-CENTURY BANKS SPECIAL?

Competition from other intermediaries and markets has recently called into question banks' distinctive role in the financial system. Gorton and Pennacchi (1993) find some evidence for an unbundling of banks' two functions, making loans and creating deposits. They posit that this development is a result of technological changes that have lowered the cost of information production. Money market mutual funds compete with demand deposits by investing in commercial paper instead of loans, while nonbank lenders, including finance companies and revolving credit, produce loans that compete with banks' commercial and industrial loans. As these intermediaries do not tie demandable liabilities to nonmarketable assets, there appears to be little threat of panic to MMMFs from commercial paper defaults or to nonbank lenders from the failure of some of their number.⁹ As banks' special character is supposed to be embodied in their ability to collect information and monitor borrowers, the growth of MMMFs seems to imply that the market has an increased ability to make short-term credit marketable. Although late nineteenth-century banks were different in many respects from contemporary banks, having no rivals like money market mutual funds, they were forced to compete with the commercial paper markets which grew very fast in this period.

The American commercial paper market developed with the spread of the unsecured promissory note (Greef 1938). The structural defects of the banking system, dominated by small unit banks, spurred its growth on by allowing businesses an alternative to borrowing

from their local banks. Most banks were relatively small, and lending to a single borrower was usually restricted to a fraction of the bank's capital (White 1983).¹⁰ These regulations helped to stimulate the development of the deep American capital markets, where there were no regulatory restrictions on the size of an issue (White 1992). A firm requiring a large short-term loan found commercial paper an attractive alternative to borrowing simultaneously from several banks.

Borrowers in the commercial paper market were typically businesses with a rapid turnover of merchandise or working capital. A substantial net worth was required for a firm to enter this market (James 1996, pp. 222-23). While some firms in this market relied on it exclusively, most maintained lines of credit with commercial banks to meet usual short-term credit needs. Kniffin (1915, p. 463) advised that "it is good policy for a concern to borrow in the open market and reserve its home banks for emergencies. It can often obtain better rates in the broad market, and has the home bank to fall back upon when needed." The lengthening of credit terms and a growth in receivables helped to spur the development of single-name paper (Baxter 1966, p. 5). Commercial paper had maturities ranging between two months and nine months, but most commonly four months to six months. While banks might feel obliged to renew loans, paper, once granted, was paid off at maturity and thus made a good investment for excess funds. Initially both single-name and double-name notes were issued in odd amounts to mirror the exact credit demands of the firms. But banks found this practice inconvenient, and by 1890, commercial paper came to be issued in common denominations, usually \$2,500, \$5,000, and \$10,000 (Greef, pp. 75-77). By the 1890s, not only merchants but also many manufacturers were active issuers in this market. At the same time, commercial paper houses became more professionalized. They acted less as brokers between borrowers and lenders than as outright buyers, who held the paper for resale,

⁹ Gorton and Pennacchi caution that these institutions may mimic banks in that commercial paper is backed by bank loan commitments, and nonbank lenders often finance their activities by issuing puttable bonds.

¹⁰ Some larger banks were adept at innovation to provide more credit. The Philadelphia National Bank skirted around the national bank rule limiting loans to one borrower to 10 percent of capital by purchasing railroad bonds, on which there was no such limitation, with the understanding that they would be repurchased (Wainwright 1953).

and organized credit departments to evaluate borrowers. Whereas most banks had relied on the recommendations of dealers and correspondent banks, many began to set up credit departments to investigate the quality of notes.

Unlike today, commercial banks were the largest purchasers of commercial paper before the First World War. Although commercial paper houses were the rivals of banks for lending, banks bought most of the paper issued. It was a useful alternative investment for banks, paying a lower rate of interest than loans they originated, but it was safe, being carefully selected by the houses that dealt in these obligations. By 1900, banks bought approximately 95 percent of all new offerings (James 1996). The market was not very liquid. There was no secondary market before the establishment of the Federal Reserve, although some city banks might rediscount paper for their correspondents. Commercial paper was held to maturity when it was paid off with near certainty. Competition from commercial paper brokers vexed bankers, especially rural bankers who lost customers to commercial paper houses (James 1978). Pressure from these bankers led the American Bankers' Association to form a committee in 1908 to examine competition from note brokers. The committee attacked the bidding away of good customers with low rates of interest, lowering rates on commercial loans below what they believed was sustainable for banks.

Unfortunately, there is little information on the size or growth of the commercial paper market before the establishment of the Federal Reserve System. After examining the various estimates, Greef (1938) concluded that just before the founding of the Fed, total annual sales of commercial paper was somewhat less than \$2 billion, representing obligations of 2,500 to 3,000 borrowing firms. James (1978) accepts a figure of \$1.7 billion for 1912.¹¹ Assuming that commercial loans and commercial paper had the same average maturity, he computed the total volume of new loans for banks in 1912

and found that commercial paper was 5 percent of total loans. Foulke's (1931) estimate of 5 percent to 12 percent of all unsecured bank loans is in the same range. Similarly, McAvoy (1922) estimated commercial paper to be 10 percent of loans made by national banks.

In the first authoritative study of the commercial paper market, Greef (1938) found that borrowers used the commercial paper market to obtain working capital and for seasonal needs, the same reasons that firms borrowed from banks. He noted that most firms borrowed through commercial paper dealers and from banks at the same time or "rotated" their open-market paper and bank loans. Even if open-market rates were well below the cost of bank loans, Greef found that firms were careful to maintain satisfactory average balances and open lines of credit with banks. Greef observed that coordinated borrowing from banks and the commercial paper market offered firms advantages in raising short-term capital. First, the cost of open-market borrowing was usually below the cost of bank loans, even after adjustment for commissions to dealers and retention of unused bank balances. Open-market borrowing also gave firms bargaining power with their banks and an ability to "clean up" bank loans, when desirable, and borrow larger sums than an individual bank could supply. Contradicting modern banking theory, Greef saw no major disadvantages to open-market borrowing relative to bank borrowing, and he left their coexistence largely unexplained.

In spite of the different character of the commercial paper market in the nineteenth and early twentieth centuries, it stood in a similar relationship to banks then as it does today, and as posited by contemporary banking theory. Firms raised money in both markets, but they appear to have resorted to banks for many of the same reasons firms are alleged to do so today. Firms needed to maintain open lines to their banks, which they would use presumably when they needed their credit-worthiness verified. As Moulton (1918, p. 720) pointed out, commercial paper was

¹¹ The commercial paper market shrank rapidly in the 1920s. In 1929, the ratio of open-market paper to bank loans was 0.66. It experienced a postwar revival, becoming more important than it had been in the pre-World War I era. Commercial paper's share of all short-term lending rose from 2 percent in 1966 to 15 percent in 1991 (James 1996, pp. 232-49).

not automatically renewable. If a firm could not pay off a note when it was due, it could look to its bank for credit instead of the open market. The real-bills doctrine claimed that firms should not be wholly dependent on outside finance for working-capital needs; hence the recommendation for annual clean-ups. Although clean-ups followed the precepts of real bills, they also may be explained by modern theory—and grudgingly seen by some contemporaries—as a useful way to subject the firm to regular checkups to signal their creditworthiness. Banks were thus performing much the same function as they do today.

TWO CASE STUDIES

The paucity of aggregate quantitative data on vital lending characteristics makes it difficult to evaluate actual lending practices in the late nineteenth century. There was a great deal of variation in banking operations, depending on the region, the location of the bank, and the size of its operations. Under the National Banking System, large city banks' portfolios were dominated by deposits they held as part of other banks' reserve requirements, which they largely invested in brokers' loans. In Boston, Lamoreaux (1994) found that the Merchants National Bank had 95 percent of its loans backed by stocks and bonds, and the Second National Bank had 79 percent. The one bank for which she found detailed records, the Suffolk National Bank, had 46 percent of its loan portfolio in collateral loans (almost entirely brokers' loans) and 54 percent in short-term loans on personal security (mostly commercial paper). One of the largest banks in New York City, the National City Bank (Cleveland and Huertas 1985), handled the financing of major corporations and entered investment banking. Other prominent banks, the First National Bank and National Bank of Commerce, were allied with investment banks and life insurance companies in the flotation of new securities (North 1954)—arrangements derided as the "Money Trust." These roles of the larger

banks are worth contrasting to the vast majority of banks, which served as correspondent banks. While the large money-center banks served as reserve banks for the rest of the banking system, other banks concentrated more on commercial and industrial loans to local business.

While the two case studies presented here offer only a partial picture of bank lending, they do show the unique role of banks as lenders to small borrowers who had limited access to other sources of credit. In making and monitoring its loans, the Bank of A. Levy, a commercial bank, constantly observed the local farmers and businessmen. The bank was familiar with all aspects of local economic activity and did not require financial balance sheets, which could not have been produced in any event. The Emigrant Savings Bank, a mutual savings institution, carefully mapped out the properties on which it offered mortgages in New York and had a great familiarity with its ethnic clientele through its large base of savings deposits.

The Bank of A. Levy

In 1885, the Bank of A. Levy began operation as a small, rural, private commercial bank in Ventura County, California (White 1997). It grew rapidly with the expansion of local agriculture and took out a state charter in 1905, surviving to become a prominent local institution that was finally absorbed by First Interstate in 1995. The surviving loan book I examined covered the period from August 1892 to October 1894 and included 330 loans. Only 22 of the loans had been purchased by the Bank, the rest being made directly by the Bank. The small percentage, 6.7 percent, of indirect lending is similar to estimates for all banks. These purchased loans do not appear to have been obtained from a commercial paper house, and they were held to maturity. The total value of all loans was \$124,120.

The loans made by A. Levy were typical for a country bank, as described by James (1978). The average loan was small,

Table 5

Maturity of Bank of A. Levy Loans

Days	Stated Maturity (number of loans)	Percent	Actual Maturity (number of loans)	Percent
1	283	85.8	0	0.0
2 to 30	6	1.8	25	7.9
31 to 60	6	1.8	34	10.7
61 to 90	5	1.5	29	9.1
91 to 120	7	2.1	28	8.8
121 to 365	19	5.8	128	40.4
365 +	4	1.2	73	23.0
Total	330	100.0	317	100.0

under \$400, but loans ranged from \$5 to more than \$5,000. The loans were all unsecured promissory notes. Achille Levy's bookkeeper wrote that Levy was a disciple of the "character loan" method. If he decided that an applicant was of good character, a loan was forthcoming (Carroll 1958). If the borrower's reputation was flawed, the offer of thousands of dollars worth of collateral could not persuade Levy to make a loan. Levy carefully monitored his customers' activities, not only by observing their banking activities but also by traveling around the county on horseback, recording information in his pocket notebook.

Although Levy's lending violated the real-bills doctrine by the use of single-name, unsecured promissory notes, it followed the spirit, in that all the loans were nominally short term. Almost 86 percent of all loans were one-day loans. These loans would appear to have been quite liquid, since Levy could call them for repayment at a day's notice. The typical term of Levy's borrowed funds for 1895 (the closest year with available data) was one day—the same as the nominal maturity of the loans. Thus, there appeared to be no maturity mismatch with one-day loans funded by demand deposits and one-day bills payable.

However, the bank's loans were automatically rolled over. There is no

evidence that they were called before the borrower was ready to repay the debt. Table 5 shows the stated loan maturities and the actual maturities for the 317 loans for which there was information. The average actual maturity for a loan was 279 days, or about nine months. Forty percent lasted between four months and one year, and 23 percent had actual maturities of more than one year. There was only a small actual maturity mismatch because the average actual maturity for 1895 bills payable was 199 days. The actual loan maturities were at the high end of the estimates for the length of loans; few writers suggested that banks make loans for more than one year. Levy clearly provided a continuous source of working capital for local farmers and merchants. Although there were 330 loans, 71 borrowers accounted for 70 percent of the funds. Thirty-five borrowers obtained credit twice in this two-year period, and several individuals obtained anywhere from five to eight loans. The pattern of loan rates suggests that Levy monitored his customers closely, lowering the rate once they proved themselves by repaying the first loan and raising the rate if they were observed to borrow too heavily.

Was A. Levy's bank special? In Ventura County, several other banks competed for customers. None of Levy's customers could have entered the national commercial

paper market, although they might have sold a promissory note to someone locally. Achille Levy, the banker on horseback, was very much like the banker of modern theory whose role is to overcome the asymmetry of information between borrower and lender. His intimate knowledge of local business—gained from his daily contacts, monitoring of customer accounts, and frequent travels around the county—enabled him to do a close evaluation of loan prospects. He knew enough about his clients that he could offer them unsecured credit and have only seven out of 330 loans fail to pay him back in full. The fact that most of these were one-day loans, for which, in theory, full payment could have been demanded the next day, may have disciplined some borrowers. Levy did not attempt to earn funds from a maturity mismatch. In principle, to the extent that his depositors and creditors might run on his bank, he could have liquidated his loans very rapidly.

The Emigrant Savings Bank

The Emigrant Industrial Savings Bank (EISB) was incorporated in 1850, thanks largely to the efforts of John Hughes, Catholic Bishop of New York. Hughes prevailed on a group of 18 prominent citizens, most of them Irish-born, to organize a safe deposit institution aimed at encouraging thrift among poor Irish immigrants. The EISB was one of a score of mutual savings banks set up in New York before the Civil War and founded with strong philanthropic motives. While these banks did encourage the savings habit, they also were operated on a sound commercial basis (Olmstead 1976).

The EISB was one of the largest mutual savings banks in New York City. It was limited by its charter to invest in bonds and mortgages worth double the amount lent. About half its earning assets were invested in mortgages. The mortgages are recorded in the Bond/Mortgage Principal and Interest books deposited in the New York Public Library. To date, in a joint study with Cormac O'Grada, we have col-

lected information on all the mortgages made between 1866 and 1877. While the depositor base remained firmly Irish and Irish-American, loans were not as restricted. They were made to individual home buyers, to developers, and to religious organizations. Loans to religious organizations accounted for 7 percent of the total and 25 percent of the value of loans made by EISB. The developers are hard to identify, except when they took out multiple loans on adjacent plots. According to this method of identification, 16 percent of the loans, accounting for 10 percent of the value of all loans, were given to developers. Most loans were made for Manhattan property, and the average loan size was \$10,574 for the 894 loans examined.

One striking feature of the EISB loans, in contrast to modern mortgages, was the lack of any provision for amortization and the absence of any stated maturity date. For most loans, borrowers simply paid interest until they were able to pay off the balance, although some made irregular payments on the principal. For approximately the first dozen years, the mortgage rate was set at 7 percent, the legal maximum; but when interest rates declined, the semiannual payments were reduced to the new rate. Borrowers appear to have paid off the loans when they saved up enough to pay the principal. Table 6 shows the maturities for the 894 mortgages made over 11 years. While 29 percent of the mortgages by number and 32 percent by value lasted more than 20 years, a very large fraction had maturities of under 10 years or even five years. No one group of borrowers—residential, religious, or commercial and industrial—stood out as taking shorter or longer mortgages.

The EISB was a careful and prudent investor, and there were very few defaults on its mortgages. All properties were local and appear to have been carefully identified, recorded, and examined before any loan was made. Mortgages were given for only 50 percent of the value of the property, offering the bank ample protection. Bolles' recommendations for prudential savings-bank management

Table 6

Maturity of EISB Mortgage Loans

Years	Number of loans	Percent	Dollar value (thousands)	Percent
0 to 4	183	20.5	1,795	23.4
5 to 9	214	23.9	2,814	36.7
10 to 14	155	17.3	1,481	19.3
15 to 19	85	9.5	921	12.0
20 and over	257	28.7	2,444	31.9
Total	894	100.0	7,660	100.0

appear to have been carefully followed. Although the maturities for the 1895 census study were, on average, shorter for the whole nation, the average for the Northeast—six years—falls within the modal range for the EISB. The EISB was faced with a maturity mismatch because it was funded with savings deposits that were typically kept open only a few years. However, the EISB invested a substantial fraction of its portfolio in call loans and bonds, which allowed it to meet rapid decline in deposits in periods of panic.

CONCLUSIONS

Exemplifying the characteristics that are believed to make banks “special” today, banks in the late nineteenth century were the dominant intermediaries. For small, medium, and even many large borrowers, banks were the only financial institutions that offered credit. In the absence of well-established methods of accounting to measure business performance, banks’ intimate knowledge of local business and local business conditions was essential to collecting information and monitoring borrowers. The modern attributes of banks as special lenders were formed principally during the post-Civil War shift from two-name to single-name paper, when banks began to concentrate on the analysis and monitoring of the borrower rather than the examination of the underlying transactions and endorsers of the bills. This development also increased the

number of loans backed by collateral, since lenders who knew less about their borrowers needed some kind of protection. Although banks did not have as many near-competitors as they have today, they did compete with the commercial paper market in the creation of short-term credit. Attempting to provide customers with sufficient capital, in spite of regulatory restrictions and real-bills strictures, banks offered considerable medium-term loans, although their maturity periods were shorter than those of most such loans in the middle of the twentieth century.

The long twentieth-century decline of commercial banks from their position of preeminence has been told, partly, as the result of regulatory disadvantage. However, the decline may also be explained as the consequence of technological improvements, including established accounting standards for business and specialized management procedures for assessing borrowers’ financial information. The rise of credit analysis services and the building of credit departments in banks and commercial paper houses improved banks’ ability to gauge the creditworthiness of borrowers. These developments set the stage for further improvements in information collection by other intermediaries and markets. The dominant position of commercial banks among financial intermediaries in the late nineteenth century may thus be interpreted as the best solution to the asymmetric information problem between borrowers and lenders when there were few technologically

feasible alternatives. The twentieth-century decline in the prominence of banks as intermediaries can be traced back to the development of alternative markets and the improvement of information collection that began during the National Banking Era.

REFERENCES

- Agger, Eugene E. "The Commercial Paper Debate," *Journal of Political Economy* (July 1914), pp. 661-83.
- Baxter, Nevins. *The Commercial Paper Market*, Bankers Publishing Company, 1966.
- Berger, Allen N., Anil K. Kashyap, and Joseph M. Scalise. "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been," *Brookings Papers on Economic Activity* (1995:2), pp. 55-200.
- Berlin, Mitchell, and Loretta Mester. "Debt Covenants and Renegotiation," *Journal of Financial Intermediation* (June 1992), pp. 95-133.
- Bhattacharya, Suddipto, and Anjan Thakor. "Contemporary Banking Theory," *Journal of Financial Intermediation* (October 1993), pp. 2-50.
- Board of Governors of the Federal Reserve System. *Federal Reserve Bulletin*, Board of Governors of the Federal Reserve System, various issues.
- Bolles, Albert S. *Practical Banking*, 5th and 7th eds., Homans Publishing Company, 1888, 1890.
- Boyd, John J., and Mark Gertler. "U.S. Commercial Banking: Trends, Cycles and Policy," *NBER Macroeconomics Annual 1993*, Olivier Jean Blanchard and Stanley Fisher, eds., MIT Press, 1993, pp. 319-68.
- Brief, Richard P. "The Origin and Evolution of Nineteenth-Century Asset Accounting," *Business History Review* (Spring 1966), pp. 1-23.
- Calomiris, Charles, and Charles Kahn. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review* (June 1991), pp. 497-513.
- _____. "Regulation, Industrial Structure, and Instability in U.S. Banking: An Historical Perspective," *Structural Change in Banking*, Michael Klausner and Lawrence J. White, eds., Business One Irwin, 1993, pp. 19-116.
- Carroll, Russell. "The Bank of A. Levy," *Ventura County Historical Society Quarterly IV* (November 1958), pp. 10-13.
- Cleveland, Harold van B., and Thomas F. Huertas. *Citibank 1812-1970*, Harvard University Press, 1985.
- Coffin, George M. *The ABC of Banks and Banking*, S.A. Nelson, 1901.
- Diamond, Douglas W. "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," *Journal of Political Economy* (August 1991), pp. 689-721.
- _____. "Liquidity, Banks and Markets," *Journal of Political Economy* (October 1997), pp. 928-56.
- Deshmukh, Sudhaker D., Stuart I. Greenbaum, and George Kanatas. "Interest Rate Uncertainty and the Financial Intermediary's Choice of Exposure," *Journal of Finance* (March 1983), pp. 141-47.
- Fama, Eugene F. "What's Different About Banks?" *Journal of Monetary Economics* (January 1985), pp. 29-39.
- _____, and Michael Jensen. "Separation of Ownership and Control," *Journal of Law and Economics* (June 1983), pp. 301-25.
- _____, and _____. "Agency Problems and Residual Claims," *Journal of Law and Economics* (June 1983), pp. 327-49.
- Foulke, Ray A. *The Commercial Paper Market*, Bankers Publishing Co., 1931.
- Goldsmith, Raymond W. *Financial Intermediaries in the American Economy Since 1900*, National Bureau of Economic Research and Princeton University Press, 1958.
- Gorton, Gary, and George Pennacchi. "Money Market Funds and Finance Companies: Are They the Banks of the Future?" *Structural Change in Banking*, Michael Klausner and Lawrence J. White, eds., Business One Irwin, 1993, pp. 173-214.
- Greef, Albert O. *The Commercial Paper House in the United States*, Harvard University Press, 1938.
- Homans, Issac Smith. *The Bankers' Common-Place Book*, Bankers Magazine, 1857.
- Jacoby, Neil H., and Raymond J. Saulnier. *Term Lending to Business*, National Bureau of Economic Research, 1942.
- James, Christopher M. "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* (December 1987), pp. 217-35.
- James, John A. *Money and Capital Markets in Postbellum America*, Princeton University Press, 1978.
- _____. "The Rise and Fall of the Commercial Paper Market, 1900-1929," *Anglo-American Financial Systems: Institutions and Markets in the Twentieth Century*, Michael D. Bordo and Richard Sylla, eds., Irwin, 1995, pp. 219-59.
- Johnson, Alexander B. "A Treatise on Banking," *The Bankers' Common-Place Book*, Issac Smith Homans, ed., Bankers Magazine, 1857.
- Kniffin, Jr., William H. *The Practical Work of a Bank*, Bankers Publishing Company, 1915.
- _____. *The Savings Bank and Its Practical Work*, 3rd ed., Bankers Publishing Company, 1918.
- Lamoreaux, Naomi. *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England*, Cambridge University Press and National Bureau of Economic Research, 1994.

- Lummer, Scott L., and John J. McConnell. "Further Evidence on the Bank Lending Process and the Capital-Market Response to Bank Loan Agreements," *Journal of Financial Economics* (November 1989), pp. 99-122.
- Miranti, Jr., Paul J. "Associationalism, Statism, and Professional Regulation: Public Accountants and the Reform of the Financial Markets, 1896-1940," *Business History Review* (Autumn 1986), pp. 438-68.
- McAvoy, Walter. "The Economic Importance of the Commercial Paper House," *Journal of Political Economy* (February 1922), pp. 78-87.
- Moulton, H.G. "Commercial Banking and Capital Formation," *Journal of Political Economy* (May, June, July, November 1918), pp. 484-508, 638-63, 705-31, 849-81.
- Myers, Margaret G. *The New York Money Market, Volume I: Origins and Development*, Columbia University Press, 1931.
- North, Douglass C. "Life Insurance and Investment Banking at the Time of the Armstrong Investigation," *Journal of Economic History* (Summer 1954), pp. 209-28.
- Olmstead, Alan L. *New York City Mutual Savings Banks, 1819-1861*, University of North Carolina Press, 1976.
- Snowden, Kenneth A. "Mortgage Rates and American Capital Market Development in the Late Nineteenth Century," *Journal of Economic History* (September 1987), pp. 671-92.
- _____, and Walid Bu-Saba. "Mortgage Loan Duration and Renegotiation in the Late Nineteenth Century," unpublished manuscript.
- Steiner, W.H. "The Commercial Paper Business," *Federal Reserve Bulletin* (August-September 1921), pp. 920-26, 1052-57.
- Thakor, Anjan V. "Maturity Transformation," *The New Palgrave Dictionary of Money and Finance*, Vol. 2, London: Macmillan, 1992, pp. 678-80.
- Wainwright, Nicholas B. *History of the Philadelphia National Bank: A Century and a Half of Philadelphia Banking, 1803-1953*, William F. Fell Co., 1953.
- Westerfield, Ray B. *Banking Principles and Practice*, Ronald Press Co., 1924.
- Willis, Henry Parker. *The Federal Reserve System: Legislation, Organization and Operation*, Ronald Press Co., 1923.
- Wheelock, David C. "Is the Banking Industry in Decline? Recent Trends and Future Prospects from a Historical Perspective," this *Review* (September/October 1993), pp. 3-22.
- White, Eugene N. *The Regulation and Reform of the American Banking System, 1900-1929*, Princeton University Press, 1983.
- _____. "The Effects of Bank Regulation on the Financing of American Business, 1860-1960," *Finance and the Enterprise*, Vera Zamagni, ed., Academic Press, 1992.
- _____. "California Banking in the Nineteenth Century: The Art and Method of the Bank of A. Levy," unpublished manuscript.
- U.S. Comptroller of the Currency. *Annual Reports*, Government Printing Office, various years.
- U.S. Department of Commerce. *Historical Statistics of the United States: Colonial Times to 1970*, U.S. Bureau of the Census, 1975.

Naomi R. Lamoreaux is a professor of economics and history at the University of California-Los Angeles and a research associate at the National Bureau of Economic Research.

Commentary

Naomi R. Lamoreaux

One of the hallmarks of Eugene White's scholarship is his knack for using detailed historical examples to raise large, thought-provoking questions. "Were Banks Special Intermediaries in Late Nineteenth Century America?" is no exception. White combines case studies of two small financial institutions, the Bank of A. Levy and the Emigrant Savings Bank, with other information on nineteenth-century banking theory and practice to highlight banks' "specialness" during that period—that is, their unique ability to serve as delegated monitors for savers. He then argues that a fall in the cost of information eroded this specialness in the twentieth century. The result, according to White, was a steep decline in banks' share of the assets of financial intermediaries.

My goal in this comment is less to criticize White's argument than to amplify and recharacterize it—to make stronger his case that, by the early twentieth century, banks had lost many of the attributes that allowed them to perform the role of delegated monitors more effectively than other kinds of financial institutions. I will begin by countering White's case studies with the example of a late nineteenth-century institution, the Suffolk Bank of Boston, that had largely abrogated its position as delegated monitor. I will then use what is known about the evolution of managerial practice within banks in the late nineteenth and early twentieth centuries to argue that the Suffolk example is more representative of historical trends than the cases White describes. Banks' earlier informational advantage had derived from the imbeddedness of their officers in the communities within which they did most of their lending. As credit markets lost their local character,

these advantages disappeared, and banks increasingly had to rely on the same general information sources as other financial institutions.

THE SUFFOLK COUNTER EXAMPLE

White uses the examples of the Bank of A. Levy and the Emigrant Savings Bank to highlight the information services that, he argues, were at the heart of banks' specialness in the late nineteenth century. The former institution was a small rural commercial bank in Ventura County, California. Its president, Achille Levy, knew borrowers personally and traveled around the county on horseback in order to monitor their activities. The second, a mutual savings bank in New York City, was run by prominent members of the Irish immigrant community who presumably were personally acquainted with the mortgagees to whom they lent the bulk of the bank's funds. White makes no pretense that these cases are representative, but he does imply that they capture in important ways the kind of information-gathering facilities that made banks special.

The problem is that it is easy to offer counterexamples. One bank that appears to have behaved very differently was the Suffolk National Bank of Boston. During the first half of the nineteenth century, as Rolnick, Smith, and Weber describe in their contribution to this volume, Suffolk was the most important bank in Boston and exerted what was in effect regulatory authority over the notes issued by all the banks in the New England region. This regulatory role ended before the Civil War, however, and by the late nineteenth century, the Suffolk was just one of a considerable number of large banks in the city of Boston. A brief run of its lending records is extant from the turn of the century, and close examination reveals that very few of Suffolk's loans derived from activities of

the sort that White would label as special. For example, 46 percent of the Suffolk's portfolio (by value) consisted of collateral loans, 84 percent of which were granted to brokers and other intermediaries who dealt in the securities markets. These loans were backed by securities that bank examiners regarded as readily marketable. That is, in lending on the basis of this collateral, the Suffolk Bank was not making use of any special informational advantage; it was only accepting securities that were generally perceived by the investment community to be of high quality. The other 54 percent of Suffolk's portfolio consisted of short-term loans based on personal security, but approximately two-thirds of these (by value) were notes purchased on the commercial-paper market. Only 19 percent of the bank's portfolio consisted of notes backed by personal security (one- or two-name paper) that were discounted for the benefit of signatories who were customers of the bank. In other words, only 19 percent of the bank's portfolio consisted of loans for which the bank might perform some special informational role (Lamoreaux 1994, p. 128).

TRENDS IN BANK MANAGEMENT

Suffolk, of course, is just one example, but it is an example that I would argue was representative of larger trends in the banking system. Unfortunately, it is impossible to demonstrate this claim by analyzing a large sample of banks' loan portfolios. Very little information of this sort is extant. Instead, it is necessary to approach the problem more indirectly—by thinking about what is known about banks' management structures and practices during this period.

Nineteenth-century banks typically had very lean managerial hierarchies. For example, national banks were governed by a board of directors, one of whose members was elected president. The daily affairs of the bank were generally run by a cashier, who might work alone or, depending on the size of the institution, might be assisted by one or more tellers and clerks, and per-

haps a bookkeeper. For most of the century, the cashier was the chief operating officer. Presidents (like other members of the boards of directors) were usually part-time officers. They had other business interests to which they devoted their primary attention (Lamoreaux 1994, pp. 3-4).

This type of managerial structure underwent some changes over time. For example, by the end of the century it was increasingly common, especially at large urban banks like Suffolk, to have presidents who had previously served as cashiers and who devoted all their time to their banks (Lamoreaux 1994, pp. 123-4). But the important point is that, whether the chief of operations was the cashier or the president, he supervised relatively few people. In particular, there was little investment in the nineteenth century in developing the *organizational* capability to collect and process information about the creditworthiness of borrowers. To the extent that banks had an informational advantage over other financial institutions, it was a *personal* one that derived from having a chief officer who was well connected locally, had repeated dealings with the same people, and spent time (as A. Levy did) traveling around the community checking up on borrowers.¹ A bank might also gain additional information about potential borrowers by choosing for its directors people who had good knowledge of particular segments of the business community and who were willing to use this knowledge for the benefit of the bank.

The kind of information that nineteenth century banks acquired through their officers was thus local and personal. It derived from the imbeddedness of these men in the communities from which most of the institutions' borrowers were drawn. A lender who was not similarly imbedded did not have access to information of comparable quality. Such a lender might subscribe, for example, to the reports of credit agencies like the R.G. Dun Company, which had corresponding agents, often lawyers, located in communities throughout the country. These agents gossiped with local merchants, kept their ears open, and reported any news that might affect the creditworthiness of

¹ Thus William Goddard felt he had to resign from the presidency of the Providence National Bank in Providence, Rhode Island, when he was afflicted with "an obstinate lameness [that] so tethers me to the spot . . . that I feel disqualified from seeking elsewhere the information regarding the credit of borrowers, which I regard as of the highest importance to the successful management of a bank." Quoted in Lamoreaux (1994), p. 103.

potential borrowers. There is reason to believe, however, that for much of the century this kind of information was inferior in quality to that obtained by local bank officers for their institutions. After all, bankers had access to all the same sources of information as the agents of the credit agencies. In addition, they had the direct knowledge that came from their own private dealings with borrowers.

Even in this period, however, there were important limits on the kinds of information that banks were able to collect. For example, it was not generally considered appropriate to ask borrowers for financial statements. Even bankers, therefore, had only impressionistic evidence of their borrowers' net worth. Like White's Achille Levy, they based their lending decisions as much on their assessment of an applicant's character as on precise information about income and liabilities.

THE DECLINE OF LOCAL LENDING

This limitation on information collecting would become more important over time. By the end of century, the informational advantage that bankers derived from their local imbeddedness and from their repeat dealings with borrowers was increasingly inadequate (and regarded as so by contemporary observers), especially in the most economically developed and urban areas of the nation.² Part of the problem was the trend toward single-name paper that White discusses in his essay. Loans on personal security were more risky than they had been earlier because they were backed only by the wealth of the maker, not the maker plus one or more endorsers approved by the bank. A more important problem for our purposes was the growing tendency for businesses to borrow from more than one financial institution and also to float commercial paper on the market. As a result, it was now more difficult for bankers to get a good sense of a borrower's financial position just from their own repeated dealings with the individual (James 1978, pp. 55-59; Lamoreaux 1994, pp. 89-90).

Banks dealt with this problem by moving to require formal, sometimes audited, financial statements from borrowers. They also began to invest in information-gathering capabilities, creating new credit departments whose business was to keep track of customers' creditworthiness. These developments came relatively late. As White tells us, financial statements were not in common use, even in large banks, until the 1890s; in small banks the delay was much longer. Credit departments were also first organized in the 1890s. As late as 1899, only 10 banks had them, and they were all in New York.

I would like, however, to question whether these investments in the organizational capability to collect information about borrowers really did much good in the sense of allowing banks to recapture their informational advantage over other kinds of financial institutions. After all, any lender could require a financial statement. Moreover, to the extent that borrowers sought loans from multiple institutions and floated commercial paper on the market, banks were not particularly well placed to assess the truthfulness of these statements. As a result, banks ultimately had to depend on external sources of verification such as independent audits, information collected by credit agencies, and (later) tax returns—sources of information that were available on the same terms to other financial institutions.

There was, of course, another possibility. Banks could have embraced what Charles Calomiris and others have called "universal banking" and developed long-term relationships with the companies to which they lent funds, taking equity positions in the firms and naming directors to their boards. In that way, they could have gained inside information about income and net worth and also, perhaps, some say over the companies' managements. But U.S. banks did not go this route.³ Instead, they coped with their growing information problems in a number of alternative ways: by requiring customers to maintain deposits with them worth a certain percentage of their credit line;⁴ by keeping borrowers on a short leash and forcing them to renew their

² See, for example, Rhodes' *Journal of Banking* (February 1893, p. 137; April 1893, p. 377; and June 1893, p. 585) and *Bankers' Magazine* (January 1893, pp. 525-26; August 1898, pp. 286-87; and September 1898, p. 384 and pp. 413-22). See also Cannon (1891), pp. 535-36.

³ It is beyond the scope of this commentary to explore the reasons that U.S. banks did not practice universal banking. For two alternative views, see Calomiris (1995) and Lamoreaux (1994).

⁴ As one banker wrote in Rhodes' *Journal of Banking*, "the best paper to accept is that offered by firms or individuals who are in the habit of carrying balances with their bank from whom the accommodation should be obtained. There appears to me *no* better means to determine the amount of risk a bank incurs than by regulating its loans according to the average balance carried" (May 1893, pp. 486-89). See also Forgan (1920), p. 7.

loans frequently;⁵ and even, as the Suffolk Bank essentially did, by giving up the whole idea of maintaining an informational advantage and engaging in the kinds of lending that did not require special knowledge (for example, lending to stock-market brokers on the security of readily marketable stocks or buying commercial paper on market). Suffolk's strategy, by the way, was not particularly profitable, and the bank's stockholders voluntarily reduced its capital in the early twentieth century. Banks' other coping mechanisms also proved ultimately impracticable as pressures from regulators and competition from other lenders forced them to give up the idea of compensating deposits and to lengthen lending terms in the twentieth century (Lamoreaux 1994, pp. 101-2, pp. 136-7, p. 163).

In the end, therefore, banks' loss of specialness was not so much a result of a fall in the cost of information as of a lack of advantage in collecting the kinds of information needed to assess the creditworthiness of borrowers operating in the geographically wider markets of the twentieth century. Information costs are undoubtedly a part of this story, but not, I think, in the way White originally intended. The story I would tell would be one that emphasized the decline in local lending. Banks were special in the nineteenth century because, unlike most other financial institutions, they were located close to their borrowers. As local lending declined, banks' informational advantage disappeared. Not surprisingly, their market share also dropped as a consequence.

James, John A. *Money and Capital Markets in Postbellum America*, Princeton University Press, 1978.

Lamoreaux, Naomi R. *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England*, Cambridge University Press, 1994.

⁵ Bankers felt that short-term loans helped to discipline borrowers. As one put it, "borrowers will use the proceeds of loans which they are to repay in a few months more wisely than might be the case if the payment were indefinitely deferred." Moreover, because the information in a borrowers' financial statement reflected current conditions, which were liable to "change radically for the worse" with the passage of time, such a statement was clearly "a basis for short time credit only." Coman (1907); and American Institute of Banking (1916).

REFERENCES

- American Institute of Banking. *Loans and Investments*, New York: privately printed, 1916.
- Calomiris, Charles W. "The Costs of Rejecting Universal Banking: American Finance in the German Mirror, 1870-1914," *Coordination and Information: Historical Perspectives on the Organization of Enterprise*, Naomi R. Lamoreaux and Daniel M. G. Raff, eds., University of Chicago Press, 1995, pp. 257-315.
- Coman, E. "Requisites of a Good Loan," *Practical Problems in Banking and Currency*, Walter Henry Hull, ed., Macmillan, 1907, pp. 66-77.
- Cannon, James G. *An Ideal Bank*, privately printed, 1891.
- Forgan, James B. *A Good Note*, Chicago: privately printed, 1920.

Caroline Fohlin is an assistant professor of economics at the California Institute of Technology. The author is grateful to Lance Davis, Michael Edelstein, Barry Eichengreen, John Latting, Peter Temin, and David Wheelock as well as conference participants at the St. Louis Federal Reserve Bank Annual Economic Policy Conference for helpful comments and discussions, and to the National Science Foundation for funding this research.



Banking Systems and Economic Growth: Lessons from Britain and Germany in the Pre-World War I Era

Caroline Fohlin

Regulation of the financial system has occupied economists and policymakers since the beginning of financial history. Such attention has been warranted because of the crucial role of these institutions in economic life. There remains, however, much disagreement over the ways financial and real variables interact and the extent to which financial development can promote economic growth. Modern growth theory has made strides in modeling the relationship between financial development and economic growth, but a causal relationship is difficult to verify empirically.¹

Despite the burgeoning research on finance and growth, the importance of financial-system structure has yet to be determined. Much of the debate over banking reform in the United States hinges on the assumption that certain types of financial systems allocate an economy's resources more efficiently than others. There is a widespread sense in the United States and Great Britain that the universal banking systems of Germany and Japan have given those countries an advantage in industrial development and economic growth over much of the past 150 years.

The structure of the German banks, in particular, has been viewed as a key ingredient in Germany's industrial development before World War I. Universal banking, because it combines all phases of finance in one institution, is thought by many to have yielded economies of scope and greater efficiency. Such efficiency has been argued in turn to have increased the volume and reduced the costs of finance, thus promoting industrial investment.² Furthermore, German banks are often assumed to have maintained close, long-term relationships with industrial firms. Equity positions are thought to have aligned the incentives of banks and firms and encouraged multi-period optimization of their behavior. In contrast, a long line of detractors has chastised the British banks for avoiding engagement with domestic industry and leaving firms to seek financing from other sources. Firms' resultant recourse to securities markets is argued to have served investors' short-term profit motives at the expense of long-term growth.³ As a result, the banks have been blamed for the apparent underperformance of the British economy since the late nineteenth century.

Two lines of historical investigation may shed light on the continuing debates over the relative efficacy of German and British banking systems. The first step is to determine whether the German banks offered the advantages that have been ascribed to them; the second step is to ascertain whether the provision of these services by universal banks fueled economic growth. In comparing the two systems, however, it is important to acknowledge that the British banks were not prohibited from combining functions or from pursuing long-term relationships with industrial firms. Thus, research on the real effects of financial structure must accept that, if the British banks' organization and activities were suboptimal for industrial growth, such inefficiency stemmed from market failures of one sort or another: rationing

¹ Greenwood and Smith (1997) offer what may be the most reasonable compromise on the question of causality: a model in which financial markets arise after some period of real development, and in which the expansion of those markets fuels further real growth. A logical implication of this model is that exogenous creation of a financial system with advanced features may not spur real growth. The problem then for implementing development policy is determining how to get poor countries to the point at which financial systems will arise endogenously.

² Most recently, Calomiris (1995) has advanced this idea.

³ For a review of the literature on British banking and industrial development, see Collins (1991). Also see Capie and Collins (1992). For a critical appraisal of the British banking system, see Edwards (1987).

relatively low-return or high-risk ventures or failing to perceive or act upon favorable prospects.

This study uses aggregate bank balance sheet data to investigate systematic differences in the financial makeup and activities of universal and specialized banks. By explicitly comparing British and German banks, it takes steps toward quantifying the possible disparity in financial-system growth effects over the decades leading up to World War I. Financial systems are thought to influence both the quantity and quality of investment. Thus, this paper first measures the rate of expansion and the ultimate magnitude of capital mobilized by British and German banks. The study then investigates the makeup of the banks' asset portfolios and estimates the extent of direct involvement by the two types of banks in equity ownership.

The findings suggest that, compared to British banks, German banks maintained at least as much liquidity relative to their short-term liabilities, mobilized a smaller share of the economies' capital, and held approximately the same (small) proportion of their assets in the form of nongovernment securities. Furthermore, the German banks seem to have held only a limited number of industrial equities in their portfolios and often did so merely because of insufficient markets for new issues.

The results offer insights into both differences and similarities in the organization of banking in Germany and the United Kingdom, specifically, and into the historical importance of financial structure, more generally. The findings suggest that the gulf between specialized and universal banking in terms of their influence on economic growth and industrial development is less than commonly believed.

THE LINK BETWEEN BANK STRUCTURE AND GROWTH

The primary purpose of banks is to mobilize otherwise idle resources for use in productive investment. A wide array of theoretical models has appeared in the growth and development literature in the

past decade to formalize the link between financial-system functioning and the growth of the real economy.⁴ In comparison to the traditional growth models—in which output was seen as a function of capital, labor, and disembodied technological progress—the current models provide a richer framework for interpreting the potential impact of financial systems. For their motivation, nearly all appeal to the observed correlations between financial-system development and industrial growth uncovered by economic historians and development economists during the 1960s and '70s.⁵

Pagano (1993) provides a simple way to summarize the newer models of finance and growth. Using several simplifying assumptions, the model yields the growth rate of output per capita as a function of three variables: savings rate, return on investment, and costs of intermediation. Thus, financial institutions may enhance economic growth by raising the total quantity of financial capital available to entrepreneurs, improving the quality (productivity) of investments, and increasing the efficiency of intermediation (lowering costs) between the sources and uses of funds.

This framework can help in comparing the effectiveness of the German and British banking systems, but further refinement is required to clarify the ways financial institutions affect the variables in the growth formula. The following sections take some first steps at comparing the impact of specialized and universal banking systems on the quantity and quality of investment.

QUANTITY OF INVESTMENT

Banks influence the accumulation of physical capital by directing funds to entrepreneurs who wish to invest. Such capital mobilization proceeds in two stages: capital collection through deposit-taking or sales of equity shares, and fund dispersal through loans or advances. By repeating this process, the banking system multiplier expands the money supply and redistributes the economy's capital. These banking functions increase the share of resources targeted to productive investment.

⁴ For an overview of some of the literature, see Pagano (1993) and Galetovic (1996). Greenwood and Smith (1997) provide more technical details.

⁵ Cameron (1967), Goldsmith (1969), McKinnon (1973), and Shaw (1973) are the standard references.

The German universal banks are credited with mobilizing significant amounts of capital from the public and thereby promoting industrial growth. The British banks, by comparison, are typically presumed to have participated less aggressively in the accumulation of funds. Total assets of financial institutions as a share of gross national product grew substantially in both Britain and Germany between 1860 and 1913, but Goldsmith's (1969) figures indicate that this ratio expanded more in the latter than in the former. Furthermore, while Britain's ratio exceeded Germany's in 1860, the British lagged the Germans by 1900. The gap grew to over 50 percent by World War I.⁶ Nonetheless, the Goldsmith data indicate that the British deposit banks accounted for a greater share of their country's GNP than did the German universal banks at each point in the pre-war era.

In Germany, virtually all of the functions relating to corporate finance fell under the purview of the universal banks. The British financial system largely separated investment banking, brokerage, and commercial services; thus, comparing the German universal banks to the British deposit (commercial) banks underestimates the share of corporate financing institutions in the British economy. Nonetheless, at 50 percent to 60 percent, deposit banks and private banks accounted for twice the share of total financial institution assets in Britain than did the universal banks (of both joint-stock and private forms) in Germany.

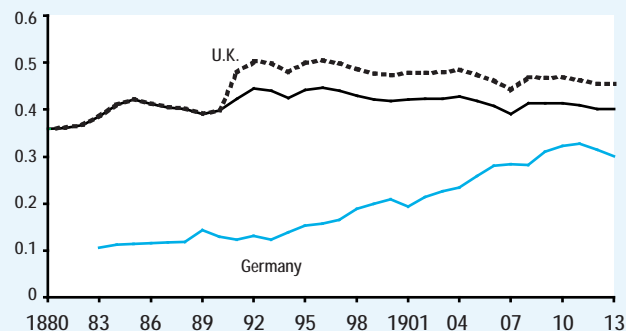
Given the traditional emphasis on the universal banks' role in promoting industrialization and economic growth in Germany, the universal banks' share in both financial assets and GNP seems relatively small. Furthermore, the sharp increase in the German universal banks' share between 1900 and 1913—especially compared to the more gradual changes from 1860 to 1880—raises doubts about the causal link between universal bank expansion and industrial growth.

In contrast, British deposit banks' share of assets grew most rapidly between 1860 and 1880 and then leveled off. While some

Figure 1

Total Joint-Stock Bank Liabilities Less Total Cash Holdings, United Kingdom and Germany, 1880-1913

Liabilities Less Cash/GNP



NOTES: The British data come from the *Economist* series as reported in Sheppard (1971) and include private banks, starting in 1891. The solid line represents an estimation of the joint-stock banks' liabilities from 1880 to 1913, based on the ratio of private to joint-stock banks in 1891, but that ratio likely declined significantly between 1891 and 1913. The German data report only joint-stock banks for the whole period. Since the private banks accounted for a greater share of bank assets in Germany, the omission of private banks may exaggerate the British lead. Even if estimated figures for the German private banks are added, however, some gap in liabilities less cash still remained as late as 1913. Furthermore, the denominator for the German series is net national product, and the ratio may therefore overestimate bank liabilities as a share of GNP. The GNP/NNP data come from Mitchell (1978).

SOURCES: United Kingdom, Sheppard (1971) and Mitchell (1978); Germany, Deutsche Bundesbank (1976) and Mitchell (1978).

of the differences in these patterns likely stem from divergent timing of industrialization, part may arise from the varied structure and practices of the British and German banks. The German universal banks are widely believed to have internalized the secondary market in securities, and indeed a significant portion of trading took place through their offices. The continued expansion of universal banks, therefore, may represent the expansion of the market for securities.

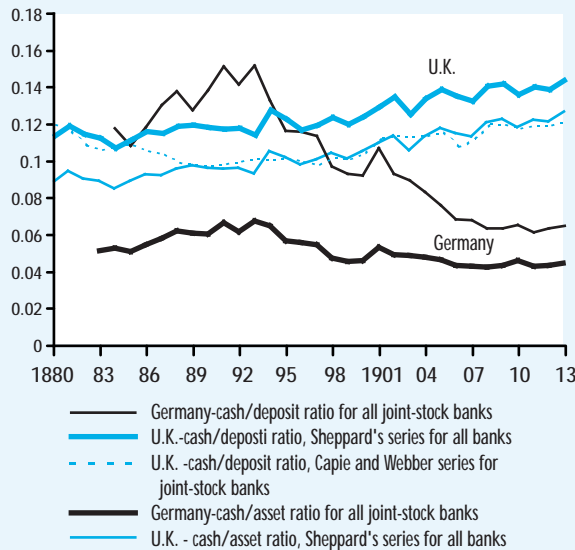
Another measure of the volume of capital mobilized by the British and German commercial banks is represented by total bank assets less cash. As a share of GNP, liabilities less cash in the British deposit banks exceeded that in the German universal banks by a significant margin from 1883 until after the turn of the twentieth century (Figure 1). While the gap between the United Kingdom and Germany seems to have virtually disappeared by the outset

⁶ The pattern reversed after the war, and as of 1963, Britain led Germany again by a substantial margin.

Figure 2

Cash-to-Asset and Cash-to-Deposit Ratios, United Kingdom and Germany 1880-1913

Cash/Total Assets and Cash/Total Deposits



NOTES: Sheppard's series for all United Kingdom banks includes private banks after 1890. Capie and Webber's data on U.K. joint-stock banks include all Irish and Scottish banks.

SOURCES: United Kingdom, Sheppard (1971) and Capie and Webber (1987); Germany, Deutsche Bundesbank (1976)

While the commercial banks clearly represented a greater share of the economy in the United Kingdom than in Germany, the universal banks may have expanded available capital at a faster rate. Indeed, the faster growth rate of German bank liabilities compared to British bank liabilities suggests this might have been the case.

The ultimate impact of the banks' activities depends inversely upon the proportion of the system's assets retained as cash reserves. In a simple model of a monetary economy, the total nominal money stock is a function of the nominal monetary base (currency plus reserves), the ratio of bank deposits to currency, and the cash-to-reserve ratio.⁹ Financial intermediaries maintain partial control over both the reserve ratio and the deposit-to-currency ratio. For example, banks can raise the deposit-to-currency ratio by encouraging individuals to deposit their savings or to buy equity shares in the bank.

Two measures offer some insight into the banks' roles in multiple expansion. The money multiplier is a negative function of the cash-to-deposit ratio. Where banks are financed by equity or private capital in addition to deposits, a more relevant ratio may be the cash-to-liabilities ratio. In comparing the British and German cases, we find both ratios informative.¹⁰

Germany's cash-to-liability ratio ranged between 5 percent and 6 percent in the late 1880s and early 1890s, but it declined considerably after 1893. This decline in the universal banks' cash-to-liability ratio coincided with the growth of their liabilities less cash. Over the same period, the British banks seem to have maintained considerably higher cash-to-liability ratios, and the gap appears to have widened after 1893 (Figure 2).

Theoretically, at least, the cash-to-liability ratio affects monetary expansion, interest rates, and economic growth. Yet banks' holdings of cash are not exogenous, and differences in funding methods between British and German banks help explain part of the gap in cash-to-liability ratios. Particularly in the nineteenth century, British deposit banks financed a much

of World War I, the series again diverged during the war and its aftermath.

Figure 1 also reveals a clear difference in growth rates of liabilities less cash per capita in the two countries. When we take into account the shift in coverage for the British series, the trend for that country is relatively flat. The German figures, in contrast, indicate gradual expansion before 1894, but rapid growth thereafter.⁷ The disparity in growth rates may be explained by the different patterns of industrial development in the two countries; however, the late development of the German joint-stock banks is somewhat surprising. Joint-stock universal banking seems to have taken off more than 40 years after the first universal bank was founded and after the industrialization pushes of the mid-nineteenth century.⁸

⁷ Regression of the log of liabilities less cash on a time trend yields an estimated annual average growth rate of 8.6 percent in the post-1894 period, as opposed to a rate of 5.1 percent before 1894.

⁸ To some extent, the apparently late take-off by the universal banks is due to the switch to the joint-stock form. Private banks were more prevalent before 1894 than after. Inclusion of the private banks would flatten the trend some, but it is not clear that the private banks provided the same services in the same way as the later joint-stock banks.

⁹ See Champ and Freeman (1994) and sources cited there.

greater share of their operations with deposits than did the German universal banks, and United Kingdom provincial banks (at least to some extent) also issued notes. Universal banks were prohibited from issuing their own notes.

Given the divergent liability structures of German and British banks, the cash-to-deposit ratio offers greater insight into the banks' participation in maturity transformation.¹¹ Among the German universal banks, cash-to-deposit ratios followed a similar, though more extreme, pattern than cash-to-liabilities ratios, rising over the late 1880s and declining after 1893. Until the last years of the nineteenth century, German cash-to-deposit ratios exceeded the United Kingdom ratios, and the gap reached as much as 10 percentage points around 1891.

The variability of cash ratios is also important. The largest German banks kept aggregate ratios as low as 7 percent and as high as 22 percent during this period. The British deposit banks, by comparison, held relatively steady cash ratios throughout the end of the nineteenth and the start of the twentieth centuries. As a group, the British joint-stock banks maintained cash balances between 10 percent and 15 percent of deposits.

Though neither the British nor the German banks were bound by minimum reserve requirements in the pre-World War I period, cash ratios often still depended on factors outside the banks' control. Even in the absence of regulations, central monetary authorities may tacitly impose a ratio on commercial banks. In Britain, the apparent floor at 10 percent, while certainly not proof of the central bank's role, is consistent with the notion that the Bank of England held some sway over the banks' minimum cash ratio. The ratios for Germany, however, suggest no successful suasion by the Reichsbank. Bankers, politicians, and economists often debated the need for a required reserve, but little was done toward imposing regulations like those enacted in the United States.¹²

Clearly, many of the factors involved in determining the cash-to-deposit ratio fall outside the purview of the banks. At the

same time, banks do retain significant control over their investment portfolios, and the riskiness of those portfolios must also affect the banks' assessment of their need for cash reserves. A bank's structure and activities may measurably influence the composition of its asset portfolios, and the different levels of bank specialization may therefore partly explain the somewhat divergent patterns of cash ratios—and thus of capital expansion—of British and German banks.

QUALITY OF INVESTMENTS

Banks' role in mobilizing capital is intimately tied to their involvement in the utilization of funds. Through decisions about how to lend and invest funds, banks can influence the quality of capital formation. As with capital mobilization, the structure of the German universal banks is thought by many to have offered advantages over the British system in promoting the efficient use of financial capital. The literatures on German and British banking have suggested that the British banks invested rather conservatively, while the German banks opted for riskier strategies. Such risky investment, it is argued, channeled funds into high-growth and high-return industries and helped promote Germany's industrialization.¹³

For influencing the quality of investment, the crucial organizational advantage of the German banks is their supposed long-term direct participation in industrial firms. By holding industrial shares, the banks are thought to have monitored and even controlled the firms they financed. The British banks, in contrast, are traditionally accused of having little to do with industry and are criticized for taking a short-term, arms-length approach to industrial lending.

There are several theoretical reasons why bank equity holdings may increase the efficiency of investment. Many of these hypotheses originate in the idea that asymmetric information between borrowers and lenders poses extra costs and creates inefficiencies in the selection and funding of investment projects. Cost reductions may result from imposing discipline on

¹⁰ The data for the two ratios come from two sources: Sheppard's (1971 [1873]) compilation of *The Economist's* series and Capie and Webber's (1985) newer estimates. The latter only provide cash and deposit figures, so the cash-to-liabilities ratio cannot be calculated from this source.

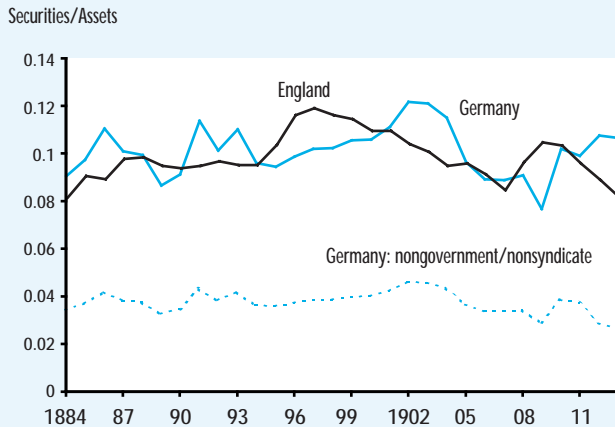
¹¹ That is, changing short-term liabilities into longer-term assets.

¹² Most proposals recommended required cash holdings of 1 percent to 5 percent of the sum of deposits and current account balances to be held at the Reichsbank. Defenders of the German joint-stock banks claimed that the British banks held much slimmer reserves than the Germans. The British banks were accused of padding their reserves for their semi-annual statements of account. See Riesser (1910, 1911) and sources cited there. Goodhart (1972) also discusses the reporting practices of British banks.

¹³ Tilly (1986) points out that, given that the main clientele of the universal banks appears to have been large, older, publicly traded enterprises, the banks may not have been actively involved in risky, innovative investment in general.

Figure 3

Nongovernment Securities as a Share of Total Assets, United Kingdom and Germany 1884-1913



NOTES: Figures for the United Kingdom include England and Wales only and include private banks after 1890. Figures for Germany include great banks only (large Berlin banks). See explanation in text.

SOURCES: United Kingdom, Sheppard (1971) ; Germany, Deutsche Bundesbank (1976).

resolving agency problems associated with multiperiod financial contracts. Explicitly motivated by modern perceptions of the German and Japanese banking systems, this model shows that the efficiency of inside investing hinges on the use of fixed fraction contracts. In such arrangements, the investor receives a fixed percentage of project returns and finances that same proportion of future investments.¹⁵

According to these theoretical arguments, banks that hold equity stakes in firms improve the firms' investment efficiency. Thus, the relative extent of equity holdings in the portfolios of British and German firms offers one way to assess banks' direct involvement in raising investment quality.¹⁶ Figure 3 compares nongovernment securities for the two countries. Because of uncertainties about valuation and reporting, these figures should be viewed as approximations.¹⁷

The German banks show no consistent tendency toward higher securities holdings than the British banks. Indeed, according to these estimates, the range was nearly identical in the two countries (7 percent to 12 percent for the German banks and 8 percent to 12 percent for the British). The figures, it should be noted, provide as conservative an estimate as possible of the German and British nongovernment securities holdings. The Scottish and Irish deposit banks held higher levels of investments than did their English and Welsh counterparts, and the largest of the German universal banks held more of their assets in securities than did the provincial banks. Therefore, the fact that Figure 3 still shows the British banks' securities holdings on par with the Germans' provides a strong indication that, despite the measurement difficulties, the British banks held a position in nongovernment securities similar to that of the German banks.

Such a finding would fall in line with expectations if one thought that the two types of banks were roughly similar. The predominance of underwriting and brokerage functions among the universal banks, however, should have led to higher levels of securities holdings at German banks, compared to British commercial banks.

management, overseeing investment planning and outcomes, optimizing risk-taking by firms, and aligning banks' and firms' incentives for long-term benefits.¹⁴

John, et. al. (1994) models the effects of equity ownership on firms' risk-taking and shows that investment efficiency increases with the proportion of bank financing held in the form of equity. Imperfect oversight of investment choices and outcomes creates incentives for firms to use borrowed funds in an excessively risky manner. When banks maintain veto power over the use of funds, pure debt holdings induce them to minimize their risks in order to guarantee a fixed return. Equity holdings, in contrast, encourage banks to seek higher firm valuation. Thus, the greater the banks' equity holdings in the firm, the higher the banks' incentives for efficient tradeoff between risk and expected return.

In related work, Admati and Pfleiderer (1994) also demonstrate the potential importance of equity stakeholders in

¹⁴ Myers and Majluf (1984) analyze many of these theoretical issues and provide a formal model of the potential suboptimality of investment under asymmetric information.

¹⁵ Repeated interaction naturally adds the problem of renegotiation. Admati and Pfleiderer (1994), Persons (1994), and Dybvig and Zender (1991) all address this question.

¹⁶ The remainder of this section borrows heavily from Fohlin (1997c).

¹⁷ Disaggregation for the German figures begins only in 1912. The figures for the years before that are estimated on the basis of the lowest holdings of government securities between 1912 and 1920 as well as on the detailed account of one of the great banks between 1896 and 1899. The proportion for great banks ranged from 17.6 percent to 28.6 percent of total securi-

Thus far, the numbers for the German banks have included securities holdings resulting from their underwriting and brokerage business.¹⁸ A significant portion of the universal banks' total investments arose out of their involvement in underwriting consortia (or syndicates). These participations therefore include some shares that remained on the banks' books only temporarily and because of the banks' inability to place the shares. It is useful to compare the estimates for the largest German banks to the securities holdings of British institutions engaged in investment banking. Cottrell (1985) shows that, at least in some cases, British investment banks held more than half their assets in the form of illiquid investments. By contrast, the German universal banks reported liquidity coefficients (the ratio of immediately available or quick assets to total liabilities) of 85 percent in 1893. These figures gradually declined by more than 20 points over the ensuing 15 years.¹⁹ Naturally, these banks cannot be compared directly with the German universal banks, but the forgoing examples do support the notion that the great banks in particular, because of their active engagement in investment banking, should be expected to have held a significantly greater share of their assets in the form of securities than did the British deposit banks. Comparison with the British investment banks also underscores the potential inconsistency in the idea that universal banks could hold substantial long-term (illiquid) engagements with industrial firms and still operate a commercial business on the order of the British deposit banks.

To understand how important the banks' direct investment in industrial companies may have been for the growth of the economy, it is useful to combine the data on bank investments with those on bank assets relative to the economy as a whole. Table 1 reports the results of this calculation and indicates that the nongovernment securities holdings of universal banks ranged between 2 percent and 4 percent of GNP for the three decades preceding World War I. Even if the estimates are only approximately correct,

Table 1

Bank Holdings of Non-Government Securities/GNP

	1880	1900	1913
Germany	0.022	0.027	0.040
Britain	0.044	0.063	0.058

SOURCE: Calculated from Deutsche Bundesbank (1976) and Goldsmith (1969).

the banks' holdings of nongovernment securities accounted for a very small share of the economy. The German banks' share did increase after 1880, but their holdings of nongovernment securities still amounted to only 4 percent of GNP by 1913. Furthermore, the biggest part of the increase came after the major push of industrialization in Germany.

The British banks' holdings of non-government securities were also low relative to GNP; in contrast to the German case, however, the banks' securities share of GNP rose between 1880 and 1900 and then leveled off. Given the measurement difficulties already discussed, and the likelihood that securities accounted for a greater share of the economy in Britain than in Germany, it is best not to overemphasize the differences between the German and British numbers.²⁰ Nonetheless, these calculations cast doubt on the idea that the banks' holdings of securities provided a significant stimulus to either the German or the British economies during the last half of the nineteenth century.

It is often claimed that the British banks held only gilt-edged securities in their portfolios, and that the German banks participated more actively and directly in risky, start-up ventures. The official figures, however, do not allow specific types of securities to be distinguished. Such distinctions, unfortunately, depend on spottier evidence from individual banks. German bank records for the pre-1880 period are generally unavailable. Nonetheless, some details are available for two of the earliest German joint-stock universal banks, the Disconto-

ties held between 1912 and 1920. Given that this period covers World War I, it would be natural to expect that government securities might comprise a higher proportion of securities than they did in the preceding years. In the one detailing of bank securities holdings that I could find for the period before 1900 (Bank für Handel und Industrie, a great bank), government securities amounted to 24 percent to 55 percent of total securities (in the period 1896-99). Thus, since I am trying to err on the side of finding high rates of nongovernment securities holdings, 17 percent seemed a conservative enough estimate of the proportion of all great bank securities held in the form of government securities.

¹⁸According to Riesser (1910), the largest universal banks earned approximately half of their gross profits from underwriting and brokerage services.

¹⁹Riesser (1911), p. 655, discusses the banks' liquidity at length.

²⁰Banks' equity holdings may actually represent a greater proportion of share capital in Germany than in the United Kingdom. Such figures are interesting from the point of view of the banks' involvement in the secondary market for shares, but the banks' ultimate impact must be measured against the economy as a whole.

Gesellschaft (DG) and the Darmstädter Bank (Bank für Handel und Industrie, or BHI).²¹

The DG held no securities in its first four years, but the proportion of securities holdings rose to around 12 percent of assets in 1856 and grew rapidly over the following few years. The bank seems to have unloaded securities during the boom years of the early 1870s, but it then took on extremely high shares of securities during the middle of that decade. While the bank's holdings continued to fluctuate throughout the remainder of the nineteenth century, the proportion of securities followed a generally downward trend toward the end of the period.

Between 1856 and 1865, two mining companies accounted for the vast majority of DG's industrial holdings, averaging around 11 percent of bank assets during this period. Direct participation arose out of the bank's intention to convert the firms into joint-stock companies, but because of the thin market for the securities, DG was forced to hold these companies' shares until the bank could extricate itself in the more favorable market of the late 1860s and early 1870s.

The remainder of DG's securities portfolio was held in relatively conservative investments: government debt, railway shares and bonds, and other priority bonds and shares. With the exception of a few minor holdings of shares, the DG confined its participation in industry to three companies (the two already discussed plus another mining concern). Indeed, the bank's holdings of industry stocks amounted to between zero percent and 3 percent of its assets for the years in which disaggregated data are available (1852-65).²²

Tilly (1967) shows in his discussion of the early industrial promotion activities of the Bank für Handel und Industrie, another of the great banks, that while the bank was energetic in such activities in its first four years, it had difficulty placing shares at reasonable prices. By the early 1860s, BHI had extricated itself from this side of the business and had turned to railway and government finance. Thus,

it can hardly be argued that even the early activities of the great banks included extensive, direct involvement in industrial companies.

Economic historians can pick up this story in the 1880s by using evidence from BHI.²³ Holdings of industrial shares amounted to less than 1 percent of BHI's assets for most of the 1880s and '90s, and even at its peak, the proportion of industrial shares to assets reached only 1.3 percent (in 1882).²⁴ Moreover, BHI reported substantial holdings of only 12 different companies between 1882 and 1897 and no more than seven firms in any one year. Together, these data provide further support for the notion that the great banks invested a relatively small portion of their portfolios in long-term stakes in industrial firms.²⁵

As for the securities holdings in Britain, Goodhart provides some details for three British commercial banks (Metropolitan Bank, London and Midland, and Union Bank). Nearly all of the investments reported consisted of British, colonial, or foreign-government securities or railway stocks and bonds. Given his warnings about the banks' desire to hide any investments in industrial firms, it is impossible to tell for sure what industrial shares the banks may have held. Edelstein, however, has provided more general estimates of U.K. securities holdings, and those results indicate an expansion of industrial holdings between 1871 and 1913. Industrial concerns and railways, both foreign and domestic, accounted for 37 percent of all securities holdings in 1871 and 62 percent by World War I. Home company holdings alone increased from 4 percent to 17 percent of all U.K. holdings over the period.²⁶

For the period between 1883 and 1907, Davis and Huttenback (1986) find that the financial community owned around 5 percent of U.K. share value and averaged 4 percent stakes in those companies. In addition, public companies, some of which may have been banks, held nearly 4 percent of domestic share capital. The banks might be expected to have participated to some extent in these investments, though firm

²¹ See Dabritz (1931) on DG and *Saling's Börsen-Jahrbuch* on the Darmstädter Bank.

²² Dabritz (1931).

²³ BHI published unusually detailed accounts of its securities holdings, and until 1899, *Saling's* reproduced the information in its series on Berlin-listed companies.

²⁴ Fohlin (1997c).

²⁵ While the experiences of two banks may not necessarily be generalized to the population as a whole, these two banks do represent a significant proportion of the great banks.

²⁶ Edelstein (1982), p. 48.

proof of such a contention is apparently unavailable. Yet even if the British banks held no industrial shares, the evidence for DG and BHI suggests that the German universal banks were not far ahead on this count.

It is important to note that the banks' ownership of shares, at the margin, may have provided important injections of liquidity or signals of quality for newly public firms. In a thin market for industrial securities, and in cases of lumpy investments, such holdings may permit firms to invest when they otherwise would not have. Thus, small and transient equity stakes may increase the quantity of investment, even if they do not have the qualitative, efficiency effects that long-term holdings are thought to have. Since such equities may not have made it onto the banks' books, though, it is difficult to estimate the ultimate impact of transient holdings.

Share ownership represents only the most direct kind of involvement in industrial firms. The banks may have also participated indirectly in companies, either through proxy voting of customers' shares or through positions in the firms' supervisory boards.²⁷ Because of their combination of underwriting, brokerage, and commercial services, the German banks probably obtained greater control of industrial shares than did the British banks. Since shares taken as collateral or simply held as a service to customers would not appear in the banks' balance sheets, and since firms did not have to reveal their shareholders, it is virtually impossible to quantify the extent of proxy voting by the German banks.

It is possible to quantify board positions, and such data suggest that the bank directors held positions in relatively few companies. Approximately 23 percent of German joint-stock companies had a private banker or bank manager on their supervisory boards, but only half of these attached companies received representation from the great banks.²⁸

Proxy votes and supervisory board positions may have enabled banks to monitor their investments and even control the use of bank funds. From a theoretical perspective, however, it is unclear whether such indirect

participation yields the same kind of incentive effects as direct ownership. In theory, at least, systems in which banks exert control over investment decisions but do not align their incentives with those of the firms through equity stakes force firms into excessively safe and thus inefficient investment programs. So the German system of proxy voting and interlocking directorates may have increased bank control and oversight, but it may have led to more internal financing and fewer risky investments.

CONCLUSION

The financial system may promote real growth of the economy by enhancing the quantity, quality, or efficiency of investment. Using evidence on bank financial structure, this study has compared the contributions of the British and German banking systems in the first two of these areas. The analysis yields no compelling evidence that one system consistently or significantly outperformed the other in raising the quantity or quality of investment.

The findings indicate that the German universal banks, despite their broader involvement in corporate finance, accounted for a markedly smaller proportion of the economy than did the British banks. The gap of the 1880s, much of which may have been due to the later onset of industrialization in Germany than in Britain, only began to diminish after 1894 and never fully disappeared. The universal banks may have, however, expanded their available capital at a faster rate, since they invested or lent a greater share of their total liabilities than did the British banks. The disparity in cash-to-liability ratios, however, stems from the heavily deposit-based financing of the British banks. Until the late 1890s, the German banks actually maintained more conservative coverage of short-term liabilities than did the U.K. banks. Only with the serious onset of the deposit business in the mid-1890s did the German cash-to-deposit ratios begin their steady decline.

The German banks are frequently credited for their active participation in industry, and bank equity positions in

²⁷ The German supervisory board comprises shareholders' representatives. Currently, this body must also represent the firm's workers.

²⁸ Fohlin (1997a, b) discuss the prevalence, sectoral distribution, and determinants of interlocking directorates between banks and firms.

firms are thought to improve the quality of investment. The findings show, however, that the universal banks held only a small share of their portfolios in the form of industrial equities. Evidence from two of the largest universal banks suggests that the universal banks held stakes in only a few firms and often did so for lack of demand for their shares. Based on the theoretical work on bank equity stakes, this article also argues that if the German banks wielded greater control over firms (through board positions, for example) than did the British banks but took no greater equity stakes in those firms, then the German system of relationship banking may actually have led to relative underinvestment in risky projects.

This study has raised the possibility that the German banks' choices of investment and reserve holdings were constrained by the extent of the secondary market in securities. It is possible that the dominance of universal banking in Germany may have hampered the growth of complementary financial institutions. In the German case, however, regulation that encouraged the growth of universal banking also may have inhibited the development of securities markets. Thus, the existence of universal banking, per se, may not curtail the functioning of securities markets.

Clearly, further work on this subject is required. But if specialized and universal systems of finance generally provide similar quantities and qualities of investment, then cost-efficiency may prove to be the crucial determinant of the relative growth effects of the two systems. Universal banking may yield economies of scale or scope compared to a specialized system, but these economies may also lead to excessive concentration, market power, and inefficiency in the banking sector. In addition, the internalization of the secondary securities market within the banking system may hamper both the efficient distribution of financial capital and the market for corporate control. Such factors bear directly on the costs of finance, and such costs influence economic growth.

This study narrows the perceived gulf between the British and German banking

systems; and it indicates quite strongly that without a significant period of real development, financial institutions can offer only limited benefits for economic growth.²⁹ It may still be true that the German economy has outperformed its British counterpart over the past century, but this study suggests that differences in banking structure are probably not the cause. Such findings may prove useful for policymakers both in forecasting the effects of banking deregulation in the United States and in formulating development programs in other parts of the world.

REFERENCES

- Admati, Anat R., and Paul Pfleiderer. "Robust Financial Contracting and the Role of Venture Capitalists," *The Journal of Finance* (June 1994), pp. 371-402.
- Calomiris, Charles. "The Costs of Rejecting Universal Banking: American Finance in the German Mirror, 1870-1914," *Coordination and Information: Historical Perspectives on the Organization of Enterprise*, Naomi Lamoreaux and D. Raff, eds., University of Chicago Press, 1995.
- Cameron, R. *Banking in the Early Stages of Industrialization*, Oxford University Press, 1967.
- Capie, Forrest, and Michael Collins. *Have the Banks Failed British Industry?* London: Institute of Economic Affairs, 1992.
- Capie, Forrest, and Alan Webber. *A Monetary History of the United Kingdom, 1870-1982*, London: George Allen & Unwin, 1985.
- Champ, Bruce, and Scott Freeman. *Modeling Monetary Economies*, John Wiley & Sons, 1994.
- Collins, M. *Banks and Industrial Finance in Britain 1800-1939*, London: MacMillan, 1991.
- Cottrell, P. L. *Investment Banking in England, 1856-1881*, Garland, 1985.
- Däbritz, Walther. *Gründung und Anfänge der Disconto-Gesellschaft Berlin*, Leipzig: Verlag Von Duncker & Humblot, 1931.
- Davis, Lance, and Robert A. Huttenback. *Mammon and the Pursuit of Empire*, Cambridge University Press, 1986.
- Deutsche Bundesbank. *Deutsches Geld- und Bankwesen in Zahlen 1876-1975*, Frankfurt: Fritz Knapp Verlag, 1976.
- Dybvig, Philip, and Jaime Zender. "Capital Structure and Dividend Irrelevance with Asymmetric Information," *Review of Financial Studies*, (4(1) 1991), pp. 201-19
- Edelstein, Michael. *Overseas Investment in the Age of High Imperialism: The United Kingdom, 1885-1914*, Columbia University Press, 1982.

²⁹ Tilly (1967), pp. 114-15, argued similarly 30 years ago.

REVIEW

MAY/JUNE 1998

- Edwards, George T. *The Role of Banks in Economic Development*, London: Macmillan, 1987.
- Fohlin, Caroline. "Universal Banking Networks in Pre-war Germany: New Evidence from Company Financial Data," *Research in Economics* (September 1997a), pp. 201-25.
- _____. "The Universal Banks and the Mobilization of Capital in Imperial Germany," Social Science Working Paper No. 1008, California Institute of Technology, 1997b. Forthcoming in *Finance and the Making of Modern Capitalism*, P. Cottrell, G. Feldman, and J. Reis, eds., Sclar Press.
- _____. "Bank Securities Holdings and Industrial Finance before World War I: Britain and Germany Compared," *Business and Economic History* (Winter 1997c), pp. 1-13.
- Galetovic, Alexander. "Finance and Growth: A Synthesis and Interpretation of the Evidence," *BNL Quarterly Review* (March 1996), pp. 59-82.
- Goldsmith, Raymond. *Financial Structure and Development*, Yale University Press, 1969.
- Goodhart, Charles. *The Business of Banking, 1891-1914*, London: Weidenfeld & Nicolson, 1972.
- Greenwood, Jeremy, and Bruce D. Smith. "Financial Markets in Development, and the Development of Financial Markets," *Journal of Economic Dynamics and Control* (January 1997), pp. 145-81.
- Harris, Milton, and Artur Raviv. "The Theory of Capital Structure," *The Journal of Finance* (March 1991), pp. 297-355.
- John, Kose, Teresa John, and Anthony Saunders. "Universal Banking and Firm Risk-Taking," *Journal of Banking and Finance* (March 1994), pp. 307-23.
- McKinnon, Ronald. *Money and Capital in Economic Development*, The Brookings Institution, 1973.
- Mitchell, Brian R. *European Historical Statistics 1750-1970*, Columbia University Press, 1978.
- Myers, Stewart, and Nicholas Majluf. "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* (June 1984), pp. 187-221.
- Pagano, Marco. "Financial Markets and Growth: An Overview," *European Economic Review* (April 1993), pp. 613-22.
- Persons, John. "Renegotiation and the Impossibility of Optimal Investment," *The Review of Financial Studies* (Summer 1994), pp. 419-49.
- Riesser, Jakob. *Die deutschen Großbanken und ihre Konzentration*, Jena: Verlag von Gustav Fischer, 1910. English translation: *The German Great Banks and their Concentration*, The National Monetary Commission, Washington: Government Printing Office, 1911.
- Shaw, Edward. *Financial Deepening in Economic Development*, Oxford University Press, 1973.
- Sheppard, David K. *The Growth and Role of UK Financial Institutions, 1880-1962*, London: Methuen, 1971.
- Tilly, Richard. "Germany (1815-1870)," *Banking in the Early Stages of Industrialization*, Cameron, R., ed., Oxford University Press, 1967.
- _____. "German Banking, 1850-1914: Development Assistance for the Strong," *Journal of European Economic History* (Spring 1986), pp. 113-51.
- Verlag für Börsen- und Finanzliteratur A.-G. (various years), *Saling's Börsen-Jahrbuch*, Leipzig.

Peter Temin is the Elisha Gray II Professor of Economics at the Massachusetts Institute of Technology.

Commentary

Peter Temin

In this article, Fohlin continues her exploration of the role of the German banking system during pre-World War I industrialization—a system that has been celebrated among economists and economic historians for many years. A generation ago, Gerschenkron (1962) argued that Germany needed banking to mobilize savings in the nineteenth century because of its relative backwardness, which was manifest in a poorly developed capital market and a scarcity of savings. When universal banks mobilized savings, this stimulated the growth of heavy industry. The German banking system has also been acclaimed by economists who have argued that the integrated, universal banks of Germany and other countries provide the most efficient intermediation.

Fohlin does not start from Gerschenkron, although I—as one of Gerschenkron's students—will. She adopts a separation that Pagano (1993) used in a survey of models of financial institutions and economic growth. He proposed a simple linear model in which growth was equal to the product of the level of technology, the savings rate, and the cost of financial intermediation. Taking the first of these as given, he argued that we can evaluate financial institutions by their effects on the volume of savings and by their efficiency in intermediation. Increasing the volume of savings often is termed mobilization, while more efficient intermediation equalizes the rate of return across the economy and yields greater output for a given amount of savings. Fohlin calls these outcomes the quantity and quality of investments, and she discusses them in turn.

Gerschenkron emphasized the role of banks in mobilizing savings. Fohlin shows that the ratio of financial assets to GNP was higher in Britain than in Germany before World War I. Banks also had fewer assets relative to GNP in Germany than in Britain, although their assets rose more quickly and passed those of the British banks before the war. Fohlin concludes from these data and from data on bank reserve ratios that financial institutions in Germany did not play the role in mobilizing savings that they played in Britain, or at least not until Germany had caught up with Britain in many dimensions.

But is Britain the right standard of comparison for the question of savings mobilization? Gerschenkron, were he here, would have admitted happily that Germany had fewer financial intermediaries than the British. That was why it needed universal banks, in his view. In other words, the relevant comparison is not with Britain, which did not need universal banks, but with a less developed country that could not mobilize its savings. From the perspective proposed by Gerschenkron, one possible conclusion from Fohlin's data is that the German banks were doing a very good job of mobilizing savings, thereby bringing Germany into the same ballpark as the more advanced British economy.

Fohlin turns next to the efficiency of intermediation. The distinction she makes is clearer in theory than in practice, because the mobilization of savings is itself a measure of efficient intermediation. Fohlin looks at the particular question of supplying capital to industrial firms and argues that German banks held no more equity in industrial firms than their British counterparts. This is an important finding and a thought-provoking conclusion.

One obvious point Fohlin makes is that other countries that did not have the German banking structure—specifically, Anglo-Saxon countries—were not precluded by banking regulation from having this structure. British banks could have

organized themselves the way German banks did in the nineteenth century if they had wished. The same is true of American banks, some of which did so. It is instructive to combine this observation with Fohlin's data.

If the British did not adopt the German banking structure, why not? Two answers are possible: The first, in the spirit of Gerschenkron, is that the British banking structure was preferable to the German. The German banking structure, in other words, was a second-best system forced on the Germans by their relative backwardness. The large British capital market allowed banks to specialize for reasons Adam Smith would have recognized. But German banks, in the more constricted German capital market, did not have this luxury. By this metric, it is impressive that the German banks did almost as well as the British, that they were catching up to the British before the war.

This view suggests that current attention to German universal banking may be misplaced—that nineteenth-century German institutions were used only because conditions would not allow better ones to flourish. It suggests that comparison of the British and German capital markets in the late nineteenth century will reveal greater, rather than smaller, costs of intermediation. Even if such a difference is not apparent, there is no reason to expect to find that German capital markets worked better than British ones.

This view, however, conflicts with that of Calomiris (1995), who argued that, in comparison to the cost of intermediation in American banks, the cost in Germany was very low. He did not compare Germany and Britain as Fohlin did, so his work does not bear directly on this issue. But if the costs of intermediation were similar in the two Anglo-Saxon countries without universal banking—or with very limited universal banking—then his conclusion is relevant. Fohlin did not collect data on the cost of intermediation, so she does not consider this aspect of efficiency. But if the cost of intermediation by issuing equity in Germany was low, then it is surprising that there was not more intermediation.

There is, however, another possible answer to the question of why Britain did not have universal banks. Relying on the concept of complementarity of Milgrom and Roberts (1990), we may conclude that British conditions may not have been conducive to German-style banking. This choice may not have been due to the advantages and disadvantages of the respective banking structures, but rather to how well each complemented the other institutions in its own country. German banks would not have worked well in Britain, just as British banks would not have worked well in Germany.

This is a more complex situation. In a world filled with complementarities, there is no way to evaluate specific institutions and practices in isolation. Each of them is good or bad in specific contexts; they do not stand isolated in the world. To the extent that British or German banking was located within such a web of complementarities, it does not make sense to compare them with each other independent of the rest of the capital market.

Complementarities are relevant here because the German equity market was not as well developed as the British in the late nineteenth century—and it may not be as well developed even now. German firms relied more heavily on debt than British firms; that is, they were more highly leveraged. The composition of debts in the economy as a whole therefore must have differed in the two countries. Fohlin separates assets into debt and equity as she describes bank assets, but not for the economy as a whole.

For example, Fohlin shows that banks in Britain and Germany held approximately the same small share of their assets in the form of equity in nonfinancial firms. Since the equity market was less developed in Germany, German firms relied more heavily on debt financing than did their British counterparts. Edwards and Ogilvie (1996), using data from Goldsmith (1985), calculate that shares of domestic joint-stock companies were 12 percent of national assets in Britain in 1895 and 2 percent in Germany. These shares had changed slightly to 10

percent and 3 percent, respectively, by 1913. The share of business equity held by banks in Germany then may have been larger than in Britain.

Alternatively, many observers have argued that German banks provided capital to industrial firms through what has been translated as the “current account.” This was a deposit with automatic overdraft privileges. The interest rate on overdrafts was 1 percent above the market rate; on deposits, it was 1 percent below the market rate. If this practice was used widely in Germany but not in Britain, then the comparison of equity holdings would reveal less about German banks than about British banks.

This discussion leads to another point Fohlin makes. As this is a new point, I need to back up and work back to the issue of the “current account.” The literature typically assumes that the great German banks wanted to dominate the economy, just as the Money Trust and J. P. Morgan are alleged to have wanted to do in the United States. But why did they want to do this? Fohlin assumes they wanted to exert this influence to make money—like any other firm in the economy.

This assumption leads Fohlin to ask how German banks could have made money from dominating industrial firms. They could have held the equity of these firms, which would have provided a direct line to their growth and profits. But Fohlin found that the German banks did not hold much equity—certainly not more than the British banks. If they exerted influence over managers, they must have done so by voting shares deposited with them but not owned by them, or by force of personality. How, then, did they make their money?

As I have noted, German banks loaned extensively through the “current account.” This was a debt instrument, not equity, in which the interest rate was fixed and rather low; there was little money to be made here. German bankers were not, so far as I know, active in politics. They were not men of great wealth independent of their banks, and their wealth was not increased by the action of borrowing firms. German bankers lived well, but then so did British bankers.

Fohlin therefore has posed a question that the literature on banking needs to take seriously. Were the German universal banks that have been both glorified and vilified over the years simply banks along the lines that we know them, loaning money to firms and earning money by choosing good risks? Or were they like the man pulling the levers behind the curtain in *The Wizard of Oz*? And if they were pulling the levers of industrial power, how were they being paid for doing so?

REFERENCES

- Calomiris, Charles. “The Costs of Rejecting Universal Banking: American Finance in the German Mirror, 1870-1914,” *Coordination and Information: Historical Perspectives on the Organization of Enterprise*, Naomi Lamoreaux and Daniel Raff, eds., University of Chicago Press, 1995.
- Edwards, Jeremy, and Sheilagh Ogilvie. “Universal Banks and German Industrialization: A Reappraisal,” *Economic History Review* (August 1996), pp. 427-46.
- Gerschenkron, Alexander. *Economic Backwardness in Historical Perspective, A Book of Essays*, Harvard University Press, 1962, pp. 5-30.
- Goldsmith, Raymond W. *Comparative National Balance Sheets*, University of Chicago Press, 1985.
- Milgrom, Paul, and John Roberts. “The Economics of Modern Manufacturing: Technology, Strategy, and Organization,” *American Economic Review* (June 1990), pp. 511-28.
- Pagano, Marco. “Financial Markets and Growth: An Overview,” *European Economic Review* (April 1993), pp. 613-22.

Larry Neal is a professor of economics at the University of Illinois at Urbana-Champaign.*



The Financial Crisis of 1825 and the Restructuring of the British Financial System

Larry Neal

Today's financial press reports regularly on evidence of systemic risks, financial fragility, banking failures, stock market collapses, and exchange rate attacks throughout the global financial network of the 1990s. To a financial historian, these reports simply reprise similar concerns and risks in numerous episodes of financial innovation and regime change in the past. True, the 1990s have the peculiar feature of emerging markets among newly independent states that are trying to market either their government debt or securities issued by their former state enterprises. But this situation does not eliminate the relevance of past episodes; it merely limits it to fewer periods. The period after World War I had many of the same problems, for example, although policymakers then subsumed them largely under the issues of whether, when, and how to return to the pre-war gold standard that had created a much more benign financial system worldwide. Policymakers of that time were much more interested than their modern counterparts in exploring lessons from the past.

For example, William Acworth's classic study, *Financial Reconstruction in England, 1815-1822*, was published in 1925. He argued convincingly that the severe deflationary policy followed by the government and the Bank of England after peace in

1815 had prolonged and deepened unnecessarily the economic troubles accompanying the transition from a wartime to a peacetime economy. Nevertheless, the British government and the Bank of England pursued much the same strategy after World War I, again taking six years after the peace treaty to resume convertibility—and at the prewar standard. Again, monetary ease that followed resumption led to a surge of prosperity, speculative ventures in the capital markets, and eventual collapse of the financial system. The difference was that in 1825-26, there was a systemic stoppage of the banking system, followed by widespread bankruptcies and unemployment, while in 1931 there was abandonment of the gold standard, followed by imperial preference and worldwide movements toward autarky. So much for the lessons of history!

As pessimistic as Acworth was in assessing the consequences of Britain's first return to the gold standard in 1821, the consequences of the ensuing monetary expansion and speculative boom that ended in the spectacular collapse at the end of 1825 proved to be not so dire in the long run for the British economy. The policy changes that affected the monetary regime—the exchange rates, the structure of the banking sector, the role of the Bank of England and the management of the government's debt—while minor in each particular and slow to take effect, were cumulatively effective in laying the basis for Britain's dominance in the world financial system until the outbreak of World War I. This outcome contrasts sufficiently with the disappointing pattern of British economic progress during the twentieth century after both World War I and World War II that perhaps we should take a fresh look at the economic and financial transition after the Napoleonic Wars. What caused the problems identified by Acworth that culminated in the stock market crash of 1825 and the English banking system's failure to

* The author acknowledges with gratitude the support of the University of Illinois during his sabbatical leave of 1996-97 as well as the Guggenheim Foundation and the British Fulbright Commission, for research efforts on this project. This version of the paper has benefited not only from the comments of the discussant, Michael Bordo, and participants at the conference, but also from comments and suggestions made during seminars at the University of Illinois (especially Lee Alston and Salim Rashid), Indiana University (especially George von Furstenberg and Elmus Wicker), and the Research Triangle in North Carolina (especially Douglas Fisher, Judith Klein, and Gianni Toniolo). Remaining shortcomings and misinterpretations of the facts are, of course, the sole responsibility of the author.

withstand its impact? More important, why did the British government's relatively modest reforms prove to be so effective in the long run? Perhaps we can glean more useful lessons for today's policymakers than previous historians have been able to provide.

The argument developed in this paper is that the common element in all the problems of Britain's first return to gold arose from the pressures of coping with vastly increased informational uncertainties within the existing structure of English institutions.¹ These problems started with the Treasury itself, confronted by the difficulties of servicing the huge government debt accumulated during the Napoleonic Wars and deprived of its primary source of revenue, the income tax. They continued within the Bank of England, forced now to take on new responsibilities while searching for new sources of revenue to replace its wartime profits. They were compounded by the response of the London capital market, which produced a bewildering array of new financial assets to its customers to replace the high-yielding government debt now being retired. All this left the London private banks and their corresponding country banks—as well as their customers in agriculture, trade, and manufacturing—floundering in the resulting confusion. The government's piecemeal reforms, introduced during the crisis of 1825 and its immediate aftermath, provided smoother patterns of tax collections and interest disbursements, established Bank of England branches throughout England, stimulated country bank competition with joint-stock companies outside of London, and eliminated the Bubble Act of 1720. Even the bankruptcy laws began to be rewritten in 1831.²

These disparate reforms made marginal improvements in the efficiency of information gathering and processing by the government, the central bank, the banking system, and the stock market while preserving the separation of functions among them. Maintaining these “firewalls” among the types of institutions making up the financial sector of the British economy diminished the immediate impact of the

reforms, but it enabled them to become increasingly effective over time. True, crises continued to arise throughout the rest of the century as the British economy was subjected to repeated shocks of wars, famines, frauds, and foreign defaults. But the evolving financial sector of the British economy surmounted each crisis with increasing confidence, and all the while these firewalls were preserved. The firewalls meant that relationships among financial intermediaries and financial markets had to be maintained by short-term contracts in a competitive market environment rather than by regulations imposed by centralized authority with long-term rigidity.

The focal point for these new market relationships was the market for discounted commercial bills that arose rapidly in importance after the crisis of 1825.³ Once again, as in earlier crises and in those that were to follow until World War I, the British financial sector was able to find a market solution to the problems created by its relatively inefficient and disparate financial institutions. In the longer run, the flexibility of response provided by the combination of markets and financial intermediaries coexisting in the British financial system enabled it to withstand exogenous shocks and to finance expansion of the real economy. To elucidate and elaborate this argument, I analyze, in turn, the shock to the financial system of shifting from wartime to peacetime finance in 1821, the financial crisis that occurred at the end of 1825, the Bank of England's efforts to pick up the pieces, and, finally, the rise of a market in discounted commercial bills that put things right again—for awhile. The lessons of each episode highlight the importance of appraising the financial system as a whole, rather than focusing on what appears to be its weakest link. In retrospect, it seems critical to allow information to flow freely among the various parts of the system in order that markets may form to price and intermediate risk. At the time, the Bank of England refused to divulge important information and remained aloof from market activity until it was forced to act, usually too late. Only gradually were

¹ Note the emphasis on English, rather than British, institutions. The Scottish and Irish banks avoided the Panic of 1825 almost entirely, a fact that caused much soul-searching among the English at the time.

² Duffy (1985), ch. 1.

³ King (1936), ch. 2.

these lessons learned; now is not the time to forget them.

THE SHOCK: FROM WARTIME TO PEACETIME FINANCE IN 1821

In the expansion of war finance that the Napoleonic Wars induced in Britain, all parts of the British financial system prospered. At the top, the Treasury benefited from increased taxes, especially the income tax, as well as the expanded market for its debt, both long-term, funded debt and short-term, unfunded debt. The Bank of England profited throughout the Napoleonic Wars as the government's agent for fiscal transfers both at home and abroad throughout the most expensive war fought in history to that time. It increased its annual dividend to 12 percent from 7 percent in 1805 (reduced back to 10 percent in 1807), greatly enlarged its staff, built new facilities at its location on Threadneedle Street, and expanded its note issue as well as its advances to merchants and manufacturers in London.⁴ The business of the London private banks expanded at the same time that foreign merchants fleeing the extortions of Napoleon's troops brought their affairs to London.⁵ Country banks multiplied in great number and profited by issuing small-denomination banknotes to replace metal coinage in the domestic circulation after the Bank of England suspended convertibility in February 1797, and the restrictions against issuing small-denomination notes were suspended in March 1797.⁶ In short, the entire British financial sector enjoyed prosperity on the basis of war finance.

True, the commercial crisis of 1810 brought the Bank of England's prosperity—and arrogance—under close scrutiny by its enemies and led to the Bullion Report of 1810. By undermining the intellectual authority of the Bank's directors, the Bullion Report provided the courage needed for subsequent governments to constrain the Bank's power and to overrule its recommendations on monetary matters if that became politically popular. The Bank's

practical autonomy, however, remained intact as the government still relied on it for managing its remittances and, especially, its recurrent issues of debt—both long-term, funded debt (perpetual annuities comprised mainly of 3 percent consols) and short-term, unfunded debt (one-year Exchequer bills bearing daily interest). The Treasury at this point was the Bank's strongest defender against the criticisms of the Bullionists, arguing that the needs of war finance justified the fall in the exchange rate of the paper pound.

As a result, for three years after the signing of the peace treaty in Paris in 1815, the government acquiesced to the Bank's various arguments that resumption of cash payments should be delayed—whether until the exchanges had stabilized, or the bond market had strengthened, or foreign trade had picked up, or its gold reserves were increased. Finally, in 1819, the government initiated a bill to force the Bank to resume convertibility, after initial experiments in 1817 at limited convertibility of Bank notes had succeeded without any harmful consequences. Even so, the Bank managed to make the transition as difficult as possible, first by amassing a large stock of gold, which helped keep up the price of gold in the markets, and then by withdrawing the notes from circulation that the government used to repay £10,000,000 of Exchequer bills that had been held by the Bank. Further, it refused to lower its rate of discount on bills and notes even as its loan business to the private sector declined. The resulting price deflation intensified both agricultural and manufacturing distress but enabled the Bank to resume full convertibility of notes into coin in May 1821 and to skip almost entirely the intermediate step of limiting convertibility to ingots of 60 ounces, as proposed by Ricardo. While, at the time, Ricardo criticized the Bank's directors as “indeed a very ignorant set,”⁷ it appears to later historians that the Bank was responding angrily to the government's efforts to use the Bank to support its short-term debt financing while taking away the Bank's power to control the level of its own liabilities.⁸

⁴ Clapham (1945), vol. 2, ch. 1.

⁵ Chapman (1984), p. 4.

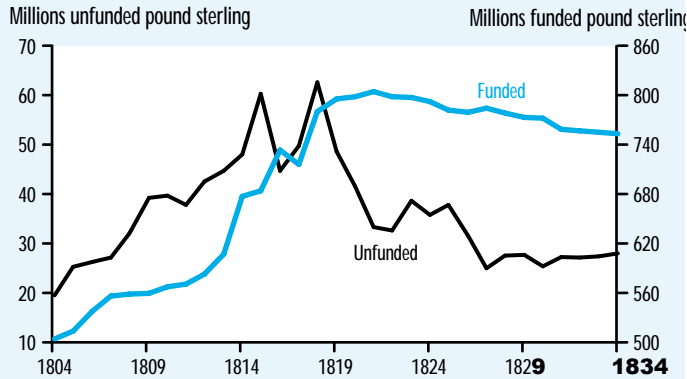
⁶ Pressnell (1956), pp. 142-44.

⁷ Letter to Malthus of July 19, 1821, in Works, IX, p. 15 as cited in Fetter (1965), p. 98, n. 11.

⁸ Hilton (1977), p. 54.

Figure 1

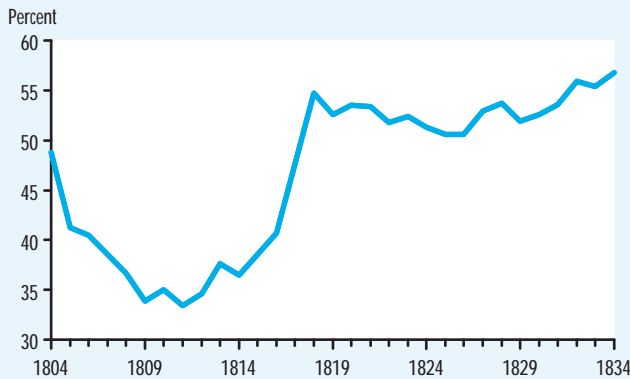
National Debt of the United Kingdom Nominal amounts, funded and unfunded



SOURCE: Mitchell (1988), p. 601.

Figure 2

British Debt Service/Revenue



SOURCE: Mitchell (1988), pp. 581, 587.

The elimination of the income tax in 1816 brought an end to the mutually agreeable arrangements between the Treasury and the Bank that had existed during the war. The fall in tax revenues meant a sharp rise in the ratio of tax revenues that the government had to devote to servicing the huge debt accumulated during the war. Figures 1 and 2 show clearly the rise in government debt during the war, the ease with which the mounting debt was serviced while the income tax existed, and then the constraint upon the government's peacetime budget created by the continuing debt service. In the absence of an emerging revenue source,

it was a serious shock to the Treasury to lose an income stream that had amounted to nearly 20 percent of its total gross income in 1816 (£14.6 million) and had virtually vanished by 1818.⁹ This was the shock that forced readjustment throughout the entire British financial system, from the Treasury right down to the country banks.

The Treasury confronted this situation with a variety of ploys. One was to raise the price of its long-term bonds in the London Stock Exchange so that new debt at lower interest rates could be issued in order to reduce its expenditures on debt service. It preferred to reduce this form of expenditure rather than cut back on traditional sinecures of the royal family and the landed aristocracy or reduce further the army and navy. Expenditures had to be cut not only because the repeal of the income tax had reduced revenue, but also because of the fear of further losses of revenue that might follow from reductions in various customs duties and excises. Counterarguments that both foreign trade and domestic commerce would increase in response to lower tax rates enough to generate the same revenues as before failed to persuade a timorous government. A few experiments were tried, some of which proved successful, but in the prevailing disturbed monetary conditions, any reductions in protection levels were vehemently opposed by manufacturing interests. The government was forced to find its budget balance in reduced debt service. By 1821, it became increasingly possible to do this.

Figure 3 shows the course of prices for the major government "stocks," namely the price of 3 percent consols, Bank of England stock, and East India Company stock, over the period 1811-31. The price of consols, with their constant £3 interest payment each year, reflects inversely the default risk-free yield on long-term debt. Its pattern shows clearly the increasing pressures of war finance during the Napoleonic Wars and the rocky road traversed by the British debt overhang in the decade-and-a-half following Waterloo. In the period encompassing the resumption of specie payments, from late summer 1820

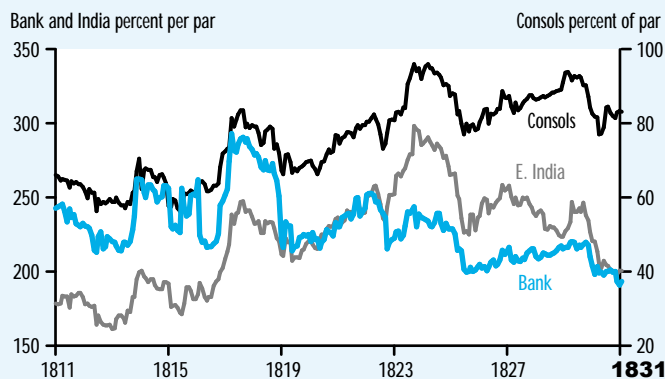
⁹ Mitchell (1988), p. 581.

to late 1822, the price of all three securities rose. With their dividend rates maintained at wartime levels, this meant the market yields on each fell for first-time investors. The actual market yields available to investors in “the funds,” as they were known at the time, are shown in Figure 4. There was clearly a period of marked recovery from the trade crisis of 1819, when it was finally determined that Bank of England notes would again be convertible into gold at the pre-suspension mint par. A check occurred, however, at the end of 1822 that lasted until the spring of 1823. Then the upward course in price (and fall in yields) resumed for a year, leveling off from March to September 1824. The government’s success in managing its debt service problem after resumption led to unusually low interest rates, especially in 1824, the year preceding the boom and bust of 1825.

The charts of prices and yields for “the funds” illustrate nicely the problems created by the transition from wartime to peacetime finance. The price patterns of the three major securities available to risk-averse British investors changed their relationship from moving in synchrony to diverging unpredictably. The capital stock of both the East India Company and the Bank of England was invested in permanent government debt, on which the government paid regular interest. Typically, the two chartered monopolies passed this interest payment through to their shareholders along with some part of the profits obtained from their business activities. The dividends declared by the two had increased over the eighteenth century but rose to all-time highs during the Napoleonic Wars. The Bank’s business as the remitting agent for the government’s war finance has already been mentioned. The East India Company gained from absorbing all the Asian trade previously serviced by the French and Dutch East India companies while the hostilities lasted. However, it was assessed a huge annual sum by the government, purportedly as compensation for the naval and military services the government provided for the protection of the East India Company’s trade.

Figure 3

**Price Levels of “The Funds”
London Stock Exchange**



SOURCE: Neal (1990), data appendix.

Figure 4

**Yields on “The Funds”
End of month**



SOURCE: Derived data from Figure 3.

A crossover in the prices of Bank of England and East India Company stock emerged clearly at the beginning of 1823 and widened through 1824. Part of the decline in the Bank’s stock was certainly due to its decision in 1823 to drop its semi-annual dividend from £5 per £100, which had remained constant from 1807 through 1822, to £4 per £100. It remained there through 1838 before dropping again.¹⁰ The Bank was steadily withdrawing from its discount business, husbanding reserves, and fending off Parliamentary pressures to resume convertibility. The East India Company, meanwhile, was in its final phases as a trading company in the period 1813-33

¹⁰ Clapham (1945), v. II, p. 428.

and faced with a mounting problem of encroachment by noncompany English traders in the exports of Indian goods to Britain. To counter this, the company was allowed to maintain its monopoly on all British trade with China. It was in the 1820s that the company's import of trade goods from India began to feel the pressure of competition—in 1826-27, they imported no goods whatsoever from India.¹¹ So it was the prospects of the continued China monopoly, and the earnings on monopoly pricing of tea for British consumers, that raised its market value in the early 1820s and the decline in Indian trade that lowered it in the mid-1820s.

In the early period from 1811 until 1819, by contrast, the London stock market had established a stable price relationship among the three securities. The market yield on East India Company stock was always the highest of the three. Presumably, this situation reflected the higher risk associated with the stock. The directors succeeded in keeping the dividend rate high at a steady 10 percent per annum throughout this period, but there was always a high risk that the government would either increase its charges on the company or reduce its source of profits, say by returning Ceylon and Indonesia to the Dutch. The much lower market yield on Bank stock reflected the perception that the Bank's business with the government was assured and even less risky than the government's financial affairs. The Bank stock's market yields were always lower than those available from the 3 percent consols, at least until 1819. This is not as counterintuitive as it may first appear, because the amount of Bank stock was fixed by terms of its most recent charter, while the supply of "Three Per Cents" kept changing unpredictably with the shocks to the government's finances.

All this changed, however, with the Resumption Act of 1819. The Bank's stock was assessed by the market to be then as risky as that of the East India Company. The success of actual resumption in full in 1821 appears to have reassured the market that it was less risky than the stock of the East India Company, whose fate was still a matter of intense discussion and dispute.

At times, Bank stock even appeared less risky than consols. The crisis of 1825 disrupted further the price and yield patterns. Thereafter, consols were clearly judged the safest security, East India Company stock became priced with a higher risk premium yet, and Bank stock was priced with a risk premium that seems to have risen steadily toward the fateful year of 1833, when its charter was up for renewal.

It may be helpful to put this argument, derived from visual inspection of the price and yield charts, in terms more familiar to modern financial analysts. The visual evidence is that the three major components of "the funds" were co-integrated in the period up to 1819 and presumably for a number of years before 1811. At some point in the period of conflict between the Bank and the government over the timing and terms of resumption of cash payments, from 1819 to 1821, the co-integrating relationship was broken. Table 1 presents the results of some formal testing of the statistical hypotheses implied by this argument.¹² The top panel demonstrates that the prices of all three securities probably followed random walks, both during the period 1811-20 and the period 1821-30. This is reassuring evidence that the market was at least weakly efficient in pricing each security. That is, there was no obvious trading rule that investors could use to make consistent profits by knowing when prices would rise or fall.

The second panel shows the results of Dickey-Fuller tests to see if there existed co-integrating relationships between each pair of securities in each subperiod. These indicate that co-integration did exist between 3 percent consols and both Bank of England stock and East India Company (EIC) stock in the first subperiod, 1811-20. This is sensible, as the dividends for both the Bank and the EIC rested in large part on passing through the interest payments each company received from the government. However, no co-integration existed between Bank of England and East India Company stock. This is also reasonable, because each company's additional earnings above the interest payments received from the government were

¹¹ Clapham (1967), p. 487.

¹² I am grateful to Marc Weidenmier for his expertise in carrying out this analysis for me.

Table 1

Co-Integration of the Funds and Market Index on the London Stock Exchange: 1811-20 and 1821-30

Panel A. Integration Diagnostics

1811-20	D-F Test	ADF	1821-30	D-F Test	ADF
Bank of England	-2.01	2.08	Bank	-1.35	-1.26
EIC	-1.02	-1.11	EIC	-1.40	-1.50
Threes	-1.68	-1.54	Threes	-2.30	-2.20

Panel B. Dickey-Fuller Tests for Co-Integrating Regressions

1811-20	D-F Test	1821-30	D-F Test
Bank of England vs. EIC	-2.15	Bank vs. EIC	-1.80
Threes vs. Bank of England	-2.65*	Threes vs. Bank	-2.18
Threes vs. EIC	-4.27***	Threes vs. EIC	-2.01

Panel C. Johansen Tests for Co-Integrating Vectors

1811-20	λ MAX	λ TRACE	1821-30	λ MAX	λ TRACE
Bank of England vs. EIC	5.33	7.08	Bank vs. EIC	5.02	7.09
Threes vs. Bank of England	19.27**	26.70***	Threes vs. Bank	5.59	8.10
Threes vs. EIC	15.35*	26.93***	Threes vs. EIC	5.09	6.81

NOTE: The Dickey-Fuller statistics reported under the integrating diagnostics and the co-integrating regressions are the t-statistics to test if the residuals are stationary. Critical values are based on James Hamilton (1994), *Time Series Analysis*, Princeton University Press, Table B-6, Case 2, p. 763.

Critical values for the Johansen statistics are taken from Osterwald-Lenum (1992).

* denotes 0.10 or less probability that there was a unit root,

** a 0.05 probability, and

*** a 0.01 or less probability.

determined independently of each other. But even the co-integration of each company's stock with consols disappeared in the second subperiod, 1821-30.

Because the length of each time period is relatively short by the standards of time series statistics, and the Dickey-Fuller statistics are relatively inefficient for small samples, the third panel uses the Johansen technique for testing for the existence of a co-integrating vector for each pair of securities. Again, it shows that such vectors likely did exist in the first subperiod between consols and both Bank of England and East India stock, but not between

Bank of England stock and East India stock, while no co-integration among any of the funds is evident in the second period. This reaffirms my argument that the transition from war finance to peace finance disrupted all the relationships within the entire structure of the British financial system, especially from 1821 on.

THE CRASH: FROM LATIN AMERICAN BONDS TO COUNTRY BANKNOTES

Eventually, the government managed to bring the government budget back into

balance and even run a small surplus, thanks mainly to reductions in the armed forces, especially the withdrawal of occupation forces from France after 1818. But in the period immediately following resumption of the gold standard, the government continued to make payments into the Sinking Fund, which was used to make periodic purchases of long-term debt at market prices and retire it. In effect, the Treasury was running open-market operations that increased liquidity in the economy. It did this by issuing Exchequer bills to the Bank and then using its credits with the Bank to retire some of the funded, long-term debt, mainly consols. Encouraged by the possibilities of retiring high-interest debt and reducing expenditures in this way, the government overreached in 1823. At the end of that year, the government converted £135 million of its 5 percent bonds to 4 percent bonds. It then continued to take advantage of monetary ease early in 1824 by converting £80 million of the 4 percent bonds to 3½ percent.¹³

This had a double-barreled effect, according to traditional accounts. On the one hand, British investors were disappointed to be receiving lower yields on their holdings in “the funds.” “Even in that day ‘John Bull could not stand two per cent.’”¹⁴ On the other hand, the Bank of England was now obliged to buy back the “deadweight” part of the annuity yielding 3½ percent that the government had issued to cover its expenditure on naval and military pensions but had failed to sell to the public. The Bank had ample reserves to accomplish this, having accumulated bullion for minting into coins to replace the £1 and £2 notes it had issued during the paper pound period (1797-1821). In fact, as late as October 1824 the Bank’s reserves amounted to fully one-third of its liabilities, and by February 1825 it had increased its holdings of public securities by 50 percent from the low of February 1822.¹⁵

This increase meant the Bank was also conducting open-market operations, inadvertently and unwillingly, that added to the monetary ease by placing cash in the public’s hands in exchange for the government securities they previously held. This was

done at the same time the Bank was drawing down its excessive gold reserves, a process that also increased public liquidity. John Easthope, a member of Parliament and a stockbroker, in his testimony to the Committee on the Bank of England Charter in 1832, argued that while the increase in the Bank’s note issue before 1825 was not so large, it should have been decreased in light of falling gold reserves.¹⁶ The episode he referred to was very likely the operations of Nathan Rothschild, who took advantage of the falling price of gold in Britain to borrow a large amount from the Bank to sell in France in November 1824.¹⁷ Later, in mid-1825, when the Bank became concerned about its falling reserves and the fall of stock prices, Rothschild agreed to repay the loan, restoring the gold in installments spread over the months of June, July, August, and September.¹⁸ The result was exceptional monetary ease in 1824 and into 1825, and then contraction in mid-1825, helping to bring on a payments crisis for country banks.

As Easthope argued, this was not the behavior one would want from a bank devoted to public service, although it was understandable behavior for a bank more concerned about the dividends it could pay to its stockholders than the general state of the monetary regime. On this point, the Bank’s defense was that the exchanges had turned against Britain in 1825, so it was necessary then to contract its note issue and restore its gold reserves. Yet the evidence produced by the Bank itself for the committee indicates that the exchange rate was never seriously threatened (see Figure 5), at least no more than in earlier and later fluctuations that were not accompanied by financial panics. Indeed, such fluctuations as occurred may have created profit opportunities for the House of Rothschild, which the Bank was only too happy to share in part without taking the risks incurred by Rothschild.

The dysfunction of the financial system created at the top by the separation of operations and objectives between the Bank and the Treasury spread even further, affecting the country banks. Confronted by the dis-

¹³Gayer, Rostow, Schwartz (1975), vol. I, p. 185.

¹⁴Gayer, Rostow, Schwartz (1975), vol. I, p. 185.

¹⁵Pressnell (1956), p. 480.

¹⁶Great Britain (1968), p. 469, item 5790.

¹⁷Bank of England, Committee of the Treasury Minute Book, Oct. 29, 1823, to April 12, 1826, fo. 117. Rothschild on Nov. 30, 1824, requested a loan of £300 or £500,000 of bar gold at 77/10 1/2 per oz. and paid 3 1/2 percent per annum with collateral of stock. “As I may require about £225,000 value of Bar gold tomorrow, I beg to mention it to you, in order to facilitate the delivery.” The Bank’s Court of the Treasury complied with this application.

¹⁸Bank of England, Committee of the Treasury Minute Book, May 26, 1825, fo. 161.

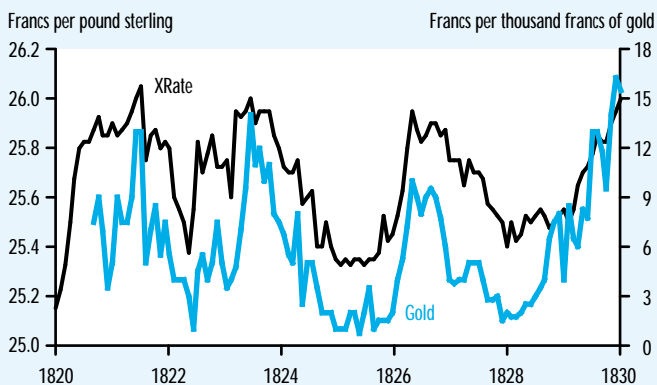
stress caused by severe and unanticipated deflation in 1819-21, the Treasury did not wish to renew its reliance upon the Bank for buying Exchequer bills, as it had done in 1817 in order to finance public works projects in the manufacturing districts and Ireland.¹⁹ Instead, it allowed the country banks to continue to issue notes of small denominations, deferring their elimination from circulation for 10 years. Instead of disappearing from the money supply in 1823, as previously provided in legislation of 1816 (which mandated their termination within two years after the Bank resumed cash payments), such notes were allowed to continue circulating until 1833.

The country banks, already providing necessary finance to manufacturing districts throughout England by the second half of the eighteenth century, found their business prospects greatly enhanced during the Napoleonic Wars.²⁰ Part of the reason was the expansion of heavy manufacturing in the Midlands and South Wales, part was the growth of foreign trade from outports other than London, and much was due to the role of country banks in remitting to London the government's revenues from the land tax, the stamp tax, and the income tax while it was in effect. The end of the war reduced the basis for all these activities and eliminated the income tax. Moreover, the continuing threat of cash resumption by the Bank of England meant that the profitable note-issue business would have to be wound up and replaced by some other form of revenue.

Into the breach stepped the stockjobbers and brokers operating in the London stock exchange. Their business, too, was greatly enhanced by the incredible increase in government debt issued during the wars of 1793-1815. It was interrupted briefly by the crisis of 1810, which foretold the difficulties the stock exchange traders would face when the war ended. In 1811, the response of stock traders was to enlarge greatly the list of securities available for investors in the London stock exchange. Canal stocks were especially favored, although a few other joint-stock companies were listed—iron-tracked railways, docks, waterworks, and a few gasworks. Trading

Figure 5

London's Exchange on Paris and Premium on Gold at Paris



SOURCE: Great Britain (1968), Appendix 97, pp. 110-11.

in most of these public-utility stocks was quite limited, however, as most shareholders preferred to hold them for their value as long-term assets and for their voting power. The various forms of government debt remained the most lucrative source of commission and speculative income for traders.

Latin American Securities

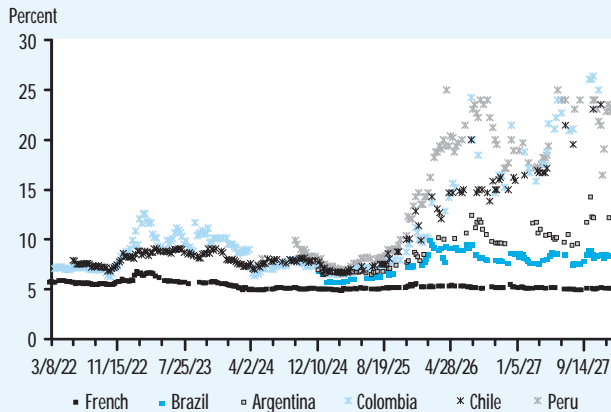
The withdrawal of foreigners from the British national debt after the war, however, removed one class of customers that had been most active in trading, while the rise in the price of government bonds reduced their attractiveness as sources of interest income to the *rentier* classes. The traders on the stock exchange began to develop a variety of new assets to maintain their customer base and their personal incomes. New government issues that mimicked in form the British 3 percent consol were offered by the peacetime governments in France, Prussia, Spain, Denmark, Russia, and Austria. The military successes of the revolting Spanish American colonies stimulated offerings of government bonds from the new Latin American states as well, followed by stocks in newly privatized mines. Many more gasworks were listed as every community in England rushed to provide its residents and businesses the gas light-

¹⁹ Hilton (1977), pp. 82-87.

²⁰ Pressnell's classic study (1956) remains the standard work on English country banks.

Figure 6

Yields of Latin Bonds



SOURCE: *Course of the Exchange*, Friday quotes.

ing that was proving so successful in London. A number of insurance companies were created when entrepreneurs saw that the existing companies seemed especially able to profit from the ease of credit and the lack of attractive alternative assets to government debt.

But the most attractive assets offered were those from Latin America, following the success of the French 5 percent *rentes*. Following the final defeat of Napoleon at Waterloo in 1815, capital flowed back to the Continent from Great Britain. Foreign holdings of British debt diminished rapidly, the price of consols rose as the supply diminished, and prices of Bank and East India stock rose in tandem. British investors used to safe returns ranging between 4 and 6 percent for the past 20 years now found their options limited to yields between 3.5 and 4.5 percent. The opportunities for investment in new issues of French 5 percent *rentes* were more attractive than continuing their holdings in consols. Figure 6 shows that the *rentes* maintained a steady return over 5 percent throughout the crisis period and offered a stable alternative to the British funds. Baring Brothers and Co., by its successful finance of Wellington's army in 1815, had established itself as the dominant merchant bank in England. By undertaking the flotation of the first two issues of French *rentes* sold to pay the reparations and sup-

port Wellington's occupation forces, Barings became the "Sixth Power" in Europe, according to the Duc de Richelieu.²¹ From February to July 1817, Barings disposed of three loans, the first two at a net price of 53 percent of par for 100 million francs each and the third at 65 percent of par, which raised 115 million francs. Yet, according to the historian of the Baring firm, no disturbance in the British trade balance or in French reserves seems to have occurred—the inflow of capital to France from Britain resulting from the issue of *rentes* seems to have been offset by indemnity payments and army contracts from France to Britain.²² (What the historian has missed, of course, is the fall in the exchange rate of the British pound that occurred at the time; the pound was still floating after the suspension of convertibility in February 1797.) From this success for British investors in foreign investment with the French *rentes*, it has traditionally been argued, came increased enthusiasm for other forms of investment, first in the bonds issued by the new government of Spain established in 1820, and then in the bonds issued by the new states emerging in Latin America.²³

The collapse of Spanish control over its American empire during the Napoleonic Wars led to a variety of independent states being formed out of the former colonies by 1820. Battling one another for control over strategic transport routes, mainly rivers and ports, and over state enterprises, mainly mines, each appealed to foreign investors as a source of government finance and as a means to substitute foreign expertise and technology for the vanquished Spanish. Their government bonds and their mining shares found a ready market in the London Stock Exchange, which had become the dominant marketplace for finance capital in the world during the Napoleonic Wars. The loan bubble of 1822-25 ensued, eventually giving British foreign-bond holders their first experience with defaults by sovereign states. None of the new Latin American states emerging from the remains of the Spanish empire (Brazil remained part of the Portuguese empire) found the means, whether by exports or taxes, to service the debts they had incurred in London. Mean-

²¹ Jenks (1927), p. 36.
See also Ziegler (1988), pp. 100-11.

²² Jenks (1927), p. 37.

²³ While the focus for foreign loans was mostly on Spain and Spanish America, Greece received a loan and much-needed publicity for its then-premature efforts to break away from Turkish rule. More than 50 years later, when the Greek government was attempting to assure the international community it would go on a gold standard, part of its commitment was to resume payment on these initial bonds!

while, the net proceeds they had received after the bonds were sold at discount—and after they had paid large commissions up front—to the London investment houses were dissipated rapidly in military conflicts with neighboring states.²⁴

From 1822, when both Chile and Colombia floated bond issues with London agents, an increasing number of Latin American governments tried to find the means for financing their transition to independence from the flush pockets of British investors. The bonds they issued, in terms of the amounts actually paid up, as distinguished from the amounts actually received by the governments, were the largest single category of new investment in the London capital market in this period.²⁵ It is true, even so, that the amount was small relative to the remaining sum of the British government's funded debt—£43 million compared with £820 million.²⁶

Figure 6 compares the prices of several bond issues of the emerging South American states, as given in James Wetenhall's semiweekly *Course of the Exchange*.²⁷ At the peak of the stock market boom, there was surprising convergence in the prices of all the Latin American bonds. It was only in the ensuing two years that information on the fiscal capacity of the individual governments and their respective economic bases enabled the London market to distinguish among them. Mexico and the Andean countries were clearly marked to be disaster cases by the end of 1828, while already Argentina and Brazil were demonstrating their attractiveness to British investors, an allure that would increase until the Baring crisis of 1890.

The pricing pattern of foreign government bonds displayed in Figure 6 is a classic illustration of the so-called “lemons” problem that can occur in emerging financial markets. In this case, it appears that investors in the London market priced the Latin American bonds at a substantial discount so that the typical 5 percent or 6 percent yield on par value could provide a substantial risk premium compared with both the British funds and the now-seasoned and solid French government

debt. Until further information came in from newspapers or merchants' letters from the respective countries concerning their fiscal situation and credit arrangements, however, they all looked much alike, and all were priced at punitively low levels. This discouraged higher-quality governments, perhaps Brazil, from issuing debt until the House of Rothschild had assured itself that adequate provision was forthcoming for servicing it. But it also encouraged lower-quality governments, perhaps Peru, to issue debt early on. Indeed, at one point in October 1822, it induced the Scottish adventurer, Gregor McGregor, to issue bonds from an imaginary government of Poyais, presumably located around Honduras. On October 29, 1822, the official *Course of the Exchange* quoted Poyais scrip for 6 percent bonds at 81 1/2 percent of par, compared with Peru's 6 percent bonds at 86 3/4, Chile's at 84, and Colombia's at 86!

Only as more information came in or as investors began to pull out of higher-risk investments and seek safer, better-quality assets did price differences begin to show up. This change began to occur in the fall of 1825 for the new government issues from Latin America; it did not affect the now-seasoned and secure French *rentes* at all. While the history of the various bond issues is extremely colorful, it appears that Leland Jenks' assessment of many years ago is still fundamentally correct—their main effect was to enrich some issuing agents and impoverish or imperil others, including the redoubtable Barings. Jenks notes that the typical arrangement mimicked that devised by the Goldschmidts for the Colombian loan of 1824, whereby “[t]hey received a commission for raising the money, a commission for spending it, and a commission for paying it back.”²⁸ On the other hand, the most recent historian of Barings argues that they lost money on the Argentina loan by buying back large amounts of it in a futile effort to maintain the market price of the bonds and lost even more on the ill-advised investments in Mexico of Francis Baring, the second son of Alexander Baring.²⁹ In the case of both the Rothschilds and the Barings, however, it appears that the sums risked were relatively

²⁴Dawson (1990) provides a readable account of this episode, but Marichal (1989) puts it into a longer-run Latin American perspective. Brazilian bonds never went into default, which is why their prices remained the highest among the Latin American bonds in the late 1820s. They were, in fact, the only ones issued by the Rothschilds. None of their government bond issues for Austria, Belgium, Naples, Prussia, or Russia defaulted in this period (Doubleday, p. 281).

²⁵Gayer, Rostow, and Schwartz (1975), vol. I, p. 189.

²⁶Gayer, Rostow, and Schwartz (1975), vol. I, p. 408, fn. 8, and Mitchell (1976), p. 402. These are nominal values in each case, but government debt was then trading at close to par, so its market value was roughly the same.

²⁷Beginning probably in January 1825, Wetenhall apparently also began publishing a daily stock price list (No. 171 was for July 8, 1825), with slightly different coverage than that provided in his officially sanctioned, twice-weekly price sheet (which was No. 11,131 for July 8, 1825)—a bit of circumstantial evidence for the information-asymmetry theory, but I have located only one issue of the daily list for this period.

²⁸Jenks (1927), p. 49.

²⁹Ziegler (1988), pp. 102-07.

small and the risks generally appreciated even by an inexperienced British public. We have to look elsewhere for an explanation of the 1825 speculative bubble and collapse, perhaps in the new domestic companies that were formed.

Domestic Securities

As the London stock market had proved attractive for the new issues of debt by the restored European governments and the revolutionary Latin American governments, by 1824 a much wider variety of newly formed joint-stock corporations offered their shares to London investors. In the words of a contemporary observer, “bubble schemes came out in shoals like herring from the Polar Seas.”³⁰ The success of three companies floated to exploit the mineral resources of Mexico—the Real del Monte Association, the United Mexican Company, and the Anglo-American Company led to flotations of domestic projects in early 1824. In February 1824, the Barings and Rothschilds cooperated to found the Alliance British and Foreign Life & Fire Insurance Company. It enjoyed an immediate, enormous success.³¹ In March there were 30 bills before Parliament to establish some kind of joint-stock enterprise, whether a private undertaking for issuing insurance or opening a mine, or a public utility such as gas or waterworks, or a canal, dock, or bridge. In April there were 250 such bills.³²

The limitation of joint-stock enterprises to these fields arose from the limitations, first, of the Bubble Act of 1720, which forbade joint-stock corporations from engaging in activities other than those specifically stated in their charters; second, of common law, which made stockholders in co-partnerships with transferable shares (i.e., unincorporated joint-stock enterprises) liable in unlimited amount, proportional to their shares in the equity of the company; and, third, of the limited liability and ease of transfer for shareholders in mines created on the “cost-book” system.³³ They were subject only to calls up to the capitalization authorized by the cost-book, which required

neither deed, charter, nor act of Parliament to establish. Despite the resistance of Parliament to incorporating new companies with limited liability, the speculative mania continued with new projects floated daily. Speculation was encouraged on the possibility that an enterprise might receive a charter, based on the connections in Parliament of its board of directors.

The extent of the speculative fervor and its lack of permanent effect was spelled out by a contemporary stockbroker, Henry English, and his analysis has remained authoritative to this day. Briefly, English listed 624 companies that were floated in the years 1824 and 1825. They had a capitalization of £372,173,100. By 1827, only 127 of these existed with a capitalization of £102,781,600, of which only £15,185,950 had been paid in, but the market value had sunk even lower to only £9,303,950.³⁴ But even at the height of the enthusiasm for new issues, the total capital paid in had amounted to no more than £49 million.³⁵ Compared with the stock of government debt available (£820 million), this amount was still almost as limited in scale as the investments in Latin American securities. Perhaps we have to look still further for an explanation of the events of 1825. The role of the country banks, in particular, needs to be examined.

The Country Banks

The expansion of the economy continued through 1823 and 1824. By April 1825 at the latest, the stock market boom reached its peak (Figure 7),³⁶ and the resulting drop in collateral values, combined with a contraction by the Bank of England in its note issue, began to create jitters in the money market. By July, city bankers were beginning to be more cautious. In September, reports of difficulties by country banks in Devon and Cornwall began to appear. All country banks were then faced with the seasonal strain that occurred each autumn. Government tax revenues were required to be remitted to London in the autumn before interest payments on government debt were made in December. This caused more country banks to fail in October and November

³⁰ Hunt (1936), p. 30, quoting a letter to *The Times*, April 20, 1826.

³¹ Hunt (1936), p. 32.

³² Hunt (1936), p. 32.

³³ Burt (1984), pp. 74-81 describes the cost-book system and its advantages for investors at this time.

³⁴ As reproduced in Hunt (1936), p. 46.

³⁵ Gayer, Rostow, and Schwartz (1975), vol. 1, p. 414.

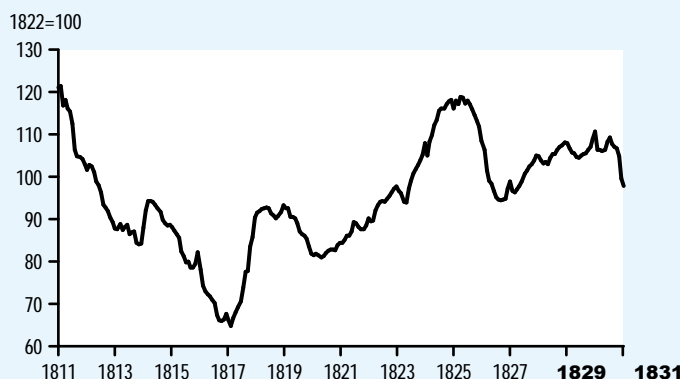
³⁶ According to my own value-weighted index of 50 of the most important stocks traded on the London Stock Exchange, the peak occurs in March. Gayer, Rostow, and Schwartz (1975), using different weights for the same stocks, put the peak in April, although the actual peak if mine stocks are included occurs in January 1825.

in 1825. When the major London banks of Wentworth, Chaloner, & Rishworth and Pole, Thornton & Co. failed on December 8 and 13, respectively, and forced dozens of correspondent country banks to suspend payments, a general run began on country banks. These banks, in turn, came to their London banks for cash, and the London banks turned to the Bank of England. Finally, the directors of the Bank woke up to the crisis and began to discount bills and notes for their customers as fast as they could with diminished staff and resources. The pressure on the Bank lasted for the rest of December, depleted their bullion reserves, and forced them to issue small £1 and £2 notes again but did not force them to suspend payments as they had feared.

The credit collapse led to widespread bank failures (73 out of the 770 banks in England and even three out of the 36 in Scotland)³⁷ and a massive wave of bankruptcies in the rest of the economy, reaching an unprecedented peak in April 1826.³⁸ The Bank of England and the London private banks joined forces for once by blaming both the speculative boom and the subsequent credit collapse on excessive note issue by the country banks. They argued that the ease of note issue had encouraged the more careless or unscrupulous partners in country banks to invest in high-risk, high-return financial ventures such as the Poyais scrip that were being offered on the London capital market. The historian of British country banks, L.S. Pressnell, discounts this factor as the driving force both in the boom and in the timing of the collapse. Relying on evidence supplied by Henry Burgess, secretary of the Committee of Country Bankers, to the Bank Charter Committee of 1832, Pressnell notes that many country banks did increase their note issue substantially between July 1824 and July 1825. Burgess' unweighted index of the indexes of note issues provided to him by 122 country banks for the month of July in each year from 1818 through 1825 gave an overall average increase of 6.7 percent in the final year before the crisis, while 50 of the banks showed increases of more than 10 percent.³⁹

Figure 7

London Stock Price Index



SOURCE: Compiled by author from quotes for 50 companies in the *Course of the Exchange*.

Figure 8 shows, however, that the final level, reached in July 1825, was barely above the initial level of July 1818, which had fallen sharply until 1822. No doubt the country banks expanded their note issue in the years immediately preceding the crash. But much of this expansion was simply restoring note issue that had been reduced in response to Parliament's acts of 1816 and 1819. What is missing, of course, is evidence on the extent to which the initial withdrawal of notes was compensated for by an increase in demand deposits. If there was a one-to-one compensation (which is highly unlikely), then the expansion of note issues in 1824 and 1825 may have helped fuel the speculative fires burning on the London stock exchange. However, the expansion may also have been compensated by a reduction of deposits. Burgess' figures were collected from banks operating in 1830, which clearly had not been among the unfortunate firms that disappeared in the aftermath of the crisis. If those firms were much more aggressive than the survivors that appear in Burgess' large sample, then the country banks may remain indicted as a major contributing cause of the crisis of 1825.

Pressnell gives balance sheets from a handful of country banks that were operating in this period and whose records have survived. The bank Barnard & Co. of Bed-

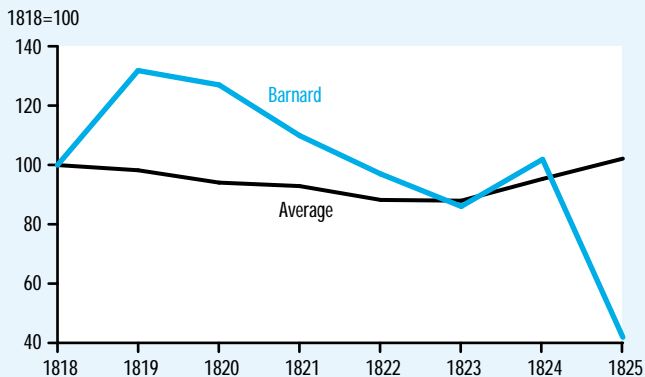
³⁷ Kindleberger (1984), p. 83.

³⁸ Gayer, Rostow and Schwartz (1975), vol. 1, p. 205.

³⁹ Pressnell (1956), pp. 480-81.

Figure 8

Average Issues of 122 Country Banks and of Barnard & Co. of Bedford



SOURCE: Great Britain (1968), pp. 414-16 and Pressnell, pp. 512-13.

ford had an unusually rich set of accounts covering the entire period from 1800 to 1845. On the asset side of the balance sheet, this bank increased its cash holding substantially in 1821-23, and then greatly in 1824. By the end of 1825, however, its holdings had fallen from £106,559 to only £31,201, the lowest level since the crisis year of 1810. While the bank had begun to place surplus funds with a London bill-broker in 1823, this account remained quite small until the 1830s. Total assets fell sharply in 1825, from £152,585 to £109,079, but they fell less than the cash account. The difference came primarily in a doubling of the bank's balance with its London correspondent, from £33,877 to £66,256.⁴⁰ Apparently, this bank was one of the solvent banks whose surplus funds could be channeled to others through the intermediation of its London bank.

On the liability side of the Barnard & Co. bank, the note issues followed much the same path as the average shown by Burgess for his sample of 122 country banks (see Figure 8). The most striking difference occurs in 1825, but this is mostly explained by the fall that must have occurred in the note issue of all the country banks between July, for the average of the 122 banks, and December, for Barnard & Co. As far as Barnard's deposits are concerned, they fell as well from 1818 through 1823, but not as

much as note issues. Deposits rose in 1824 more than note issues, and although they fell in 1825 as well, they ended the year of 1825 at a higher level than note issues. This was a bank that stayed clear of the speculative frenzy going on in London, weathered the storm and survived to prosper afterwards. Its good fortune was due, no doubt, to the large loss sustained by the founder, Joseph Barnard, the one time he did place funds in speculative issues available in London. That loss occurred in the crisis of 1810, and Barnard's "warning to those who may succeed me" from that incident was apparently heeded in 1825.⁴¹

If the record of accomplishment of Barnard & Co. may be dismissed as unrepresentative of the "problem" country banks, we can also examine the accounts of a country bank that did most of its business by note issue and that failed in the wave of bankruptcies occurring in December 1825. Figure 9 shows the gross level of £1 notes issued over the period 1817-25 of one of the unfortunates—the country bank of Sarah Crickett in Chelmsford, Essex County.⁴² These do not take account of notes that may have been retired, but by plotting the highest number found for each date (notes were issued weekly) on a semi-logarithmic scale, we can get a sense of how this bank, which seemed to rely more on note issue than deposit accounts, responded to the vagaries caused by the Bank of England's return to the gold standard in 1821.

At first glance, this bank shows quite a different pattern of note issuing from that of the successful banks shown in Figure 8. At the outset of business in 1817, it increased its issue of £1 notes very rapidly (it's interesting that these were still outstanding in 1826 when the holders turned them in to the Bankruptcy Commission), as a startup bank might be expected to do. But then it increased issues rapidly again in 1819, when it made sense for country banks to start withdrawing their notes, given that the Bank of England had resumed cash payments, and the Act of 1816 mandated that country banknotes under £5 should cease entirely two years later. The steady rise of notes in the early 1820s does not show any

⁴⁰ Pressnell (1956), pp. 512-13.

⁴¹ Pressnell (1956), pp. 433-34

⁴² Public Record Office, B3/1008 and B3/1010-1029 contains the files of the Bankruptcy Commission for Sarah Crickett and her **bank**.

similar acceleration until the end of 1825, when the crisis was breaking.

Given the bank's location in one of the richest agricultural districts of England, and the prevalence of small tradesmen and farmers among its noteholders, it may be that the note surges shown in 1817 and 1819 reflect local harvest conditions more than responses to the changes occurring in the London money market. They do occur in the fall of those years. It must be emphasized that these totals are cumulative and take no account of notes that may have been withdrawn when presented to the bank, so they are not comparable to the net issues outstanding, shown in Figure 8.

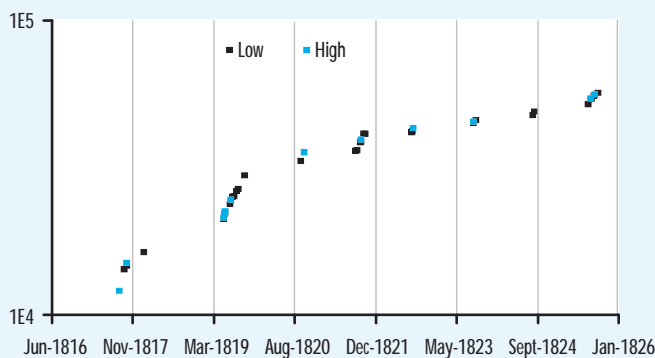
By the end of the Bankruptcy Commission for Crickett's Bank in the mid-1830s, 18 shillings in the pound (i.e., 90 percent) of the claims had been deposited in the assignee's account. Given the small sums claimed by most creditors, the length of time taken by the Bankruptcy Commission, and the location of the assignee's account at the Bank of England in London, much of the funds available for payment were not disbursed—a situation that was convenient for the commissioners and the assignee, who could then cover their charges very easily from the account. But for our purposes, the apparent willingness of so many note holders to retain their notes for long periods of time, plus the bank's basic soundness when its claims and assets were finally realized by the Bankruptcy Commission, indicates that this particular failure was an unfortunate victim of circumstances, not a contributor to the crisis.

The Bank of England

To understand the internal causes of the crisis of 1825, therefore, we must turn back to the role of the Bank of England—in particular, the relationship between its activities as a potential lender of last resort and the wave of bankruptcies that disrupted English commercial life for years following the crisis of 1825. This ground was covered many years ago by Norman Silberling (1923). He simply counted the number of bankruptcy commissions opened as recorded in

Figure 9

Chelmsford Country Bank One Pound Notes Issued

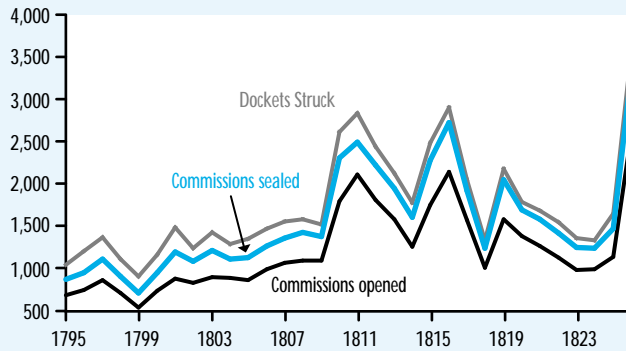


SOURCE: Public Record Office, B/3/1029. Sample figures extracted by author.

the *London Gazette*. These have some weaknesses as discussed in Mitchell (1976, pp. 245-46), Duffy (1985, pp. 331-35), and Marriner (1989), but they are still useful as general indicators of the incidence and timing of bankruptcy over regions and industries. The problems arise from British bankruptcy law, which confined the possibility of bankruptcy to firms engaged in trade, excluding farms, factories, and the other professions. The latter were covered by the much harsher law of insolvency, but in case of difficulty they did what they could to come under bankruptcy law. To do this, they had to be engaged to a significant extent in trade, stop payment on debts amounting to over £100, and refuse in front of witnesses to pay a legitimate creditor. The creditor would then petition with other major creditors to open a commission; this was "striking a docket." If the Bankruptcy Court judged that the creditors had a legitimate case, they would "seal a commission," which would authorize a trio of commissioners to begin collecting evidence of the bankrupt's assets and liabilities. As this was an expensive procedure, which could last for years and eat up the remaining assets of the bankrupt in commissioners' fees, mutual efforts were often made to settle the dispute before the proceedings began. Once they began, the "commission opened." Figure 10, from Duffy (1985), shows the

Figure 10

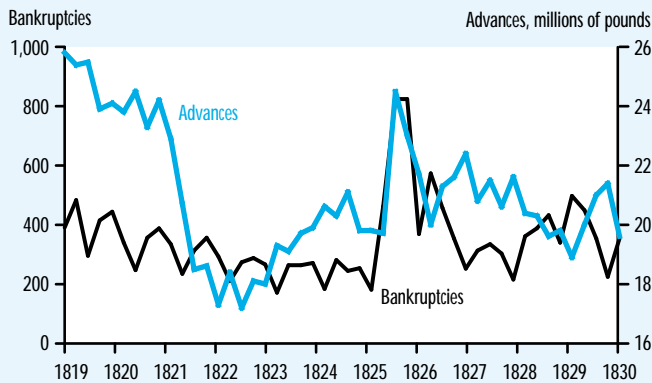
Bankruptcy Records Compared



SOURCE: Duffy (1985), p. 399.

Figure 11

Bankruptcies and Bank of England Advances



SOURCE: Silberling (1923).

comparison of the pattern of advances with that of banknote issue, prices, and bankruptcies, he concluded that advances were a much better barometer of prices and business conditions than banknote issues and, moreover, that in general the claim of the Bank's officers that they followed a real-bills doctrine—responding passively to the demands of business for credit on realized trade contracts—was justified. The exceptional decrease in advances after 1819, driven by the Bank's determination to accumulate sufficient bullion to validate the resumption of convertibility of its banknotes into specie at the pre-war par in terms of gold, did not show up in bankruptcies.⁴³

Closer examination of the relationship between advances and bankruptcies from 1819 through 1830, shown in Figure 11, shows possible encouragement of speculative movements in 1823 and 1824 but moderation in 1825 until the Bank responded to the crisis at the end of the year by increasing the total of advances enormously in the first quarter of 1826. Afterwards, Silberling's figures show a distinctive inverse pattern, which is so short in duration that it could again be consistent with the real-bills story, especially if we allow a lag of six months to a year from the actual credit restriction to the recorded opening of a bankruptcy commission.

Parliament collected evidence in the years afterward to determine the pattern of bankruptcies. Table 2 distinguishes town and country bankruptcies opened within the total of commissions sealed from 1822 through 1833. Again, 1826 shows up as the crisis year, but what is striking here is the much more dramatic jump in the country bankruptcies, a situation that continued afterwards with a consistently higher number of bankruptcies for country banks. Moreover, bankruptcies of banks located within 65 miles of London totaled only 38 from February 7, 1824, to March 22, 1832, compared to a total of 116 for banks located outside the 65-mile radius from London. Only 12 of the London banks failed in the crisis period from December 13, 1825, to March 11, 1826, while 52 of the country banks failed from December 12, 1825, to

⁴³Doubleday rants about the widespread distress created from passage of Peel's Act in 1819 until its final full effect in May 1823, "but, in fact, his prime example of distress . . . calculated to tear in pieces, almost, the heart of every just and sensible man that reads it," deals with the loss of a country estate purchased with wartime profits by the son of a trader who went bankrupt in 1822. Clearly, this was not a general condition.

annual numbers of docketed struck, commissions sealed, and commissions opened. Regardless of which measure of financial distress is taken, the crash of the London stock market at the end of 1825 resulted in record numbers of business failures.

The 1825 spike is all the more anomalous for coming at the end of a period of declining numbers of bankruptcies, with no major changes in trade direction or policy evident, much less any sign of renewed warfare. From 1794 on, Silberling constructed quarterly totals of the advances made by the Bank of England to its private customers and the government. From his

Table 2

**Bankruptcy Commissions Sealed (total) and Opened
(town and country): 1822-32**

Year	Commissions Sealed	Town Commissions Opened	Country Commissions Opened
1822	1,419	468	534
1823	1,250	532	396
1824	1,240	574	396
1825	1,475	683	448
1826	3,307	1,229	1,220
1827	1,688	671	742
1828	1,519	601	620
1829	2,150	809	910
1830	1,720	661	748
1831	1,886	692	770
1832	1,772	643	740

SOURCE: British Parliamentary Papers, 1833, XXXI, p. 342.

March 11, 1826. These bankruptcy records indicate further that the financial panic was transmitted through the credit channels of Great Britain, radiating out from the London capital market, and had its final impact in the trade and industry of the countryside through the liquidity crunch exerted upon the country banks.

Picking up the Pieces

The question naturally arises: Could the Bank of England have prevented this financial disaster, say, by acting earlier and as a monopoly bank bearing more responsibility to the public than to its stockholders? It must bear part of the blame for the expansion of the money supply that apparently arose in 1823-24 and especially for failing to offset the monetary expansion occurring elsewhere. But if, as Duffy suggests, it was the Bank's drawing account activity rather than its note issue that played the strongest role in easing or constraining the credit conditions in the London money market, then the Bank of England can be no more culpable than the country banks. The sums advanced from

the Drawing Office plummeted after resumption of cash payments in 1821, and the Bank of England restricted drawings through most of 1825, never rising to the pre-resumption level until the first quarter of 1826. But this analysis simply casts the Bank of England in the role of just another bank, albeit much larger and more influential. If it was supposed, through its ability, to combine up-to-date, authoritative information from the worlds of finance, commerce, and government policy, it might be expected to have played an earlier, more constructive role. In fact, the evidence from the minutes of the Court of Directors of the Bank indicate that the Bank was taken by surprise and responded with much too little, much too late.

The Bank of England

The first mention of the crisis occurs on December 8, 1825, when "The Governor [Cornelius Buller] acquainted the Court that he had with the concurrence of the Deputy Governor [John Baker Richards] and several of the Committee of Treasury afforded assistance to the banking house of Sir Peter Pole, etc."⁴⁴ This episode is

⁴⁴Bank of England, TVC3/11 G4/48, fo. 150.

described in vivid detail by the sister of Henry Thornton Jr., the active partner of Pole, Thornton & Co. at the time. On the previous Saturday, the governor and deputy governor counted out £400,000 in bills personally to Henry Thornton, Jr., at the Bank without any clerks present.⁴⁵ All this was done to keep it secret so that other large London banks would not press their claims as well. A responsible lender of last resort would have publicized the cash infusion to reassure the public in general. Instead, the run on Pole & Thornton continued unabated, causing the company to fail by the end of the week. Then the deluge of demands for advances by other banks overwhelmed the Bank's Drawing Office.

Table 3 shows the breakdown of the Bank of England's discounts by branch of trade. I have ordered them by the largest amounts disbursed in the quarter ending in December 1825, when "Bankers" dominate. However, as late as November, the bankers were not unusually present in the Bank of England's offices. Indeed, it appears that the merchants engaged in the trades with "Hamburg, France, Spain, Portugal, South America, the Baltic, and General Merchants" were especially pressing in their demands upon the Bank in the quarter ending in June 1825. No other branch of trade showed unusual demands until the final month of 1825. But this alone should have warned the Bank of repercussions that would follow. If it was the South American merchants who accounted for the bulk of the increased demands for accommodation in June, this gave the Bank much better warning than could have been available to any country banker that remittances from South America were in disarray. This would affect the disbursement of dividends upon mining stocks as well as interest on government bonds. Instead of reacting to this information in a constructive way, however, the Bank decided it would be risky to advance funds on some categories of collateral, kept its rate of discount high compared with the rest of the market, and raised its rate of discount back to 5 percent in early December 1825, when demands became increasingly urgent. In

the interim, the Bank chose to respond to the lack of discounting business by cutting costs. The number of clerks in the Drawing Office had fallen from 17 to 11 by February 1825, and of these 11, four were regularly sent to serve in other departments.⁴⁶

The Bank of England's first proactive response at the level of the Court of Directors did not appear until January 12, 1826. At that meeting the court appointed a committee to report on the practicality and expediency of establishing branch banks. The very next week, the committee reported "Branch banks would be highly expedient." The reasons it gave, however, were quite revealing of the ruling mentality among Bank of England leaders at the time. The benefits were listed first for the Bank of England and then for the general public (see Table 4). The practicality was not an issue, given Scotland's experience for 80 years, not to mention the success of the Bank of the United States, the Bank of Ireland, and the recently established Provincial Bank of Ireland.

In this report, the Bank of England was clearly responding belatedly to the government's decision to force it to open branches and to promote large, joint-stock banks. The week after the report was laid before the court, the governor presented to the directors the letter he had received from Lord Liverpool, First Lord of the Treasury, and Mr. Frederick Robinson, Chancellor of the Exchequer. The arguments laid out in the letter show that the government, in this instance, was determined to work around the Bank rather than through it. The Liverpool-Robinson letter began with the assertion, "there can be no doubt that the Principal Source of it [the recent distress] is to be found in the rash spirit of Speculation which has pervaded the Country for some time, supported, fostered, and encouraged by the Country Banks."⁴⁷ So, the letter continued, it seemed advisable to repeal the authority of the country banks to issue small notes and return to a gold circulation. This action would spread pressure on the exchanges over a wider surface and make it felt earlier—a clear reference to the Bank's negligence in 1825. But this alone would not suffice; after all, a similar

⁴⁵ Forster (1956), p. 117.

⁴⁶ Bank of England, C 35/2 4783/2, No. 2, "Special Discount Committee from 12 Feb. 1811 to 26 Jan. 1830 inclusive," fo. 159. Later, the committee recommended a further reduction in the number of clerks (fos. 164-65).

⁴⁷ Bank of England, TVC3/11 G4/48, fos. 201-2.

Table 3

**Amount of Each Branch of Trade in Discounts
(thousands of pounds sterling)**

Branch	Mar	Jun	Sept	Nov	Dec
Bankers	273	595	608	699	3,408
Hamburg, Fr., Sp., Port, S. Amer., Baltic and general merchants	411	1,809	1,094	1,238	2,955
Tea dealers, grocers, and sugar refiners	275	334	324	470	959
Russian merchants and dealers in hemp and tallow	46	95	154	243	733
Blackwell Hall factors and warehousemen woolen drapers	188	337	400	441	701
Linen drapers and Manchester warehousemen	220	300	363	413	594
West India merchants	120	156	196	242	559
Irish merchants, factors, dealers	114	191	201	272	551
Hop merchants	113	130	144	145	503
North American merchants	55	65	164	184	308
Silk men, mfrs. gauze weavers	137	185	226	247	297
Wine and brandy merchants	147	229	158	200	290
Corn factors	137	195	135	135	293
Dry salters	75	118	167	122	234
El agents and merchants	19	93	13	68	226
Leather sellers, factors, tanners	177	254	259	190	224
Stationers	110	141	182	162	210
Timber merchants	81	85	148	160	200
Scotch factors and merchants	58	67	67	51	154
Totals (42 branches in all)	3,080	5,865	5,588	6,324	14,430

SOURCE: Bank of England. C 36/16 TVF 3/25 "Account of the Principal Amounts Discounted in Bills and Note per month for the years 1825 and 1826."

convulsion had occurred in 1793 when there were no small notes and Scotland had "escaped all the convulsions which have occurred in the Money Market of England for the last thirty-five years, though Scotland during the whole of that time has had a circulation of One-Pound Notes." In the past, the Bank of England "may have been in Itself and by Itself fully equal to all the important Duties & Operations confided to it," but "the rise of country banks alone shows it is no longer up to the tasks required from the increased wealth and new wants of the Country."⁴⁸

The government proposed two remedies: The Bank should establish branches of its own, and it should give up its exclusive privilege to issue notes within a certain distance from the Metropolis. The first suggestion was impracticable, in the government's view, and it was obvious that Parliament would never agree to an extension of the Bank's privileges in London. All in all, the government's proposed legislation would remove pressure from the Bank, and it would still have the government's business and be the only establishment at which the dividends on the national

⁴⁸Bank of England, TVC3/11 G4/48, fo. 204.

Table 4

Report of the Bank of England's Committee on Branch Banking**Benefits to the Bank of England:**

- 1) Increase circulation of Bank of England notes.
- 2) Increase Bank's control of whole paper circulation "and enable it to prevent a recurrence of such a convulsion as we have lately seen."
- 3) Provide large deposits.
- 4) Protect the Bank against competition of "large Banking Companies" if the government should encourage them.

Benefits to the General Public:

- 1) Provide more secure provincial circulation.
- 2) "Disasters arising from the sudden expansion and contraction of the currency would not so often occur."
- 3) Increase security and facility of transmission of money.
- 4) Provide secure places of deposit "in every quarter of the Kingdom."

SOURCE: Bank of England. TVC3/11 G4/48 "April 13, 1825, to 6th April, 1826, Minutes of the Court of Directors," folio 194.

debt would be paid. With this condescending argument, the letter concluded, "so we hope the Bank will make no difficulty in giving up their privileges, in respect of the number of Partners in Banking as to any District [left blank] Miles from the Metropolis."⁴⁹

Clearly, the Bank had failed to meet the recent challenges adequately, and the government was determined to create competitive banks that might better serve the public and, presumably, the government. The Bank's response was understandably churlish, which Liverpool informed them on January 25 he regretted, but he was determined to move ahead, merely asking if the government had any amendments to propose to the bill pending in Parliament to permit joint-stock banking. He did then accede to encouraging them to set up their own branches as well. Thus, the Bank went ahead with establishing branches, gradually dispersing seven of them into the industrial cities of Manchester, Gloucester, Swansea, Birmingham, Liverpool, Bristol, and Leeds, starting in 1828, and adding Exeter, Newcastle, Hull, and Norwich in 1829, when small note issues by the remaining country banks ceased. By the time of the Bank Char-

ter Committee in 1832, the branches at Manchester and Birmingham were clearly the most dominant in terms of note issues and bills discounted.⁵⁰

The Commercial Bill Market

Wilfred T. C. King, in his classic study of the London discount market, identified the crisis of 1825 as bringing about "changes in the banking structure which were responsible for every major influence upon market evolution in the succeeding twenty years."⁵¹ His analysis of the crisis follows very much the lines above, adding only the additional factor that a series of good harvests had made the country banks in agricultural districts especially flush with funds. In terms of the conditions in the money market, however, the effects were limited in duration. By June of 1826, the money market rate had fallen well below 5 percent, and the Bank of England was no longer besieged with requests for re-discounting of bills. Of more interest to King were the implications for the development of the bill market in London from four changes in the financial structure that occurred in response to the crisis. These were: 1) the beginnings of joint-stock banking, 2) the establishment of

⁴⁹Bank of England, TVC3/11 G4/48, fo. 215.

⁵⁰Bank Charter Committee Report, Appendix No. 46, p. 47.

⁵¹King (1936), p. 35.

Bank of England branches, 3) the cessation of re-discounting by the London private banks, and 4) the assumption of some central banking functions by the Bank of England.⁵²

The new joint-stock banks had to function outside London (thanks to the resistance of the Bank of England) and they had to compete with existing country banks by attracting deposits rather than issuing notes. King does not explain why this was so, noting only that those joint-stock banks that began business by issuing notes gave them up after a few years. It appears that this development arose in large part because the Bank of England branches refused to do business with joint-stock banks that did issue notes.⁵³ Given that their business was necessarily local and that they had no notes to redeem, the new joint-stock banks kept minimum reserves, relying upon re-discounting bills of exchange to obtain cash when needed to meet withdrawals of deposits. They also had a strong preference for short-term loans in the form of bills, rather than government securities, as had been the case earlier.⁵⁴ As the country banks wound up their small-note business, they also turned increasingly toward deposits and the behavior of joint-stock banks, as described by King. King concludes that it was the period from roughly 1830 until the 1860s or 1870s that the bill market became the most important way in which domestic credit was distributed within Great Britain.⁵⁵

The second change identified by King, the establishment of branches by the Bank of England, also promoted the rise of the bill market. While initially the Bank's branches would seem to be serious competitors to the local banks, they limited their lending activities strictly to commercial bills and then only to very short-term and highest-quality bills, as approved in London. This limitation effectively kept business intact for the existing local banks, save that their commissions on discounting bills were reduced by the knowledge among their customers that the Bank of England branches did not charge commissions. But the facility of making remittances to London and receiving credits back from London through the Bank's

branches helped local bankers use the London bill market more cheaply. A bill drawn locally could now be sent directly to a bill broker in London, who would be instructed to pay the proceeds into the Bank of England for the credit of the local bank at the branch bank. Moreover, a trader in Leeds could pay or receive money in Birmingham through the medium of the Bank's branches, for the "simple charge of postage of a letter."⁵⁶ In short, the branches of the Bank of England greatly improved the payment mechanism that underlay the smooth functioning of the bill market.

The third change noted by King was the withdrawal of London private banks from re-discounting after the 1825 crisis. The run upon the Bank of England—as well as its obvious reluctance to hold too much reserves in gold, which was not earning income for its stockholders—convinced the London banks they should not rely on the Bank of England exclusively for cash in times of pressure. Instead, they turned to providing call loans to bill brokers, who could, in turn, increasingly become bill dealers. Instead of delaying discounting of bills in London until a matching buyer had been found for the bills offered for sale, larger firms could now purchase the bills immediately, using funds on deposit with them by the London private banks.⁵⁷ Only a few firms were as yet large enough to be able to risk this next step, moving from brokering to dealing in bills. Even those like Gurney's probably would not have done it then had not the market rate of discount fallen below the usury limit of 5 percent. Had it been at or above the usury limit, there would have been no possibility of making a profit from strict dealing.

The final step in completing the new structure did not occur until 1830, when the Bank of England opened its re-discount facilities to the bill brokers. Even this was not sufficient to overcome the informational asymmetries that could still arise in the market and that lay at the heart of later crises when the emerging bill market was abused opportunistically. The remaining problem was the Bank's continued refusal to discount at market rates, meaning that it

⁵² King (1936), p. 38.

⁵³ Testimony by Henry Burgess, the Secretary of the Association of Country Banks to the Committee on Bank of England Charter, 5324-26, in Great Britain (1968), pp. 427-28.

⁵⁴ Pressnell (1956) later confirmed this tendency, even for country banks, pp. 415-34.

⁵⁵ King (1936), p. 41.

⁵⁶ Testimony of William Beckett to the Committee on Charter of Bank of England, 1436-38, in Great Britain (1968), p. 101.

⁵⁷ King (1936), p. 64.

was unaware of emerging imbalances in the demand and supply of bills of exchange until a large excess demand for cash showed up at the Drawing Office, as in December 1825. Only when the practice of maintaining fixed discount rates at the Bank was foresaken in the crisis of 1847 did the role of the "Bank Rate" come to play its key regulating role in the British financial system. But the information flows that had arisen through the medium of the bill market enabled the London banks to keep closer tabs on the conditions of the country banks, whether they were in agricultural or industrial districts, essentially through the intermediation of the London discount houses. Further improvements in the management of information flows within the entire financial structure were elicited in response to later financial crises, caused by new, unanticipated shocks encountered as the global economy of the nineteenth century was created.

POLICY LESSONS?

The evidence of the bankruptcies certainly suggest that problems of adverse selection in the London credit markets arose in intensified form during the 1824-25 bubble on the London stock market. Combined with the evidence on changing yield spreads for East India Company stock compared with Bank of England stock, and especially with the evidence of the initial bundling and then wide dispersion of yields on the various Latin American government bonds, it lends support to the hypothesis that the problem of information asymmetry, always present in financial markets, became especially severe in the London markets in the years leading to the crash of 1825.

Asymmetric information is the term applied to the usual situation in which borrowers know more about the actual investment projects they are carrying out than do the lenders. Lenders, knowing this, charge a premium proportional to the uncertainty they feel about the borrowers in question. This situation, in turn, creates an adverse selection problem, in which

higher-quality borrowers are reluctant to pay the high interest rates imposed by the market, while lower-quality borrowers are willing to accept the rates and to default if their ventures fail. In an expanding market, which the London stock exchange certainly was in the boom years of 1806 to 1807 and again in the early 1820s, the availability of loanable funds at premium rates will attract lemons to the market (say, Mexican mines) and discourage borrowing by sound enterprises (say, Brazilian diamonds). Borrowers turn back to internal sources of funds or to a compressed circle of lenders who know their superior quality and are willing to extend credit at lower rates.

In the case of British firms in the 1820s, the compressed circle of knowledgeable, low-interest lenders was the web of country banks that had arisen in the past three decades. The continued credit access of high-quality firms, however, depends in each case upon the continued liquidity of the small, local financial intermediaries. Their willingness to continue lending at preferential rates is limited increasingly by the risk of withdrawals by depositors who wish to participate in the high-interest, high-risk investments available in the national financial market. A financial boom of the kind normally experienced before financial crises can discourage real investment, therefore, and intensify the lemons problem as high-quality borrowers withdraw from the loanable funds market.⁵⁸ It can also place increasing pressure on local financial intermediaries that specialize in monitoring credit to local enterprises. It cannot be mere coincidence that the collapse of the bubble of 1825, according to one account, was set off by the refusal of a country bank in Bristol to honor the request of a Mr. Jones to redeem in gold its notes that he presented.⁵⁹ The *coup de grâce* occurs when higher-risk borrowers are asked to provide collateral for additional loans, and the financial collapse decreases the value of their collateral. The outcome is a general wave of bankruptcies.

Under public pressure, the Bubble Act was repealed in June 1825. In July 1826, joint-stock banks were allowed to establish

⁵⁸ Mishkin (1991), pp. 70-75, gives a detailed exposition of the various routes by which increases in asymmetric information may exacerbate adverse selection, monitoring, and moral hazard problems, especially if a banking panic limits the ability of financial intermediaries to serve a monitoring function.

⁵⁹ Doubleday, pp. 288-89.

beyond a 65-mile radius of London without limitation on the number of partners (the previous limit had been six). Both actions were counterproductive, if we take as given the traditional story that the entire episode was yet another example of irrational speculative bubbles derived from crowd behavior in which investors acted first too optimistically and then too pessimistically in response to fragments of information. On the basis of the information-processing story told above, however, we can conclude that both actions were constructive. Repeal of the Bubble Act sped up the Parliamentary process of granting corporate charters, limiting the speculative period during which uncertainty over the prospects of passage of the proposed charter dominated price movements in the initial share offerings. Moreover, repeal did not mean that shareholders were granted limited liability in the new joint-stock enterprises; unlimited liability remained in principle. Supplementary legislation in 1826 specified, moreover, that Parliament could determine for each charter the extent of liability of the shareholders. With these changes, Parliament both encouraged the continuation of the corporate charter business, which must have been profitable to large number of the members of Parliament, and discouraged overpricing of the subscription shares while the incorporation bill was in progress.

The collapse of country banks was one of the last examples of a banking panic in the British banking system. As Mishkin (1991) argues for U.S. banking panics, bank failures removed from the capital markets the principal monitors who could effectively distinguish borrowers by their quality without resorting to credit rationing or arbitrarily high prices for credit. Bank failures worsened the informational problems in the British capital markets. Creating joint-stock banks within which the country banks would become branches instead of correspondents helped restore this critical monitoring function to the British system. In the peculiarities of the 1826 Act, this was done by linking the various country banks within the structure of a joint-stock

bank headquartered in London. But the London headquarters performed no banking function. Its role was to process and diffuse information to the various branch offices located beyond the 65-mile radius from London.

The results of the financial crisis of 1825 were beneficial for the British government. The funded debt continued to decline, after a small rise in 1827, throughout the remainder of the century. The government's gross income remained high and comfortably above gross expenditures, save for the years 1827 and 1828, when it dropped slightly below.⁶⁰ The comfortable financial situation gave Britain the lowest interest rates on its debt of any European government throughout the nineteenth century—a great advantage whenever it became necessary to mobilize resources for armed conflict anywhere in the world.⁶¹ It also laid the basis for continuing political reform, culminating in 1834, and economic reform, culminating with the repeal of the Corn Laws and the Navigation Acts in the 1840s, and the promotion of limited liability joint-stock corporations in the 1850s and 1860s.⁶²

REFERENCES

- Acworth, Angus Whiteford. *Financial Reconstruction in England, 1815-1822*, London: P.S. King and Son Ltd., 1925.
- Burt, Roger. *The British Lead Mining Industry*, Cornwall: Dyllansow Truran, 1984.
- Clapham, John Harold. *The Bank of England, A History*, Vol. II, "1797-1914," Cambridge: University Press, 1945.
- _____. *An Economic History of Modern Britain*, Vol. I, "The Early Railway Age, 1820-1850," 2nd ed., Cambridge: University Press, 1967.
- Dawson, Frank Griffith. *The First Latin American Debt Crisis. The City of London and the 1822-25 Loan Bubble*, Yale University Press, 1990.
- Doubleday, Thomas. *A Financial, Monetary and Statistical History of England, from the Revolution of 1688 to the Present*, 2nd ed., London: Effingham Wilson, 1858-59.
- Duffy, Ian P.H. *Bankruptcy and Insolvency in London During the Industrial Revolution*, Garland Publishing Co., 1985.
- Evans, George Heberton. *British Corporation Finance, 1775-1850: A Study of Preference Shares*, Johns Hopkins Press, 1936.

⁶⁰ Mitchell (1976), pp. 392, 396, and 402.

⁶¹ See Neal (1992a), pp. 84-96, for a comparison of the British interest rates with the rest of the world.

⁶² See Neal (1992b).

- Fetter, Frank W. *Development of British Monetary Orthodoxy, 1797-1875*, Harvard University Press, 1965.
- Forster, Edward M. *Marianne Thornton, A Domestic Biography, 1797-1887*, Harcourt, Brace and Company, 1956.
- Gayer, Arthur D., W. W. Rostow, and Anna Jacobson Schwartz. *The Growth and Fluctuation of the British Economy, 1790-1850*, 2 vols., Irish University Press Series of Harper & Row, 1975.
- Great Britain. *British Parliamentary Papers, Monetary Policy, General*, Vol. 4, "Session 1831-32," "Report from the Committee of Secrecy on the Bank of England Charter," Shannon Ireland: Irish University Press, 1968.
- Great Britain. Public Record Office, Bankruptcy Commissions, B/3.
- Hamilton, James. *Time Series Analysis*, Princeton University Press, Table B-6, Case 2, p. 763.
- Hilton, Boyd. *Corn, Cash, Commerce: The Economic Policies of the Tory Governments 1815-1830*, Oxford University Press, 1977.
- Hubbard, R. Glenn, ed. *Financial Markets and Financial Crises*, University of Chicago Press, 1991.
- Hunt, Bishop Carleton. *The Development of the Business Corporation in England, 1800-1867*, Harvard University Press, 1936.
- Jenks, Leland H. *The Migration of British Capital to 1875*, Barnes and Noble, 1927.
- Kindleberger, Charles P. *A Financial History of Western Europe*, George Allen & Unwin, 1984.
- _____. "British Financial Reconstruction, 1815-22 and 1918-25," *Economics in the Long View*, Charles P. Kindleberger and Guido di Tella, eds., *Essays in Honour of W.W. Rostow*, Vol. 3, "Applications and Cases, Part II," New York University Press, 1982, pp. 105-20.
- _____. *Manias, Panics, and Crashes. A History of Financial Crises*, Basic Books, 1978.
- King, Wilfred T.C. *History of the London Discount Market*, George Routledge & Sons, 1936.
- Kynaston, David, and Richard Roberts, eds. *The Bank of England: Money, Power & Influence 1694-1994*, Oxford: Clarendon Press, 1995.
- Marichal, Carlos. *A Century of Debt Crises in Latin America: From Independence to the Great Depression, 1820-1930*, Princeton University Press, 1989.
- Marriner, Sheila. "English Bankruptcy Records and Statistics before 1850," *Economic History Review* (August 1980), pp. 351-66.
- Mishkin, Frederic S. "Asymmetric Information and Financial Crises: A Historical Perspective," *Financial Markets and Financial Crises*, R. Glenn Hubbard, ed., University of Chicago Press, 1991, pp. 69-108.
- Mitchell, Brian R. *Abstract of British Historical Statistics*, Cambridge: Cambridge University Press, 1976.
- _____. *British Historical Statistics*, Cambridge: Cambridge University Press, 1988.
- Morgan, E. Victor, and W.A. Thomas. *The Stock Exchange, Its History and Functions*, London: Elek Books, 1962.
- Neal, Larry. "The Disintegration and Re-integration of International Capital Markets in the 19th Century," *Business and Economic History* 21, 1992a, pp. 84-96.
- _____. "Commandite System," *The New Palgrave Dictionary of Money and Finance*, Murray Milgate and John Eatwell, eds., London: Macmillan, 1992b.
- _____. *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason*, Cambridge: Cambridge University Press, 1990.
- Osterwald-Lenum, Michael. "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics* (August 1992), pp. 461-72.
- Pressnell, Leslie S. *Country Banking in the Industrial Revolution*, Oxford: Clarendon Press, 1956.
- Silberling, Norman J. "British Prices and Business Cycles, 1779-1850," *Review of Economics and Statistics*, V, Suppl. 2 (October 1923), pp. 223-61.
- Smart, William. *Economic Annals of the Nineteenth Century, 1821-1830*, Vol. II., Augustus M. Kelley, 1911, reprinted 1964.
- Ziegler, Philip. *The Sixth Great Power: A History of One of the Greatest of All Banking Families, the House of Barings, 1762-1929*, Alfred A. Knopf, 1988.

Michael D. Bordo is a professor of economics at Rutgers University and a research associate at the National Bureau of Economic Research.

Commentary

Michael D. Bordo

England's stock market crash and banking panic of 1825 provide a fascinating story that is considerably relevant to today's policy issues. The story has resonance for three reasons: (1) The crisis was probably the first example of an emerging market-induced financial crisis. (2) It offers an early lesson on the importance of timely lender-of-last-resort intervention by the monetary authorities. (3) It provides valuable insights on the role of information in credit markets, as Larry Neal emphasizes in his paper.

WHAT HAPPENED IN 1815?

At the end of the Napoleonic wars in 1815, the Bank of England began following the deflationary policies required to restore specie convertibility at the pre-1797 suspension parity, leading to successful resumption in 1821. Following resumption, the British economy began a period of rapid expansion, characterized by both an export boom and an investment boom. The opening up of the newly independent states of Latin America stimulated a boom in exports. At the same time, important infrastructure projects (e.g., gas lighting, canals, and railroads) stimulated investment expenditures. The sale of stocks to finance these ventures, in addition to gold and silver mines (some real, some fictitious) in Latin America, and sovereign government debt (initially European and later Latin American) propelled a stock market boom. The Bank of England's easy monetary policy fueled the stock market boom and economic expansion. The Bank was also flush with high gold reserves amassed in the drive to resumption. These aided the British government in servicing and converting some

of its debt to lower yield issues. The increase in the Bank of England notes and deposits in turn served to increase the British monetary base. The country banks then freely issued notes to finance both economic activity and stock market speculation. The stock market boom became a bubble as investors bid up the prices of real and imaginary stocks (e.g., bonds from the imaginary South American Republic of Poyais). Asymmetric information led to adverse selection, and legitimate firms found it more difficult to obtain finance, except at premium rates. Banks infected with the euphoria let down their guard and made risky loans.

As always happens, the bubble burst. It is unclear what caused the April 1825 collapse, but the Bank of England had in March sold a very large block of Exchequer bills, presumably to "contract the circulation" (Clapham 1945). The Bank in succeeding months continued to follow a cautious policy. The collapse of stock prices triggered commercial failures. By autumn (a season of normal financial stress), a number of country banks also failed. When several important London banks failed (e.g., Henry Thornton's bank), a full-fledged panic ensued in early December. The Bank of England then reversed its discount policy and began acting as a lender of last resort. The Bank was saved at the last minute from suspension of convertibility by gold flows from France. However, although the Bank's discount policies were very liberal, it acted too late to prevent massive bank failures, contraction of loans, and a serious recession in early 1826. The English crisis then spread to Europe and also to Latin America, prompting a general default on its sovereign debt.

In the aftermath of the crisis, blame was placed on the country banks for fueling the stock market boom and on the Bank of England for not policing them. Neal views several institutional changes that began in 1826 (e.g., creating branches

for the Bank of England, permitting joint stock banks of issue beyond a 65-mile radius of London, and prohibiting small-denomination country banknotes) as setting the stage for a new financial order in the nineteenth century.

Neal's presentation of the tale, which differs somewhat from my rendition, is convincing, but parts of his story are not clear. Neal views the deflation of 1815-20 as unnecessary. To back up this view, he would need to make the type of purchasing power parity calculations that Officer (1981) did for the United States after the Civil War. Also, it is not clear from Neal's narrative exactly what triggered the crash. Neal's rendition of the story is similar to the Minsky (1977) and Kindleberger (1978) version or the Diamond and Dybvig (1983) view, which asserts that no identifiable trigger exists, and the crash may have been a random event such as a sunspot. It is also not clear why Neal devotes several pages to co-integration tests on the yields of the three funds traded on the London market. Does the break in co-integration shown between consols, East Indian, and Bank of England stock tell us that the environment for private sector enterprises has become more risky? Finally, his emphasis on the legislation that followed the crisis may be only a sideshow. I believe 1825 was just a preview for a number of other crises to occur in the next 40 years.

NEAL'S LIST OF "USUAL SUSPECTS"

Neal discusses at great length the list of "usual suspects" as possible causes of the crisis: Latin American debt issues, country banknote issues, and the Bank of England. He dismisses the first two as causal factors and attaches more weight to the third factor.

Speculation in Latin American debt cannot explain the collapse of the stock market bubble because, Neal argues, "the sums risked were relatively small and the risks generally appreciated even by an inexperienced British public." Neal bases this conclusion on the experience of the

Rothchilds and the Barings. However, if these key players were not unduly exposed, surely others were because they bought the stock on the expectation of further appreciation and had less accurate information than did the Barings and Rothchilds. The highly speculative Latin mining stocks and sovereign bonds made up a very significant fraction of the shares issued.

The country banks also could not be blamed, Neal argues. Data from two failed banks show that note holders were willing to hold onto their notes for long periods and were eventually largely compensated for their losses. These facts suggest that country banks were victims of circumstances and not contributors to the crisis. I agree with Neal when he proposes that the country banks could not have issued their notes in a vacuum, and that the key determinant of the growth of their liabilities was rapid expansion of the monetary base. But the country banks were surely an exacerbating factor because of their inherent weak structure, which in turn was related to the regulations that governed their operations. The prohibition on joint stock banking outside of London, the limit of six on the number of bank partners, and their unit banking character constrained the size of country banks. Also, their ability to diversify risk made country banks prone to easy failure in the face of big shocks.

Neal is correct that the Bank of England is the main culprit. Expansionary monetary policy fueled the boom, tight money ended it, and the Bank acted as lender of last resort too late to prevent massive bank failures from creating real economic distress.

The value added of this paper is not so much the retelling of the sordid (thrilling) tale but the author's emphasis on the role of information at every stage of the cycle. The lending boom in the upswing was rife with poor information, adverse selection, and careless surveillance, as is the case today in Latin America and Southeast Asia. More information on why some country banks were sound and others were not would be of value, as would information

on the role of the Bank of England in surveillance and supervision.

TWO RELATED THEMES

I conclude my comments by focusing on two themes that underlie the paper but that Neal does not analyze: the macroeconomic background and generality of the crisis, and the lessons for the lender-of-last-resort function.

The Macroeconomic Background

The 1825 crisis is of more than antiquarian interest because it was a global event, and because it contained many elements of the crises that occurred during the subsequent century.

The crisis contained three unifying forces that occurred in most of the historic crises: (1) monetary shocks; (2) price-level variability and financial distress; (3) real shocks.

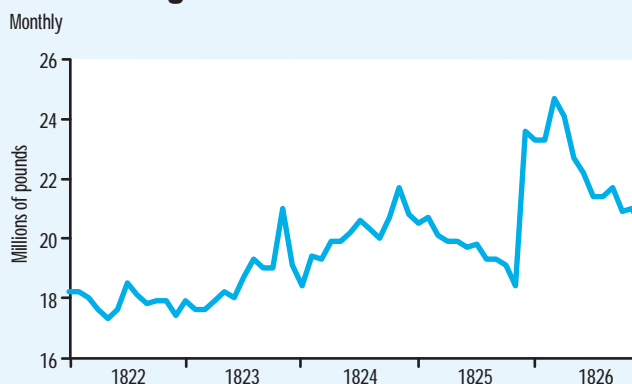
Money. Expansionary monetary policy fueled the boom and created the 1825 crash, as shown in Figure 1. As Neal argues, the Bank followed a liberal policy to accommodate the government's fiscal demands. The expansion in the monetary base (notes shown in Figure 1 and deposits in the Bank shown in Figure 2) created the conditions that allowed country banks to expand their note issues.

At the same time, expansionary monetary policy in the gold standard environment was creating the seeds of its own reversal, as rising domestic prices (Figure 3) led to a trade deficit (Figure 4); an external drain of specie, as manifest in a decline in the Bank's bullion reserves (Figure 5); and a decline in the price of Paris Bills on London (Figure 6). The Bank began tightening early in 1825 (Figure 1), and the stock market (including mining stocks) peaked in January (Figure 7). All other stocks peaked in April.

Price-Level Variability and Financial Distress. Price-level instability is closely related to banking instability. According to one hypothesis (Schwartz 1988), rising prices may contribute to banking instability by increasing misperceptions about current and prospective real returns and possibly by creating an environment in which misman-

Figure 1

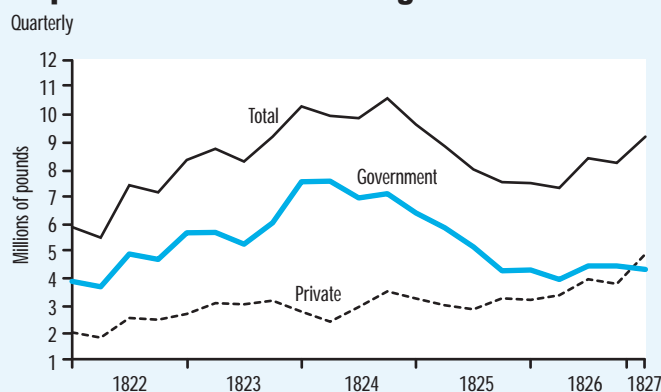
Bank of England Notes in Circulation



SOURCE: Gayer, Rostow, and Schwartz (1953), Table 135.

Figure 2

Deposits in the Bank of England



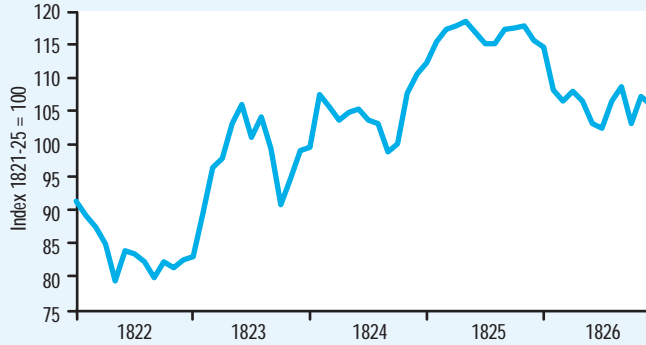
SOURCE: Gayer, Rostow, and Schwartz (1953), Table 146.

agement and fraud are more likely to persist. Unexpected disinflation promotes financial instability by adversely affecting financial intermediaries' balance sheets. As a result of unanticipated disinflation, the real value of nominal debt rises and, without complete contracts, can lead to an increase in bankruptcy and banking distress. Figure 8 shows the U.K. annual inflation rate from 1821 to 1991. As can be seen by the arrows in Figure 8, virtually all banking panics occurred at inflection points of the inflation rate. The first arrow points to the crisis of 1825. The panics ceased after 1866 when the Bank of England learned to act as a proper lender of last resort.

Figure 3

Wholesale Price Index

Domestic commodities, monthly

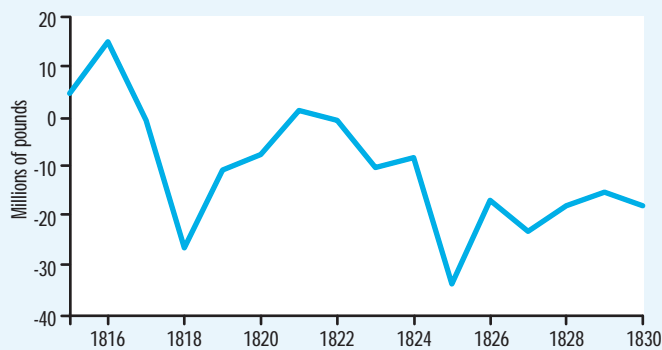


SOURCE: Gayer, Rostow, and Schwartz (1953), Table 40.

Figure 4

The Trade Balance

Annual

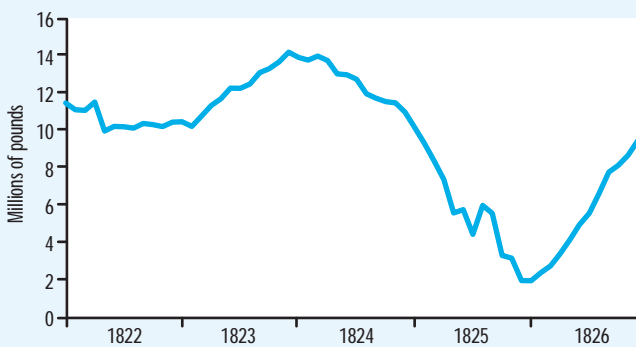


SOURCE: Gayer, Rostow, and Schwartz (1953), Table 114.

Figure 5

Bullion in the Bank of England

Monthly



SOURCE: Gayer, Rostow, and Schwartz (1953), Table 154.

Real Shocks. The background to the crisis of 1825 was not strictly monetary, however. Big real shocks or displacements, as Irving Fisher (1932) termed them, also occurred. These shocks included the massive investment in infrastructure, consolidation of the industrial revolution in England after the upheavals of the Napoleonic wars, opening up of trade, and foreign investment (first with the continent of Europe and then with the newly emerging countries of Latin America). Gayer, Rostow, and Schwartz (1953, Chapter IV) describe some of the details of the investment boom. These investments in turn required finance and monetary accommodation.

The Role of the Bank of England as Lender of Last Resort

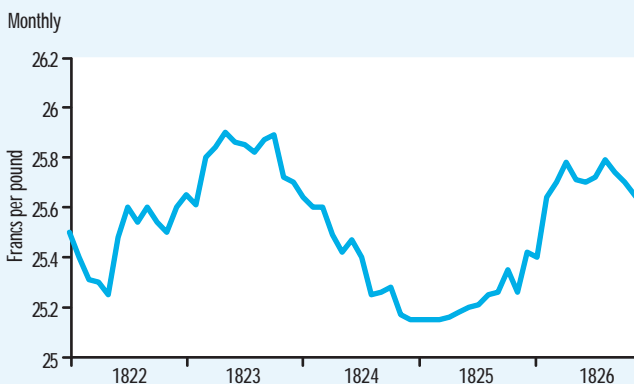
The Bank of England in 1825 was a public bank, not a central bank. The Bank had three loyalties: its shareholders, the British government, and its correspondent commercial bankers. In the first half of the nineteenth century, the Bank of England was learning to balance these three roles. During the suspension period, the Bank made considerable profits from its issue of inconvertible banknotes. With resumption, profits declined; hence there was less incentive to discount freely on unprofitable paper. Indeed, the Bank had not yet adopted Bagehot's (1873) "Responsibility Doctrine" of acting in the public interest first to allay a banking panic or to prevent a stock market crash from spilling over into the monetary system. Henry Thornton basically laid down all the strictures for proper lender-of-last-resort action in 1802 (Humphrey 1989), but the Bank of England did not really "get it" until after the Overend Gurney crisis of 1866 (Schwartz 1986). Also, by following easy money to aid the government in its debt service and loan conversions, the Bank had not yet established the independence needed to follow a stable monetary policy.

Thus, 1825 was just one of a series of crises—1837, 1847, 1857, and 1866—in which the Bank of England did not act properly as a lender of last resort. As Neal tells it,

the Bank's reluctance to lend early in 1825—when signs of stress were looming, merchants in the Latin American trade were failing, and the Bank raised the discount rate and cut back on advances later in the summer—likely exacerbated the crisis. Therefore, when the Bank finally did act in December, it was much too late to prevent a large number of banks from failing. True, the Bank did not have to suspend specie payments. Gold from the Banque de France saved the Bank of England. But had the Bank of England lent earlier and prevented the bank failures, it would also have prevented a serious recession. The Bank would likely have received permission from the government to temporarily suspend payments. The Bank Act of 1844 later institutionalized the practice of the Bank's requesting a letter from the Treasury granting permission to suspend payments, which the Bank did in later crises, e.g., 1847 (Dornbusch and Frenkel 1984). The lesson that central banks have learned since Bagehot (1873) is to lend freely, in a timely manner, and on the basis of any acceptable collateral, but to lend at a penalty rate. The Bank apparently followed the penalty rate part of the rule (seen in a rise from 4 percent to 5 percent in December 1825) but did not lend freely nor in a timely manner. Today, most central banks have learned Bagehot's rule, and they do not make the mistake the Bank of England made in 1825 (e.g., the Federal Reserve's response to the 1987 stock market crash). The problem today is that many central banks have learned the lesson too well and now follow the "too big to fail" doctrine, which leads to new problems under the rubric of moral hazard. For emerging countries today, however, the 1825 crisis still has resonance. The problems of adverse selection during lending booms echo the story Neal tells. Structural problems in the banking system, poor oversight and regulation (papered over by rising prices), and profits taken during the boom are revealed when the crash occurs.

Figure 6

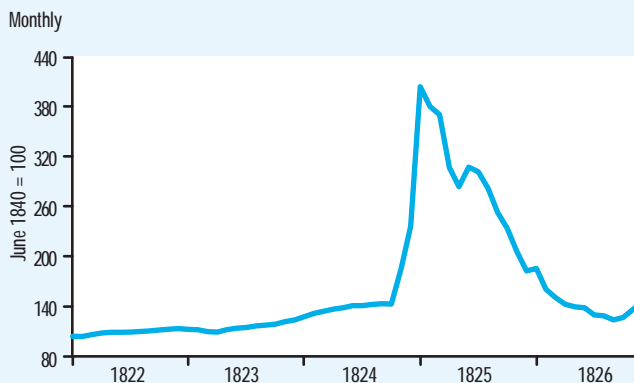
Price of Paris Bills on London



SOURCE: Gayer, Rostow, and Schwartz (1953), Table 194.

Figure 7

Total Index of Share Prices

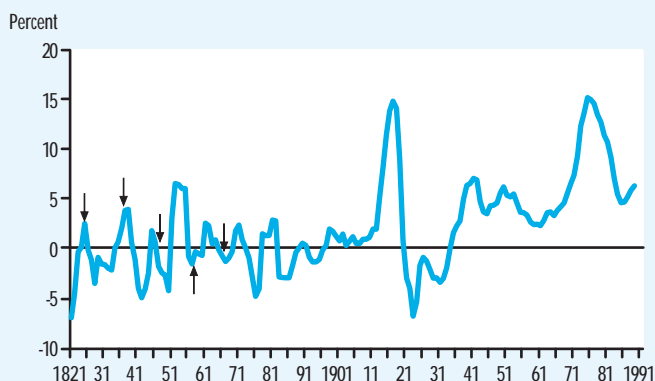


SOURCE: Gayer, Rostow, and Schwartz (1953), Table 9.

Figure 8

U.K. Inflation Rate

Annual. Five-year-centered moving average of year-to-year change in log price index.



SOURCES: For 1821-79: Mitchell (1962); for 1880-1990: National Bureau of Economic Research.

REFERENCES

- Bagehot, Walter. *Lombard Street: A Description of the Money Market*, H.S. King, 1873.
- Clapham, John H. *The Bank of England, A History*, Vol. II, 1797-1914, Cambridge University Press, 1945.
- Diamond, Douglas W., and Philip H. Dybvig. "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* (June 1983), pp. 401-19.
- Dornbusch, Rudiger, and Jacob A. Frenkel. "The Gold Standard and the Bank of England in the Crisis of 1847," *A Retrospective on the Classical Gold Standard, 1821-1931*, Michael D. Bordo and Anna J. Schwartz, eds., University of Chicago Press, 1984, pp. 233-71.
- Fisher, Irving. *Booms and Depressions*, Adelphi, 1932.
- Gayer, Arthur D., W. W. Rostow, and Anna Jacobson Schwartz. *The Growth and Fluctuation of the British Economy 1790-1850*, Oxford Clarendon Press, 1953.
- Humphrey, Tom. "Lender of Last Resort: The Concept in History," Federal Reserve Bank of Richmond *Economic Review* (March/April 1989), pp. 8-16.
- Kindleberger, Charles. *Manias, Panics and Crashes: A History of Financial Crises*, Basic Books, 1978.
- Minsky, Hyman. "A Theory of Systemic Fragility," *Financial Crises: Institutions and Markets in a Fragile Environment*, Edward I. Altman and Arthur W. Sametz, eds., Wiley, 1977, pp. 138-52.
- Mitchell, Brian R. *Abstract of British Historical Statistics*, Cambridge: Cambridge University Press, 1962, pp. 470, 474.
- Officer, Lawrence A. "The Floating Dollar in the Greenback Period: A Test of Theories of Exchange Rate Determination," *Journal of Economic History* (September 1981), pp. 629-50.
- Schwartz, Anna J. "Financial Stability and the Federal Safety Net," *Restructuring Banking and Financial Services in America*, Rose Marie Kushmeider and William S. Haraf, eds., American Enterprise Institute, 1988, pp. 34-62.
- _____. "Real and Pseudo-Financial Crises," *Financial Crises and the World Banking System*, Forrest Capie and Geoffrey E. Wood, eds., St. Martin's Press, 1986, pp. 11-31.
- Thornton, Henry. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, F. A. Hayek, ed., Augustus M. Kelley, 1802.

Richard Sylla is the Henry Kaufman Professor of the History of Financial Institutions and Markets and a professor of economics, Stern School of Business, New York University. This research was supported by the Alfred P. Sloan Foundation.



U.S. Securities Markets and the Banking System, 1790-1840

Richard Sylla

A fact underappreciated about the rise of the United States in the world economy is that a modern, “world class” financial system—by the standards of the time—emerged virtually at the beginning of the nation’s history and provided a solid underpinning for the country’s subsequent growth and development. This paper explores the emergence of that financial system. It emphasizes the mutual support between the banking system, which has been well studied by financial historians, and securities markets, which have been relatively neglected. Distinctive features of the U.S. banking system depended on the existence of securities markets, and before long, distinctive features of U.S. money and securities markets depended on developments in the banking system.

BANKING SYSTEMS AND FINANCIAL SYSTEMS

The “Anglo-American” or “Anglo-Saxon” pattern of financial organization features a functional division of labor and a balance among three main, interrelated sectors: the banking system, the money market, and the securities market. This pattern is often contrasted with the “Continental European” or “German” pattern, in which banks dominate the financial system while the money and securities markets are relegated to minor, secondary roles. The advantages and disadvantages of each pattern of organization relative to

the other are much studied and debated. Also discussed are the questions of whether today’s globalization of finance (which is less unprecedented than many believe) will bring about a convergence of financial systems and, if so, in what direction. Financial historians have become interested in an additional question: why the two different patterns of financial organization emerged in history. Thus far, however, they have only scratched the surface in attempts to answer it.

A reason for the limited progress in understanding why systemic differences emerged in history is that, while Anglo-Saxon and German systems may compete with each other in the real world, in the world of financial historians—be it the Anglo-Saxon, the Continental European, or any other division—the German model seemingly has carried the day. This is not meant to imply that financial historians have weighed the evidence and decided that the German bank-based pattern of financial organization is best, although some on both sides of the Atlantic would agree with such a contention.¹ Rather, it is meant to suggest that banks everywhere have received the lion’s share of attention from financial historians. I would hazard the guess, based on some years of experience, that there are 25 or 50 dissertations, articles, and books on the history of banks and banking for each one on the history of money and securities markets.

No doubt there are many reasons for the overwhelming attention financial historians devote to banks and banking. Among them is the obvious one that, in Continental Europe, banks were by far the dominant financial institutions during the past two centuries, so that other components of financial organization merited less study. But why is the emphasis on banking history much the same among Anglo-American scholars? Here I think an explanation would include several points. One is that, even in Great Britain and the United States,

¹ For a “pro-German bank” view from the American side of the Atlantic, see Calomiris (1995). German universal banks, closely involved with the firms they finance, are thought to have done a better job of monitoring firm managements than “arms-length” Anglo-Saxon banks, and to have raised capital for firms at lower costs than those experienced in Anglo-Saxon systems in which commercial and investment banking often were separated.

banks were important—perhaps even very important—sources of finance in the early stages of economic modernization during the late eighteenth and early nineteenth centuries. Another is that banknote and deposit liabilities served most of the functions of money, so that governments—regarding control over money as an element of sovereignty—had both political and economic reasons to become concerned with licensing and regulating banks and their money creation.

A by-product of governmental concern with banking had an obvious role in drawing the attention of scholars: Banks left many tracks in the public documents that are the grist of historians' mills. Moreover, since the governments overseeing banking were many, and banks as organizations were even more numerous, there were manifold topics for research, from the history of banking in country, region, or state A, B, or C, to the history of the bank of X, Y, or Z.

Securities markets have not attracted so much interest from historians. Although organizations—including banks—participate in them, and some of these organizations (such as the New York Stock Exchange) came to symbolize them, securities markets are not particular organizations but fundamental institutions or economic processes. Governments were therefore less concerned in the past with their supervision. Apart from instances of public borrowing and the debt management it entailed, government documents report little about securities markets in comparison to the voluminous records concerning banks and banking. One can get a sense of why capital markets have been neglected by historians by imagining what banking history would be like if banking were mentioned in historical records only when a government took out or paid back a bank loan. Gone would be discussions of the politics of bank chartering, of the periodic need to reform the banking system, of the need or lack of need for a central bank, of the monetary and macroeconomic effects of the expansion and contraction of bank credit.² Gone, in short, would be much if not most of banking history. The balance of financial

historiography is, however, being righted. Securities markets do have a recorded history, but one that usually is not well documented in public records. Information is one of the most important inputs and outputs of these capital markets, but historically it was of far more use to day-to-day participants in the markets than to governments. Therefore, it appeared for the most part in newspapers and other private periodicals rather than in the government documents that have been so conveniently collected and placed in numerous libraries for the use of scholars and others. Decades ago, a few historians culled information on capital market activity from such private periodical sources to study particular eras, usually in conjunction with research on business cycles (e.g., Smith and Cole, 1935). And there are some landmark studies distinguished for shedding light on the breadth and depth of securities market history over long periods of past history. In the latter category, there is work of Cowles (1939) on U.S. stock prices from the 1870s to the 1930s—the forerunner of comprehensive modern stock price averages and the progenitor of much subsequent work on the historical behavior of stock prices. There is also the work of Macaulay (1938) on U.S. interest rates, bond yields, and stock prices back to 1856. More recently there is the work of Neal (1990) on the rich but neglected quantitative history of British and Dutch capital markets from the late seventeenth to the early nineteenth century, and that of Davis and Cull (1994) on international financial flows to and from the United States in the century before 1914.

United States securities markets had a rich quantitative history before 1856 and 1871, the dates when Macaulay and Cowles began to document and analyze it. Smith and Cole (1935) drew attention to this history from the 1790s to 1860, in developing stock and bond price series and indexes of prices back almost to the start of the government under the Constitution. During the past decade or so, these have been used by scholars to study U.S. stock and bond returns over two centuries. Smith and

² There is a good example in American banking history. Until the twentieth century, many "private" (unincorporated) banks existed in the United States, but we know little else about them in the aggregate. See, for example, Sylla (1975, 1976).

Cole, however, treated the financial-markets aspect of their work almost as a curiosity because they doubted the early financial series had much to do with business conditions and because the newspapers they studied published far fewer financial asset prices than commodity prices.

Perkins (1994) was more circumspect. His study of the development of American public finance and the financial services sector during the eighteenth and early nineteenth centuries led him to the conclusion that by the time of the War of 1812, the United States possessed a complex, articulated financial organization—centered in a banking system and other capital market institutions. Perkins did not compare the U.S. system to those of other countries, but his account, placed in context, suggests that the early U.S. financial system rivaled that of any other country at the time. Given that barely three decades earlier there were almost no U.S. banks or organized domestic securities markets, this was a considerable achievement, one that Perkins himself rather underemphasized.

The establishment of a modern financial system at the start of U.S. history is also important for understanding the country's rapid growth throughout the nineteenth century. Historians have long regarded the drivers of that growth as being manufacturing technologies, transportation innovations, and the opening of the trans-Appalachian west for settlement and integration into the national and world economies. But each of these developments, which emerged mostly after 1815, relied in important ways on the financial system established earlier. Manufacturing technologies, transportation innovations, and extensive lands to be settled were available in many parts of the world in the early nineteenth century. But nowhere were they exploited as early and as well as in the United States. The telling economic difference circa 1820 between the United States and, say, Canada, Mexico, Brazil, or Argentina was in financial organization, where the Americans were way ahead. Earlier, around 1780, when the Americans had no banks or organized

securities markets and were awash in nearly worthless paper money, financial differences between the United States and its new world neighbors were less noticeable. In finance as in political organization, key changes occurred in the United States between 1780 and 1820, and the political and financial changes were very much related to one another.

THE FEDERALIST FINANCIAL REVOLUTION

During the 1780s, merchant groups organized three banks—the first ones in the United States—in Philadelphia, New York, and Boston. Two received corporate charters from their state legislatures, but New York's bank waited until 1791 for this privilege. These were isolated, local banks; there was no banking system. States serviced the debts they had incurred in the War of Independence, sometimes by raising taxes and sometimes by printing bills of credit, fiat paper money that had long colonial-era precedents. The domestic U.S. debts incurred by Congress during the war were essentially unserved “junk” bonds, with interest payments due settled by issuing more IOUs because the Confederation Congress lacked the power to raise revenue through taxation. Foreign debts were serviced by means of new loans from European investors, who had an interest in rolling over their previous American loans and probably hoped for a favorable turn in the new nation's finances.

Adoption of a new constitution in 1788 laid the basis for reforming the financial system. The new framework of American government was mainly the work of Nationalists who, because they built up financial and other powers of the new federal government while reducing those of states, came to be called Federalists. In the new arrangement, the states lost, among other powers, the right to print fiat paper money. But they did not lose the right, which they already had exercised in two instances, to charter banks that could issue paper money and deposit credits convertible into a monetary base of specie. And charter banks is

Table 1

**U.S. State-Chartered Banks:
Numbers and Authorized Capital, by Region and Total, 1790-1835**

(Capital in millions of dollars)

Year	New England		Mid-Atlantic		South		West		United States	
	No.	Capital	No.	Capital	No.	Capital	No.	Capital	No.	Capital
1790	1	0.8	2	2.3					3	3.1
1795	11	4.1	9	9.4					20	13.5
1800	17	5.5	11	11.9					28	17.4
1805	45	13.2	19	21.7	6	3.5	1	0.5	71	38.9
1810	52	15.5	32	29.4	13	9.1	5	2.2	102	56.2
1815	71	24.5	107	67.1	22	17.2	12	6.4	212	115.2
1820	97	28.3	125	74.2	25	28.6	80	28.4	327	159.6
1825	159	42.2	122	71.2	32	33.3	17	9.4	330	156.6
1830	186	48.8	140	73.8	35	37.3	20	10.5	381	170.4
1835	285	71.5	189	90.2	63	111.6	47	35.0	584	308.4

SOURCE: J. Van Fenstemaker, *The Development of American Commercial Banking, 1782-1837*. Capital data are rounded, so components may not add to total. Kent State University Press, 1965, Tables 4, 12, 13, 14, 16, 17, and A-1.

what the states did. From three such banks in 1790, their numbers rose to 28 in 1800, 102 in 1810, 327 by 1820, and 584 by 1835. During the 1790s, all of these banks were in the New England and Middle Atlantic states; in 1835, more than 80 percent of the state banks were in this same northeastern region (see Table 1).

Equally remarkable financial developments came at the federal level, or were prompted by federal actions. During George Washington's first presidential administration, 1789-93, Treasury Secretary Alexander Hamilton proposed, and the Federalist-dominated Congress enacted, a comprehensive program of financial reforms. Federal authority was exercised in 1789-90 to raise revenue from customs duties and domestic excise taxes. This revenue was pledged to pay interest in hard money on national and assumed state debts that were restructured in 1790 into three new issues of Treasury securities—6 percent and 3 percent issues that paid interest immediately, and a deferred 6 percent issue that would pay interest commencing after 10 years. These issues funded some \$65 million of domestic debt; the \$12 million of additional debts owed to foreigners, chiefly the French gov-

ernment and Dutch investors, was provided for separately. For perspective, the total national debt of about \$77 million was approximately 40 percent of estimated GNP at the time.

Two more measures of 1791 rounded out the program of financial reform. Congress enacted Hamilton's proposal for a Bank of the United States to aid federal financial operations and exercise leadership in developing a U.S. banking system. Whereas the few state banks then existing were capitalized at \$1 million or less, the Bank of the United States was capitalized at \$10 million (25,000 shares of \$400 each), one-fifth of which was subscribed by the federal government and four-fifths by private investors. The latter could use the new Treasury securities to pay for up to three-fourths of their bank shares, with the other fourth to be paid in specie. By design, the federal debt supported the bank, and the bank the debt. Headquartered at Philadelphia, the capital from 1790 to 1800, the bank was authorized to open branches in other cities throughout the nation, and it quickly did. This prompted state leaders to charter more banks, lest the new federal government co-opt

the banking business.³ The last major reform measure was enacted in 1791 to establish a mint for coining gold and silver into a dollar-denominated monetary base.

Most accounts of the Federalist financial program concentrate on public finance and politics. The credit of the national government, which was essentially bankrupt under the confederation, rocketed in the estimation of investors after 1791. Whereas evidences of public debt sold in 1788 for 10 to 20 cents on the dollar in sporadic, unorganized transactions before ratification of the Constitution, as the new federal government began to organize itself and adopted Hamilton's plans, the restructured debts rose toward par, and even above par for the 6 percent-coupon securities, by 1791-92 (Ferguson 1961, chpts. 12, 15). The consolidation of political and financial power at the federal level was troubling to anti-Federalists and even a few nationalists, notably James Madison. Agrarian in outlook and state-oriented in politics, the anti-Federalists had little use or respect for commercial elites, banks, factories, stock-jobbers, and securities speculators (in contrast to land speculators). Under Thomas Jefferson's leadership, they formed a political opposition to the Federalists that styled itself as "Republican" or "Democratic Republican."

The Federalists, led by Hamilton, had a different vision. Based on their experience of the 1780s, they viewed state governments as parochial and divisive of the nation. The states were as likely to interfere with as to promote a unifying national government and diversified, nationwide economic development. The Federalists' goal was to overcome state parochialism, to build a national government that would command the respect of Americans and foreign nations, and to use that government to foster energetic national economic development. Their means was to give Americans, and possibly foreigners, a recognizable stake in the new government's success. Long ago the great historian Charles Beard (1915, p. 131)

summed up what he saw as Hamilton's insights and statecraft:

Hamilton's measures were primarily capitalistic in character as opposed to agrarian . . . and constituted a distinct bid to the financial, commercial, and industrial classes to give their confidence and support to the government in return for a policy well calculated to advance their interests. He knew that the government could not stand if its sole basis was the platonic support of genial well wishers. He knew that it had been created in response to interested demands and not out of any fine spun theories of political science. Therein he displayed that penetrating wisdom which placed him among the great statesmen of all time.

The anti-Federalist and Republican opposition, however, saw Hamilton's measures as a sell-out to a relatively wealthy commercial minority living in cities at a time when most citizens of the country were tillers of the soil. Thus the contours of American political life—states' rights vs. federal authority, agrarians vs. capitalists, the ordinary people vs. the business elites—that have persisted to the present day were born in the lines of battle drawn up over the Federalist financial program.

The effects of the Federalist program involved more, however, than public finance and politics. Directly and indirectly, as financial historians and other students of the era have noted, it promoted the development of the U.S. banking system. Less noted have been its effects on securities market development. As old evidences of Revolutionary War debt were exchanged for some \$65 million of new, interest-paying federal securities starting in late 1790 (to which was added \$10 million stock in the Bank of the United States a year later), active and regularized trading markets for these "national market" issues emerged in major cities, particularly New York, Philadelphia, and Boston. They were joined as time went on by more and more local issues. Securities market prices began to be reported regularly, usually at weekly intervals, in the newspapers of the day. Using these sources,

³ The Bank of New York was founded by Hamilton and others in 1784, but its requests for a charter were rebuffed by the New York legislature, controlled by anti-Federalists, until 1791, when the Bank of the United States came into being and "threatened" to open a branch in New York City. The anti-Federalist legislature quickly countered the threat by granting a state charter to the Bank of New York, thereby insuring that at least some part of banking in the state would be under its control. See Wright (1996).

two collaborators and I gathered a database of end-of-month prices of U.S. debt securities from 1790 to the 1830s and analyzed the data in a recent working paper (Sylla, Wilson, and Wright 1997). Here, as background for exploring interactions of the securities markets with the emerging U.S. banking system, I summarize four key findings of that working paper that are derived from analysis of monthly securities prices covering four decades.

Domestic Intermarket Arbitrage

The New York, Philadelphia, and Boston securities markets showed evidence of pricing efficiency and intermarket integration from the beginning, despite slow communication times (one to two days between New York and Philadelphia, roughly a week between New York and Boston) and varying intercity exchange rates. In 1791 and 1792, as the markets were in their infancy, prices of the same security were about the same in each city, and they moved up and down together from month to month. The data give a strong impression of intermarket arbitrage, a point confirmed by archival evidence we and others uncovered.⁴ Although the capital was in Philadelphia, which was the nation's largest city and considered to be its leading financial center, New York even then appeared to have the most active securities market. New York market participants, some of whom likely were acting as agents of European investors, had their own agents in Philadelphia and Boston who bought and sold securities for them whenever those markets appeared to offer an advantage over New York prices. Judging by these findings, the U.S. securities markets were capable of allocating capital with a high degree of efficiency as early as the 1790s, when they first emerged to provide organized trading in federally sponsored securities issues.

Efficient Pricing of Hybrid Securities

Hamilton's 6 percent coupon-deferred security was something like a zero-coupon

bond for the 10 years between 1791 and 1800. If the securities markets priced efficiently, the difference between its market price and the market price of the 6 percent coupon issue that paid interest throughout the period would be the present value of the stream-of-interest payments promised by the 6 percent coupon issue but not the deferred 6 percent security. Lacking a market rate of interest to calculate present values, we "backed out" an internal rate-of-return series that equated each month's price difference between 6 percent coupon and deferred securities to the remaining stream-of-interest payments that the 6 percent coupon security had promised up to 1801, when the two issues became equivalent. This series, although it is an implicit short-term rate, tracks fairly well the yields of the interest-paying 6 percent coupons. Implicit yields of the zero-coupon deferreds were in the 5 percent to 10 percent range, in keeping with other American yields during the 1790s. The infant U.S. securities market could price a hybrid, zero-coupon security with efficiency.

Encouraging Capital Inflows from Overseas

As Hamilton forecasted in the 1790 Report on Public Credit that outlined his proposals for funding U.S. debts, the new securities that resulted from Congress' adoption of his plan proved attractive to overseas investors, and, in buying them, the overseas investors transferred capital to the United States. Blodget (1806), relying on Treasury and other data, estimated that by 1803 more than half of the debt of \$81 million (which included \$11.25 million in U.S. securities paid to France that year for the Louisiana Purchase) was held in Europe, mostly by English and Dutch investors. Blodget also found that more than three-fifths of the stock of the Bank of the United States had found its way to European hands. It seems clear that one attraction of American securities to European investors was the ready markets they commanded in U.S. cities. Parallel markets in U.S. issues

⁴ A collection of surviving letters from Stephen Higginson, a prominent Boston businessman, to Leroy & Bayard, New York merchants, shows that from 1790 to 1794, Higginson acted as the latter's agent, buying securities for them at Boston when they could be obtained on more favorable terms than at New York. The letters are in the Gratz Collection at the Pennsylvania Historical Society. Werner and Smith (1991, pp. 43 and 226n) cite similar evidence showing that other New Yorkers had agents doing the same thing at Philadelphia.

developed in London and Amsterdam, and newspapers on both sides of the Atlantic began to report on the “latest” (usually two months earlier) prices of the same securities across the sea. U.S. securities-market development during the 1790s thus paved the way for a flow of capital from Europe to America that would reach huge proportions in the internal improvement era of the 1820s and 1830s, and in the railroad age that followed.

Parity in International Capital Markets

Yield levels and fluctuations of U.S. 3 percent securities, the majority of which were owned by European investors in 1803, and those of a similar British security, the famous “consol 3s,” were very similar for much of the period from 1800 to 1830. An exception to this yield similarity came during the War of 1812, when the British issue rose to a price premium over the similar U.S. issue. Although Britain had its own financial problems during the Napoleonic War era, problems in the United States during the War of 1812 were even worse. States’ rights and state banking forces conspired in 1811 to prevent recharter of the Bank of the United States, an action that crippled federal financial management when war came the following year. Banks suspended specie payments outside of New England, and the Treasury was forced into printing near money and borrowing on onerous terms. Despite the financial chaos of suspension years 1814-17, we found that the securities markets priced with efficiency, adjusting prices in Boston, New York, and Philadelphia to reflect prevailing exchange rates between the local currencies of these cities. The larger lesson remains, however, that the debt of the United States, an “emerging market” of the 1790s, apparently could compete on close to equal terms in the eyes of investors with “established” British government debt by the first decades of the nineteenth century. If so, the United States was perhaps the most successful emerging market in the long history of international capital markets.

The banking system and securities markets that grew out of the Federalist

financial revolution did not sit easy among their opponents, who took charge of national affairs after 1800. They were widely attacked and sometimes undone (as in the case of the First Bank), but were eventually accepted or reinstated (as in the case of the Second Bank, founded in 1816 and then undone by Andrew Jackson in the 1830s, necessitating some decades later the “Third Bank” of the United States, which is known more familiarly to us as the Federal Reserve System). Acceptance of the Federalist financial program, even if halting, was predicated on utility. President Jefferson, for example, was among the first to discover the utility of his opponent Hamilton’s financial architecture when he found France eager to accept Treasury paper in return for the Louisiana Territory. United States credit had become so good that Napoleon’s government could easily raise cash by selling U.S. securities to European private investors. Beard was right: The institutions that sprang up out of Federalist financial policies were well calculated to serve the interests of ever-growing numbers of Americans, including Thomas Jefferson.

BANKS AND THE SECURITIES MARKETS

Apart from government itself, the sector that benefited most from early U.S. securities markets was banking. Banks in the United States, unlike most banks in other countries at the time, were corporations that raised their banking capital by issuing equity securities, which were made all the more attractive to investors by the emergence of active trading markets in the 1790s. Moreover, almost as soon as these markets emerged, securities—both government debt and corporate stock—became useful as collateral for bank loans and objects of bank investment. Early in 1790, for example, the Massachusetts Bank accepted illiquid state securities and old U.S. securities as collateral for bank loans at only 25 percent of par value. When the new 6 percent securities appeared later that year, they could be collateralized at 50 percent of par. A year later, their collateral value had risen to 90 percent of par, and

Table 2

**Corporate Stock as Percentage of Financial Assets:
Four Countries, 1800-50**

Period	USA	Great Britain	France	Germany
1800-15	10	3	0	n.a.
1850	18	11 (est.)	6	3

SOURCE: Raymond W. Goldsmith, *Comparative National Balance Sheets*, University of Chicago Press, 1985, Appendix A. For Great Britain in 1850, Goldsmith reports a combined percentage for corporate bonds and stock. I make an estimate of the stock share based on his 1895 and 1913 data for Britain, which give separate corporate bond and stock shares.

by mid-1792 they were accepted at par value as loan collateral (Davis 1917, vol. 2, p. 65). Colonial and Confederation America had not solved the problem of illiquidity in investment, most of which took the form of real and tangible personal property. The synergies of banks and securities markets released in the Federalist financial revolution led to an outpouring of liquid financial assets and in short order made this long-standing drag on U.S. economic potential disappear.

Corporate stock, like government debt, became a bankable asset. In his study of comparative national balance sheets, Goldsmith (1985) relied entirely on the amount of bank stock for his estimate of U.S. corporate stock in the early years of the nineteenth century. There were, of course, other forms of corporate stock—insurance, transportation, and even manufacturing company stocks—at the time. But banks were undoubtedly the largest component of the early U.S. stock market. Despite their limitations, Goldsmith's national balance sheet data for the first half of the nineteenth century are, in a comparative context, revealing. One can derive from them various ratios, including the ratios of corporate stock to financial assets for various countries at roughly the same dates. Table 2 presents that ratio for the United States, Great Britain, France, and Germany at the beginning and midpoints of the century. Somewhat surprisingly, at both dates the United States led the world in the proportion of financial assets held in the form of corporate stock. This is another indication of the impact of the Federalist revolution on the financial habits of Americans.

Aggregated data such as Goldsmith's are suggestive, but newspaper sources, with a more micro perspective, provide a more detailed (if still far from complete) picture of the extent of U.S. securities market development, including the market for corporate stock, in the early decades of the republic. Table 3 shows the newspaper listings of securities regularly quoted in leading U.S. markets from 1797 to 1817, the year the New York Stock Exchange was formally organized. New York provides the two-decade chronology in the table, with glimpses of the market there in 1797, 1801, 1811, and 1817. The New York, Boston, Philadelphia, and Baltimore listings in mid-1811 give a cross-section of securities-market development in leading cities by that year. All the listings are taken from one New York newspaper, which indicates that New Yorkers even then were rather interested in what was going on in markets other than their own. Such regularly published newspaper lists are, of course, the tip of the iceberg of U.S. securities-market development. Many more companies formed and issued stocks that did not make it into the weekly quotation lists of newspapers, presumably because the stocks were closely held or inactively traded compared to listed securities, or perhaps because the papers, which consisted of just a few pages—largely paid ads—could not afford to publish too much free material.

In New York, the list of state banks rose from one to eight from 1797 to 1817; all were New York City banks. In 1811, there were five listed state banks in New York, three in Boston, four in Philadelphia,

Table 3

Securities Listed and Quoted in Leading U.S. Markets, 1797-1817

New York, 1797

U.S. 6 percent bonds
U.S. 3 percent bonds
U.S. Deferred 6 percent bonds
Bank of United States
Bank of New York

New York, 1801

U.S. 6 percent bonds
U.S. 3 percent bonds
U.S. Navy 6 percent bonds
U.S. 8 percent bonds
Bank of United States
Bank of New York
Manhattan Bank
New York Insurance Co.
Columbian Insurance Co.
United Insurance Co.

New York, 1811

U.S. 6 percent bonds
U.S. 3 percent bonds
Bank of United States
Bank of New York
Manhattan Co. Bank
Merchants Bank
Union Bank
Mechanics Bank
New York Insurance
Columbian Insurance
United Insurance
Marine Insurance
Commercial Insurance
Phoenix Insurance
Eagle Insurance
Mutual Insurance
Ocean Insurance
New York Firemen Insurance

Boston, 1811

U.S. 6 percent bonds
U.S. 3 percent bonds
Massachusetts Bank
Union Bank
Boston Bank bond
Late Bank of U.S.
Boston Marine Insurance

Fire and Marine Insurance
State Notes

Philadelphia, 1811

U.S. 6 percent bonds
U.S. 3 percent bonds
Louisiana 6 percent bonds
Bank of U.S.
Bank of Pennsylvania
Bank of North America
Bank of Philadelphia
Farmers & Mechanics Bank
Ins. Co. of Pennsylvania
Ins. Co. of North America
Union Insurance
Phoenix Insurance
Delaware Insurance
Marine Insurance
United States Insurance
Lancaster & Susqueh'a Insurance
American Fire Insurance
Schuylkill Bridge shares
Delaware Bridge shares
Lancaster Turnpike shares
Germantown Turnpike shares
Cheltenham & Willow Grove Tpk shares
Chestnut Hill & Springhse Tav'n shares
Frankford Turnpike shares
Water Loan
City Loan
Masonic Loan

Baltimore, 1811

U.S. 6 percent bonds
U.S. 3 percent bonds
Louisiana 6 percent bonds
Bank of U.S.
Maryland Bank
Baltimore Bank
Union Bank of Baltimore
Mechanics Bank
Farmers Bank
Columbia Bank
Potowmac Bank
Farmers & Merchants Bank
Commercial & Farmers Bank
Franklin Bank

Marine Bank

Baltimore Insurance shares
Maryland Insurance shares
Marine Insurance shares
Chesapeake Insurance shares
Union Insurance shares
Fire Insurance
Reistertown Road stock
Fredericktown stock
York stock
Falls stock
Union Manufacturing
Water stock

New York, 1817

U.S. 6 percent bonds
U.S. 3 percent bonds
Louisiana 6 percent bonds
U.S. 7 percent bonds
Yazoo/Mississippi (U.S.)
NY State 6 percent bonds
NY State 7 percent bonds
Corporation 6 percent bonds (NYC)
Bank of U.S.
Bank of New York
Manhattan Co. Bank
Merchants Bank
Mechanics Bank
Union Bank
Bank of America
City Bank
Phoenix Bank
United Insurance
New York Insurance
Fireman Insurance
Ocean Insurance
Phoenix Insurance
American Insurance
Pacific Insurance
Mutual Insurance
Washington Insurance
Eagle Insurance
Globe Insurance
National Insurance
Spanish Dollars
Dobloons

NOTE: Securities quotations usually accompanied by quotations for inland and foreign exchange. Also gold and silver coins when specie payments were suspended.

SOURCE: *New York Price Current*, issues of Jan. 2, 1797; Feb. 28, 1801; June 29, 1811 (for four cities), and Dec. 24, 1817.

and no less than 11 in Baltimore, although two of these banks—Columbia and Potomac—were chartered by the District of Columbia. The American separation of financial and political centers likely arose because the country had securities markets before it established its permanent capital.

Having active stock markets encouraged investors to own bank stocks. The other side of this coin was that it made it easier for corporate banks to form and to raise equity capital. Starting from next to nothing in 1790, the United States experienced, mostly under the auspices of state-chartered banking corporations, the most rapid spread of banking institutions of any country over the next decades. Table 1, which is based on the painstaking archival and documentary research of Fenstermaker (1965), presents by regions the number and authorized capital of state-chartered banks at five-year intervals from 1790 to 1835. State-chartered banks increased from three to 584 during the 45-year period, while authorized capital increased from \$3 million to \$308 million. Some increases in authorized capital should be treated with skepticism, for two reasons. First, the table itself indicates unusually large increases between 1830 and 1835, particularly in the South and West. Visionary schemes there, coupled with the ease of access to European capital that was a product of the very capital market developments under discussion, made it possible for the states of Louisiana, Mississippi, and Tennessee, and the territories of Florida and Michigan to raise large sums for banking improvements by selling state-guaranteed bonds to foreign investors. In less than a decade, the bubbles burst and many states defaulted on, some even repudiating, their debts. The New England and Middle Atlantic states, where even as late as 1835 more than four out of five state banks were located, provide an indication of steadier, more orderly banking development.

A second reason for skepticism is that authorized capital was seldom the same as capital paid in. Banks requested more capital in their charters than they intended to start with, to avoid political complications

that might arise from petitioning state legislatures for increases. For the earlier dates in Table 1 there is precious little information on capital actually paid in. By 1825, 1830, and 1835, Fenstermaker (1965, Table 10) was able to gather balance-sheet information for a majority of the banks. His data indicate that, for these banks alone, paid-in capital was 50 to 70 percent of the total authorized capital for all U.S. state-chartered banks. A recent history of banking in New York state reproduces a table giving the paid-in capital of 11 city and 11 country banks in 1828 (Hubbard 1995, p. 72); together these were about half of the banks the state had chartered. By matching the banks with Fenstermaker's Appendix A giving the authorized capital of the same banks, I found that the country banks had paid in 59 percent of their authorizations, and city banks 67 percent. Since the latter were substantially larger, for all 22 banks the ratio of paid-in to authorized capital was nearly the same, 66 percent. Interestingly, three of the New York City banks had paid-in capitals larger than the amounts authorized in their original charters; as their banking businesses grew, they found it prudent to increase their capitalizations.

Pending more analysis, a conservative estimate of the ratio of paid-in to authorized capital for the U.S. banking system in its early decades would appear to be about 0.6. Applying that factor to Table 1's authorized bank capital in 1800 gives \$10.4 million for the 28 state banks then in existence (and \$20.4 million by adding the capital of the First Bank, all of which was paid in). In 1825, the corresponding figure is \$93.7 million for 330 state banks (and \$128.7 million adding the \$35 million in capital of the Second Bank).

I chose the dates 1800 and 1825 because there are corresponding estimates of the banking capital of England and Wales in those years. Cameron (1967, pp. 32-33) estimates that England and Wales in 1800 had £5.5 million (\$25.9 million) of capital invested in banks, not counting the "Rest," or surplus capital available for banking, of the Bank of England, and £9.8 million (\$46.1 million) if the Bank of England is included. Comparing

the two countries, we can conclude that by 1800 the United States, whose banking system was barely a decade old, had nearly half the banking capital of England and Wales, roughly on the order of its population in relation to that of England and Wales. The Federalists had effected rapid change. Given their fate, one wonders whether it might have been too rapid for their political fortunes.

To avoid possible misunderstanding, let me make clear that I do not mean to imply that either banking capital or the number of banks is the best or even appropriate measure of the importance of banks to a country's economy. My point is simply that for earlier periods of financial history, such as the one discussed here, such measures are the only ones currently available for making cross-national comparisons.

Carrying the comparison to 1825, Cameron gives the banking capital figures for England and Wales in that year as £8.5 million (\$40 million) without the "Rest," and £11.4 million (\$53.6 million) if the "Rest" of the Bank of England is included. The corresponding U.S. figures for 1825, it will be recalled, are \$93.7 million and, including the Second Bank, \$128.7 million. This comparison may come as a surprise to historians who were brought up on the stylized facts of Britain as the world's banking and financial leader of the early (and later) nineteenth century, and the United States as a new country attempting, in halting ways, to establish an orderly banking and financial system. To my knowledge, no one before has drawn attention to the point that leaps out of the comparative data, namely that as early as 1825 the United States, with a population approaching that of that of England and Wales, apparently had 2.4 times the latter's banking capital.

The comparison lends some perspective to the effects of the Federalist financial revolution on early U.S. economic development. England, to be sure, was a wealthy, rapidly modernizing country, but in 1825 it still had quite restrictive laws limiting banks to six partners and not limiting their liability. Of banks in England and

Wales, only the Bank of England possessed a corporate charter and limited liability. The United States in most senses was less wealthy than England, but it grew economically even more rapidly by leveraging what wealth it had, in large part by means of corporate banks with limited liability.⁵ That is part of the point, of course: The United States obviously was more liberal than England in its approach to banking. Some fraction of U.S. banking capital in 1825, to be sure, was supplied by British investors, but it is unlikely to have exceeded 10 percent of the total.⁶ If those securities had been repatriated and the proceeds invested in English banks, the United States, with a similar population in 1825, would still have had about twice the banking capital of England and Wales.

The comparison can be carried still further. Table 4 gives the authorized banking capital, 1790-1835, of the four cities—Boston, New York, Philadelphia, and Baltimore—identified in Table 3 as having active securities markets in 1811. In 1825 the 50 state-chartered banks of the four cities had fully a third of the authorized capital of the 330 state banks then existing, \$53.5 million of \$156.1 million. Applying the 0.6 factor, which likely is too conservative, yields an estimated paid-in capital of \$32.1 million in the four cities. The state banks of the four cities thus had 60 percent of all the banking capital of England and Wales in 1825. If the \$35 million of capital of the Second Bank, which did a considerable part of its business in the four cities, is added to the four-city banking capital, the total, \$67.1 million, considerably exceeds the England and Wales figure of \$53.6 million.

These comparisons suggest that something quite significant for future economic development occurred in the first decades of U.S. history. An effective, efficient securities market emerged immediately as the Federalists transformed the national debt from junk paper to high-grade securities and established a large national bank. The presence of the securities market and the Bank of the United States then encouraged states to charter banks and other corporate enterprises with increasing liberality as

⁵ By the 1830s, writers in England and America were debating which nation had the better banking system. For a discussion of the issues and the respective views, see Sylla (1985).

⁶ In the late 1820s, foreign investors held about a fourth of the stock of the Bank of the United States, which was capitalized at \$35 million. President Jackson used foreign ownership for xenophobic effect in his battle with the bank, even though foreign stockholders could not vote their shares. But it is unlikely that foreigners held much stock in state-chartered banks. The 1830s demonstrated that foreign investors were willing to purchase state debt issued to establish banks, but not, it seems, stock in the banks themselves.

Table 4

State Banks and Authorized Bank Capital in Boston, New York, Philadelphia, and Baltimore, 1790-1835

Year	Boston		New York		Philadelphia		Baltimore	
	No.	Capital	No.	Capital	No.	Capital	No.	Capital
1790	1	0.8	1	1.0	1	2.0	1	0.3
1795	2	1.6	1	1.0	2	5.0	2	1.5
1800	2	1.6	2	3.0	2	5.0	2	1.5
1805	3	2.8	3	4.3	3	7.0	3	4.5
1810	3	4.8	4	6.3	4	8.3	8	8.2
1815	6	7.9	8	15.8	8	11.8	9	9.7
1820	8	8.9	9	16.3	8	11.8	10	10.2
1825	16	13.2	16	20.4	8	11.8	10	10.2
1830	23	17.3	21	22.6	9	12.0	8	8.2
1835	34	23.5	26	24.6	12	17.6	10	11.2

SOURCE: Derived from J. Van Fenstermaker, *The Development of American Commercial Banking, 1792-1837*, Kent State University Press, 1965, Appendix A.

time went on. The inherent appeal of the corporate form, particularly its limitation of stockholders' liability, and the liquidity the securities markets gave to corporate stock, encouraged domestic investors to take up the ever-growing stock offerings.

Banking corporations in the New England and Middle Atlantic regions were the leading sector of this development. Because American wealth at the time of the Federalist financial revolution was illiquid—tied up in real estate, slaves, commercial ventures, and the like—the first banks, in addition to financing domestic and international commercial expansion, also provided accommodation loans to purchase stock, both in themselves and in other enterprises. The emergence of domestic securities markets gave liquidity to such stock. International markets for U.S. securities also helped to fund early U.S. banks. For if Samuel Blodget (1806) was roughly correct in his estimate that in 1803 nearly half of U.S. securities, mostly federal debt and Bank of the United States stock, had been sold to foreign investors, then it is likely that a great deal of the funds that went into early state-bank and other corporate stock offerings came from the proceeds of those sales. With their successful emerging market, Americans were able to sell large amounts of their

highest-quality securities to overseas investors and as a result to gain funding for domestic investments. These banking-securities market interactions provide an explanation for the rapid spread of banking in the American Northeast. By 1830, the New England and Middle Atlantic regions, which then contained 43 percent of U.S. population, had—according to the estimates of Table 1—fully 86 percent of the nation's banks and 72 percent of its banking capital. By the third and fourth decades of the nineteenth century, there was probably no place in the world as “well banked” and “security marketed” as the northeastern United States. Banks and securities markets complemented each other, of course, and it is probable that they had much to do with the Northeast's rapid transportation and manufacturing developments.

Lamoreaux's recent analysis of banking development in New England, where the business developed extensively in the early decades of the nineteenth century, describes the bank-capital market interaction in full sway:

By securing a charter for a bank, [entrepreneurs] obtained a vehicle that, almost if by magic, could assist them in raising funds. First, the incorporators subscribed for a controlling

interest in the new bank's stock; then, when payment for the stock became due, they borrowed the requisite sum from another institution. Such loans were not difficult to obtain, because they were essentially riskless. As soon as the state's examiners had satisfied themselves that the new bank's capital had actually been deposited, the investors could borrow back the money they had tendered for their stock (using the stock itself as security for the loan) and repay the original debt.

. . . The main source of funding for banks during this era was the sale of bank stock, for which savings institutions, insurance companies, charitable associations, and private individuals proved willing purchasers. Thanks to this market, the original investors were usually able to sell off some of their share holdings once their bank had been in operation for a few years. They could then use the proceeds from these sales to repay their stock loans at the bank Over time, as the bank established a market for its securities, the proprietors could raise additional funds by increasing the bank's capitalization and selling new shares (Lamoreaux 1994, pp. 19-20).

As New England's banks proliferated and became intimately bound up with banker-entrepreneurs' industrial ventures, a process Lamoreaux documents in detail, New England led the way in U.S. industrial development. The financial system that sprang up out of Federalist measures in the 1790s quickly became an essential underpinning of the country's modernization and growth.

BANKS AND STATE FINANCES

A securities market standing ready to finance bank IPOs was not the only reason U.S. banks proliferated. Another reason state-chartered banks grew rapidly in num-

bers is that states derived considerable revenues from banks. Initially, these revenues came from investment in bank stock.

When banks were chartered, states commonly reserved the right to subscribe at par to bank stock, and as banks proved profitable, the states exercised these rights. They thus obtained dividend revenues that, given their limited budgets, allowed broad-based property taxes to be kept low or even to be eliminated. States also demanded and received bonus payments at the time charters were granted or renewed (Sylla, Legler, and Wallis, 1987).

The practice of making the state an investor in banks presented something of a dilemma. On the one hand, more bank charters with bonus payments and stock reserved for the state could lead to increased revenues. On the other hand, if a large number of banks were chartered and competed with each other, then bank profits and dividend rates might fall (Schwartz 1947; Wallis, Sylla, and Legler, 1994). States that remained investors resolved the dilemma by not chartering "too many" banks. In this they were heartily encouraged by the banks they had already chartered, who had an interest in limiting competition in their business.

This was the solution in most states during the two decades after 1790, which likely accounts for the steady but moderate spread of banking in those decades. Many requests for charters in these years were rejected or delayed. Hence, private (unincorporated) banking flourished, and states attempted to restrict this "unauthorized" competition for their chartered institutions by passing restraining laws, usually to ban private-banknote issues (Sylla, 1976). Outside of New England, this pattern continued into the 1810s and 1820s, and even later, and it led to a certain amount of legislative corruption and political cronyism, as might be expected when the demand for bank charters greatly exceeded the supply. New York, Pennsylvania, and Maryland banking were especially vulnerable to problems of this sort. New York's free-banking law of 1838, marking the end of legislative chartering of

individual banks in the state, was a reform measure designed to stop corruption and cronyism by opening up banking to any and all who met legally prescribed rules. It became a model for other states and, in 1863, for the federal government's National Banking System.

The New England states, beginning with Massachusetts in 1812, hit upon another solution that had the effect of reconciling state financial interests in banks with the excess demand for charters. Massachusetts ended its practice of investing in bank stock and replaced it with a tax on bank capital. Under this system, the more bank capital there was, the greater were state tax revenues, so the Massachusetts legislature, followed by other New England states, began routinely to grant charters whenever they were requested. That is a major reason why, as Table 1 indicates, in New England the numbers of banks and, to a lesser extent, the amount of capital invested in them outpaced banking growth elsewhere. Because it was in the fiscal interest of New England state governments to charter more banks, the region effectively had free banking through the liberal use of legislative chartering well before free banking by means of administratively granted charters arrived in New York in 1838.

New England's experience is indicative of what might have happened elsewhere had there been a better alignment of incentives among legislators and bankers. It is evident that the capital markets stood ready to supply more capital for banks—many bank IPOs were oversubscribed throughout the era. The U.S. banking system did develop rapidly in its first decades, but it could have developed even more rapidly than it did in states where legislators and bankers sometimes had a mutual interest in limiting bank entry.

INCREASING FINANCIAL SYNERGIES: BANKS AND SECURITIES-MARKET LOANS

Myers, the historian of the New York money market, long ago drew attention to one aspect of the intimate connection of

banks and securities markets that grew up early in U.S. history:

The most distinctive feature of the present-day money market in New York is the call loan The demand loan secured by stocks and bonds is a peculiarly American product, and it is important not only by reason of that fact, but also because it has always been closely linked with other parts of the money market. Upon the supply side it has been intimately connected with the reserves upon which the entire banking structure of the nation rested, so that banks were dependent upon the call loan market for funds in times of crisis. On the demand side, it formed the basis for the investment market, securing the funds with which it operated through the medium of call loans and building up the technique of stock trading around them [I]ts relation to bank reserves was not assailed until the passage of the Federal Reserve Act in 1913, and its position in the speculative transactions of the Stock Exchange is still untouched (Myers 1931, p. 126).

Myers was not certain when the call-loan innovation developed, although she notes that it was well established under that name in the 1840s. In the 1830s, New York City newspapers published rates for temporary loans on stocks (Myers, chpt. VII). By the 1820s at the latest, New York City banks were holding substantial net balances of out-of-town banks, for purposes of bank-note redemption and to provide New York City exchange for their customers, and the City's banks reported loans on securities collateral to a near-identical amount (Myers, p. 128). It is likely that the practice of out-of-town banks keeping balances in New York City was nearly as old as the banking system. Wright (1996, p. 321) reports that almost as soon as it was organized in 1803, Albany's New York State Bank deposited \$40,000, a substantial chunk of its resources, in two New York City banks to provide for note redemption. It is likely that other country banks of New

York State did so, too, and for the same reason—to give their notes greater currency.

By the first years of the nineteenth century, New York was emerging as America's leading port city. Imports arriving there were distributed throughout the country, which meant that out-of-town merchants needed New York exchange to pay for the goods. That is why banknotes were routed to New York City, why out-of-town banks found it convenient to keep redemption funds there, and why out-of-town bankers' balances in the city in excess of what was needed to redeem country bank notes were useful in providing bank customers with drafts payable in New York.

Another attraction was that New York banks were able to pay interest to bankers on balances held with them. Here the presence of the nation's most active securities market was critically important. The securities market, as Myers noted, became a source of demand for loans to carry investments in stocks and bonds, which served as liquid collateral for loans from the city's banks. If a city bank needed to, it could call in securities market loans, and the borrower could either arrange a new loan or dispose of securities on the market to meet the call. Since New York City banks could lend out-of-town bankers' balances on liquid securities collateral, they could afford to pay interest on them, which made keeping balances in New York City banks attractive to out-of-town banks. Such balances became still more attractive when some states (and in 1863, the United States, for National Banks) enacted reserve requirements and allowed banks to count New York City balances as reserves.

With the development of securities trading, which was funded with growing amounts of bankers' balances, the U.S. banking system in a sense returned a favor. When banks were first becoming established, the securities market funded them by providing capital.⁷ Then, as banks concentrated their reserves in money centers, particularly in New York City, money-center banks found that short-term loans and call loans on securities collateral were a good use of those

funds because of the liquidity the securities market imparted to the collateral. These synergies of American banks and securities markets were well established by the 1840s but were being established throughout the previous four decades.

CONCLUSIONS

Historical and policy conclusions emerge from the foregoing account of banking and securities market interactions from the first years of U.S. history. Economic historians of the United States have known for some time now that the American economy was growing at modern rates by 1840, as well as that such modern rates of increase in per capita product were not evident before 1790. In the intervening five decades, something—or some things—happened to accelerate U.S. economic growth. Lingering obsessions with “the industrial revolution” and “the transportation revolution” made factories, canals, and railroads among the candidates for things that happened to accelerate U.S. growth. Sometime earlier, cotton and cotton exports had had their day, too, but they had been found wanting and dismissed from the list of things. It is worth noting, however, that factories came mainly after 1815, as did canals, and that railroads were not a factor until the 1830s, and perhaps not a significant factor until the 1840s and 1850s when, we now know, U.S. growth rates already had become modern. Could it be, then, that financial change was the jump-starter of economic change? It is a strong possibility, it seems, because the other candidates—factories, canals, railroads, even cotton—all relied on modern financing methods. Those financing methods were not present until the 1790s, well before the other candidates came on stream. It remains to be seen whether further analysis of the Federalist financial revolution and its impacts on American economic life will persuade others that U.S. economic growth was “finance-led.”

Since this is a policy conference, the main policy lesson I would derive from early U.S. financial experience is the impor-

⁷ A by-product of bank access to securities market capital in early U.S. banking was that capital accounts were larger items relative to deposit and note liabilities than was common in other countries and, later, in U.S. bank liability structures.

tance of taking a broad view of financial development and paying attention to the manifold ways in which components of a financial system, such as banks and securities markets, can complement and reinforce one another. Banks are important, but they are hardly the whole story, as historians have sometimes implied, of modern-era financial development. Applying the lessons from U.S. history of the emerging market of two centuries ago, we might conclude that governments need to get their own finances in order, to turn national debts into national blessings (as Hamilton in 1781 presciently argued they might be), to establish solid monetary and payments systems and a central bank to oversee them, and to align the public and private interests in fostering the development of private financial institutions and markets. Since all these things were fairly well done in the United States two centuries ago—and many, perhaps, even earlier in the Dutch Republic and Britain—without the help of a World Bank and an IMF, it ought not be so difficult to do them now, given the political will and the presence of such dedicated international institutions. History, at least, offers encouraging examples.

REFERENCES

- Alfred Cowles III, and Associates. *Common Stock Indexes*, 2nd ed., Principia Press, 1939.
- Beard, Charles A. *Economic Aspects of Jeffersonian Democracy*, Macmillan, 1915.
- Blodget, Samuel. *Economica: A Statistical Manual for the United States of America*, self-published in 1806; reprinted by Augustus Kelley, 1964.
- Calomiris, Charles. "The Cost of Rejecting Universal Banking: American Finance in the German Mirror, 1870-1914," *Coordination and Information: Historical Perspectives on the Organization of Enterprise*, Naomi R. Lamoreaux and Daniel M.G. Raff, eds., University of Chicago Press, 1995, pp. 257-315.
- Cameron, Rondo et al. *Banking in the Early Stages of Industrialization: A Study in Comparative Economic History*, Oxford University Press, 1967.
- Davis, Joseph S. *Essays in the Earlier History of American Corporations* (2 vols.), Harvard University Press, 1917.
- Davis, Lance E., and Robert J. Cull. *International Capital Markets and American Economic Growth, 1820-1914*, Cambridge University Press, 1994.
- Fenstermaker, Joseph Van. *The Development of American Commercial Banking, 1782-1837*, Kent State University Press, 1965.
- Ferguson, E. James. *The Power of the Purse: A History of American Public Finance, 1776-1790*, University of North Carolina Press, 1961.
- Goldsmith, Raymond W. *Comparative National Balance Sheets: A Study of Twenty Countries, 1688-1978*, University of Chicago Press, 1985.
- Hubbard, J. Timothy William. *For Each, the Strength of All: A History of Banking in the State of New York*, New York University Press, 1995.
- Lamoreaux, Naomi R. *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England*, Cambridge University Press, 1994.
- Macaulay, Frederick R. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*, National Bureau of Economic Research, 1938.
- Myers, Margaret. *The New York Money Market: Origins and Development*, Columbia University Press, 1931.
- Neal, Larry. *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason*, Cambridge University Press, 1990.
- Perkins, Edwin J. *American Public Finance and Financial Services, 1700-1815*, Ohio State University Press, 1994.
- Schwartz, Anna J. "The Beginning of Competitive Banking in Philadelphia, 1792-1809," *Journal of Political Economy* (October 1947), pp. 417-31.
- Smith, Walter B., and Arthur H. Cole. *Fluctuations in American Business, 1790-1860*, Harvard University Press, 1935.
- Sylla, Richard. "Early American Banking: The Significance of the Corporate Form," *Business and Economic History*, Second Series (Vol. 14, 1985), pp. 105-23.
- _____. "Forgotten Men of Money: Private Bankers in Early U.S. History," *Journal of Economic History* (March 1976), pp. 173-188.
- _____. *The American Capital Market, 1846-1914: A Study of the Effects of Public Policy on Economic Development*, Arno Press, 1975.
- _____, John B. Legler, and John J. Wallis. "Banks and State Public Finance in the New Republic: The United States, 1790-1860," *Journal of Economic History* (June 1987), pp. 391-403.
- _____, Jack W. Wilson, and Robert E. Wright. "America's First Securities Markets: Emergence, Development, and Integration," Working paper presented at Cliometrics Society Meetings, Toronto (May 1997) and NBER Summer Institute (July 1997).
- Wallis, John J., Richard Sylla, and John B. Legler. "The Interaction of Taxation and Regulation in Nineteenth Century U.S. Banking," *The Regulated Economy: A Historical Approach to Political Economy*, Claudia Goldin and Gary Libecap, eds., University of Chicago Press, 1994, pp. 121-44.
- Wright, Robert E. *Banking and Politics in New York, 1784-1829*, UMI Dissertation Services, 1996.

Kenneth A. Snowden is an associate professor of economics at the University of North Carolina at Greensboro.

Commentary

Kenneth A. Snowden

We have learned a great deal over the past two decades about the principal components of the American financial system between 1790 and 1840—banks, nonbank intermediaries, and the market for corporate and government securities. In his paper, Richard Sylla assesses this literature and argues for a new interpretation of financial development during the early national period. Sylla is well qualified for the task because he, along with collaborators, has contributed much of what we know about the structure and performance of the U.S. securities market and the finances of state and local governments before 1840. Sylla integrates the histories of the banking system and the capital market to better explain the broad contours of early U.S. financial development and its impact on the nation's growth. In doing so, Sylla articulates an interpretation that informs, challenges, and stimulates both financial historians and students of economic growth.

Sylla's argument has three parts. First, he observes that modern scholars have focused too heavily on the early development of the U.S. banking system and examined this component of the financial system in isolation. The reality, he argues, is that banks and the securities market had close ties from early on in the United States and grew up together. Sylla's second point is that the Federalist financial policies provided the initial impetus for interrelated financial development in the 1790s and continued to support it through the early decades of the nineteenth century. As a result, the United States was blessed early on with a financial system that rivaled the best in Europe and surpassed all other New World competitors. This leads to Sylla's

third and most important point: Economic growth in the United States was financed before 1840.

Sylla's argument is modest in neither goal nor scope. He believes that U.S. financial-market development was rapid, pervasive, and growth-enhancing before 1840, and that government policy was instrumental to the process. This big picture should stimulate renewed interest in financial development during the early national period of U.S. history and will certainly shape that investigation. In this comment, I offer my thoughts concerning these future directions by considering the strengths and weaknesses of Sylla's three major themes.

INTERRELATED FINANCIAL DEVELOPMENT

Sylla's central insight is that the development of the U.S. banking system and that of the securities market were closely linked and mutually reinforcing during the early nineteenth century. Like many important historical generalizations, the point seems obvious once made—after all, Hamilton, Gallatin, Girard, and others figure prominently in historiographical treatments of both banks and federal debt management policies for this period. Sylla looks beyond the accomplishments of great men and for the first time clearly focuses our attention on the functional connections between the markets in which they participated.

Sylla (1975) alerted us two decades ago to one element of the linkage between banks and the securities market that was still at work during the post-Civil War era. He argued then that the National Banking Act cemented correspondent relationships between country and city banks and effected a flow of reserves from rural areas to the New York City call loan market. He noted that the historical roots of this mechanism date back at least to the 1830s and represent one way in which the devel-

opment of the banking system supported the securities market.

Sylla (1998) is more concerned with the opposite direction of impact—how the development of the securities market stimulated bank growth. The point is made by connecting two previously separate bodies of evidence regarding financial development during the early national period. The first concerns the performance of the securities market between 1790 and 1810. Sylla, Wilson, and Wright (1997) have recently shown that the market for U.S. government securities quickly became modern in structure and efficient in performance. This result is important in its own right because it expands the existing literature on historical securities-market performance. We already know, for example, that American stocks were priced efficiently by the 1870s (Wilson and Jones 1987; Snowden 1987) and that they displayed “modern” patterns of returns and volatility as far back as 1802 (Schwert 1990). So the four results that Sylla (1998) cites appear to extend the “early security market efficiency thesis” even further back in time. More to the point for the argument of this paper is the fact that the market for government debt was already deep, broad, and sophisticated as early as the 1790s.

A second set of facts emerges from a re-examination of Festernmaker's data on U.S. banks. Sylla is careful to point out the warts in these numbers, as well as the difficulties that arise when one compares American and British bank capital for this period. Nonetheless, the results of his analysis are persuasive and startling: U.S. bank equity grew rapidly in the first few decades of the nineteenth century and by 1830 was probably twice as large as the banking capital of England and Wales.

Sylla draws these two generalizations together by asking which types of financial assets, other than federal debt, were traded in the precocious American securities market. Most important, we learn, was bank stock—most prominently, shares of the First Bank of the United States (BUS) and the equity of large, state-chartered urban

banks. Remarkably, these securities were traded in all major regional exchanges by 1811, even though the nation's commercial banking sector was still in its infancy. From this evidence, Sylla concludes that the development of the securities market and banks were intertwined and mutually reinforcing from 1790 to at least 1830. Early in the period, the stocks of the federal government and the federally chartered Bank of the United States dominated, while the securities market became deep, broad-based, and efficient. State-chartered, incorporated commercial banks then used the market to grow rapidly.

I hope and expect that Sylla and others will continue to refine our understanding of the bank/securities market nexus. Two potential lines of inquiry strike me as particularly intriguing. First, it is important to assemble at least rough estimates of the growth in banking services (loans, discounts, and note issues) to complement Sylla's estimates of the growth in bank capital. U.S. banks tended to operate at relatively high, and potentially variable, capital ratios during this period, and we will ultimately want to focus on variations in the output of American banks across space and time. Additional data like these will be hard to come by and, in my opinion, are unlikely to substantially change the estimates of aggregate bank growth that Sylla (1998) presents. These data would, however, provide a clearer picture of how widely and deeply the benefits of bank development were distributed across the economy.

A more exhaustive analysis of bank stock transactions would also provide better focus on how bank development benefits were distributed across the economy. The evidence presented in his Table 3 successfully establishes Sylla's basic point: An active market in bank stocks was operating between 1800 and 1820. But this picture is both more and less compelling than Sylla's discussion indicates. On the plus side, the bank stocks that were listed and quoted in 1811 for the Baltimore, Boston, New York City, and Philadelphia markets include every bank that had been chartered by 1810 in these four cities (Perkins 1994, p. 274). But

in these same markets not one stock was listed from the group of 80 banks operating outside the major money centers by this date. Their absence offers little support for Sylla's claim that the equity transactions involving large urban banks represent only a "tip of the iceberg" that extended to smaller banks in smaller cities. I doubt, moreover, that further analysis of the financial pages of big city papers will change the picture very much because the two later dates (1811 and 1817) represented in his Table 3 were both years of very rapid bank growth as the industry adjusted to the disappearance of the First Bank of the United States and then to the establishment of its successor (Walsh 1940, p. 123).

To be fair, Sylla cautions us to expect that dealings in the equity of smaller banks would have been more infrequent and less publicized than those of their big-city counterparts. But we will need more evidence to establish that the bank/securities market connection was pervasive and not narrowly confined to major Eastern trade centers. Perhaps one could identify wider impacts that resulted from the securities market dealings of the largest banks: Did owners of smaller institutions use these markets to diversify their bank-related investments? Did large banks raise funds to subsidize the development of their correspondents? Alternatively, one could undertake the painstaking task of culling through the histories of smaller bank enterprises in different regions to construct a systematic picture of how their equity was initially raised and how often that equity was then traded.

These suggestions for future work are not intended to detract from Sylla's accomplishment. To the contrary, they reflect my belief that Sylla has refocused and enlivened the debate on financial-market development during the early national period by clearly exposing the connection between banks and the securities market.

THE FEDERALIST FINANCIAL REVOLUTION

Government policy plays a key role in Sylla's analysis and conclusions. He argues

that modern, private financial arrangements emerge only if the public sector provides an infrastructure of sound government debt and solid monetary arrangements. There can be no substantial disagreement with this general historical observation. But financial development is ultimately driven by the private sector's real demand for intermediation and the legal and informational constraints that determine the types of financial contracts that can be written and enforced. I would have liked to have seen much more about these influences in Sylla's paper. But Sylla takes the view that the monetary and fiscal arrangements chosen by the public sector determined the pace and character of early American financial development. He also asserts that the Federalist Revolution represented the critical watershed in the process.

I cannot disagree that government policy deserves a place at center stage. In 1770, the colonists were still prohibited from organizing banks and forced to seek mercantile credit from British agents. The American Revolution swept this structure aside but created financial exigencies that drove government fiscal and monetary policies for at least a decade. And then, only five years after the cessation of hostilities, a constitutional convention began rewriting the laws of the land. These laws included the most basic fiscal and monetary rules of the game. A settlement committee began to apportion the overhanging war debt among the state and federal governments. By 1790, the American financial system had endured two decades of continuous turmoil and uncertainty in the basic public policies that condition the private sector's ability to develop and implement its own financial innovations.

From this perspective, an obvious achievement of the Federalist program was a sorely needed measure of institutional stability. But Sylla seems to argue that Hamilton's revolution was not simply permissive in character: Federalist policies provided a particular direction for private-sector financial development and were designed to accelerate the process. I

am ambivalent about Sylla's position on this last point because I have trouble following this part of his argument. When he cites Beard at some length, for example, he endorses the idea that Hamilton and his allies resolved a fundamental controversy over alternative models of financial structure in favor of the "financial, commercial and industrial classes" who opposed "state parochialism." But even though Beard clearly believed that Hamilton was responding to specific "interested demands" rather than general political theories, Sylla chooses not to identify the particular private financial arrangements or interest groups that he believes Hamilton had in mind.

At times in the discussion, in fact, it appears that the Federalist program was imposed on an environment in which no coherent private financial interests had emerged. Commercial banks had appeared in the 1780s, but "these were isolated, local" institutions. And even though private investors held and traded large amounts of risky state and federal debt during that decade, "there was no organized capital market." These characterizations may be historically accurate, but they provide little information about the forces that drove financial innovation and development in the private sector prior to the Federalist Revolution, however modest the gains may have been.

I found Sylla's discussion of the eventual impacts of the Federalist program to be much clearer. The reason, I believe, is that policy is directly connected here to the bank/securities market connection discussed earlier. Under Hamilton's program, the federal government funded \$77 million of securities in 1790 and 1791, or a debt equal to 40 percent of gross domestic product. The taxing power of the new constitutional government secured these bonds, which were made even more attractive by the provision that \$6 million of the debt could be used to subscribe to the stock of the BUS. Together, these policies created a much deeper and more active securities market than would have emerged in their absence.

FINANCE-LED ECONOMIC GROWTH

Sylla concludes his paper with a discussion of the most important implication of his analysis: American economic growth was finance-led. The argument is appealingly simple. He notes that attempts to explain the onset of growth by identifying a specific "leading sector" have been tried and have generally failed. This leads to the possibility that all of the "real" candidates—cotton, factories, canals, and railroads—were stimulated by the rapid pace of financial development after 1790.

Financial historians, and I am one, are predisposed to conclusions in which financial markets matter. But we know, and I am sure Sylla is aware, that it will take much more to make the case for finance-led growth than simply proposing it as the best, new candidate for the role of a leading sector. He provides no specific guidance to those who will take up this challenge, nor will I. For now it is enough to observe that Sylla has identified the bank/securities market connection as the driving force of early American financial development, has characterized the pace at which this mechanism worked, and has explained how it was enhanced by government policy. He has given us much to think about and left us with much work to do.

REFERENCES

- Perkins, Edwin J. *American Public Finance and Financial Services, 1700-1815*, Ohio State University Press, 1994.
- Schwert, G. William. "Indexes of U.S. Stock Prices from 1802 to 1987," *Journal of Business* (July 1990), pp. 399-426.
- Snowden, Kenneth. "American Stock Market Development and Performance, 1871-1929," *Explorations in Economic History* (October 1987), pp. 327-53.
- Sylla, Richard. *The American Capital Market, 1846-1914: A Study of the Effects of Public Policy on Economic Development*, Arno Press, 1975.
- Sylla, R., J.W. Wilson, and Robert E. Wright. "America's First Securities Markets: Emergence, Development and Integration," unpublished manuscript (1997).


REVIEW

MAY/JUNE 1998

Walsh, John Joseph. *Early Banks in the District of Columbia, 1792-1818*, Catholic University of America Press, 1940.

Wilson, Jack W., and Charles P. Jones. "A Comparison of Annual Common Stock Returns: 1871-1925 with 1926-85," *Journal of Business* (April 1987), pp. 239-58.

Arthur J. Rolnick is a senior vice president and director of research at the Federal Reserve Bank of Minneapolis, Bruce D. Smith is a professor of economics at the University of Texas at Austin, and Warren E. Weber is a senior research officer at the Federal Reserve Bank of Minneapolis and professor of economics at the University of Minnesota.*



Lessons from a Laissez-Faire Payments System: The Suffolk Banking System (1825–58)

Arthur J. Rolnick, Bruce D. Smith, and Warren E. Weber

Should the Federal Reserve maintain its strong presence in the U.S. payments system? Or should the Federal Reserve exit and allow “the market” to produce its own mechanism for making payments? While U.S. history is replete with examples of payments systems that appear inefficient and suggest a role for government, some recent research on payments systems in the United States argues that private markets are capable of producing safe and efficient payments arrangements.

The classic, often-cited example of a privately created and well-functioning payments system is the Suffolk Banking System that existed in New England between 1825 and 1858 (see, for example, Whitney 1878, Lake 1947, Redlich 1947, and Calomiris and Kahn 1996). The Suffolk Bank of Boston operated the first regionwide note-clearing system in the United States. A result of the System was that the notes of all New England banks circulated at par throughout the region. The System’s achievements have led some (Lake 1947, p. 206, and Calomiris and Kahn 1996, p. 795) to conclude that unfettered competition in the provision of payments services can—and, in the absence of govern-

ment intervention, likely will—produce an efficient payments system. In this paper, we argue that a closer examination of the history of the Suffolk Banking System calls into question this conclusion.

Before the Civil War, U.S. paper money consisted almost entirely of state banknotes—liabilities of the bank of issue that were redeemable in specie on demand. Locally, banknotes could be exchanged at par because they were redeemable on demand. But once they circulated beyond the community of the bank of issue, the notes typically were exchanged at a discount.

In the normal course of business, virtually every bank received the notes of other banks, a fact that is apparent from the balance sheets of individual banks during this period. For example, in Maine and Massachusetts, 98 percent of all individual bank balance sheets show the bank holding notes of other banks. In New York and Pennsylvania, the fraction is between 85 and 90 percent. Thus, during this period, banks had a substantial need to clear obligations among themselves.

In the mid-1820s, the Suffolk Bank created in New England an arrangement for banknote clearing that, at the time, was unique in the United States. The Suffolk Bank started a net-clearing system for banknotes. The Suffolk System operated as follows: Members of the System were required to keep an interest-free deposit at Suffolk (or at one of the other Boston member banks). Suffolk then accepted and net-cleared all the banknotes its members deposited at par. By the early 1830s, most banks in New England had become members, and, because of Suffolk’s par-clearing policy, notes issued by members of the System were exchanged at par throughout the region.

What is most remarkable about the Suffolk Bank is that for more than 25 years, it earned extraordinary profits and was the only net clearer of banknotes in

* The authors thank the Baker Library, Harvard Business School, for the materials provided from its Suffolk Bank Collection, and Ed Green, Jamie McAndrews, and Randy Kroszner for their very constructive comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

New England. Why was Suffolk so profitable? And why did it take so long for another provider to enter the market? Our answers to these questions are based on Suffolk's having benefited from large economies of scale and scope and from finding ways, including some help from government, to protect its market share.

We find, therefore, that the Suffolk Banking System may not support the case for a laissez-faire approach to the payments system. The history of the Suffolk Banking System suggests that note clearing is a natural monopoly. And there is no consensus in the literature about whether or not the unfettered operation of markets in the presence of natural monopolies will produce an efficient resource allocation.¹

We proceed as follows. In the next section, we present the history of the Suffolk Bank as it evolved from an ordinary Boston bank into a note-clearing bank for all New England. Then, we document the Suffolk Bank's extraordinary profits by showing that it was more profitable than any other bank in New England during the period that the Suffolk Banking System was in operation, and we argue that the Suffolk Bank had a monopoly on the note-clearing business in New England. Following this, we interpret the Suffolk Banking System's history, and we suggest that the note-clearing business may have been a natural monopoly. We also suggest ways that the Suffolk Bank was able to maintain its extraordinary profits for so many years before a new entrant was able to drive it out of business. In the concluding section, we draw some lessons from the Suffolk Banking System and recommend further lines of research.

THE HISTORY AND EVOLUTION OF THE SUFFOLK BANKING SYSTEM

Origins, 1818–25

Before the Civil War, virtually the entire circulating medium of the United States consisted of privately issued banknotes. These notes were issued primarily by state banks that operated

according to provisions of the charter granted by the state in which they were located. For the most part, banknotes were redeemable in specie on demand, although penalties for nonredemption were often minimal.

By the early 1800s, the Commonwealth of Massachusetts had chartered several banks located not only in Boston, but also in other parts of Massachusetts and in the province of Maine. The banks of Boston soon became concerned about the quantity of country banknotes (also known as *foreign money*) circulating in Boston (Redlich 1947, pp. 67–68). The banks thought that the extensive circulation of country banknotes was limiting their banknote business and reducing their profits.

In 1803, the Boston banks mutually agreed to stop accepting foreign money from their customers in an attempt to increase the banks' share of total Boston note circulation. The result of this collusion, however, was much different from what the banks of Boston expected. Instead of driving country banknotes out of circulation, the take-no-notes policy led others (known as *banknote brokers*) to take up the business of buying and redeeming country banknotes. After 1803, a person in Boston who received a country banknote could sell it to one of the city's brokers. The brokers made a profit by buying notes at a discount and transporting them back to the banks of issue for full redemption in specie. Consequently, despite the boycott by the city banks, country banks were still successful at getting their notes to circulate in Boston. According to Mullineaux (1987, p. 887), between 1812 and 1844, more than half the notes circulating in Boston were country banknotes.

In time, the success of the note-brokering business (and the lack of success in driving country banknotes out of circulation) led some Boston banks to reconsider their policy of not accepting foreign money. Indeed, the Boston banks established their own note-brokering operations some time after 1804, and the discount on country banknotes was driven down to 3 percent in Boston (Lake 1947, p. 184).

¹ For an example of a paper that demonstrates how a Pareto-efficient equilibrium can exist in a market with a monopolistic provider of one or more goods, see Edlin, Epelbaum, and Heller (1996).

In 1814, the New England Bank (of Boston) introduced an important modification in note-brokering arrangements. The New England Bank followed the strategy of purchasing the notes of country banks and allowing country banks to redeem them at the market rate of discount if they kept a permanent, non-interest-bearing deposit with the New England Bank. The activities of the New England Bank and other note brokers drove the average discount on country banknotes down to 1 percent by 1818.

In 1818, the Suffolk Bank became the seventh bank to be chartered in Boston. Shortly after starting operations, Suffolk entered the note-brokering business. Suffolk's note-brokering activity was much like the New England Bank's. Suffolk bought country banknotes from merchants, individuals, and other banks at a discount. Suffolk would then permit a country bank to repurchase its notes at the same discount paid by Suffolk—on two conditions: One was that the country bank maintain a permanent, non-interest-bearing deposit of \$5,000 with the Suffolk Bank. The other was that the country bank maintain an additional non-interest-bearing deposit as a redemption fund. Suffolk sent the notes of nonparticipating country banks—country banks that refused to make such deposits—home for full redemption.

Shortly after Suffolk entered the market for country banknotes, the discount on country banknotes declined from 1 percent to 0.5 percent. Because Suffolk's competitors were attracting most of the business (by 1820 only a handful of country banks were holding permanent deposits with Suffolk), Suffolk began to question the value of this business. By the end of that year, Suffolk decided to end the purchase of notes of nonparticipating banks. Suffolk found that the cost of returning notes of nonparticipating banks was not much less than the discount at which the notes were purchased. Competition had made note brokering hardly profitable (Redlich 1947, p. 72).

In April 1824, Suffolk devised a new strategy for dealing with country banknotes. It formed a coalition with the six

other Boston banks to export country banknotes with the goal of eliminating foreign money from the city of Boston. Each coalition member contributed between \$30,000 and \$60,000 for a total of \$300,000. This fund was to be used by Suffolk to purchase country banknotes at "the same or less discount than the New England Bank, or other banks in Boston, received it, and should send it home for redemption" (Whitney 1878, p. 15). Such purchases were to continue indefinitely until country notes ceased to circulate in Boston. As with earlier attempts to drive foreign money out of Boston, this attempt was also unsuccessful.

The System in Operation, 1825–58

The failure of its note-presentment strategy did not lead the Suffolk Bank to exit the foreign money business. To the contrary, it was soon to become the dominant player in this market. In May of 1825, the coalition of city banks, having all but given up on driving country banknotes out of Boston, suggested that Suffolk allow other banks to deposit all their country banknotes with Suffolk, which would establish a system to net clear the banknotes it received. No longer would Suffolk merely buy country banknotes in order to send them back to the issuing bank for redemption. Instead, Suffolk would accept and clear at par all country banknotes that participating banks chose to deposit. By 1826, most of the city banks had withdrawn from the coalition and had become members of the Suffolk Bank's note-clearing business, the Suffolk Banking System (Suffolk Bank 1826; Mullineaux 1987, p. 890).

The Suffolk Bank's note-clearing business was similar in many ways to its old note-brokering business. As before, to participate in the system, a country bank had to maintain a permanent, non-interest-bearing deposit with Suffolk or with another Boston member of the Suffolk Banking System: For each \$100,000 of capital, the bank had to hold \$2,000 on deposit. And, as before, a country bank had to maintain an additional non-interest-

bearing deposit that was, on average, sufficient to redeem its notes received by the Suffolk Banking System. Boston banks had to hold only a permanent, non-interest-bearing deposit. This deposit was initially set at \$30,000 but was gradually reduced to \$5,000.

A major innovation was associated with this new arrangement. Banknotes were cleared by netting the accounts of participating banks. Prior to this time, no net-clearing system for banknotes had been established in the United States.² For example, the (Second) Bank of the United States, which dealt heavily in the notes of state banks, practiced gross clearing, simply presenting each state bank's notes for redemption in specie. In addition, Suffolk offered loans—in effect, overdraft privileges—to members of the System. As we will argue, these innovations made the business attractive to all participating banks and ultimately very profitable.

The netting of banknotes worked as follows: Each day, the notes deposited by participating banks at Suffolk were sorted, and the following day, the net amount was posted to the account of the appropriate bank. The notes of nonparticipating banks were sent to the issuing bank for redemption as quickly as possible.

The process of net clearing had value to Suffolk Banking System members because it lowered the cost of redeeming banknotes. Because fewer notes had to travel back to the issuing bank for redemption, less specie had to be physically shipped among banks at a time when such shipment was relatively costly.

The net clearing of banknotes opened up another business to Suffolk. Suffolk became a major lender to other banks. As a net clearer, Suffolk offered the analog of overdraft privileges (at a price). Moreover, by holding member bank deposits and clearing member banknotes, Suffolk could establish strong relationships with banks and likely had an advantage over other potential lenders in monitoring banks' activities. In short, we think that Suffolk was able to exploit economies of scope in combining its clearing and lending activities.

By the end of 1825, Suffolk had to make some adjustments to its business. Because Suffolk had more than \$1,183 in losses due to deficiencies (counterfeit and irredeemable banknotes), it entered into a special agreement with the head of its foreign money department. “[I]n consideration of \$1,050 per annum, in addition to his regular salary, he should give bonds to indemnify the bank for all deficiencies, counterfeits, mutilated or uncurrent bills in his department” (Whitney 1878, p. 18). This agreement, while modified over time, lasted for the life of the business. The agreement is of some significance in the history of the Suffolk Bank, because it indicates that Suffolk paid to shed much of the risk associated with its day-to-day clearing operations.

In its early stages, the Suffolk Banking System was relatively small in both its clearing and its lending activities. By the end of 1825, the Suffolk Bank was receiving about \$2 million a month in country banknotes. This volume of note clearing was dwarfed by the Suffolk Bank's later activities. For instance, the Suffolk Bank cleared \$9 million a month in 1841, \$20 million a month in 1851, and close to \$30 million a month by 1858 (Trivoli 1979, pp. 15, 21). To put these numbers in perspective, monthly clearing in 1825 amounted to approximately one-half of the stock of notes in circulation in Massachusetts; in 1841 and 1851, it was equal to the entire stock of notes circulating in Massachusetts; and in 1858, it was slightly less than one-and-a-half times the stock of notes circulating in Massachusetts.

During its first years as a net clearer, Suffolk earned relatively low profits from this role. Until 1833, Suffolk's dividends (which are routinely used as a measure of profits; see Calomiris and Kahn 1996) were no higher than those of an average bank in either Boston or Massachusetts. According to Redlich (1947, p. 75), the earnings from note clearing were so low initially that “the organization was in danger of being discarded by about 1830.”

By the early 1830s, however, the Suffolk Banking System's membership had grown dramatically. By 1836, close to 300

² There was net clearing in other countries by this time, however. See, for example, the discussion of note clearing in Scotland in Kroszner (1996).

banks—the vast majority of banks in New England—were members of the Suffolk Banking System. And while participation in the System was voluntary—members did receive the benefits we have mentioned—state governments also created some additional incentives to join the System. In 1842, a Vermont law gave a substantial tax advantage to banks that were Suffolk Banking System members. And a Massachusetts law passed in 1843 prohibited banks from paying out the notes of other banks, which also gave banks incentives to clear notes through the Suffolk Banking System.

The increase in the size of the Suffolk Banking System eventually turned into a healthy increase in profits for the Suffolk Bank. Before 1825—that is, before the Suffolk Bank got into the note-clearing business—its annual dividend averaged 6.5 percent. Between 1826 and 1830, it fell slightly to 6.0 percent. Between 1830 and 1840, however, Suffolk's average annual dividend jumped to 7.4 percent. Between 1840 and 1850, the average annual dividend was more than 8 percent, and between 1850 and 1855, it was 10 percent.³ Moreover, in 1839, Suffolk paid out of its growing surplus a one-time 33.3 percent dividend (Whitney 1878, p. 41). [In 1852, Suffolk once again accumulated a large surplus, but, according to Whitney (1878, pp. 41–42), the surplus was not divided among the stockholders because it was stolen by the bank's bookkeeper.] As we discuss below, Suffolk's profits were impressive not only relative to its past performance, but also relative to all other banks in New England.

Demise, 1858–60

While Suffolk's earlier attempts at note brokering and note presentment were disappointments, its note-clearing business proved very popular and profitable. The Suffolk Banking System grew and prospered for more than three decades. The political situation changed in the early 1850s, however, and a competitor emerged that, in a surprisingly short

period, drove Suffolk out of the note-clearing business.

Opposition to the Suffolk System developed soon after Suffolk started its note-clearing business, but some 30 years passed before another note-clearing business emerged (Lake 1947, pp. 192–93). In 1826, a convention of country banks met in Boston to discuss a coordinated effort to oppose Suffolk, but no agreement was reached. Ten years later, a group of country banks opposed to Suffolk's control of the market tried to obtain a charter for a new bank for the sole purpose of establishing a note-clearing system that would compete directly with the Suffolk Banking System. Members of the group argued that Suffolk was essentially charging too much for the services rendered, and they wanted an alternative. They proposed that a new note-clearing bank be established and that the stock of this new venture be held only by member banks, so that all members of the system could share in the profits. But opponents of the new bank prevailed.⁴ The opponents argued that there did not appear to be a need for another note-clearing business, that the Suffolk System was working well, and that until the country banks acted as a group to request another, no action should be taken. Such a concerted request was not forthcoming until almost 20 years later (Lake 1947, pp. 193, 195).

Starting in the late 1840s, Suffolk started to shift (or attempt to shift) more of its costs and risks to member banks. In 1849, Suffolk adopted the policy of refusing to receive notes for redemption “unless they were assorted into two packages, one containing Boston bills only, and the other issues of other banks” (Whitney 1878, p. 41). Suffolk thereby shifted some of its operating costs onto member banks. However, much more significant were three events related to the Suffolk Bank's net-clearing business.

Throughout the operation of the Suffolk System, Suffolk had sent all Rhode Island notes to the Merchants' Bank of Providence, which then cleared them with the Rhode Island banks. In 1852, Suffolk imposed a new minimum charge of 50

³ Parenthetically, we do not think that the latter increase in dividends is attributable to changes going on within the Suffolk Banking System. The California gold discoveries led to some inflation. All short-term nominal interest rates in New England seem to have risen at this time (Homer and Sylla 1991).

⁴ According to Kroszner (1996), the request for a charter was tabled in the state legislature by supporters of Suffolk.

cents per \$1,000 of country money received from the Merchants' Bank. This action induced the Rhode Island banks to revive the proposal for the formation of a competitor to Suffolk whose stock would be owned by member banks (Lake 1947, p. 193). This proposal did not yet take off, but it would shortly.

It was also the case that Suffolk had always been exposed to some default risk on the notes it held between the time the notes were deposited and the time they cleared. Suffolk was even potentially exposed to similar risks on notes that were deposited by System members with other Boston banks (Whitney 1878, p. 46). In 1853, the Exchange Bank (of Boston) refused to redeem the notes of two Connecticut banks whose notes it had originally taken. The Exchange Bank had deposited the notes with Suffolk, and the issuers of the notes had defaulted. As a result, Suffolk reminded other Boston banks of its long-held policy "that the notes of country banks would be received only on condition that all notes would be redeemed by the agent banks" (Lake 1947, p. 194). A dispute with the Exchange Bank ensued in which the Exchange Bank claimed that it could not agree to Suffolk's terms, because it was illegal for it to guarantee the liabilities of a third bank. Suffolk's response was to notify the correspondents of the Exchange Bank that it would not accept their notes in the future. As a result, at least some of the Exchange Bank's correspondents transferred their deposits to Suffolk. The Exchange Bank was then soon to become an important supporter of a Suffolk competitor (Lake 1947, p. 194).

Finally, in 1853, Suffolk announced that it would receive no foreign money after noon each day "because the labor of sorting the bills was so great that the clerks . . . had to work late at night to complete their labors" (Lake 1947, p. 195). In response, the other Boston banks threatened to withdraw their deposits with Suffolk and form a new bank unless Suffolk took country notes until 2 p.m. They argued that "the Suffolk Bank was obtaining profits large enough to enable it to employ enough clerks to handle all

country bills received" (Lake 1947, p. 195). On this issue, Suffolk conceded.

In 1855, a charter was granted to the Bank of Mutual Redemption (BMR). This bank was intended to clear notes and make loans to member banks—as Suffolk did—and, moreover, its stock was to be owned entirely by banks that were members of the system. Apparently, the support of the Exchange Bank was instrumental in the granting of a charter to the BMR (Redlich 1947, p. 75).

Despite the support of the Exchange Bank and the Rhode Island banks for a Suffolk competitor, the BMR had difficulty raising enough capital to begin operations. Indeed, it did not succeed in raising the necessary capital to open its operations until 1858. Nevertheless, when the BMR opened, 143 banks (roughly half the banks in New England) were stockholders (Dewey 1910, p. 95).

The BMR operated much as the Suffolk System did. It required the maintenance of a permanent deposit and a clearing balance. But, unlike Suffolk, the BMR paid interest on its deposits at a rate of 3 percent per year.

The reaction of Suffolk to the entry of the BMR into the note-clearing business was at first combative. Suffolk initially intended to fight the BMR and began by refusing to redeem the notes of BMR members through the BMR. Suffolk's argument in doing so was that the BMR held no deposit with Suffolk and, hence, that banks clearing through the BMR were not entitled to the same treatment as Suffolk System members. Hence, notes issued by members of the BMR and received by Suffolk were sent to the issuing bank for immediate redemption.

In its opening salvos with the BMR, Suffolk was supported neither by the other Boston banks nor by the Commonwealth of Massachusetts. On October 11, 1858, the BMR was admitted to the Boston clearinghouse. "On the same day the [Massachusetts] Bank Commissioners . . . formally advised the Suffolk Bank . . . that it should either continue to receive the bills of all the banks which had withdrawn their deposits

and to present them at the BMR or it should decline to receive from its depositors the bills of such banks" (Lake 1947, pp. 200–01). The lack of support from the Boston banks and the attitude of the state bank commissioners apparently averted an open fight between Suffolk and the BMR.

Suffolk's next step was quite different. On October 16, 1858, Suffolk announced that it would withdraw altogether from the foreign money business. This announcement does not appear to have been an idle threat, because Suffolk did leave the business in 1860. And Suffolk's proposed withdrawal from its note-clearing activities apparently was a threat with teeth. Because the BMR could not handle anything like the entire volume of note clearing in New England, "the bank presidents asked the Suffolk Bank to continue receiving country money until February 28, 1859. They were met with a brusque refusal. Finally, a compromise was reached by which the banks were to make arrangements individually with the Suffolk Bank or Mutual Redemption bank. Under the terms made by the Suffolk Bank country money would be received for a charge of twenty-five cents per \$1,000" (Lake 1947, pp. 202–03). The 50 cents per \$1,000 that Suffolk charged the Merchants' Bank of Providence in 1853 thus appears to have exhibited a large monopoly-pricing element. Indeed, even the 25 cents per \$1,000 charge seems high relative to Suffolk's average costs, which, according to Whitney (1878, pp. 53–54), were 10 cents per \$1,000 cleared.

This was the end of the Suffolk Banking System and the beginning of the BMR. The operation of the BMR apparently benefited the country banks, whose note circulation rose (while that of the Boston banks fell) from 1858 to 1859. The BMR, however, was not profitable, and it ceased to pay interest on deposits when Suffolk halted its own note-clearing operations in 1860. The BMR did not pay its first dividend until October 1860 and then only at the (semiannual) rate of 2 percent.

THE SUFFOLK BANK'S PROFITABILITY

In another paper (Rolnick, Smith, and Weber 1997), we use annual data on bank dividends and prices of Boston bank stocks to document several facts about the profits of the Suffolk Bank relative to those of other Massachusetts banks. In this section, we summarize those results and present evidence that the Suffolk Bank appears to have been a monopolist in the provision of note-clearing services.

In Rolnick, Smith, and Weber (1997), we show that the Suffolk Bank's profits appear fairly similar to those of other Massachusetts banks until 1833. From 1834 until 1858, however, the Suffolk Bank was consistently more profitable than any other Massachusetts bank. Several kinds of evidence support these conclusions. One kind is aggregate evidence on dividend payments. In Rolnick, Smith, and Weber (1997), we show that until 1833, the Suffolk Bank paid dividends at a rate comparable to the average (or the median) of those paid by other banks in Massachusetts. However, from 1834 to 1858, Suffolk consistently paid dividends at a rate that was 2 percentage points higher than the typical rates paid either by other large Boston banks or by Massachusetts banks in general. This aggregate evidence is supported by a bank-by-bank comparison of dividend rates over the period from 1834 to 1858. This comparison indicates that although there were some years in which a small number of banks paid dividends at rates equal to or even slightly higher than those paid by the Suffolk Bank, no bank did this consistently. Further, those banks whose dividends occasionally rivaled the Suffolk Bank's were almost exclusively small, non-Boston banks.

Rolnick, Smith, and Weber (1997) also looks at prices of the stock of Boston banks during this period. These data come from Martin (1886), who compiled the yearly high and low stock prices of bank stocks in the Boston stock market. For each year from 1834 to 1858, with only the exception of 1839 and 1840, the lowest price paid for shares of Suffolk

Figure 1

**Due to Other Banks
Suffolk As a Percentage of Total
Massachusetts**

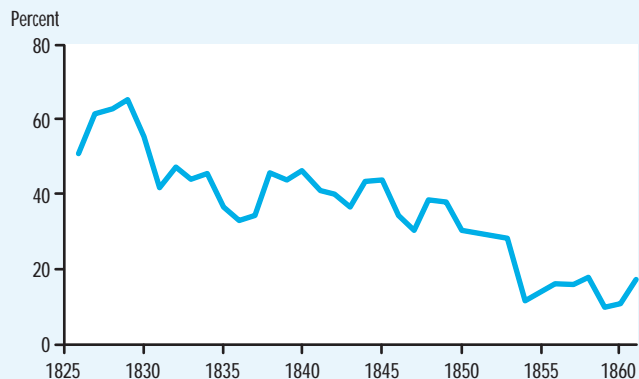
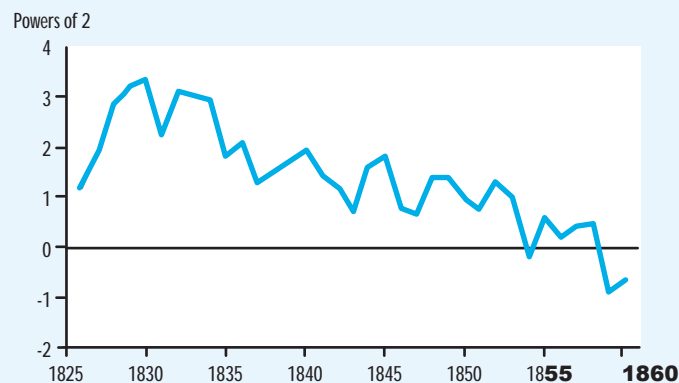


Figure 2

**Due to Other Banks
Ratio of Suffolk to Next Largest**



Bank stock was higher than the highest price paid for the shares of any other bank in Boston.

These findings allow three important points to be made with respect to the Suffolk Banking System. First, to borrow Whitney's phrase, "the [Suffolk] business was very remunerative" (1878, p. 41). Second, the fact that the Suffolk Bank routinely earned higher profits than other large Boston banks suggests that, by the early 1830s, the Suffolk Bank was acting alone in the net-clearing business, rather than as the representative of a larger coalition of Boston banks. Moreover, when the Suffolk Bank first began to earn unusual

profits in 1833, there was no corresponding increase in the profits either of other large Boston banks or of Massachusetts banks in general. Third, Suffolk's profits were *always* high. Thus, its high average profits cannot be viewed as compensation for some unusual risks it was taking.

We now present evidence that the Suffolk Bank had substantial market power and may have been a monopolist in the provision of note-clearing services, at least during the period from 1834 to 1858. We begin by establishing that the Suffolk Bank was by far the largest holder of interbank deposits.⁵ We show this in Figures 1 and 2.⁶ In Figure 1, we show the Suffolk Bank's share of the interbank deposit market. From 1828 to 1854, the Suffolk Bank consistently held between 30 and 50 percent of all "due to other banks" held by Massachusetts banks. In Figure 2, we plot the ratio of the Suffolk Bank's holdings of "due to's" to the next largest Massachusetts bank. The vertical axis in this figure is in terms of powers of 2, so that zero indicates that Suffolk's holdings are equal to those of the next largest bank, 1 indicates that Suffolk's holdings are twice as large as those of the next largest bank, and so forth. From this figure, we see that in most years, the Suffolk Bank's holdings of such deposits were at least twice as large as those of the next largest bank.

Next, we show that the identity of the banks that ranked below Suffolk in terms of the volume of interbank deposits changed frequently over time. We show this in Table 1, where we show the banks that ranked among the top five in terms of the volume of "due to" annually from 1825 until 1860. As expected from Figure 2, the Suffolk Bank virtually always has the largest amount of "due to's." However, no other bank consistently held a large share of the interbank deposit market. Up to 1845, the New England Bank and the State Bank were the banks that most frequently ranked behind the Suffolk Bank in terms of the share of the interbank deposit market. However, after 1840, those two banks were replaced in the rankings by the Merchants' Bank of Boston and the Globe Bank, and after 1850, the Bank of

⁵ Clearly, it was necessary for banks to hold deposits with a bank that was performing clearing services on their behalf.

⁶ The sources for the data used in all figures is given in Rolnick, Smith, and Weber (1997).

Table 1

Ranking of Massachusetts Banks by Amount Due to Other Banks, 1825–60

	1st	2nd	3rd	4th	5th
1825	City	New England	Other	Union	M&M/Tremont
1826	Suffolk	City	New England	M&M/Tremont	Other
1827	Suffolk	City	New England	Union	Other
1828	Suffolk	New England	Union	Other	State
1829	Suffolk	State	New England	Union	Other
1830	Suffolk	City	State	New England	Other
1831	Suffolk	New England	State	Globe	Other
1832	Suffolk	New England	Other	Union	State
1833	Suffolk	Globe	Other	State	Other
1834	Suffolk	State	City	Globe	Merchants'
1835	Suffolk	Merchants'	Other	New England	State
1836	Suffolk	Merchants'	Other	State	New England
1837	Suffolk	Merchants'	State	New England	Other
1838	Suffolk	Globe	Merchants'	New England	State
1839	Suffolk	Globe	Merchants'	New England	Other
1840	Suffolk	Merchants'	Globe	State	Other
1841	Suffolk	Merchants'	Globe	Other	Other
1842	Suffolk	Merchants'	Globe	Other	State
1843	Suffolk	Merchants'	Globe	State	New England
1844	Suffolk	Merchants'	Globe	Other	New England
1845	Suffolk	Merchants'	Globe	Other	Other
1846	Suffolk	Merchants'	Globe	Other	State
1847	Suffolk	Merchants'	Globe	State	Other
1848	Suffolk	Merchants'	Globe	M&M/Tremont	Other
1849	Suffolk	Merchants'	Globe	New England	M&M/Tremont
1850	Suffolk	Merchants'	Other	New England	Commerce
1851	Suffolk	Merchants'	Commerce	Globe	New England
1852	Suffolk	Commerce	Merchants'	M&M/Tremont	Exchange
1853	Suffolk	Merchants'	Commerce	Other	M&M/Tremont
1854	Merchants'	Suffolk	Commerce	Exchange	Globe
1855	Suffolk	Commerce	Merchants'	Other	Other
1856	Suffolk	Merchants'	Commerce	Exchange	Globe
1857	Suffolk	Merchants'	Commerce	M&M/Tremont	Other
1858	Suffolk	BMR	Merchants'	Commerce	Globe
1859	BMR	Suffolk	Merchants'	Other	Commerce
1860	BMR	Suffolk	Commerce	Merchants'	Exchange

Commerce displaced the Globe Bank in the rankings.⁷

We have already argued that a bank engaged in net clearing on a large scale

might easily exploit economies of scope by also acting as an interbank lender. Rolnick, Smith, and Weber (1997) documents that the history of the Suffolk Bank is

⁷ All of these banks were located in **Boston**.

indeed consistent with this idea. Between 1833 and 1858, the Suffolk Bank consistently held at least 15 to 20 percent of all interbank loans. Moreover, the large increase in the Suffolk Bank's profits coincided with a substantial increase in its position as an interbank lender. Indeed, in 1833, the Suffolk Bank held three times as many interbank loans as any other Massachusetts bank. In contrast, in 1831, the Suffolk Bank had interbank loans approximately equal to those of several other banks. This fact clearly suggests that the Suffolk Bank's profits derived, at least in part, from the exploitation of economies of scope in interbank lending.

AN INTERPRETATION

In this section, we attempt to interpret the facts we have just summarized and to answer the question, Why did it take over 25 years for another New England bank to enter Suffolk's market? We begin the interpretation with an observation that has been made by many other historians of the Suffolk Banking System: Suffolk was a monopolist.⁸ We also think that Suffolk was a relatively sophisticated monopolist. Its pricing practices involved a two-part tariff from 1826 on and even more elaborate nonlinear pricing schemes (and price discrimination) at later points.

These pricing practices seem to have made Suffolk very effective at garnering surplus. The data indicate that while Suffolk's profits rose dramatically in 1833, this was not true for other banks in Boston or Massachusetts. The data are therefore consistent with the notion that whatever surplus accrued to members of the Suffolk Banking System was primarily captured by Suffolk itself.

Moreover, we think that the Suffolk Banking System was a natural monopoly. It is not hard to construct arguments that there are economies of scale in net clearing and that these can be captured fully only by a system with a single net clearer. It is also not hard to construct arguments that the agent doing net clearing has cost advantages as a provider of overdrafts and as an interbank lender. Thus, we think

there is a strong presumption that Suffolk was able to exploit both economies of scale and economies of scope in its activities. And, indeed, the Suffolk Bank became unusually profitable only as it began to fully exploit both types of economies.⁹

This history of the Suffolk Banking System is, of course, fully consistent with this view. Suffolk was not an unusually profitable bank until it became a large enough player in both note clearing and interbank lending. And at least equally telling is the observation that Suffolk was not willing to split its market with the BMR. The failure of the market to sustain two net clearers is, in our minds, very suggestive of natural monopoly.

We should emphasize at this point that the presence of a monopoly—either natural or otherwise—in no way necessarily implies that any economic inefficiencies were associated with the operation of the Suffolk Banking System. Indeed, as shown by Edlin, Epelbaum, and Heller (1996), the presence of a monopolist that can engage in price discrimination and levy two-part tariffs is often fully consistent with Pareto efficiency.

In addition, if Suffolk was a natural monopoly, there is another important question. If the Suffolk experiment were repeated at another time and in another place, would we expect the Suffolk outcome to be replicated? Or, more generally, would we expect the market to produce an efficient outcome? The answer to this question can hardly be an unequivocal yes. There are many reasons, some of which are reviewed in Sharkey (1982), why the market might not produce an efficient outcome in the presence of a natural monopoly. And even an unchallenged monopolist with great powers of price discrimination and with the power to engage in nonlinear pricing need not attain an efficient allocation of resources under all cost conditions, as noted by Edlin, Epelbaum, and Heller (1996).

In general, the ability of the market to produce an efficient outcome with a natural monopoly depends strongly on cost and demand conditions in the market and on the relative strategic positions of

⁸ Whitney (1878), Lake (1947), Redlich (1947), and Bodenhorn (undated) all conclude the same thing.

⁹ Of course, there may be economies of scale and scope only over certain ranges of activity, as noted by Sharkey (1982). At some point, congestion costs may reverse decreasing average costs.

other potential market participants.¹⁰ Thus, even if one views the Suffolk experience as supportive of the notion that the free market can be an efficient provider of payment services, we do not see that one can conclude that the free market *will* lead to the efficient provision of payment services under any possible configuration of market conditions.

All of this leaves us with two final questions: How was Suffolk able to deter the entry of a competitor until 1858? And how was the BMR able to enter in 1858 and drive Suffolk out of the note-clearing business?

With regard to the first question, we think it is useful to view the industrial organization of note clearing in New England as the outcome of a game played between the Suffolk Bank and potential rivals. Through the historical accident of being asked by the other large Boston banks to be the net clearer, Suffolk was handed the position of the incumbent in the industry. Several models of industry organization in the presence of a natural monopoly exist. Although the underlying game in each of these models differs, a general implication is that the incumbent monopolist will be able to earn monopoly profits over an extended period if it enjoys some type of strategic advantage over potential entrants.¹¹

One form of such a strategic advantage is some kind of barrier to entry. In the case of the Suffolk Bank, one could think of a barrier to entry as the cost that a potential entrant would have to bear in trying to sign up banks for a rival net-clearing network. These costs are sunk because they would have to be borne by the potential entrant even if the rival never actually entered the note-clearing business. Of course, Suffolk would have already borne these costs, so they would not be relevant to its decision regarding whether or not to continue in the business. Another form of strategic advantage is the threat of predatory pricing. In the case of the Suffolk Bank, predatory pricing could have consisted of offering interest on deposits should a rival have entered. Note that even though Suffolk never engaged in offering interest on deposits, such a threat still could have been

implied. The fact that we have no record of such a threat may simply mean that the implied threat was successful. In that case, offering interest on deposits would have been out-of-equilibrium behavior because entry by a rival would never have occurred.

With regard to the second question, we think it is useful to continue to think in terms of the game described above. From the viewpoint of the relative strategic advantages in a game between an incumbent and potential entrants, the BMR was a potential entrant unlike any existing bank because its charter permitted its stock to be owned only by banks. In other words, the BMR was a rival that would be owned by its customers. This situation would change the nature of the game because now the rival would have a strategic position that was different from that of previous potential challengers. Its position might also be interpreted as lowering the sunk costs faced by the potential entrant, because one bank could see other banks' commitments to joining the competing system through their purchases of stock in the BMR.

Two other points are of interest with regard to the entry of the BMR. One is Suffolk's reaction, which ultimately was to withdraw from the net-clearing business. This is consistent with our interpretation of net clearing as a natural monopoly. The other is what the BMR did with regard to offering interest on deposits. When the BMR first entered the market, it offered interest on deposits. Once the BMR had driven Suffolk out of the market, however, it adopted Suffolk's strategy of not paying interest on deposits. This is consistent with our interpretation of temporarily paying interest on deposits as predatory pricing behavior.

CONCLUSION

Between 1825 and 1858, the Suffolk Bank of Boston operated the first region-wide note-clearing system in the United States. The Suffolk Bank, chartered by the Commonwealth of Massachusetts in 1818, evolved from an ordinary Boston bank into

¹⁰ See Bagwell and Ramey (1996) for an interesting discussion of how even an entrenched monopolist with a large productive capacity can lack the strategic wherewithal to deter entry.

¹¹ See, for example, the model in the papers by Dixit (1980) and Ware (1984) in which the incumbent enjoys the strategic advantage of being able to make a capacity commitment before the potential entrant. In the Bagwell and Ramey (1996) reformulation of the model, the strategic advantage goes to the potential entrant, however.

a note-clearing bank for all of New England. We document that it earned extraordinary profits for over 25 years and that it had a monopoly in the interbank deposit and loan markets. From this we infer that it also had a monopoly on note clearing. Our interpretation of Suffolk's history suggests ways that Suffolk was able to maintain its extraordinary profits for so many years and also suggests that the note-clearing business may have been a natural monopoly. The latter observation is of some importance because there is no consensus in the literature about whether or not the unfettered operation of markets in the presence of natural monopolies will produce an efficient resource allocation.

Future research should focus on whether or not the Suffolk Banking System was truly unique. Some have argued that a Suffolk-type system did not exist in other parts of the country. We think it would be useful to better document the types of note-clearing arrangements that existed elsewhere to determine how they differed from the Suffolk Banking System, and if they were different, what factors would account for the observed features of different payments systems.

REFERENCES

- Bagwell, Kyle, and Garey Ramey. "Capacity, Entry, and Forward Induction," *Rand Journal of Economics* (Winter 1996), pp. 660–80.
- Bodenhorn, Howard. *The End of the Suffolk System: A Critical Re-appraisal*, manuscript, Lafayette College, Easton, Pennsylvania, undated.
- Calomiris, Charles W., and Charles M. Kahn. "The Efficiency of Self-Regulated Payments Systems: Learning from the Suffolk System," *Journal of Money, Credit, and Banking* (November 1996), pp. 766–97.
- Dewey, Davis R. *State Banking Before the Civil War*, United States Government Printing Office (National Monetary Commission), 1910.
- Dixit, Avinash. "The Role of Investment in Entry-Deterrence," *Economic Journal* (March 1980), pp. 95–106.
- Edlin, Aaron S., Mario Epelbaum, and Walter P. Heller. "Is Perfect Price Discrimination Really Efficient?: Welfare and Existence in General Equilibrium," manuscript, University of California, Berkeley, 1996.
- Homer, Sidney, and Richard Sylla. *A History of Interest Rates*, 3rd ed., Rutgers University Press, 1991.
- Kroszner, Randall S. "Comment on the Efficiency of Self-Regulated Payments Systems: Learning from the Suffolk System," *Journal of Money, Credit, and Banking* (November 1996), pp. 798–803.
- Lake, Wilfred S. "The End of the Suffolk System," *Journal of Economic History* (November 1947), pp. 183–207.
- Martin, Joseph G. *Martin's Boston Stock Market: Eighty-Eight Years*, Joseph G. Martin, 1886.
- Mullineaux, Donald J. "Competitive Monies and the Suffolk Bank System: A Contractual Perspective," *Southern Economic Journal* (April 1987), pp. 884–98.
- Redlich, Fritz. *The Molding of American Banking: Men and Ideas*, Hafner, 1947.
- Rolnick, Arthur J., Bruce D. Smith, and Warren E. Weber. "The Private Provision of Payments Services: The Suffolk Banking System Reexamined," manuscript, Federal Reserve Bank of Minneapolis, 1997.
- Sharkey, William W. *The Theory of Natural Monopoly*, Cambridge University Press, 1982.
- Suffolk Bank. Directors' Records, March 14, available at Baker Library, Harvard Business School, 1826.
- Trivoli, George. *The Suffolk Bank: A Study of a Free-Enterprise Clearing System*, Adam Smith Institute, 1979.
- Ware, Roger. "Sunk Costs and Strategic Commitment: A Proposed Three-Stage Equilibrium," *Economic Journal* (June 1984), pp. 370–78.
- Whitney, David R. *The Suffolk Bank*, Riverside Press, 1878.

Randall S. Kroszner is a professor in the Graduate School of Business at the University of Chicago.

Commentary

Randall S. Kroszner

One of the key challenges for central bankers today concerns the regulation and operation of the payments system. Faced with rapid innovation in information-processing and communications technologies, central banks are struggling to understand what role, if any, they should play in the payments system of the future. These fundamental questions have prompted researchers to examine how payments systems evolved in the period before central bank control, seeking insights into how private alternatives might operate (e.g., Mullineaux 1987; Gorton and Mullineaux 1987; Cowen and Kroszner 1989, 1990, and 1994; Selgin and White 1994; Calomiris and Kahn 1996; and Kroszner 1996 and 1997). In this spirit, Rolnick, Smith, and Weber's article in this issue (RSW) evaluates the Suffolk System, which is perhaps the most important private clearing arrangement to have developed in antebellum America.

RSW challenges the sanguine view that the Suffolk System demonstrates how the unregulated market will provide an efficient payments system. It provides new evidence that the profitability of the Suffolk Bank was greater than that of other banks in Boston and Massachusetts and that Suffolk Bank had a dominant role in the interbank borrowing and lending market. The authors then raise the possibility that this evidence suggests that the market would provide a "natural" monopoly in payment services. While I will question whether the Suffolk experience can address the monopoly issue, I believe that RSW has given us a new and important direction in historical research on payments systems, namely understanding the link between clearing arrange-

ments and liquidity provision through the interbank lending market.

RETURNS AND THE RELATIONSHIP OF CLEARING AND LIQUIDITY

To make the case about natural monopoly, RSW begins by observing that for more than two decades, no competitor emerged to challenge the Suffolk Bank's note-clearing business in New England (see also Lake 1947). Next, the article examines the profitability of the Suffolk Bank for evidence of monopoly rents. Since the data that would allow us to calculate returns on assets or equity do not exist, RSW uses data on dividend payments relative to capital as a proxy for profitability. Calomiris and Kahn (1996) used dividend payment rates to compare the average profitability of banks in Boston with those in other cities during this period and found that banks in Boston did not pay higher dividends than banks elsewhere. On average, these banks do not appear to have enjoyed supernormal profits. RSW examines the profitability of individual banks in detail and finds that, from 1834 until 1858, the dividend rate for the Suffolk Bank was consistently higher than for other banks in Boston as well as for the smaller banks in the rest of Massachusetts.

The authors then investigate what might account for the relatively high dividends for the Suffolk Bank. In doing so, they provide an extremely important and original contribution to the literature. They document the dominant role of the Suffolk Bank in the interbank borrowing and lending market. The Suffolk Bank was not only the largest holder of interbank deposits, as might naturally be expected of the note-clearing agent, but also the largest interbank lender, as the authors have shown in an earlier work (Rolnick, Smith, and Weber 1997). The Suffolk Bank thus appears to have been more than simply a note-clearing agent; it also

appears to have been a major source of liquidity in New England. The increase in Suffolk's profits coincided with the expansion of its interbank lending role.

This evidence, RSW argues, suggests that there are economies of scope in the provision of clearing and liquidity services. Suffolk's detailed information about the health and activities of member banks, gleaned from operating the note-clearing system, may have reduced its costs of monitoring loans to other banks. The information advantage Suffolk gained from note clearing led to its dominant role in the interbank funds market. Recently, a number of authors have argued that this complementarity is theoretically important (e.g., Gilbert 1993, and Rajan forthcoming), but RSW provides the first empirical documentation of such a linkage. Scope economies can be used as an efficiency rationale for having the lender of last resort operate the payments system.

It would be extremely valuable to know whether Suffolk was effectively acting as a lender of last resort. During the Suffolk System era, banks in New England were more stable than in other parts of the country (Calomiris and Kahn 1996): Bank failure rates and loss rates to depositors were lower in New England, in both normal times and during the bank panics of the late 1830s and 1857. As the authors documented in Rolnick, Smith, and Weber (1997), interbank lending by the Suffolk Bank rose during the crises. Were the interbank activities of Suffolk a key contributor to the stability of banking in New England? To whom did Suffolk lend during the crises, and on what terms? What risk exposure was the Suffolk Bank willing to incur? Could higher average returns be related to the insurance role that Suffolk may have been playing?¹ A fascinating possibility to explore in future research is whether and how well the markets provided stability through a clearinghouse that was acting as a lender of last resort.

MONOPOLY: NATURAL OR UNNATURAL?

While RSW's inquiry has shed new light on the relationship between

payments and liquidity services, the evidence does not necessarily imply that Suffolk enjoyed a "natural" monopoly or that there is a tendency for the market to produce such a monopoly. A monopoly is "natural" if, for a given market size, economies of scale (and possibly scope) are sufficiently strong that production costs are minimized when there is a single producer. That producer then can drive out all competitors in the market and obtain a monopoly. As RSW acknowledges, natural monopolies do not necessarily result in socially inefficient use of resources, but they raise that possibility (see Edlin, Epelbaum, and Heller 1996).

The Suffolk System, however, did not operate in a completely unregulated environment, and regulation may have increased the costs to potential competitors and the heights of entry barriers. The difficult task is to untangle which regulations, if any, are relevant to the development of the Suffolk System and what impact they had on its operation. Although I will not attempt such a full-scale evaluation here, I will mention some potentially important considerations.

First, the Suffolk System received some special legislative support. As RSW notes, Vermont gave tax breaks to banks that joined the Suffolk System. In addition, Massachusetts did not permit banks to pay out to their customers' notes of other banks, thereby providing an incentive for banks to use the Suffolk System for note clearing. Such government encouragement may have helped to increase the profitability of the Suffolk Bank and deter new entrants.

Second, despite the frequent use of the term "free banking" to describe mid-nineteenth century banking in the United States, entry into the banking industry was far from free. Bank charters required an act of the state legislature. By 1850, only Rhode Island had passed a "free banking" statute. This statute eased, but did not make "free," entry into banking in that state. Only after Massachusetts passed its free banking statute in 1851 was a coalition of banks able to obtain a charter for what became the Bank for Mutual Redemption (BMR), the bank that triumphed over Suffolk in 1858.

¹ RSW argues that Suffolk's apparently high return was not compensation for risk because dividend rates did not drop below those of other Boston banks in any year throughout the period. We do not have direct data on the annual profits, losses, and cash flows. The relatively steady and high dividend rate, however, might mask underlying volatility, since dividends can be paid out of a surplus accumulated precisely to smooth returns to shareholders.

RSW argues that the key to the BMR's success is that, unlike other challengers to Suffolk, it was owned by its member banks, and this cooperative structure helped to reduce the obstacles to coordinating a competing network of member banks. Twenty years earlier, however, a group of country banks had tried to obtain a charter from the Massachusetts state legislature for a banker's bank that would provide a competitor to the Suffolk System (Lake 1947). The proponents of the Suffolk Bank killed the bill, and no bank owned by other banks was permitted a charter until the BMR. Suffolk was long protected from mutual organizational forms with different cost structures that might have undercut Suffolk's monopoly. Politics, not just economics, thus appears to play an important role in preventing the emergence of rivals to Suffolk (Kroszner 1996, and Kroszner and Strattmann forthcoming).

Third, and perhaps most important, were the restrictions on intrastate and interstate branching (Kroszner and Strahan 1998). If no such prohibitions on the geographic expansion of banks existed, the payments system might have developed very differently during this period. Some banks may have chosen to operate branches in major cities and towns throughout a region or, perhaps, throughout the country (much like the Second Bank of the United States). Each branch of a single bank is likely to have accepted its own notes and checks at par, regardless of the location of an individual branch. These notes would have achieved par circulation without reliance on a common clearinghouse. One or more of the widely branched banks might then have provided clearing services for other banks' notes to compete with the Suffolk System.

The branching restrictions thus may have had an important effect on the cost of producing payment services. Without branching, only one par clearing operation may have been feasible in New England. With branching, however, other banks may have faced lower costs of entering the note-clearing business, so the market may have been able to sustain multiple clearing operations. Also, geographically diversi-

fied banks may have had less demand for an interbank lending market. Rather than rely on other banks for liquidity, well-branched banks might have been able to substitute an internal interbranch funds market for the interbank market.

FUTURE RESEARCH

RSW concludes by proposing a comparative study of payments systems that developed in different parts of the United States during the nineteenth century. Given the state-by-state variation in regulation, such an investigation may shed light on the role of regulation in shaping the Suffolk System as well as other payments systems. The authors' work on the linkage between liquidity and payments systems continues to break new ground (Rolnick, Smith, and Weber 1997).


Historical payments system research also might fruitfully extend beyond the United States and banking for insights into what the market might produce. During the eighteenth and nineteenth centuries, many clearing systems emerged in European countries that had little or no financial regulation (Cowen and Kroszner 1990 and 1994; Kroszner 1990). One of the earliest examples, which has received some attention, is the note-exchange system that developed in Scotland during the 1760s (e.g., Munn 1975; White 1984; Cowen and Kroszner 1992; Kroszner 1997). In addition, commodities futures exchanges developed private clearing and settlement arrangements during the nineteenth and early twentieth centuries (Edwards 1984; Williams 1986; Moser 1994; Kroszner 1998). These systems often adopted the mutual or cooperative form that the BMR had used successfully against the Suffolk Bank, so further study of these arrangements might help us to understand how private payments systems of the future might evolve.

REFERENCES

- Calomiris, Charles W., and Charles M. Kahn. "The Efficiency of Self-Regulated Payments Systems: Learning from the Suffolk System," *Journal of Money, Credit, and Banking* (November 1996), pp. 766-97.

- Cowen, Tyler and Randall S. Kroszner. "Scottish Banking before 1845: A Model for Laissez-Faire?" *Journal of Money, Credit, and Banking* (May 1989), pp. 221-31.
- _____ and _____. "Empirical Predictions of the New Monetary Economics: Perspectives on Velocity," *Journal of Policy Modeling* (Summer 1990), pp. 265-79.
- _____ and _____. "Scottish Free Banking." *New Palgrave Dictionary of Money and Finance*, Peter Newman, Murray Milgate, and John Eatwell, eds., London: Macmillan, 1992.
- _____ and _____. *Explorations in the New Monetary Economics*, Blackwell Publishers, 1994.
- Edlin, Aaron, Mario Epelbaum, and Walter Heller. "Is Perfect Price Discrimination Really Efficient? Welfare and Existence in General Equilibrium," unpublished manuscript, University of California, Berkeley, 1996.
- Edwards, Franklin. "The Clearing Association in Futures Markets: Guarantor and Regulator," *The Industrial Organization of Futures Markets*, Ronald W. Anderson, ed., D. C. Heath, 1984, pp. 225-54.
- Gilbert, R. Alton. "Discussion Comments," in Richard E. Randall, ed., *Safeguarding the Banking System in an Environment of Financial Cycles*, proceedings of a symposium held at the Federal Reserve Bank of Boston, November 1993, pp. 138-46.
- Gorton, Gary, and Donald Mullineaux. "The Joint Production of Confidence: Endogenous Regulation and Nineteenth Century Commercial-Bank Clearinghouses," *Journal of Money, Credit, and Banking* (November 1987), pp. 457-68.
- Kroszner, Randall S. "Banking in Hong Kong: Regulation, Stability, and the Role of Money Market Mutual Funds," *Asian Monetary Monitor* (July/August 1990), pp. 16-36.
- _____. "Comment on the Efficiency of Self-Regulated Payments Systems: Learning From the Suffolk System," *Journal of Money, Credit, and Banking* (November 1996), pp. 798-803.
- _____. "Free Banking: The Scottish Experience as a Model for Emerging Economies," *Reforming Financial Systems: Historical Implications for Policy*, Gerard Caprio and Dimitri Vittas, eds., Cambridge University Press, 1997, chapt. 3.
- _____. "Private versus Public Regulatory Structures in International Financial Markets: An Analysis of Derivative Products Companies," unpublished manuscript, University of Chicago, Graduate School of Business, 1998.
- _____, and Philip Strahan. "What Drives Deregulation? Economies and Politics of the Relaxation of Bank Branching Restrictions," unpublished manuscript, University of Chicago, Graduate School of Business, 1998.
- _____, and Thomas Stratmann. "Interest Group Competition and the Organization of Congress: Theory and Evidence from the Financial Services Industry," *American Economic Review*, forthcoming.
- Lake, Wilfred S. "The End of the Suffolk System," *Journal of Economic History*, 7 (1947), pp. 183-207.
- Moser, James T. "Origins of the Modern Exchange Clearinghouse: A History of the Early Clearing and Settlement Methods at Futures Exchanges," Federal Reserve Bank of Chicago, Research Department, Working Paper 94-3, 1994.
- Mullineaux, Donald J. "Competitive Monies and the Suffolk Bank System: A Contractual Perspective," *Southern Economic Journal* (April 1987), pp. 884-98.
- Munn, Charles W. "The Origins of the Scottish Note Exchange," *Three Banks Review* (1975), pp. 45-60.
- Rajan, Raghuram G. "The Past and Future of Commercial Banking Viewed Through an Incomplete Contract Lens," *Journal of Money, Credit, and Banking*, forthcoming.
- Rolnick, Arthur J., Bruce D. Smith, and Warren E. Weber. "The Private Provision of Payments Services: The Suffolk Banking System Reexamined," unpublished manuscript, Federal Reserve Bank of Minneapolis, 1997.
- _____, _____, and _____. "Lessons from a Laissez-Faire Payments System: The Suffolk Banking, 1825-1858," this issue of the *Review*.
- Selgin, George, and Lawrence H. White. "How Would the Invisible Hand Handle Money?" *Journal of Economic Literature* (December 1994), pp. 1718-49.
- White, Lawrence H. *Free Banking in Britain: Theory, Experience, and Debate, 1800-1845*. Cambridge University Press, 1984.
- Williams, Jeffrey. *The Economic Function of Futures Markets*, Cambridge University Press, 1986.

R. Alton Gilbert is a vice president and banking advisor at the Federal Reserve Bank of St. Louis. Robert Webb provided research assistance. The author thanks William Emmons, Ronnie Phillips, Ed Stevens, and David Wheelock for their comments on earlier drafts of this paper. Errors or omissions remain the author's responsibility.



Did the Fed's Founding Improve the Efficiency of the U.S. Payments System?

R. Alton Gilbert

The payment functions of central banks vary among nations. Some central banks provide only limited payment services, such as issuing and redeeming currency and facilitating settlement among members of payments systems by transferring reserve balances.¹ In the United States, the Federal Reserve has been a provider of payment services since the early 1900s. By 1918, the Fed had acquired a large share of the nation's check-collection activities, especially in clearing interregional checks, and the Fed had begun processing wire transfers of reserves among banks.

The Federal Reserve recently has been re-examining the appropriateness of its role in processing checks and automated clearinghouse payments (Rivlin 1997). To provide a more complete background for these deliberations, this article examines the validity of arguments for the Fed's initial entry into its payment activities. Before the Fed was established, critics of the operation of the payments system argued that interregional check collection (collecting and paying banks located in different communities) was inefficient; they maintained that indirectly routing checks to avoid exchange charges by paying banks lengthened the collection process and resulted in higher operating expenses

for banks than more direct routing from collecting banks to paying banks.

An evaluation of whether Reserve Bank services made the payments system more efficient rests on the nature of the payments system prior to the Fed's formation. Therefore, the first section of this article describes the payments system before Congress established the Fed, and the second section develops a theoretical framework for examining the effects of innovations on payments system efficiency. Subsequent sections describe the legal foundation of the Fed's collection services, trace the development of Reserve Bank services, and examine the evidence for and against the argument for improved payments system efficiency.

PAYMENTS SYSTEM OPERATION PRIOR TO 1914

This section focuses on payment instruments and methods of collection in the United States from the Civil War until the formation of the Fed in 1914 (the National Banking Era).

Payment Instruments

Prior to the Civil War, the most important means of payment was currency. The dollar value of currency in the hands of the public exceeded the value of deposits, and the dollar value of currency payments exceeded the value of payments by check. After the 1850s, in contrast, checks became more important than currency. The dollar value of deposits exceeded that of currency, and the value of transactions settled with checks exceeded that settled with currency (Spahr 1926, pp. 84-98).

By the 1870s, currency consisted of coins minted by the federal government, U.S. notes first issued during the Civil War, and national banknotes. Federal legislation in the 1860s taxed out of existence the notes of state-chartered banks. Customers of national and state-chartered banks,

¹ For instance, the Bank of Canada has this kind of limited role in the Canadian payments system, which is operated by the Canadian Payments Association. Bank for International Settlements (1989) describes the roles of central banks in the payments systems of several nations.

however, settled a large share of their payment obligations with checks and drafts drawn upon the deposit liabilities of banks.

To understand the operation of the U.S. payments system during the National Banking Era, it is necessary to distinguish between checks and drafts. Bank depositors created checks payable to those with whom they wished to settle obligations. The bank that accepted a check for collection would then seek payment from the bank on which the check was drawn. Drafts, in contrast, were written by banks.² A bank might draw a draft upon itself or upon an account that it maintained at another bank. When making a payment in a distant city, a bank customer often purchased from his local bank a draft drawn on a bank in a major financial center. The bank on which the draft was drawn would be better known to the payee than the bank that had created the draft. In addition, the costs of collection borne by the payee would be smaller for a draft drawn upon a bank in a financial center than for a check drawn upon the deposit account of the payor.

Selling drafts was a source of revenue for banks, since customers paid banks more than the face amount of the drafts. The business of selling drafts involved the expense of maintaining balances with the banks in major financial centers. When drafts cleared, generally through clearinghouses in the financial centers, the banks on which they were drawn would debit the accounts of the banks that had sold the drafts.

Banks often used transactions in local markets for domestic exchange to replenish their balances with banks in the financial centers. Banks located in various communities established these markets for trading their coin and currency with other local banks that had balances due from banks located in financial centers. The rates of exchange in these markets fluctuated over time. Sometimes there were discounts on coin and currency and sometimes there were premiums. The limits on these exchange rates were determined by the cost of shipping coin and currency among cities, a service provided by express companies.³ These domestic exchange

rates for various cities were published in the local newspapers.

While the available data on payments do not distinguish between checks and drafts, descriptions of banking practices indicate that, early in the National Banking Era, individuals and businesses generally made payments outside their communities with drafts purchased from their banks that were drawn on banks in major financial centers. Over time, it became more common for individuals and businesses to settle their interregional obligations with checks drawn upon their own accounts at their local banks.

Preston (1920, p. 566) and Jones (1931, pp. 172-73) date the use of checks for interregional payments to around 1890. One indication of when this change occurred is the timing of actions by banks in major financial centers for collecting checks drawn on banks outside of these financial centers. Spahr (1926, pp. 119-30) lists a series of proposals and actions by banks to collect out-of-town checks, beginning in 1885. In 1899, the clearinghouses of New York and Boston implemented plans for collecting out-of-town checks. The New York plan attempted to eliminate altogether the use of out-of-town checks for settlement of payments in New York City. The Boston plan, which was more successful, attempted to impose collection at par (face amount) for all checks drawn upon banks located throughout New England. From the timing of these actions, we can infer that a major shift in interregional payments—from drafts to checks—occurred around the end of the last century.⁴

Methods of Collecting Checks

As the number of check transactions during the National Banking Era grew, a system for clearing these checks among thousands of banks had to be developed.⁵ The method of collecting checks depended on the distance between the collecting and paying banks. Banks collected checks drawn upon banks in their communities through local clearinghouses or by presenting the checks at the place of

² Some drafts were created by individuals or businesses, drawn upon the bank accounts of other individuals or businesses, and deposited with the creators' banks for collection. Such drafts were not common. Banks would process them only if the payor and payee were well known to the paying and collecting banks and the banks were familiar with the business arrangement between payor and payee.

³ See Garbade and Silber (1979). Howard and Johnson (1910, p. 117) lists the cost of shipping currency between major financial centers as follows: between New York and Chicago, 50 cents per \$1,000; between St. Louis and New York, 60 cents; between New Orleans and New York, 75 cents; and between San Francisco and New York, \$1.50.

⁴ Phillips (1997) argues that the Fed's services made the payments system less efficient by encouraging bank customers to maximize float by using checks rather than drafts. A problem with Phillips' analysis is that the use of checks for interregional payments predated the formation of the Fed, although Fed services may have stimulated an even greater use of checks for interregional payments.

⁵ For more details on the history of check-collection practices and the effects of Reserve Bank payment services, see Jessup (1967), Magee (1923), and Preston (1920).

DID NONPAR BANKS EARN ECONOMIC RENTS?

Evidence suggests that banks earned economic rents from exchange charges. Nonpar banks tended to be outside of urban areas, and when their depositors sent checks to payees outside their communities, getting these checks back to their banks was slow or expensive. Jessup (1967) indicates that banks which continued to impose exchange charges several decades after the Fed launched its collection system were primarily small banks in isolated communities of states that restricted branch banking. Also, banks tended to eliminate exchange charges when par banks opened offices in

their communities. The banks that resisted the Fed's plan were primarily relatively small banks in areas remote from financial centers. One interpretation of their resistance to the Fed's plan might be that the Fed attempted to deprive banks of revenue necessary for covering their cost of making payments to the Fed. In 1918, however, the Fed offered to pay the transaction expenses for nonmember banks, if such banks agreed to pay the Fed at par. This offer did not eliminate nonpar banks' resistance to Fed efforts to establish a national system for collection at par.

business of the paying banks. Typically, banks collected these local checks quickly and at par.

Collecting checks involved more time and expense when the paying bank was located in a different community. Moreover, the nature of banking law contributed to the time and expense of interregional check collection. While banking law required that banks pay at par for checks presented at their place of business, banks could pay less than par for checks presented to them by mail or other indirect means. The rationale for this deduction from the face amount, called an exchange charge, was that paying banks could incur certain expenses in remitting payment to out-of-town collecting banks, including the cost of transporting coin or banknotes to the collecting banks. But when the staff of collecting banks or their agents personally presented the checks to the paying banks, they assumed the expense of transmitting the cash to the collecting banks.

Delays created another expense for collecting banks. Under banking law, a paying bank that received checks through the mail became the collecting agent for the bank that had sent the checks. The paying bank was therefore responsible for obtaining payment from itself. As a result, paying banks often remitted funds to collecting banks several days after receiving checks through the mail.

It was considered negligence for a bank to collect checks by sending them to paying banks through the mail (Spahr 1926, p. 104). Depositors could argue legally that a bank which mailed checks to out-of-town paying banks should absorb any exchange charges and credit the accounts of depositors at par because the bank had been negligent in its collection practices.

Collecting banks attempted to minimize delays, exchange charges, and claims of negligence by using correspondent banks to collect checks drawn upon banks located outside their communities. These correspondent banks competed for check-collection business, and in attempting to give collecting banks the best terms (quickest collection at the lowest exchange charges), they developed methods to limit the exchange charges imposed by paying banks. The correspondents developed networks of banks that acted as their agents in presenting checks over the counter to banks that set relatively high exchange charges. In collecting through correspondents and their agents, depository banks might receive less than the face amount of checks, but more than if the checks had been sent to the paying banks through the mail. Also, depositors in the collecting banks would not have legal grounds for charging the banks with negligence in their collection practices.

The indirect routing of checks to paying banks through the agents of corre-

THE ROLE OF EXPRESS COMPANIES IN CHECK COLLECTION

By using the services of express companies, banks could have collected checks at par without the indirect routing that was so common before the Fed introduced its check-collection services. Employees of the express companies could take checks to the paying banks, demand payment at par, and return to the collecting banks with the funds. But histories of check collection prior to the formation of the Fed do not mention such arrangements. An extensive discussion of the use of express companies involves the Reserve Banks' collection of checks drawn upon nonpar banks. Prior to a key ruling by the U.S. Supreme Court in 1923, the Reserve Banks accepted for collection checks drawn upon all banks, including those that refused to pay the Fed at par. One of the Fed's objectives in accepting checks drawn upon these nonpar banks was to make its collection service as useful as possible for the banks that chose to collect checks through the Fed. In some cases,

express agents were the only means available to the Fed of obtaining par collection from nonpar banks. Descriptions in Harding (1920) indicate that the Reserve Banks generally accumulated at least \$100 in checks drawn upon a nonpar bank before delivering them to an express company for collection. The Fed's justification for accumulating checks drawn upon nonpar banks was to limit collection costs; a statement by Harding indicates that express companies had a minimum charge of 10 or 15 cents per item for checks in denominations as low as \$5. Charges per item were smaller when the Fed had \$100 or more to collect from a paying bank. These minimum charges may indicate why private banks did not often use express companies for collecting of out-of-town checks. The Fed's collection practices reflected its objective of establishing collection at par as the national standard, not the objective of profit maximization.

spondents deprived paying banks of exchange charges because agents presented checks in person to the paying banks and demanded payment at par. The paying banks still benefited, however, from any delays in the collection process that may have resulted when checks were routed through correspondents and their agents.

The process of collecting checks through correspondents as a means of avoiding exchange charges led to some notorious cases of checks passing through the offices of many banks and traveling over very long distances, relative to the actual distances between the depository banks and the paying banks. More direct channels would have enabled more rapid and less expensive collection.

Correspondent banks attempted to bring order and efficiency to the collection process. Some of them made arrangements with specific banks to collect at par checks drawn upon those banks. The correspondents negotiated various arrange-

ments with the paying banks on these par lists. In some cases, the correspondents were paid at par for checks drawn upon accounts of depositors in the paying banks. In addition, the banks on these par lists often served as agents for their correspondents in obtaining collection at par from *other* banks in their communities (Vest 1940, p. 90). Other correspondents, in contrast, offered to pay exchange charges on checks that they sent to banks for collection, and to credit the accounts of these banks at par for checks received from them. The respondents paid for this service by maintaining balances at the correspondents (Tippett 1929, pp. 258-59). Some banks maintained balances with several correspondents in major cities so they would receive exchange charges on almost all of the checks presented to them by out-of-town banks (Willis 1951, pp. 7-9).

The savings in operating expenses and interbank balances that resulted from using correspondents' collection services

were limited by the large number of banks and the complexity of correspondent banking relationships. A depository bank or its correspondent had to maintain lists of paying banks for which correspondents provided par collection, and they had to route checks to the appropriate correspondents. Unless the correspondents of the depository and paying banks maintained balances with each other, a check would pass through other intermediaries with which these correspondents maintained accounts. A check might pass through several banks in the collection process. Thus, the arrangements for collecting interregional checks through correspondents encouraged indirect routing of checks and forced a complex matrix of interbank balances to facilitate the collection system.⁶

Banks in some cities attempted to cooperate in coping with the challenge of collecting out-of-town checks. For instance, the Boston clearinghouse established a plan in 1899 for par collection of checks drawn upon all banks located in New England (Hallock 1903). The plan, modeled after the earlier Suffolk system for the circulation of banknotes at par (Spahr 1926, p. 127), was largely successful.⁷ While some rural banks in New England did not join the par collection system, approximately 97 percent of checks in New England were collected at par (Spahr 1926, pp. 126-29).

Members of the New York clearinghouse, in contrast, agreed to impose high fees on customers who deposited checks drawn on out-of-town banks. Their objective was to eliminate the use of out-of-town checks for payments in New York City. The banks agreed to charge a minimum fee of 10 cents per check, with higher fees on relatively large out-of-town checks. The penalty for a bank that cheated on this agreement was \$5,000 for a first offense and expulsion from the clearinghouse for the second offense (Spahr 1926, pp. 125-26). New York firms, however, continued to accept out-of-town checks in payment, collecting the checks through correspondent banks located in other financial centers, often with longer collection periods and higher expenses than would

have been possible if the New York banks had collected the checks.

The inefficiencies of the payments system that resulted from this collusion among New York City banks might have been remedied through antitrust enforcement, rather than through government involvement in check collection. But when the Fed was founded, such collusive agreements among banks were common, and they were not subjected to antitrust enforcement. For instance, members of clearinghouses often agreed on the maximum interest rates they would pay on deposits (Cannon 1910, pp. 11-23).

The check-collection process created settlement obligations among banks in different communities. If the banks did not hold balances with each other, they generally settled among themselves with drafts drawn upon banks in major financial centers. The most important center was New York City; drafts drawn upon New York banks served as the national currency for interbank settlement.⁸ This method of settlement relied on the operation of markets for domestic exchange, through which banks in the same community traded coin and currency for balances due from banks in New York City (Hallock 1903).

PAYMENTS SYSTEM EFFICIENCY: A THEORETICAL FRAMEWORK

To examine the implications of Reserve Bank services for payments system efficiency, one needs to develop a framework for defining efficiency. The framework in this article is based on that developed by Berger, Hancock, and Marquardt (1996), hereafter BHM. Their analytical framework, derived from welfare theory, emphasizes a trade-off between risks and costs in the payments system, and the effects of innovations on this trade-off. This section describes the framework and uses it to examine issues in payments system efficiency around the time the Fed was founded.

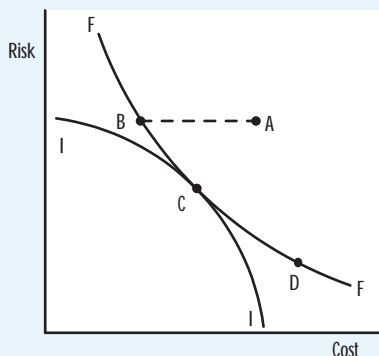
⁶ Weinberg (1997, p. 39) argues that the circuitous routing of checks prior to the formation of the Fed does not necessarily indicate that the check-collection networks operated by correspondents were inefficient. While these networks may have been efficient under the existing constraints on bank behavior and legal relationships between collecting and paying banks, such an argument does not prove that the collection system could not be made more efficient through regulatory changes (such as permission for nationwide branch banking) or government provision of clearing services.

⁷ For analysis of the Suffolk system, see Calomiris and Kahn (1996) or Rolnick, Smith, and Weber (in this issue).

⁸ Citing a survey of national banks in 1890, the Comptroller of the Currency (1890, p. 16) reported that 63.1 percent of bank drafts (dollar value) were drawn upon banks in New York City, compared with 9.4 percent on banks in Chicago and 1.6 percent on banks in St. Louis.

Figure 1

The Risk-Cost Frontier for the Payments System



Basic Features of the Framework

Figure 1 illustrates the theoretical framework. The axes measure the risk and costs borne by payors, payees, and related parties. The risk borne by any one party reflects the distribution of that party's payments-activity losses; the vertical axis in Figure 1 measures some combination of these risk measures for all the parties bearing such risks. Similarly, the horizontal axis measures the weighted sum of costs borne by all the parties.

The curve FF represents the frontier of minimum risk in operating the payments system for a given cost and the minimum cost for a given risk. The position of the curve depends upon the technology used to process and settle payments, financial techniques for monitoring and controlling risk, and the regulatory environment, all of which may be altered by innovations. Thus, BHM classifies innovations in the operation of the payments system as technological, financial, or regulatory innovations. The convex shape of the frontier reflects the usual assumption of diminishing marginal returns—as risks get lower (moving toward the horizontal axis), the marginal costs of further risk reductions increase. All of the points in the area above and to the right of the efficient frontier, FF, as well as those on the frontier, are feasible outcomes. Points off the frontier, however,

reflect inefficient choices from a social viewpoint, in that both risk and costs could be reduced. For instance, the risk and cost combination A is inefficient; the cost of operating at the same level of risk on the frontier (point B) would be lower.

Curve II reflects the social utility function. Since greater risk and greater cost reduce social welfare, shifting curve II to the left increases social welfare (lower cost for given risk). Given the frontier curve FF, the point of highest social welfare is represented by point C, where the II and FF curves are tangent. Innovations improve payments system efficiency if they shift society to a higher indifference curve.

BHM develops this framework with rather general concepts of risk and cost. This section describes the U.S. payments system prior to 1915 using the BHM framework.

The Nature of Risk and Cost in Payments System Operation

Point D on the FF frontier reflects the minimum cost associated with an arrangement in which all payments are made in cash. Risk would be low, but cost would be high, especially for interregional payments. While point D is on the FF frontier, it does not represent the point of maximum social welfare because of the high cost of minimizing risk with cash payments. Other payment arrangements, based on the use of checks and drafts, involved more risk: the possibility of losses from bad checks, from account closures for insufficient funds, and from bank failures. The observation that not all payments were settled in coin and currency prior to 1915 indicates that, to economize on the cost of settling payment obligations, individuals and firms were willing to accept more than the minimal level of risk illustrated by point D.

The Comptroller of the Currency (1890) and Garbade and Silber (1979) mention that innovations in communications and transportation over time reduced the cost of transporting coin and currency over long distances. In the framework of BHM, these technological

innovations cause the FF curve to shift to the left over time (greater efficiency of the payments system) by reducing the cost associated with minimum risk (all payments in cash). In addition, the Comptroller of the Currency (1890) indicated that the fees banks charged for drafts drawn upon banks in financial centers declined sharply over time. In terms of this theoretical framework, technological innovations tended to enhance social welfare by reducing the cost of interregional payments.

The relevant operating costs of the payments system in this analysis are as follows:

1. The cost of transporting coin and currency (one option for settling obligations).
2. The cost of collecting checks, including the cost (borne primarily by correspondent banks) of routing out-of-town checks to paying banks.
3. The cost of operating the markets for domestic exchange that existed prior to the formation of the Fed.
4. The opportunity cost of holding cash balances.

Exchange Charges and Social Welfare

Given the prominence of exchange charges in Fed policies for payments system operation, their role in the framework of Figure 1 deserves special attention. Since exchange charges were a cost to some banks and bank customers, should we include them in the measure of cost in Figure 1? If so, eliminating exchange charges would tend to shift the FF curve to the left, increasing social welfare. In contrast, banks that imposed exchange charges viewed Reserve Bank activities to promote check clearance at par as a threat to their welfare. How should our measure of social welfare reflect their view of what was in their interest?

BHM (p. 701) mentions that some participants might prefer an inefficient arrangement for the payments system (a combination of risk and cost off the FF

curve) if that arrangement would impose the extra costs or risk on other participants. Thus, changes in payment arrangements that enhance efficiency are not necessarily in the interest of each participant in the payments system.

Many banks located outside urban areas resisted the Fed's plan for national par clearance of checks, since they wanted to maintain their revenue from exchange charges.⁹ Their resistance, however, does not necessarily indicate that the Fed's system reduced social welfare.

On the other hand, eliminating exchange charges would not necessarily increase welfare. Welfare theory focuses on the implications of changes in market practices for the pricing of goods and for the quantity of market output, not on the effects of such changes on the distribution of profits among firms that sell the goods (Scherer 1970, pp. 8-38). Exchange charges affected the allocation of profits among banks, in that they were revenue to some banks and expenses to others. Exchange charges would reduce welfare if the efforts by participants to avoid such charges increased the cost of payments system operation (for a given level of risk) above the level of cost that would have been possible without the charges.

Effects of Reserve Bank Payment Services and the Discount Window on System Efficiency

The inherent risks in the payments system before the establishment of the Fed included disruptions created by banking panics. For example, when banks in major financial centers suspended currency payments in response to panics, they restricted depositors' currency withdrawals. Bank failures resulting from such panics also disrupted the system. A major purpose for establishing the Fed was to eliminate panics by providing banks with reserves through the discount window during liquidity crises (Dwyer and Gilbert 1989).

The Fed could have reduced risk in the payments system through its discount

⁹ Reserve Bank services reduced substantially the role of drafts drawn upon New York City banks in the settlement of interregional obligations. After the Fed established its collection system, however, the non-par banks located outside urban areas, not the New York City banks, opposed the Fed's collection system.

window operations without providing payment services. The purpose of this article is to assess the implications of Reserve Bank payment services in terms of risk vs. cost, independent of the effects of discount window operations.

The challenge of separating these effects of Fed policies is illustrated in Figure 1. Let us suppose that point A in Figure 1 represents the risk-cost combination for the payments system under the banking laws and regulations of 1913, with the addition of a central bank-operated discount window but no central bank payment services. This central bank would provide currency to banks that borrowed at the discount window and credit the banks' reserve accounts when they deposited currency, but it would not provide wire transfer services, interbank settlement through debits or credits to reserve accounts, or collection services. Now let us suppose that point C on the FF curve represents the risk-cost combination with all of the above, plus Federal Reserve Bank payment services. What kind of influences would have kept the private sector from moving to the efficiency frontier on its own? What was special about the Fed that enabled it to improve the efficiency of the payments system?

The Reserve Banks did not invent a new technology for check collection. Development of their check collection service was a regulatory innovation, reflecting two aspects of the legal and regulatory environment that existed prior to the formation of the Fed.

One legal barrier to efficient check collection may have been the principal-agent relationship created when banks mailed checks to other banks for collection; that is, the paying bank became the agent of the collecting bank, responsible for collecting from itself. This legal relationship created a disincentive for banks to collect interregional checks by the most direct method, the mail service. Instead, they sent checks to correspondents, who did the collection for them.

A second legal barrier to efficient operation of the system may have been restrictions on nationwide branch banking. If the

major correspondent banks had offices located throughout the nation, they could have routed checks directly to their own offices that were nearest to the offices of the paying banks and had their employees present the checks for par collection. Instead, correspondents had to develop ad hoc arrangements for getting checks to paying banks through networks of correspondents and respondents. The Reserve Banks, with their nationwide network of offices and legal authority to demand remittance at par from member banks, may have reduced the cost of interregional check collection by reducing operating expenses and shortening collection times. In addition, shorter collection times may have reduced risk by reducing the possibility that checks would be dishonored or that the paying bank would fail prior to settlement. A later section examines the evidence for and against the hypothesis that the Federal Reserve's collection services improved payments system efficiency.

Relevant Comparisons in Assessing Fed Services Efficiency

In evaluating the evidence, it is important to make appropriate comparisons. It is possible that payments system efficiency would have been more improved if Congress had permitted nationwide branch banking instead of authorizing the Reserve Banks to offer payment services. In Figure 1, if FF reflects the efficient frontier for the payments system given Fed payment services and restrictions on branching as of 1913, nationwide branch banking (with or without Fed services) might have shifted the frontier farther to the left. The evidence in this article, however, does not allow us to explore that hypothesis.

It is also possible that the Reserve Banks' entry into the payments business precluded future innovations by the private sector that would have shifted the FF curve farther to the left in the absence of Reserve Bank services. While the Fed has encouraged innovation in the operation of the payments system in recent decades (Summers and Gilbert 1996,

pp. 17-18), the evidence cited in this article is not relevant for judging whether Reserve Bank participation in the payments system has, on net, retarded or enhanced payments system efficiency over the many years since the Fed's founding.

Finally, in assessing the limitations of the data, one should note a critique of Reserve Bank payment services by Baxter (1983). He developed an economic theory of relationships among payors, payees, and their banks (assuming they use different banks for payment services). In modeling the demand for payment services, Baxter notes that a transaction involves a joint *demand* by payor and payee for method of payment and, on the *supply* side, the cooperation of two banks. Baxter notes that the arrangement that maximizes welfare is likely to involve a side payment, or "interchange fee," between the two banks. He also argues that in a payments system with many payors, payees, and banks, negotiations among these parties over the allocation of costs and fees for each transaction would be inefficient. Thus, he argues that an efficient payments system will have some standard practice for interchange fees among banks.

Baxter describes arrangements in credit card associations as reflecting efficient pricing in a segment of the payments system. In contrast, Baxter (1983, p. 571) criticizes the Federal Reserve for imposing par clearance in the check collection system:

Thus the role of the exchange fee in the process of check clearance, a commercial context in which an unregulated market solution might have been expected to work reasonably well and to yield instructive results, was aborted and continues to be suppressed by a mixture of subsidies and coercion by the Federal Reserve System.

Baxter does not attempt to prove that check collection at par was inefficient. Rather, he suggested that interchange fees in check collection might have been necessary for maximum efficiency in payments

system operations. Baxter notes, however, that prior to the formation of the Fed, members of clearinghouses cleared checks among themselves at par. The clearinghouses would have been free to set interchange fees among their members if they felt that such fees would make the payments system more efficient.

There is an important reason why clearinghouses cleared checks at par: The common law requirement that a bank pay at par when checks drawn upon its depositors' accounts were presented at its place of business. This legal standard for par collection limited pricing options for clearinghouses. Since clearinghouses are cooperatives set up to avoid the cost of bilateral exchange, they could not function effectively if members could avoid paying interchange fees by presenting checks directly to each other. Credit card associations are effective in imposing interchange fees because their members do not have the option of collecting credit card receivables from each other directly at par. The courts limit the rights of banks with credit card receivables to the rights specified by their card associations.

Baxter's framework suggests that changing banking law to eliminate the requirement for banks to remit at par for checks presented at their place of business might have led to greater payments system efficiency than did the Reserve Banks' entry into the payments system. Eliminating the requirement for payment at par might have facilitated private-sector development of check-clearing organizations similar to today's national credit card associations. These check-clearing organizations would have set the interchange fee for maximum payments system efficiency, thus shifting the FF curve to the left.

In considering the effects of Reserve Bank operations on payments system efficiency, we should note that the Fed did not have the prerogative to eliminate the requirement for paying banks to remit at par for checks presented at their place of business. While Baxter's theoretical framework is interesting, his critique of the Fed is not relevant for judging whether Reserve

Bank operations improved the efficiency of the payments system, given the legal and regulatory environment in place when the Fed was founded.

LEGAL FOUNDATION FOR RESERVE BANK COLLECTION SERVICES

The history of the Fed's role in the payments system has been shaped largely by acts of Congress and litigation. In tracing this history, it is necessary to examine the legal foundation of the Fed's role in the payments system in some detail. The Federal Reserve Act (FRA) as enacted in 1913 did not state clearly the Congressional intent for the Fed's role in the payments system. Section 13 authorizes the Reserve Banks to receive checks from any member banks drawn upon other member banks. Section 16 states that Reserve Banks shall receive checks from member banks at par, and it authorizes the Board of Governors to establish a clearinghouse for clearing checks and drafts among the Reserve Banks. Did these sections simply direct the Fed to provide payment services to member banks, or did they provide the Fed with a mandate to make the collection of checks at par the national standard for the banking industry?

These sections of the FRA have been subject to various interpretations. From its earliest beginnings, the Board of Governors interpreted them as giving the Fed a mandate to establish a national system for par clearance of checks. In a recent review of the legislative history of the FRA, however, Stevens (1996) concludes that the Fed's founders did not see a need for a government service to deal with inefficiencies in the nation's check-collection system. Rather, Stevens argues that the founders included these sections in the FRA on check collection to make reserve balances useful for member banks. In Stevens' argument, the founders were concerned that banks would resent the opportunity cost of holding idle balances at the Reserve Banks, and their resentment might undermine the Fed's effectiveness in providing an elastic

currency and acting as lender of last resort in financial crises. Stevens argues that the collection system was to be the glue that tied banks to the Fed. In drawing this conclusion, he emphasizes some statements by the founders about the need to make reserve balances useful for members. He also cites the fact that sections dealing with check collection were added to legislation for a central bank late in the legislative process.¹⁰

An amendment to the FRA in June 1917 helps clarify Congressional intent involving the Reserve Banks' check-collection services. One provision allowed nonmember banks to become clearing members of the Reserve Banks. These clearing members could present checks at their Reserve Banks for collection if they held clearing balances. Proceeds from collecting the checks would be credited to the clearing account, and the value of checks that were received by the Reserve Bank and drawn upon the clearing member would be charged to its clearing account.

A second provision of the amendment as originally proposed by Senator Hardwick would limit bank exchange charges to no more than one-tenth of 1 percent (10 cents per \$100) of the face value of a check. The proposed legislation would have permitted member banks to impose exchange charges, within the specified limit, on the Reserve Banks. While the bill was in conference (versions having passed the House and Senate), President Wilson intervened in a letter to Senator Hardwick:

I should regard such a provision [Reserve Banks absorbing exchange charges in their check collection activities] as most unfortunate and as almost destructive of the function of the Federal reserve banks as a clearing house for member banks, a function which they have performed with so much benefit to the business of the country (Vest 1940, p. 91).

Because of Wilson's intervention, the conferees modified the section on exchange charges by adding that "no such charges

¹⁰Harding (1925) presented an alternative view for why the Reserve Banks were involved in check collection. Member banks would need methods of increasing and decreasing their actual reserve balances in matching their reserves and required reserves. The Fed's check-collection services offered a method of generating debits and credits to reserve balances. This purpose is more mundane than that asserted by Stevens and less ambitious than establishing par clearing as the national standard.

shall be made against the Federal Reserve banks” (Vest 1940, p. 91). This amendment to the FRA indicates that Congress viewed the Fed’s check-collection system as more than just a means of giving member banks some value for their required reserves. If Congress had included provisions for a check-collection system in the FRA in 1913 just to make membership in the Fed attractive, why would Congress grant nonmember banks access to the clearing system in 1917?

Exchange charges against the Reserve Banks would not have undermined the purpose of giving member banks some value for their reserve balances, since industry practice involved paying exchange charges.¹¹ The Fed’s collection system could have continued to serve as the glue binding member banks to the Fed. President Wilson’s intervention, however, indicates that he considered the Fed’s par collection system a valuable service to the nation. In modifying the bill to prohibit exchange charges against the Reserve Banks, Congress assented with this view.¹²

The other major legal developments that shaped the nature of the Fed’s collection services involved litigation challenging the Reserve Banks’ efforts to establish par collection as standard practice throughout the nation. In the early years of the Fed’s collection system, the Reserve Banks accepted checks drawn upon all banks, including those that had not agreed to pay the Fed at par. The Reserve Banks used a variety of methods to collect at par from these nonpar banks, including hiring express agents to travel to the offices of the nonpar banks, present checks over the counter, and return with the funds. Some of the nonpar banks interpreted the Fed’s collection practices as attempts to harass them into agreeing to pay the Fed at par.¹³

Nonpar banks challenged the Fed’s collection practices in the courts, and a decision by the Supreme Court of the United States was announced in 1923. The Court ruled that since Congress did not require the Fed to establish a national system of par collection for checks, the Reserve Banks could not compel nonmember banks to pay them at par. In response

to this Court ruling, the Reserve Banks restricted the checks they would accept for collection to those drawn on the banks that agreed to pay the Fed at par. Banks had to use other channels for collecting checks drawn on nonpar banks.

CHRONOLOGY OF RESERVE BANK PAYMENT SERVICES

When the Reserve Banks began providing check-clearing services in 1915, the Board of Governors initially pursued what it called a voluntary collection system. The Reserve Banks would receive for collection only those checks drawn upon banks that had volunteered to join the Fed’s collection system. The banks had agreed to pay at par for checks presented by the Fed for collection, even if the Fed sent the checks to the paying banks through the postal service. However, only about one-fourth of the member banks joined the collection system. In its annual report for 1916, the Board expressed regret that the voluntary system had not been more successful and concluded that the voluntary plan would never achieve its objective of a universal par collection system for the U.S. economy. To promote this goal, the Board decided in April 1916 to change its collection plan from voluntary to compulsory for member banks. Under this new plan, the Fed required each member to remit at par for checks the Reserve Banks presented for collection, including checks sent through the mail. Member banks were not, however, required to send checks to the Reserve Banks for collection.

Introduction of Collection Fees

The Board also adopted a policy of charging banks for the collection service: Each Reserve Bank charged the depositing bank a fee per check that reflected its expenses. Initially, the fees ranged from 0.9 cents to 2 cents per check. At this time, the Fed also stopped the practice of immediately adding the value of checks to the reserve accounts of collecting banks; instead it deferred credit according to a schedule based upon estimates of the time

¹¹ Willis (1923, pp. 1062-63) asserted that passage of legislation authorizing member banks to impose exchange charges on the Reserve Banks would have destroyed the Federal Reserve check-clearing system. Willis did not, however, provide the basis for that assertion.

¹² The Board of Governors was uncertain whether this legislation granted the Fed authority to set limits on the exchange charges of all banks or only banks that were Fed members. The Board requested the opinion of the U.S. Attorney General, who held that the Board’s authority over exchange charges under this act applied only to Fed members. Because of this ruling, the Board did not use its authority over exchange charges in its attempt to make collection of checks at par the national standard for the banking industry. See Spahr (1926, p. 200).

¹³ See Harding (1920) for the Federal Reserve Board’s response to accusations by nonpar banks about the Fed’s check-collection practices.

it took the Reserve Banks to present the checks to the paying banks (Willis 1923, p. 1060). Under the prior plan, a collecting bank's reserve account was credited and the paying bank's reserve account was debited for the amount of the check when the Fed received the check, before the paying bank had a chance to see it or learn of the debit to its reserve account. Reserve Bank annual reports for 1916 indicated that members of the voluntary collection system had objected to this timing of debits to their reserve accounts.

Also during 1917 the Fed began allowing banks to use transfer drafts for settling their payment obligations with other banks. Banks with accounts at the Reserve Banks could create transfer drafts drawn upon their accounts, which were then payable immediately at any Fed office. Member banks could therefore use transfer drafts for interbank settlement instead of drafts drawn upon their accounts at banks in New York City. For example, a bank in St. Louis would write a draft drawn upon its reserve account at the St. Louis Fed and mail it to a bank in Atlanta. The Atlanta bank would get immediate credit to its reserve account when it presented the draft to the Atlanta Fed. Settlement between the two Reserve Banks would occur daily through the gold exchange fund.

Participation by Nonmember Banks

An amendment of the FRA in 1916 permitted the Reserve Banks to collect checks drawn upon nonmember banks. In June 1917, Congress amended the FRA in response to a request by the Board to permit the Reserve Banks to collect checks for nonmember banks that opened clearing accounts at their Reserve Banks. The Fed required these clearing members to pay the Fed at par for checks drawn upon them.

Only a small number of nonmembers joined the Fed's collection system.¹⁴ Tippetts (1924, pp. 632-33) concluded that a major reason why few nonmembers took this option for check collection was the nature of state reserve requirements. Nonmember banks counted balances with

other banks as part of their reserves for meeting state requirements, and banks tended to count all of their funds deposited with correspondents, including uncollected funds, as balances due from banks. The Fed's accounting system, in contrast, separated cash items in the process of collection (CIPC) from balances due from the Reserve Banks. Nonmember banks would thus tend to increase the burden of state reserve requirements by collecting checks through clearing accounts at the Reserve Banks.¹⁵ These banks could benefit from the Fed's check-collection system indirectly by clearing checks through correspondents that used the Fed's collection system.

In 1918 the Fed began operating its leased wire system for reserve transfers among banks, an electronic alternative to transfer drafts. Also, in July 1918, the Board ended the policy of charging fees to banks that deposited checks with the Reserve Banks for collection. The objective for dropping the fees was to promote use of the Fed's collection system. For the nonmember banks that agreed to pay the Fed at par, the Fed began absorbing the expenses they incurred in remitting payment to the Reserve Banks. One objective of this offer was to eliminate the argument that nonmember banks could not remit at par because of the expenses they would incur in remitting payment. In response to a ruling by the U.S. Supreme Court in 1923, the Reserve Banks began refusing to accept for collection checks drawn upon nonpar banks.

Volumes of Reserve Bank Payment Services

The Reserve Banks very quickly became major processors of payments. Table 1 presents the number of checks cleared by the Reserve Banks and the dollar value of these checks relative to the value of checks cleared through the private clearinghouses. In the period around the formation of the Fed, data were available on the dollar value of checks cleared through about 200 clearinghouses in cities around the nation,

¹⁴ Willis and Steiner (1926, p. 607) reported that on July 1, 1925, only 158 nonmember banks in seven Federal Reserve districts maintained such clearing accounts.

¹⁵ Tippetts (1924) poses the hypothesis that an option for banks to meet their reserve requirements with uncollected funds will affect bank behavior. Gilbert (1978) found evidence in the 1970s to support this hypothesis. Prior to 1980, the state-chartered banks that did not choose to be Fed members were exempt from Fed reserve requirements but subject to state requirements. Some states permitted banks to count CIPC as part of their reserves, whereas other states excluded them. Gilbert found that nonmember banks in states that excluded CIPC from reserves reported relatively low levels of CIPC; they appeared to report their uncollected funds as demand balances due from banks, a common practice prior to the Fed's formation.

Table 1

Volume of Checks Processed by the Reserve Banks

Year	Millions of Checks	Dollar Value of Checks (billions of dollars)	Value of Checks Processed by Reserved Banks as Percentage of Checks Cleared Through Clearinghouses
1915	8.8	\$ 4.7	2.9%
1916	25.8	10.9	4.5
1917	75.7	44.9	14.7
1918	154.4	105.7	32.9
1919	305.2	136.5	35.2
1920	452.1	156.5	35.6
1921	522.7	119.2	34.1
1922	584.9	150.5	39.1
1923	639.2	196.6	48.6
1924	684.0	209.1	46.9
1925	716.5	247.2	49.4
1926	758.5	261.4	51.0
1927	794.8	266.7	49.0
1928	818.5	289.0	46.4
1929	852.1	351.7	49.1
1930	834.2	311.2	57.2
1931	796.9	237.8	57.8
1932	677.0	169.2	65.4
1933	635.0	151.2	62.0
1934	754.7	171.9	65.1

SOURCES: Observations on check clearings by the Federal Reserve Banks are from the annual reports of the Board of Governors of the Federal Reserve System. Data on the value of checks cleared through clearinghouses are from U.S. Department of Commerce (1960, p. 640). The annual observations for the number of checks cleared by the Reserve Banks, and the dollar value of the checks, are adjusted to eliminate duplications in reports by the individual Reserve Banks. A duplication occurs if two Reserve Banks report the same item as one of their items handled. The annual reports of the Board of Governors for 1919 through 1926 provide data on items handled with and without duplications. Data for 1915-18 reported with duplications have been adjusted in accordance with data for 1919, and data for the years 1927-34 have been adjusted in light of the data for 1926.

including the major financial centers and many relatively small cities. While this series does not reflect the dollar value of all checks, it provides the broadest available measure of check clearings outside the Federal Reserve System.¹⁶

In 1915, the Reserve Banks processed 8.8 million checks, with a dollar value of \$4.7 billion, which was about 3 percent of the value cleared through the private clearinghouses. The volume of checks processed

by the Fed's clearing system rose rapidly after the Fed adopted its compulsory plan in 1916, rising to about 33 percent of the clearings through the private clearinghouses by 1918. Table 1 indicates that clearings through the Reserve Banks as a percentage of clearings through private clearinghouses continued to rise through 1934.¹⁷

Those familiar with the current infrastructure of the Fed's check-collection system—which includes a ground and air

¹⁶ See Garvy (1959) for a description of the data on check clearings.

¹⁷ The sharp rise in the Fed's market share after 1929 may reflect a flight to safety by respondent banks during the banking panics of the early 1930s. That is, respondent banks considered clearing checks through the Reserve Banks less risky than clearing through private correspondents. Respondent banks also behaved in this way during the 1980s; the Federal Reserve Bank of Dallas had a sharp rise in its check clearings during the Texas banking crisis (Clair, Kolson and Robinson, 1995).

transportation network to clear checks—may be surprised by the rapid development of the Fed's check collection operation. How did a new organization put a national collection infrastructure in place so quickly?

In fact, the Fed did not create a new infrastructure to collect checks. Instead, it used the existing national infrastructure for communication: the postal service. Banks mailed checks to the Reserve Banks for collection, and the Reserve Banks mailed checks to paying banks. As explained above, banking law discouraged collecting banks from sending checks directly to paying banks, but the Fed asserted its legal authority to obtain payment at par from member banks for checks sent to them through the mail (Jones 1931, p. 138). The FRA did not alter private banks' authority to obtain par collection from other banks. In terms of the theoretical framework presented above, the Fed's use of the postal service for check collection at par was a regulatory innovation.

EFFECTS ON PAYMENTS SYSTEM EFFICIENCY

To investigate whether evidence is consistent with the view that the Fed's payment services improved the efficiency of the payments system, this section looks at the effects on transaction costs, the national credit market, bank operating costs, cash assets, and the rapid adoption of Fed payment services by the banking industry.¹⁸

Transaction Costs and the National Credit Market

In the theoretical framework presented above, the Fed's payment services increase payments system efficiency if they shift the FF frontier in such a way that society is on a higher indifference curve, a higher level of social welfare. The most relevant evidence of increased social welfare would be a positive macroeconomic shock that could be attributed to the development of Reserve Bank payment services.

If the Fed's collection system had its intended effects, it would tend to reduce

transaction costs for payments across regions. In turn, reductions in the costs of interregional transactions would tend to facilitate the operation of a national capital market, rather than separate regional capital markets, each with its own balance of supply and demand for capital. Economic historians have examined the process by which the national integration of capital markets occurred by tracing patterns in regional interest rate differentials (James 1978).

An important challenge in investigating the effects of the Fed's collection system is to focus on effects that cannot be attributed to other aspects of Fed operations. To the extent that the Fed's formation promoted the integration of regional capital markets into a national market, the discount window may have been more significant than the Reserve Bank collection system in reducing transaction costs. Miron (1986) attributes a change in the seasonal pattern of interest rates after the formation of the Fed to the operation of the discount window.¹⁹ The rest of this section focuses on evidence that can be tied more directly to the operation of the Fed's collection system.

Operating Costs

In terms of the theoretical framework, evidence of increased efficiency in the operation of the payments system would include reduced operating costs for banks—for instance, smaller ratios of operating expenses to total assets and lower ratios of employees per dollar of assets. Unfortunately, such data are not available for the period around the time when the Federal Reserve was founded.

Cash Assets

One cost component in the payments system was the opportunity cost to banks of holding cash assets. Banks had two major reasons for holding cash in their vaults and balances with other banks: to facilitate payment-order processing and to meet reserve requirements. If the Fed's payment services improved the efficiency of the payments system, we should be able

¹⁸In their assessment of the Fed's role in check collection, Duprey and Nelson (1986) mention that the Fed failed to achieve its goal of universal participation in its national system of par check collection, but they conclude that the Fed did improve the efficiency of intercommunity check collection.

¹⁹The work by Miron (1986) generated controversy. See Canova (1991), Clark (1986), Fische (1991), Fische and Wohar (1990), Holland and Toma (1991), and Mankiw, Miron, and Weil (1987 and 1990). There is general agreement among these authors that the seasonal pattern of interest rates in the United States changed after the formation of the Fed, with stronger consensus that the Fed changed the seasonal pattern of interest rates after 1917.

to find evidence that the services allowed banks to operate with smaller percentages of their assets in cash, independent of the effects of changes in reserve requirements.

Fed payment services may have permitted banks to operate with lower cash ratios for the following reasons. First, cash holdings, which included uncollected funds, would decline to the extent that the Fed's check-clearing operations reduced the number of days required for checks to clear by eliminating indirect routing of checks to paying banks. Second, Fed payment services may have reduced the interbank balances that banks needed to maintain for clearing checks, since they could send checks to the Fed for collection.

An important challenge in assessing the effects of the Fed's payment services on bank cash holdings is to separate their effects on indicators of payments system efficiency from the effects of other changes unrelated to the Fed's payment services. The FRA reduced reserve requirements substantially for national banks, which were required to be Fed members. The Comptroller of the Currency (1915) calculated the reductions in required reserves for national banks as follows: 28 percent for national banks located in central reserve cities, 41 percent for banks in reserve cities, and 26 percent for national banks located elsewhere (see appendix for details). In addition, there was a reduction in required reserve ratios for Fed members in 1917. Because these changes occurred around the time the Fed developed its payment services, this section examines the effects of the Fed's payment services on the cash ratios of state-chartered banks.

About half of the states reduced their reserve requirements around the time the Fed was founded (see appendix). This division creates a type of experiment: If the cash ratios of state-chartered banks declined around the time the Fed developed its payment services, were the declines limited to banks in states that reduced their reserve requirements?

Figures 2 and 3 present the relevant ratios of cash to total assets for the years 1900 through 1930. Figure 2 presents

cash ratios for national and state-chartered banks in states that lowered their reserve requirements in the years 1913 through 1925 (see appendix for the list of states). A wide gap exists between the cash ratios of national and state-chartered banks prior to the Fed's formation. Cash ratios of national and state-chartered banks declined sharply around the time the Fed was founded. Because of the reserve-requirement reductions, it is not possible to attribute the cash-ratio reductions in Figure 2 to the introduction of the Fed's collection system.²⁰

Figure 3 presents the same ratios for national and state-chartered banks in states that did not reduce their reserve requirements in the years 1913 through 1925. The cash ratios of the state-chartered banks declined about 5 to 6 percentage points during the period of rapid growth in the Fed's check-collection system, and these ratios remained at a new lower level throughout the 1920s. This pattern of change in the cash ratios of state banks in Figure 3, which is similar to the pattern in Figure 2, can be attributed to the Fed's collection system, since these states did not reduce their reserve requirements. The impact of the Fed's collection system is estimated at about 5 percent of the total assets of banks, which were liberated for other uses.²¹

General Acceptance of Reserve Bank Payment Services

Banks chose to use the Reserve Bank payment services, even though they were free to continue using the payment arrangements that had been available prior to the formation of the Fed. Acceptance of these services was rapid and on a large scale, indicating that the Reserve Banks met a demand for efficient interregional payment services that had been unmet through private arrangements. Banks would not have chosen to use Reserve Bank payment services if those provided by private parties had been more efficient.

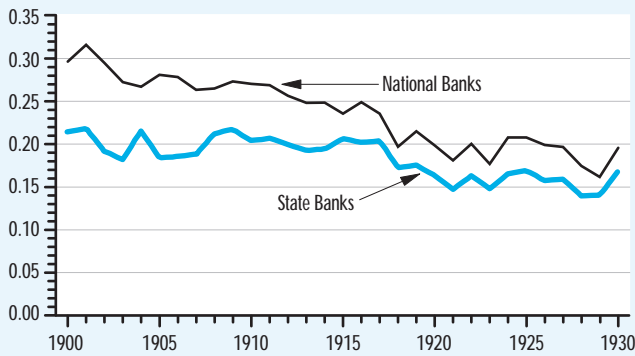
Elimination of the Markets for Domestic Exchange. Changes in the payments system that can be tied directly to Reserve

²⁰ The numerator of the cash ratio includes the value of vault cash, balances with other banks, and CIPC. The denominator is total assets. Since banks held cash primarily to serve their customers who made payments out of demand deposits, a more ideal ratio might be cash to demand deposits. Data on the deposit liabilities of banks for the period covered in Figures 2 and 3 were of relatively poor quality, especially the division of deposits between demand and other deposits. See Board of Governors (1959). For this reason, cash ratios in this article use total assets as the denominator.

²¹ Another factor that might have influenced the cash ratios of banks around the time that the Fed developed its collection service was the financing of World War I. State-chartered banks increased their holdings of U.S. government securities substantially during the war, from less than half of 1 percent in 1916 to about 9 percent in 1918. With more liquid assets, banks may have been comfortable operating with lower cash ratios than before. The ratio of U.S. government securities to assets declined after 1918, however, falling to around 6 percent by 1929. If the securities holdings of banks explained their declines in cash ratios after 1916, we would expect the cash ratio to rise again after 1918, which did not happen.

Figure 2

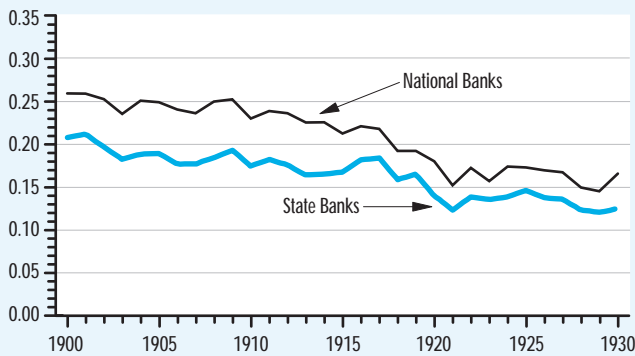
**Ratio of Total Cash Assets to Total Assets:
States That Lowered Reserve Requirements, 1913-25**



SOURCE: Board of Governors of the Federal Reserve System (1991)

Figure 3

**Ratio of Total Cash Assets to Total Assets:
States That Did Not Lower Their Reserve Requirements, 1913-25**



SOURCE: Board of Governors of the Federal Reserve System (1959).

Bank services include the elimination of two forms of economic activity: trading in domestic markets for exchange, and shipment of coin and currency among regions of the nation to arbitrage differences in domestic exchange rates. Under the Federal Reserve System, shipments of gold and currency for interbank settlement were almost wholly eliminated (Burgess 1927, p. 82). Garbade and Silber (1979, p. 15) report that by 1918 (the year the Fed opened its leased wire system for reserve transfers among banks), newspapers reported domestic exchange rates essentially at par, and by 1920 they ceased reporting exchange rates. The timing of

the elimination of the markets for domestic exchange indicates that the most important factor was the opening of the Fed's wire transfer system, although access of member banks to reserves through the discount window may have limited the variation in exchange rates among cities.

Network Effects and the Timing of Growth in Fed Payment Services. The timing of growth in the Fed's collection system is consistent with the view that the Reserve Banks satisfied a demand for more efficient interregional check collection. In 1915, the dollar value of checks cleared by the Fed was only about 3 percent of the dollar value of checks cleared through the private clearinghouses. Under the voluntary system of 1915, member banks did not have to join the Fed's collection system. If they joined, however, they had to agree to pay at par when the Fed presented checks.

The limited response from banks to the Fed's offer of a voluntary collection service reflects the nature of the payments system as a network good. Network effects are exhibited when the demand by an individual customer depends on the number of other customers who use the good.²² While many banks might have viewed a par collection system as valuable in principle, the value of such a system to each bank would depend on the number of other banks that had agreed to join. When only a small percentage of banks agreed to pay the Fed at par, others had limited interest in joining. Why should they forgo some of their revenue from exchange charges if the new collection system could make par presentment to only a small percentage of banks?

The Fed changed its collection system substantially in 1916: All member banks were required to pay at par for checks presented by the Fed, and the Reserve Banks began charging the banks that deposited checks fees to cover their collection costs. The Reserve Banks' check-clearing volume began growing rapidly immediately after these changes were implemented.

The timing of this growth does not support the hypothesis that it occurred because the Reserve Banks had suspended the prac-

²² For analysis of network effects, see Katz and Shapiro (1994) and Economides and White (1994).

tice of charging fees to depositing banks. The Fed's collection system had already grown to a relatively large share of national check collection by the time the Fed stopped charging collection fees in 1918. The Fed's share continued to rise after 1918, but at a slower pace than in the period from 1915 to 1918. The most important action to stimulate growth of its collection system appears to have been the Fed's decision to require all member banks to pay at par for checks presented by the Reserve Banks.

Demand for a Par Collection Service and the Geographic Location of Paying Banks.

Critics of the payments system prior to the formation of the Fed acknowledged that check collection was efficient where collecting and paying banks were located in the same communities, but they wanted a more efficient mechanism for interregional check collection. If these critics were correct, banks would find the Fed's collection system valuable as a means of interregional check collection, not a mechanism for local collection.

In 1918, the Fed began reporting information on the location of the banks at which the Reserve Banks presented checks for collection. Table 2 indicates that only about 10 percent of the checks were presented to banks in the cities where the Fed had offices. Table 2 also indicates that the dollar denomination of checks presented to banks in cities where the Reserve Banks had offices was much larger on average than the size of checks drawn on banks located outside the Reserve Bank cities. These differences in the average dollar size of checks, which were pronounced in each of the 12 Districts, probably indicate that a relatively high percentage of the checks presented to banks where the Reserve Banks had offices involved interbank settlements rather than checks written by bank customers, since interbank reserve transfers tended to be much larger than most checks.²³

CONCLUSIONS

Histories of the Fed's payment services generally focus on the Fed's failure to

achieve its goal of getting all banks to participate in a national system for collecting checks at par. This article uses a different standard of evaluation: effects on the efficiency of the payments system. While the goal of universal par check collection remained out of reach for the Fed in its early years, evidence from the period when the Fed was founded suggests that the Fed's services improved payments system efficiency.

This analysis suggests that Reserve Bank payment services grew in popularity because they permitted banks to operate with lower ratios of cash to total assets. The relevant evidence is for banks chartered in states that did not reduce their reserve requirements around the time when the Fed was founded. Ratios of cash to assets for these banks declined about 5 percentage points during the period of rapid development in the Fed's payment services and then remained at the new lower level. For national banks and state-chartered banks in other states, changes in cash ratios reflect changes in reserve requirements, in addition to the effects of Reserve Bank payment services.

Banks chose to use the Fed's payment services rather than the payment arrangements that were available to them prior to the Fed's formation. By 1920 the Fed's wire transfer service had eliminated the old system of interregional settlement among banks, which had involved the use of drafts on New York City banks and the markets for domestic exchange. The share of checks cleared through the Reserve Banks rose dramatically after the Board acted in 1916 to require member banks to pay the Fed at par, despite the fee per check that each Reserve Bank began charging collecting banks in that same year. The Fed's check-collection activities involved primarily interregional collection of checks, and banks found the Fed's system more attractive than the old system of collecting interregional checks through correspondents. The growth in the Reserve Banks' collections indicates that banks' demands for a national par collection service had been unmet prior to the Fed's formation.

²³ Transactions involving the Federal Reserve Bank of Boston and the Boston Clearing House Association (BCHA) illustrate the use of Reserve Bank clearing services for interbank settlement. In November 1914, the Boston Fed became a limited member of the BCHA. The manager of the BCHA opened an account at the Boston Fed for settlement among members of the clearinghouse. Members in net debit positions paid the managers of the BCHA with checks drawn upon their accounts at the Reserve Bank, and the manager paid the members in net credit positions with checks drawn upon the account of the BCHA at the Boston Fed. The checks in Table 2 presented by the Boston Fed to Boston banks included the checks written by members of the BCHA to cover their net debit positions at the BCHA.

Table 2

Local Checks and Other Checks Collected Through the Reserve Banks, 1918

District	Percentage of Checks Presented by a Reserve Bank for Collection that Were Drawn on Banks in Cities with Reserve Bank Offices	Average Denomination of Checks Drawn On Banks Located In:	
		Reserve Bank Cities	Elsewhere
Boston	10.8%	\$ 2,613.5	\$ 278.8
New York	9.0	7,336.7	595.8
Philadelphia	24.9	1,559.2	360.2
Cleveland	6.7	2,024.7	586.7
Richmond	4.4	3,311.0	480.5
Atlanta	12.1	1,278.5	325.7
Chicago	21.1	2,210.5	221.0
St. Louis	17.7	2,538.9	241.2
Minneapolis	14.7	2,682.9	197.8
Kansas City	6.2	3,128.1	389.0
Dallas	6.2	1,476.1	369.2
San Francisco	17.2	1,409.0	245.9
System	9.6	2,639.6	496.3

SOURCE: Annual Report of the Board of Governors, 1918.

REFERENCES

- Bank for International Settlements. *Payments Systems in Eleven Developed Countries*, Bank Administration Institute, 1989.
- Baxter, William F. "Bank Interchange of Transactional Paper: Legal and Economic Perspectives," *Journal of Law and Economics* (October 1983), pp. 541-88.
- Berger, Allen N., Diana Hancock, and Jeffrey C. Marquardt. "A Framework for Analyzing Efficiency, Risks, Costs and Innovations in the Payments System," *Journal of Money, Credit and Banking* (November 1996, Part 2), pp. 696-732.
- Board of Governors of the Federal Reserve System. *All-Bank Statistics, United States, 1896-1955*, 1959.
- Burgess, W. Randolph. *The Reserve Banks and the Money Market*, Harper & Brothers, 1927.
- Calomiris, Charles W., and Charles M. Kahn. "The Efficiency of Self-Regulated Payments Systems: Learning from the Suffolk System," *Journal of Money, Credit and Banking* (November 1996, part 2), pp. 766-97.
- Cannon, James Graham. *Clearing Houses*, U.S. National Monetary Commission, Senate Document No. 491, 61st Congress, second session, Government Printing Office, 1910.
- Canova, Fabio. "The Sources of Financial Crisis: Pre- and Post-Fed Evidence," *International Economic Review* (August 1991), pp. 689-713.
- Clair, Robert T., Joanna O. Kolson, and Kenneth J. Robinson. "The Texas Banking Crisis and the Payments System," *Economic Review*, Federal Reserve Bank of Dallas (First Quarter 1995), pp. 13-21.
- Clark, Truman A. "Interest Rate Seasonals and the Federal Reserve," *Journal of Political Economy* (February 1986), pp. 76-125.
- Comptroller of the Currency. *Annual Report*, Vol. 1, December 7, 1914, Washington, Government Printing Office, 1915.

REVIEW

MAY/JUNE 1998

- _____. *Annual Report*, December 1, 1890, Washington, Government Printing Office, 1890.
- Duprey, James N., and Clarence W. Nelson. "A Visible Hand: The Fed's Involvement in the Check Payments System," Federal Reserve Bank of Minneapolis *Quarterly Review* (Spring 1986), pp. 18-29.
- Dwyer, Gerald P., and R. Alton Gilbert. "Bank Runs and Private Remedies," this *Review* (May/June 1989), pp. 43-61.
- Economides, Nicholas, and Lawrence J. White. "Networks and Compatibility: Implications for Antitrust," *European Economic Review* (April 1994), pp. 651-62.
- Federal Reserve Board. *Second Annual Report*, U.S. Government Printing Office, 1916.
- Fishe, Raymond P.H. "The Federal Reserve Amendments of 1917: The Beginning of a Seasonal Note Issue Policy," *Journal of Money, Credit and Banking* (August 1991), pp. 308-26.
- _____, and Mark Wohar. "The Adjustment of Expectations to a Change in Regime: Comment," *American Economic Review* (September 1990), pp. 968-76.
- Garbade, Kenneth D., and William L. Silber. "The Payment System and Domestic Exchange Rates: Technological Versus Institutional Change," *Journal of Monetary Economics* (January 1979), pp. 1-22.
- Garvy, George. *Debits and Clearings Statistics and Their Use*, Board of Governors of the Federal Reserve System, revised edition, 1959.
- Gilbert, R. Alton. "Effectiveness of State Reserve Requirements," this *Review* (September 1978), pp. 16-28.
- Hallock, James C. *Clearing Out-of-Town Checks in England and the United States*, 1903.
- Harding, W.P.G. Report to the U.S. Senate on "State Banks in Federal Reserve System," Senate Document No. 184, 66th Congress, 2nd Session, January 26, 1920.
- _____. *The Formative Period of the Federal Reserve System*, Houghton Mifflin Co., 1925.
- Holland, A. Stephen, and Mark Toma. "The Role of the Federal Reserve as 'Lender of Last Resort' and the Seasonal Fluctuation of Interest Rates," *Journal of Money, Credit and Banking* (November 1991), pp. 659-76.
- Howard, Earl Dean, and Joseph French Johnson. *Money and Banking*, DeBower-Elliott Co., 1910.
- James, John A. *Money and Capital Markets in Postbellum America*, Princeton University Press, 1978.
- Jessup, Paul F. *The Theory and Practice of Nonpar Banking*, Northwestern University Press, 1967.
- Jones, Thatcher C. *Clearings and Collections; Foreign and Domestic*, Columbia University Press, 1931.
- Katz, Michael L., and Carl Shapiro. "Systems Competition and Network Effects," *Journal of Economic Perspectives* (Spring 1994), pp. 93-115.
- Magee, James D. "The Fight Against Par Check Collection," *The Journal of Political Economy* (June 1923), pp. 433-45.
- Mankiw, N. Gregory, Jeffrey A. Miron, and David N. Weil. "The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve," *American Economic Review* (June 1987), pp. 358-74.
- _____, _____, and _____. "The Adjustment of Expectations to a Change in Regime: Reply," *American Economic Review* (September 1990), pp. 977-79.
- Miron, Jeffrey A. "Financial Panics, the Seasonality of Nominal Interest Rates, and the Founding of the Fed," *American Economic Review* (March 1986), pp. 125-40.
- Phillips, Ronnie J. "Payments System Technology and Central Banks: Early Lessons from the Federal Reserve and the Reichsbank," mimeo., Colorado State University, May 15, 1997.
- Preston, Howard H. "The Federal Reserve Banks' System of Par Collection," *The Journal of Political Economy* (July 1920), pp. 565-90.
- Rivlin, Alice M. "Testimony before the Subcommittee on Domestic Monetary Policy of the Committee on Banking and Financial Services," U.S. House of Representatives, September 16, 1997.
- Rolnick, Arthur J., Bruce D. Smith, and Warren E. Weber. "Lessons From a Laissez-Faire Payments System: The Suffolk Banking System (1825-1858)," this issue of the *Review*.
- Scherer, F. M. *Industrial Market Structure and Economic Performance*, Rand McNally College Publishing Co., 1970.
- Spahr, Walter E. *The Clearing and Collection of Checks*, The Bankers Publishing Co., 1926.
- Stevens, Ed. "The Founders' Intentions: Sources of the Payments Services Franchise of the Federal Reserve Banks," Financial Services Working Paper 0396, Federal Reserve Bank of Cleveland, December 1996.
- Summers, Bruce J., and R. Alton Gilbert. "Clearing and Settlement of U.S. Dollar Payments: Back to the Future?" this *Review* (September/October 1996), pp. 3-27.
- Tippetts, Charles S. "The Par Remittance Controversy," *American Economic Review* (December 1924), pp. 629-48.
- _____. *State Banks and the Federal Reserve System*, D. Van Nostrand Co., Inc., 1929.
- U.S. Department of Commerce. *Historical Statistics of the United States: Colonial Times to 1957* (1960).
- Vest, George B. "The Par Collection System of the Federal Reserve Banks," *Federal Reserve Bulletin* (February 1940), pp. 89-96.
- Weinberg, John A. "The Organization of Private Payment Networks," Federal Reserve Bank of Richmond *Economic Quarterly* (Spring 1997), pp. 25-43.
- White, Eugene Nelson. *The Regulation and Reform of the American Banking System, 1900-1929*, Princeton University Press, 1983.
- Willis, H. Parker. *The Federal Reserve System: Legislation, Organization and Operation*, The Ronald Press Co., 1923.

REVIEW

MAY/JUNE 1998

_____ and William H. Steiner. *Federal Reserve Banking Practice*, D. Appleton and Co., 1926.

Willis, Parker B. "The Background of the Par Collection and Clearance Provisions of the Federal Reserve Act and Some Features of the Par Collection Controversy," mimeo, Federal Reserve Bank of Boston, September 20, 1951.

Appendix

This appendix provides details on changes in federal and state reserve requirements that occurred more or less simultaneously with the introduction of the Federal Reserve System, because such changes allow us to draw conclusions about the effects of Reserve Bank services on banks' ratios of cash to total assets. Under the Federal Reserve Act (FRA), national banks were required to become members of the Federal Reserve System, and the FRA reduced reserve requirements for national banks substantially. Just prior to FRA's passage, reserve requirements were 25 percent of total deposits for national banks in central reserve cities (three major financial centers) and reserve cities (other important regional financial centers), and 15 percent for national banks located elsewhere (commonly called the "country" national banks). Banks in central reserve cities were required to hold all of their reserves as vault cash, whereas those located in reserve cities could hold up to half of their reserves as deposits at banks in central reserve cities, and country banks could hold 60 percent with banks located in reserve cities or central reserve cities.

Under the FRA, as enacted in 1913, reserve requirements for Fed members were different for demand and time deposits. For all member banks, the reserve requirement on time deposits was 5 percent, while the requirements on demand deposits were as follows: for banks located in central reserve cities, 18 percent; for those located in reserve cities, 15 percent; and for country member banks, 12 percent. As described in the text, calculations by the Comptroller of the Currency (1915) indicated substantial reductions in reserve requirements for national banks under the FRA.

It is difficult to estimate the effect of these changes on the cash assets of national banks, for the following reasons: Under the 1913 version of the FRA, only vault cash and balances at the Fed counted as part of reserves; balances with other banks were excluded from reserves. This exclusion of interbank balances limited the effects of the

Table A1

States That Lowered Their Reserve Requirements, 1913-25

California	New Mexico
Delaware	New York
Georgia	Oklahoma
Idaho	Oregon
Illinois	Pennsylvania
Indiana	South Dakota
Kansas	Texas
Kentucky	Virginia
Louisiana	Washington
Michigan	West Virginia
Minnesota	Wisconsin
Montana	

SOURCE: White (1983).

lower reserve requirements on the demand for cash by country national banks and national banks located in reserve cities in two ways. First, any collected balances that these national banks held with other banks did not count as reserves. In addition, the new requirements eliminated the practice of counting uncollected funds as part of reserves. For banks that collected checks through the Reserve Banks, uncollected funds were classified as cash items in the process of collection and, therefore, were not counted as part of reserves.

Fed reserve requirements were reduced again in 1917. On time deposits of all members, they were reduced from 5 percent to 3 percent. On demand deposits, they were reduced for banks in central reserve cities, from 18 percent to 13 percent; for banks in reserve cities, from 15 percent to 10 percent; and for country banks, from 12 percent to 7 percent. These changes had an offsetting effect on the demand for cash by member banks: Vault cash no longer counted as reserves; only balances in reserve accounts at the Reserve Banks were officially designated reserves.

Thus, the reserve requirement changes at the federal level were large, and their

effects on the demand for cash by member banks were so complex that it would be difficult to separate them from other effects, such as the Fed's collection service. Another possible approach is to examine changes in cash ratios of state banks around the time the Fed developed its payment services. About half the states lowered their reserve requirements around the time the Fed was founded, as indicated in Table A1. These state actions can be treated as an experiment. The issue is whether the ratios of cash to assets declined for state banks in those states that did not lower their reserve requirements around the time the Reserve Banks developed their payment services. The text presents the evidence on this experiment.

John A. James is a professor of economics at the University of Virginia. The author thanks Ross Levine and Ron Michener for their discussions with him, but does not implicate them in the result.

Commentary

John A. James

Let me say straight off that Gilbert (1998) is a fine article. It is a clear and comprehensive study of both the issues and institutions of the payments system in the early twentieth century. Gilbert covers so much ground that my remarks necessarily address only a few of his topics. Other important topics, such as whether universal par clearance is a good thing, I must ignore here.

Virtually all early twentieth century writers on U.S. banking and the system of payments condemned the system of check clearing and collection that existed under the national banking system that was built on correspondent bank relationships. Jones (1931, p. 131), for example, grumbled, "Circuitous routing, pyramiding and decentralization of reserves, inelasticity of our currency and excessive collection and exchange charges, characteristic of the correspondent system, placed unbearable burdens upon business." Walter Spahr, author of the most comprehensive tome on the subject (1926, pp. 101, 103), referred to the "great defects" and "evils" associated with the old (correspondent) system of clearing and collection, as did many others.

More recent students of the American banking system have been kinder to the correspondent banking system. Although it was not without its problems, due in large part to the lack of a lender of last resort, the correspondent banking system is now recognized as a quite efficient mobilizer and allocator of funds across industries and regions (Sylla 1975). White (1983, p. 66) writes, "In spite of contemporary criticism, the correspondent banking system and the clearinghouses served the financial needs of the nation well." In view of the more

kindly light in which we see the system of interregional transfers and concentration of reserves under the correspondent banking system in the pre-Federal Reserve period, we might ask whether the pre-Federal Reserve system of payments was really as inefficient as its early twentieth-century critics claimed.

A check was defined as a bill of exchange drawn on a bank payable on demand. But, by common law, payment in full or at par was required only if the check was presented for payment at the bank against which it was drawn (the drawee bank). In-town checks posed no problem: They could be sent to a clearinghouse or presented directly by messenger. The difficulty came with out-of-town collections. As Gilbert describes, checks presented through indirect means, such as through the mail, did not have to be paid at par. Rather, an exchange charge could be deducted from the face value, reflecting in principle the cost of shipping cash to the collecting banks. The correspondent banking system played a fundamental role in the system of out-of-town collections. To avoid such exchange charges, out-of-town checks were usually collected through the correspondent system.¹ Watkins (1929, p. 105) notes that the use of correspondent channels for collection was "adopted whenever possible." An item for collection from an out-of-town bank would have been sent to the receiving bank's in-town correspondent who in turn would pass on the check to another bank in its correspondent network in the vicinity of the drawee bank. This bank might pass it on to another bank until ultimately the check would be presented to the drawee bank for payment. In return, the first bank would collect for other banks in the correspondent network.

This method for collecting out-of-town checks involved a lot of paperwork. In his report to the National Monetary Commission (1910, pp. 64-70), James G. Cannon details the *straightforward* collec-

¹ These "excessive" exchange charges did not seem all that high compared with potential handling costs. In the controversy about nonpar clearing after the Federal Reserve began its clearing and collection operations, several states passed laws explicitly allowing banks to make exchange charges. Maximum rates were fixed at 1/8 or 1/10 of 1 percent of the face value of the check (Jessup 1967, p. 11). Kniffin (1928, p. 310) quotes somewhat higher charges—1/20 to 1/4 of 1 percent. A later (1917) study put average exchange costs at \$1 per \$1,000—1/10 of 1 percent (Jones 1931, p. 177).

tion of a check drawn on an Ohio country bank and remitted to New York City through a Cleveland correspondent. To do it, “two checks had to be drawn, four letters had to be written, 8 cents in postage stamps were used, and seventy-five or more handlings of the check were involved by a score or so of clerks, in five different banks, located in three different cities.” Moreover, the check might travel a circuitous route from the depository bank to the paying bank. One of the more celebrated checks in banking history was one for \$43.56, drawn by Woodward Brothers of Sag Harbor, NY, and paid to Berry, Lohman, & Rasch of Hoboken. Cannon (1910, pp. 70-72) describes the check’s route from its depository bank in Hoboken to correspondents in Manhattan, Boston, Tonawanda, Albany, Port Jefferson, Far Rockaway, Manhattan, Riverhead, Brooklyn, and finally to the paying bank in Sag Harbor. Critics of the correspondent bank system of clearings and collections cited this example time and time again. Given the paucity of other examples, one might be suspicious of this example’s general applicability.

In any case, there does not seem to be a way to measure directly the alleged inefficiencies of the pre-Federal Reserve system of clearing and collection in terms of “unnecessary” administrative costs and long delays in collection. It’s difficult therefore to get any quantitative notion of how bad the correspondent check-clearing and collection system was.² However, even if we take the indictments by contemporary critics as having merit (one might prefer to remain a bit agnostic here), it does not follow, I believe, that any payments systems based on correspondent banking networks would have been grossly inefficient.

Consider the earlier system of making out-of-town remittances: This system was based on bank drafts, which personal checks began to displace toward the end of the nineteenth century. Within this system the process of settlement was quite simple. Maintaining accounts with correspondents in financial centers, such as New York City, allowed interior banks to sell drafts on New York funds to their cus-

tomers. If the receiving bank had an account with the same New York correspondent, payment could simply be accomplished through a book transfer. If the receiving bank had an account with a different bank, settlement between the two New York banks could simply be done through the clearinghouse. The links among financial centers and interior banks were quite extensive under the correspondent banking system, so virtually all banks had access to funds in major financial centers. An 1890 Comptroller of the Currency Report survey found that, of the 3,329 responding national banks, 3,147 banks had drawn drafts on New York City during the previous year. A 1925 survey found that 600 of 655 Georgia banks had New York correspondents, and 1,146 of 1,600 Texas banks had them as well. Only 832 of 1,896 Illinois banks had a direct New York City link, but 1,705 had Chicago correspondents (Watkins 1929, p. 141). Most interior banks had New York City correspondents, or at least ones in other regional financial centers.

The price of New York City funds was reflected in domestic exchange rates (quoted regularly in newspapers) and in commercial and financial periodicals. Sprague nevertheless observed, “There is no part of our banking machinery which has received so little elucidation as that of the domestic exchanges. Even for normal times the subject is obscure” (1910, pp. 2 and 3). This statement is still, by and large, true today.³ Similar to foreign exchange rates, domestic exchange rates fluctuated with changes in supply and demand within the bands set by the cost of shipping currency to and from New York plus lost interest on the currency in transit. If the New York City balances of a New Orleans bank began to rise because of collections there, it might sell exchange to other New Orleans banks whose New York accounts were running low. The quoted rates therefore appear to have applied to business between banks rather than being the direct charge to customers.

Based on 1859 estimates for domestic exchange, the 1890 Comptroller’s Report claimed the average exchange rate had

² Indeed, the circuitous routing of checks for collection may not have been inefficient at all. Weinberg (1997, p. 39) argues that the pattern of correspondent relationships was determined by the normal pattern of commerce. Circuitous check routings then simply indicated that there were exceptions now and then to the usual flows. In view of the existing structure, it was efficient to send these occasional items along with routine shipments even if they were not going by the most direct route.

Exchange charges might have reinforced network efficiency here by reducing incentives for depository banks to bypass the network.

³ Only two papers, of which I know, have recently examined this market: Garbade and Silber (1979), and Phillips and Swamy (1997).

decreased more than elevenfold by 1890 (p. 21). This dramatic fall was attributed to the retirement of state banknotes and the substitution of national banknotes circulating at par throughout the country. Garbade and Silber (1979, pp. 14-15) also add the fall in railroad freight rates in the early postbellum period. In 1890, the average cost of domestic exchange in the United States was calculated at 85 cents per \$1,000.00 (1/12 of 1 percent); however, the rates ranged as high as \$2.10 in Texas and \$2.00 in Arkansas, Nevada, and Arizona.

To examine more closely domestic exchange rates in the period before the Fed was founded, I've collected some figures from *Bradstreet's*, which reported them weekly from the early 1880s. I stop in 1917, when the Fed opened its wire system for reserve transfers, and domestic exchange rates essentially remained at par. Domestic exchange rates fluctuated over the course of the year with the "needs of trade," so to look at longer trends I sample the same period every year—the first week of June. Figure 1 shows the deviations from par for \$1,000 in New York City funds in various cities over time. Positive values indicate times when New York funds were at a premium locally. Negative values denote times when New York funds sold at a discount. The straight line in each graph is the trend line over the period.

First, note the levels. During the first week of June, New York exchange was generally at a premium, and on average that premium seemed to increase with distance from New York. But the interesting thing here is generally how low the premiums were: in St. Paul and Kansas City, for example, more or less around 50 cents (or .05 percent), in St. Louis about half that.

Second, notice the predominant downward trend in the exchange rates. There was no trend in Boston to be sure, where the rate fluctuated very close to par. But in Cincinnati, Cleveland, Chicago, St. Louis, Kansas City, New Orleans, Memphis, Atlanta, and spectacularly so in San Francisco, there was a downward trend. Indeed, in several cities—Boston, Cincinnati, Cleveland, St. Louis, New Orleans,

Memphis, Atlanta—the June premium seems to have settled at zero by the time the Federal Reserve was established. On the other hand, it should be noted that in a few other cities—Minneapolis, St. Paul, Omaha, Charleston, Portland—there was no distinct downward trend in June exchange rates. Their stability is a bit of puzzle, but Sprague (1910, p. 297) observed, "The quoted rates of exchange were often without much significance. The ordinary course of dealings was so completely disorganized in many places that the rates were purely nominal, representing little or no actual transactions."⁴

The data in Figure 1 are nominal values. If one adjusts for changes in the price level (using the Warren-Pearson wholesale price index), the convergence toward par is even more pronounced. This convergence in domestic exchange rates was moreover not just a June phenomenon. Figure 2 shows the price of New York exchange over time in Chicago, St. Louis, New Orleans, and San Francisco for the first weeks of February, June, October, and December. The range of exchange rates in each city decreased markedly. At the times when New York funds sold at a discount (as in Chicago, St. Louis, and New Orleans), that discount decreased during the period; similarly, at the other times of the year, the premiums decreased as well.⁵

The factors that might have caused this diminution (in absolute value terms) in domestic exchange rates over the period are not obvious. It could have been simply long-term changes in the seasonal demands for New York City funds and/or increases in the supplies of New York correspondent balances. But this seems unlikely since both premiums and discounts were decreasing over time in several cities. The less-than-universal character of the decline would argue against a general fall in shipping costs (if one takes the nondeclining observations as legitimate).⁶ Garbade and Silber (1979, pp. 4, 15) attribute the decline in the variability of domestic exchange over the course of the year in this period to correspondent banks' increasing role as market makers in exchange. This may have also had some

⁴ Also, rates in locations such as Minneapolis and St. Paul were quite different during a number of years. These apparently unexploited arbitrage possibilities might lead one to suspect that not all these series are completely reliable. Some local correspondents may not have been very assiduous in gathering information (see James 1978, pp. 255-62).

⁵ In addition, the violations of the currency shipping points declined dramatically over time. Use the 1910 figures that Gilbert quoted on the cost of shipping \$1,000 in currency between New York and Chicago (\$.50), and St. Louis (\$.60), and New Orleans (\$.75), and San Francisco (\$1.50), and the claim by Garbade and Silber (1979, p. 15) that real currency shipping costs stabilized after the early 1880s. We see then in all four cities in the first decade numerous, indeed regular, violations of the currency bands, but in the last decade only a few (St. Louis and New Orleans) other than in the Panic of 1907.

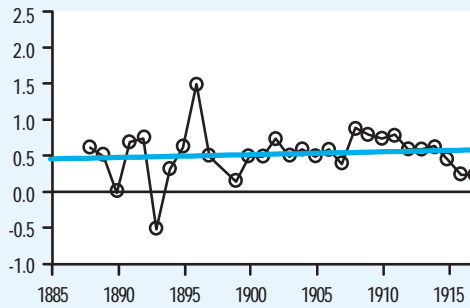
⁶ As noted above, Garbade and Silber (1979) argue that currency shipping costs barely declined after the early 1880s, but this is just on the basis of New York-Chicago freight rates. If differing degrees of railroad monopoly power between different city pairs existed, perhaps such a disperse pattern would have been possible.

Figure 1

June Domestic Exchange Rates*

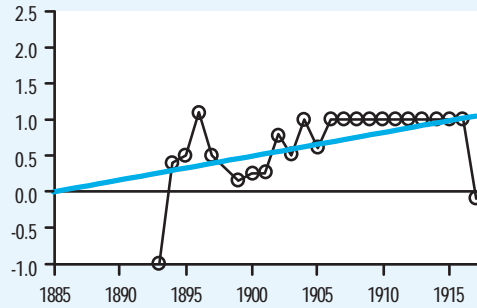
St. Paul

\$ per \$1000



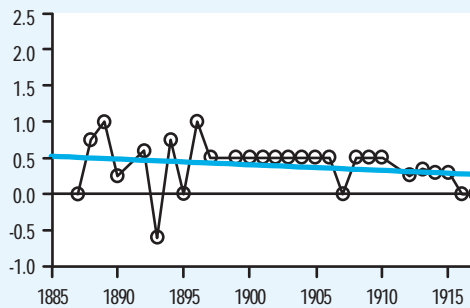
Minneapolis

\$ per \$1000



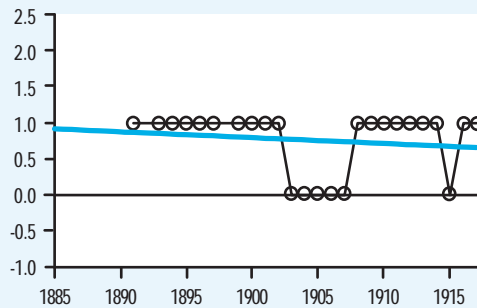
Kansas City

\$ per \$1000



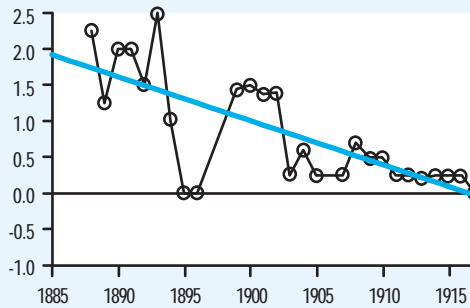
Omaha

\$ per \$1000



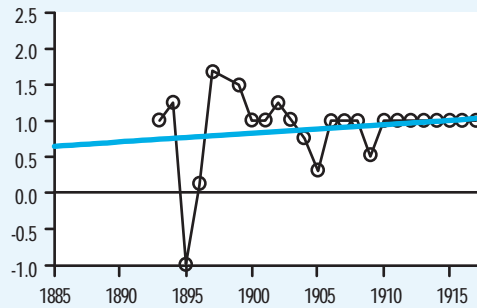
San Francisco

\$ per \$1000



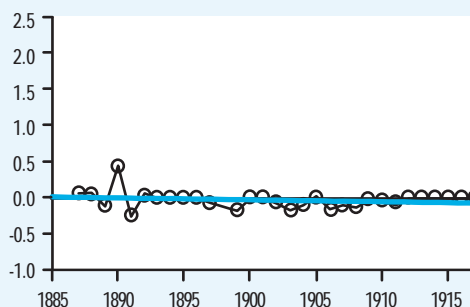
Portland, OR

\$ per \$1000



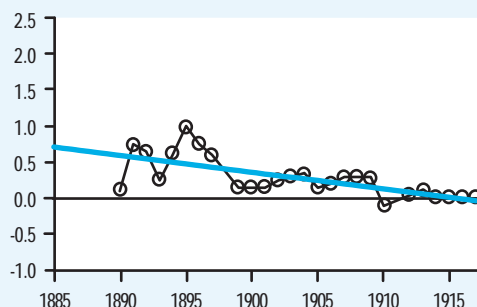
Boston

\$ per \$1000



Cincinnati

\$ per \$1000

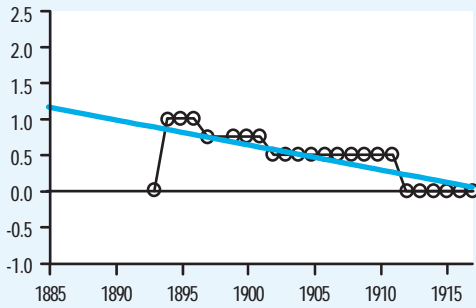


*Straight line is trend line.

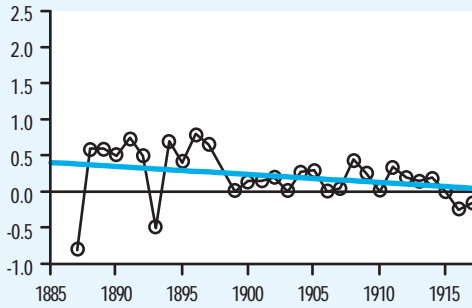
REVIEW

MAY/JUNE 1998

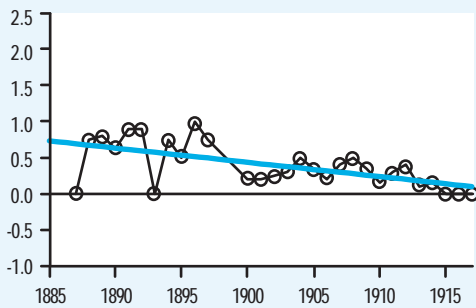
Cleveland
\$ per \$1000



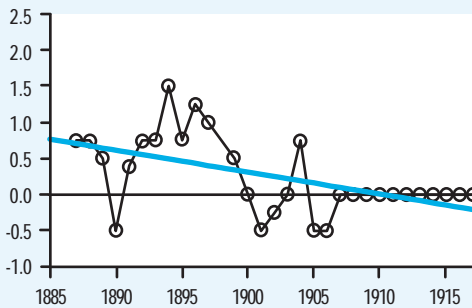
Chicago
\$ per \$1000



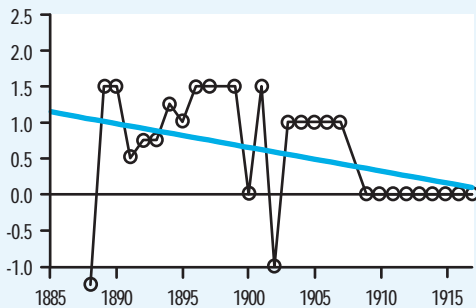
St. Louis
\$ per \$1000



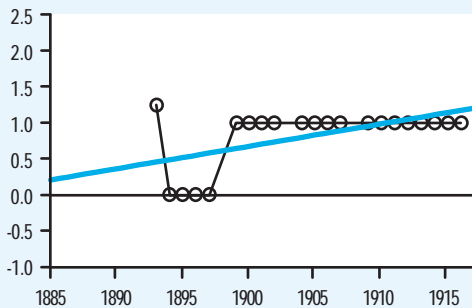
New Orleans
\$ per \$1000



Memphis
\$ per \$1000



Charleston
\$ per \$1000



Atlanta
\$ per \$1000

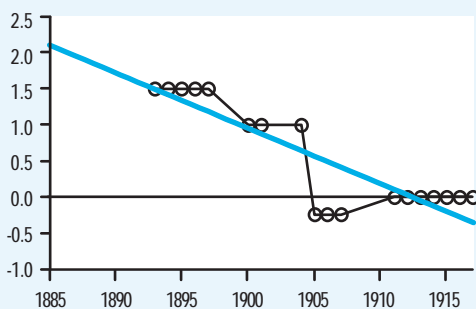
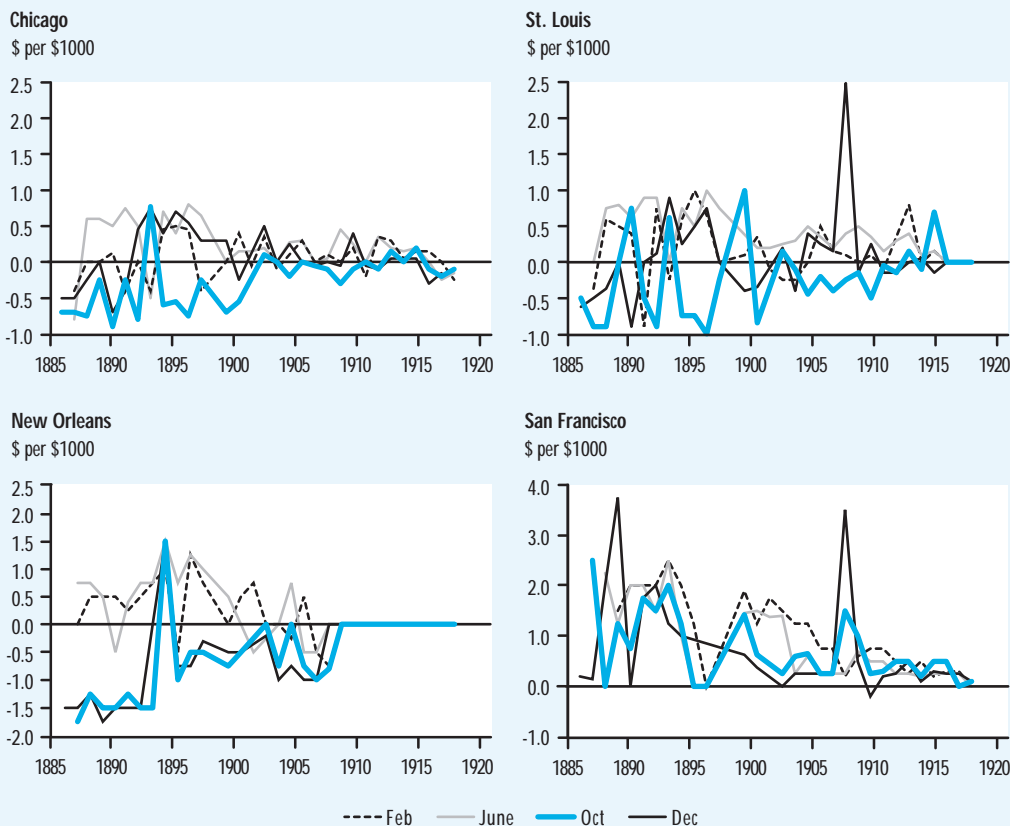


Figure 2

Domestic Exchange Rates



impact on the exchange-rate levels. In any case, domestic exchange rates in general were not high, and they fell dramatically in the decades before the Fed's establishment, in several cases settling down to zero.

In contrast to checks, which were necessarily payable at par only when presented physically to the issuing bank, the use of drafts payable in New York City or other financial centers greatly simplified the process of clearing and collecting out-of-town payments. Draft usage took advantage of the often-reviled centralization of reserves promoted by the correspondent and national banking systems to, in effect, create a nationally centralized clearing system. Administrative costs were small and, as we have seen, the costs of New York funds were not that great. Little shipping of currency around the country would have been required to settle

remittances. (However, perhaps critics of the old system may have exaggerated the amount of currency transferred in settling checks, since checks were often settled by issuing a bank draft.) Interior banks had to maintain correspondent accounts on which they generally earned 2 percent interest. If, instead, they held excess reserves at home to settle demands for payments of checks, they would have earned nothing. Gilbert agrees that collection costs "borne by the payee would be smaller with a draft drawn upon a bank in a financial center than with a check drawn upon the deposit account of the payor."

If the payments system based on drafts was really more efficient than the one based on checks, why did the latter begin to displace the former in the late nineteenth century? The advantages of checks to the payor were clear: It was simple to write

out a check; obtaining a draft required a trip to the bank. Paying by draft involved an immediate debit to one's account, while paying by check allowed the issuer to draw interest on the funds in his or her account until the check was presented for payment. In turn, banks often absorbed check collection costs rather than pass them on to the payee. (We know that banks in this period generally had some monopoly power and were earning profits.) It seems, then, that the increased convenience of checks must have outweighed any increased inefficiencies of collection.

Now let me turn to the question of whether the founding of the Federal Reserve improved check-clearing and collection efficiency in the United States. Although the indictments of the old check-clearing and collection system were essentially anecdotal, I agree with Gilbert's reasoning that the Federal Reserve most probably did improve the efficiency of the payments system. Although the Federal Reserve Banks did not devise a new technology for check collection, legal changes did solve the principal-agent problem that discouraged banks from using the most direct and efficient method of check presentation—the mail. The use of regionally centralized clearings may have reduced the necessary physical transfer of cash, but the magnitude of this effect is not clear, since checks were often settled by drafts on financial centers anyway. Shorter collection times, compared with alleged wandering checks of the earlier era, would have reduced the risk that the paying bank might default on or dishonor a check.

Although the Federal Reserve system of check clearing and collection appears superior to the old regime as it is described in the literature, empirical tests or supporting evidence are difficult to come by. Cost data for a direct test are alas not available. Gilbert's examination of cash holdings of state-chartered banks in states that did not reduce their reserve requirements is quite ingenious. Banks conserving on cash holdings did seem to represent a social benefit as well as a private one since, other things equal, a banking system with lower cash

holdings could engage in more intermediary activity. Gilbert finds that cash ratios of state-chartered banks declined 5 to 6 percentage points during the growth period of the Fed's check-clearing services and remained at that lower level through the 1920s.

The problem, of course, with using such straightforward time-series evidence is that other things could have been going on at the same time. Suppose the establishment of the Federal Reserve System led banks and the public to believe that panics were now a thing of the past. Even if state banks did not have direct access to the Fed's rediscount facilities, this general feeling of confidence could have led state banks to reduce their cash ratios permanently. On the other hand, Gilbert's sample included state banks engaged in nonpar clearings.⁷ Since it was quite possible that an agent of the Federal Reserve might appear at their banks with a large bundle of checks for collection any day, such banks might have held more cash than they had previously. In that case, the observed trend in cash holdings would understate the efficiency-enhancing effects of the Fed's clearing system. I have no idea which of these conflicting influences might dominate, so perhaps the data in Gilbert's Figure 3 are in fact a pretty good picture of what was going on.

If, however, one compares the efficiency of the Federal Reserve's clearing and collection with the earlier system (based on bank drafts), the differences narrow considerably. Both offered centralized clearing and collection: Drafts did so nationally (in New York City), and the Federal Reserve Banks did so regionally. Therefore, not much cash would have to move around the country. As we have seen, the costs of New York exchange generally were low and declining, stabilizing at par during the 1910s in several cities. Both systems reduced collection times and float over the pre-Federal Reserve check collection system, in which items in the process of collection were usually counted as reserves as soon as they were sent off for collection. In view of the sometimes leisurely process of collection, a single

⁷ Jessup (1967, p. 105) lists 7,499 nonmember banks in 35 states not on the par list as of August 31, 1919. Of the 25 states that did not lower reserve requirements, 15 had nonpar banks.

check might have served double, triple, or more duty as legal reserves. The size of this float in the early twentieth century was estimated at between one-third and one-half of deposited reserves (Jones 1931, p. 163). Giving immediate credit to banks in turn allowed customers to draw checks against uncollected funds (Preston 1920, p. 567). In contrast, credit to a bank's account at the Federal Reserve was deferred from one to eight days, depending on a schedule based on the average mail time required for the item to reach the paying bank and for the remittance to be made to the Federal Reserve Bank. Under the draft system, customers' accounts were debited as soon as the drafts were purchased, and there seemed little risk of the payments' being dishonored. The draft system solved the principal-agent problem by having the city correspondent do the clearing and collection.

As Gilbert notes, clearing and collection would also have been more efficient with nationwide branch banking than under the pre-Federal Reserve system. Such an institutional arrangement was, however, not legal in the early twentieth century. At least the draft system was a legally feasible alternative to the then-existing system. But the question of efficiency and the underlying Berger, Hancock, and Marquardt framework that Gilbert uses may be too narrow. One might ask instead whether having the Federal Reserve play a role in check clearing and collection facilitated the meeting of its primary objectives as spelled out in the Federal Reserve Act. I would think the answer here would be yes. The reduction in float, for example, allowed a more precise measurement of bank resources and cash on hand at a point in time, even if there had been no improvement in payments efficiency. Similarly, Gilbert quotes Stevens' argument that "the collection system was the glue that tied banks to the Fed," allowing it to become a more effective central banker. Subsidizing collection charges was one way in which to make up to member banks the loss of interest on legal reserves. So even if the Federal Reserve's takeover of clearing and collections would not have dramatically improved the efficiency of the payments system relative to

alternatives (such as drafts), it still would have been, on the whole, a good thing.

REFERENCES

- Bradstreet's* (1880-1917).
- Cannon, James G. *Clearing Houses*, U.S. National Monetary Commission, Senate document no. 491, 61st Congress, 2nd Session, Government Printing Office, 1910.
- Garbade, Kenneth D., and William L. Silber. "The Payment System and Domestic Exchange Rates: Technological Versus Institutional Change," *Journal of Monetary Economics* (January 1979), pp. 1-22.
- Gilbert, R. Alton. "Did the Fed's Founding Improve the Efficiency of the U.S. Payments System" this issue of *Review*.
- James, John A. *Money and Capital Markets in Postbellum America*, Princeton University Press, 1978.
- Jessup, Paul F. *The Theory and Practice of Nonpar Banking*, Northwestern University Press, 1967.
- Jones, Thatcher C. *Clearings and Collections: Foreign and Domestic*, Columbia University Press, 1931.
- Kniffin, William H. *The Practical Work of a Bank*, 7th ed., The Bankers Publishing Co., 1928.
- Phillips, Ronnie J., and P.A.V.B. Swamy. "Par Clearance in the Domestic Exchanges: The Impact of National Banknotes," U.S. Comptroller of the Currency Working Paper 97-4 (March 1997).
- Preston, Howard H. "The Federal Reserve Banks' System of Par Collection," *Journal of Political Economy* (July 1920), pp. 565-90.
- Spahr, Walter E. *The Clearing and Collection of Checks*, The Bankers Publishing Co., 1926.
- Sprague, Oliver Mitchell Wentworth. *History of Crises under the National Banking System*, National Monetary Commission, Senate document no. 538, 61st Congress, 2nd Session, Government Printing Office, 1910.
- Sylla, Richard E. *The American Capital Market, 1846-1914: A Study of the Effects of Public Policy on Economic Development*, Arno Press, 1975.
- U.S. Comptroller of the Currency. *Annual Report, 1890*, Government Printing Office, 1890.
- Watkins, Leonard L. *Bankers' Balances: A Study of the Effects of the Federal Reserve System on Banking Relationships*, A. W. Shaw Company, 1929.
- Weinberg, John A. "The Organization of Private Payment Networks," Federal Reserve Bank of Richmond, *Economic Quarterly* (Spring 1997), pp. 25-43.
- Westerfield, Ray B. *Banking Principles and Practice*, Ronald Press Co., 1921.
- White, Eugene N. *The Regulation and Reform of the American Banking System, 1900-1929*, Princeton University Press, 1983.

Joseph R. Mason is an economist at the Office of the Comptroller of the Currency.

With financial assistance from the National Science Foundation, the National Bureau of Economic Research, the University of Illinois at Urbana-Champaign, and the Federal Reserve Bank of St. Louis, Professor Charles W. Calomiris of Columbia University and Joseph R. Mason are nearing the completion of a significant compilation of individual commercial bank records from the 1930s. Mason described the project and some of their research findings at the conference. The following is a summary of his remarks.

American Banks During the Great Depression: A New Research Agenda

Joseph R. Mason

Professor Charles Calomiris and I have been working for some time now on a project, entitled *Assistance to Banks During the Great Depression*, made up of two data-coding endeavors. The first involves the encoding of data from original bank *Reports of Condition* and *Reports of Earnings, Expenses, and Dividends* (i.e., *Call Reports*) of about 6,500 state- and nationally chartered Federal Reserve member banks in existence between 1929 and 1935. The second involves the collection of bank-structure data, including mergers, acquisitions, and failures, for this same time period.

All *Call Report* data came from microfilm archived at the Board of Governors of the Federal Reserve System, the result of a 1947 initiative to preserve data that had been collected on Federal Reserve member banks since 1916. The original *Call Reports* of individual member banks were supplied by the Federal Reserve Banks and microfilmed by the Board of Governors.

Table 1 summarizes the data microfilmed by the Board of Governors covering the period from 1929 to 1935. The Fed focused its filming effort on the *Call Reports* of state Federal Reserve member banks, filming semiannual data from these banks (except the balance sheets of June 1934 and June 1935). For national banks, only the December 1929, 1931, and 1933-35 *Reports of Condition*

Table 1

Availability of Data from Board Micro Im, 1929-35

	State Member Banks		National Banks	
	Balance sheets	Income statements	Balance sheets	Income statements
Dec. 1929	X	X	X	X
June 1930	X	X		
Dec. 1930	X	X		
June 1931	X	X		X
Dec. 1931	X	X	X	X
June 1932	X	X		
Dec. 1932	X	X		
June 1933	X	X		X
Dec. 1933	X	X	X	X
June 1934		X		X
Dec. 1934	X	X	X	X
June 1935		X		X
Dec. 1935	X	X	X	X

and *Reports of Earnings* and the June 1931 and 1933-35 *Reports of Earnings* were transferred to microfilm. At each Reserve Bank's discretion, the original records were destroyed after filming and, apparently, few of these records have survived. The microfilm held by the Board of Governors is the only known comprehensive collection of the reports of individual banks before the use of magnetic tape beginning in 1961.

From the *Reports of Condition*, we encoded balance-sheet data from the main schedule, as well as information from various supplementary schedules. For example, we encoded information on bank loans and discounts, including such items as real estate loans (farm and nonfarm), loans on securities, etc. We also encoded data on bills payable and rediscounts with Federal Reserve Banks and the Reconstruction Finance Corporation, the distribution of cash and amounts due from other banks, a detailed breakdown of demand and time

deposits, and the liabilities to each bank of its officers and directors. The *Reports* also list the Federal Reserve district, state, county, and city in which each bank is located, and the number of its branch offices.

We also encoded data from the *Reports of Earnings, Income, and Expenses* of each bank. Specifically, we encoded each bank's current earnings, expenses, chargeoffs, and recoveries. One of the first results to emerge from these data concerns the relationship between deposit interest rates and bank risk. Subsequent to the Depository Institutions Deregulation and Monetary Control Act of 1980, most banking experts believed that risky banks could attract deposits by increasing interest rates, thereby placing undue burden on the safety net. Calomiris and Mason (1997) found that during the Great Depression, interest paid as a percent of demand deposits was, on average, lowest at high-risk banks. This finding confirmed a weak relationship, as indicated by aggregated data (Benston 1964). On first glance, this finding appears paradoxical, though Gorton and Pennacchi (1990) suggest that some bank depositors may be very unwilling to accept increased risk on their accounts, even if offered higher interest rates. Risk-intolerant depositors may prefer to adjust to changes in a bank's riskiness by adjusting the quantity of their balances with the bank (see also Calomiris and Wilson 1996). Our research, which found large deposit outflows from high-risk banks, confirms this assertion.

The final phase of the project involved the collection of merger, acquisition, and failure information for the member banks whose *Call Reports* were encoded. These data were obtained from *Rand McNally Bankers Directory*. We are currently building usable structure databases to link banks across time and to explain disruptions in individual bank histories. We anticipate that our project will be completed within the next year, at which time the data will be made available to researchers through the Inter-university Consortium for Political and Social Research and the Federal Reserve Bank of St. Louis.

REFERENCES

- Benston, George G. "Interest Payments on Demand Deposits and Bank Investment Behavior," *Journal of Political Economy* (October 1964), pp. 431-49.
- Calomiris, Charles W., and Joseph R. Mason. "Contagion and Bank Failures During the Great Depression: The June 1932 Chicago Banking Panic," *American Economic Review* (December 1997), pp. 863-83.
- _____, and Berry Wilson. "Bank Capital and Portfolio Management: The 1930s Capital Crunch and the Scramble to Shed Risk," Working Paper, Federal Reserve Bank of Chicago Conference on Bank Structure and Competition (1996:32d).
- Gorton, Gary B., and George G. Pennacchi. "Financial Intermediation and Liquidity Creation," *Journal of Finance* (March 1990), pp. 49-71.