

# From exceptional to normal: changes in the structure of US banking since 1920

RICHARD SYLLA  
*New York University and NBER*

A century ago the US commercial banking system was exceptional in two ways. It was by good measure the largest commercial banking system of any country. And it was different from the commercial banking systems of other leading countries in having tens of thousands of independent banks with very few branches rather than the more typical pattern of a far smaller number of banks with many branches. Today, a century later, the US system is more normal than exceptional, dominated by a small number of very large banks with extensive branch systems. This article describes the US banking-structure transition from exceptional to normal. It closes with an interesting contrast of US and European banking developments.

**Keywords:** banking structure, financial regulation, unit vs. branch banking, bank concentration

**JEL classification:** G21, G28, N11, N14

## I

Youssef Cassis succinctly posed one of the main issues of this article: ‘To what extent can banks alter the economic environment in which they are working, and to what extent are they straitjacketed by this environment? All the major debates in banking history have ultimately revolved around this question’ (Cassis 2002, p. 49). A century ago the US commercial banking system was very large in numbers and total assets, but its members were so straitjacketed by politics and regulation as to make it the odd man out when compared to the banking systems of other leading nations. Constrained commercial banking continued for half a century, during which it lost market share in finance to US money and capital markets. Then bankers began to work with politicians and regulators to loosen the straps of the straitjacket. By the end of the century the straps were undone and the jacket thrown away. The global financial crisis of 2007–9 indicated that America’s leading bankers were perhaps too successful in achieving their deregulatory goals. In the process, however, they did succeed in creating a banking structure like that of other nations. The US banking system was no longer the odd man out.

R. Sylla, Professor Emeritus of Economics, Department of Economics, Stern School of Business, New York University, 44 West 4th St, New York, NY 10012, USA; email: [rsylla@stern.nyu.edu](mailto:rsylla@stern.nyu.edu)

At its all time peak in 1921, there were 30,456 independent commercial banks in the USA, and nearly all of them were unit (or one-office) institutions. Branch banking was rare, almost forbidden to federally chartered national banks, which numbered 8,150 (and had just 72 branches, up from 5 in 1900). And apart from a few of the then 48 states, branching was little practiced by 22,306 state-chartered banks. In 1921, all state banks had just 1,383 branches, up from 114 in 1900. Total branches were 1,455 (*Historical Statistics 2006*, 3–665), implying that most branch banks had just one branch.

In having such large numbers of mostly one-office banks, the US system was unique. Other developed national economies a century ago had far fewer banks and much more developed branch-banking systems. The US system was unique in another respect. It was by far the largest commercial banking system of any country. One might think that was a result of World War I, which placed greater stresses on the economies and financial systems of the UK, France and Germany, the other large developed economies of the early twentieth century, than it did on the USA. Those differing stresses of war did have large impacts. In just a few years, for example, they transformed the United States from the world's largest international debtor nation to its largest creditor nation, and the center of international finance moved from London to New York (Sylla 2011).

Even before the war broke out in 1914, however, the US banking system measured by total deposits, the main source of funding for bank loans and investments, was by good measure the largest of any country. Michie (2003), gathering data from Mitchell (1998) on bank deposits across the world in 1913, and putting them into comparable form, showed that the US commercial banks had deposits (\$9.25 billion, 36 percent of world commercial bank deposits of \$26.1 billion) almost as large as those of the UK, Germany and France combined (\$10.2 billion, 39 percent of world). But Mitchell (1998) counted only demand deposits for the USA, leaving out \$4.43 billion of time deposits (Sylla 2006). Assuming the Mitchell data for other countries and the world are complete, and recalculating to account for the omission of a large amount of US deposits, we find that the UK–France–Germany combined share in 1913 is 33 percent of world commercial bank deposits, while the US share rises to 46 percent. Michie also gathered data on savings bank deposits, which were relatively large in some countries, particularly Germany where they substantially exceeded commercial bank deposits. Still the 1913 US share, 37 percent, of world deposits – total commercial plus savings bank deposits – exceeded the combined share of the UK, Germany and France, 36 percent. Together these four large national economies in 1913 accounted for about 75 to 80 percent of total worldwide bank deposits.

A century later, the United States continues to have a large number of independent commercial banks compared to most countries, but far fewer than it had at the 1921 peak. And most of those banks were branch banks, in sharp contrast to 1921 when almost all banks were unit banks. Table 1 indicates that in 2018 the USA had less than a sixth of the number of commercial banks it had in 1921, and more than 80 percent of these banks were branch banks in contrast with less than 2 percent in 1921. These were large changes in banking structure. They have made the US

Table 1. *US Commercial banking structure, 1921–2018: number of banks, branches, offices, unit banks and branch banks*

Year	Number of banks	Branches	Offices	Unit banks	Banks with branches
1921	30,456	1,455	31,911	29,909	547
1929	24,970	3,353	28,323	24,206	764
1933	14,207	2,786	16,993	13,623	584
1984	14,507	42,731	57,227	7,403	7,093
2008	7,087	83,040	90,117	1,714	5,363
2018	4,718	78,014	82,732	879	3,839

Sources: For 1921, 1929 and 1933, *Historical Statistics* (2006), vol. 3, series Cj251, Cj325, Cj337; ‘Offices’ is number of banks plus branches, and ‘Unit banks’ is number of banks minus banks with branches. For 1984, 2008 and 2018, FDIC, *Historical Bank Data*, 31 Dec. each year, <https://banks.data.fdic.gov>, accessed May 2020.

structure look more like that of most nations, instead of being a unique outlier as it appeared a century ago.

A goal of this article is to identify and explore the key developments and turning points that led to the long-term structural change in US commercial banking. Key turning points are given in Table 1. The 1920s were a decade of great prosperity in the USA, but that decade saw some 5,000 banks, more than a sixth of all commercial banks, disappear. The years 1929–33 marked the Great Depression, in which more than 10,000 additional banks suspended, failed or otherwise disappeared. The next half century of was one of great stability in the number of banks. The peak of bank numbers between 1921 and 2018 came in 1984, but it differed little from the number in 1933. The next quarter century leading up to the Global Financial Crisis of 2007–9 saw the number of independent commercial banks cut by more than half while the number of bank branches roughly doubled. Then in the subsequent decade, 2008–18, the USA lost another third of its banks.

Thus, in four of the five periods identified in Table 1, comprising about half of the total years covered, there were sharp declines in bank numbers. And in one, there was a half century of little change in bank numbers. We discuss below what drove the change, or a lack of change, in each of these periods. But first we ask and attempt to answer the question of just how the USA came to have such a unique banking structure a century ago. It is an integral part of the story.

## II

The main reason why the USA developed and maintained its unique banking structure for most of its history is that the states, not the federal government, maintained most of the authority to create and regulate banks. State governments jealously guarded their authority, protecting their banks from outside competition, meaning

from banks in other states, which were not allowed to enter. In many states, the state government also protected individual banks within the state from out-of-town competition by decreeing that all banks in the state had to be unit banks. Banking interests, populist politics and state governments worked together to obtain these outcomes (Calomiris and Haber 2014; Lu 2017).

Apart from the early decades, the 1790s to the 1830s, when banks received their corporate charters by means of individual acts of state legislatures, there is not much evidence that banking growth and development was impeded. Starting in the 1830s, states began to enact 'free banking', that is, general incorporation laws for banks and other enterprises, making entry into banking relatively easy for the subsequent century, when the total number of US banks grew from less than a thousand to more than 30,000. State control of banking determined the manner in which US banking developed, but it did little to limit its rapid growth.

The US federal government did not entirely defer to the states in banking. Two national banks, the First and Second Banks of the United States (BUSs), actually gave the USA nationwide branch banking under their 20-year charters, 1791–1811 and 1816–36. Both were for the most part well-managed quasi-central banks that acted as fiscal agents for the federal government, and promoted monetary, financial and economic stability. They also furnished a great deal of private credit to a rapidly growing US economy. That private credit provision contributed to their failure to have their charters renewed, as it made them vulnerable politically to the interests of state banks and state governments, which had various financial interests in the banks they chartered. Getting rid of the national banks came to be viewed by state governments and banks as a win-win-win proposition. It would rid them of both a competitor and a *de facto* regulator, and without them the state banks would likely gain the banking business of the federal government. These parochial state interests in 1811 managed to defeat the renewal of first BUS's charter by one vote in the US Senate, and in 1832 they prevailed on President Andrew Jackson to veto Congress's renewal of the second BUS's charter (Hammond 1957). After Jackson's 1832 veto, the USA would not again have nationwide branch banking until 1994. And it would not have a third central bank until 1914, when the Federal Reserve System and its regional reserve banks began to operate. Congress designed the Fed not to compete with commercial banks in providing private credit. Unlike the ill-fated BUSs, the Fed has lasted for more than a century.

One might wonder how US banking would have developed had the charter of the BUS been routinely renewed, as was the charter of the Bank of England after 1694. Canada's banking history offers an interesting and ironic contrast that is suggestive. The irony is that the first Canadian banks in the early nineteenth century modeled their own charters on Alexander Hamilton's 1791 BUS charter enacted by Congress, which allowed nationwide branch banking. The Canadians proceeded to implement that model both before and after the USA junked it in the 1830s.

Even more important was that when Canada gained its independence from the UK in 1867, it decided to make banking oversight a federal rather than a provincial function, the very opposite of what happened in the US. So as US states (and the federal government, as discussed below) chartered tens of thousands of mostly one-office banks, Canada chartered far fewer banks and allowed them to build extensive nationwide branch systems (Kobrak and Martin 2018). For two centuries Canada's banking system has had a much lower incidence of failures than the USA. Arguably Canada's large, geographically and economically diversified banks made for greater financial stability than the USA experienced with its highly fragmented banking system and less diversified banks.

The US federal government made another bold attempt to insert itself into US banking, and even take it over entirely, during the Civil War of 1861–5. It did so by creating a National Banking System, a system of federally chartered commercial banks based on earlier state free-banking models. Congress hoped that all the old state-chartered banks would shift to national charters. When that didn't happen, it levied a prohibitive tax on notes issued by state banks. The tax failed to accomplish its goal because several hundred state banks simply abandoned note issuing and continued to operate as pure deposit banks. States then proceeded to regulate banks with a lighter touch than the federal government employed to regulate national banks, leading to a massive revival of state banking. At the peak of bank numbers in 1921, state banks greatly outnumbered national banks, 22,000 to 8,000, and also held a majority of all US commercial bank assets.

Branch banking was not explicitly disallowed in the National Bank Act of 1864. But early Comptrollers of the Currency, regulatory authorities created by the legislation, read the legislation narrowly and literally as a ban on branching. National banks under the law had to specify 'the place' where they would do business at 'an office or banking house' in the city specified in their organizational documents. The terms quoted were singular, not plural, and thus were interpreted as precluding branch banking (Robertson 1995). Later Comptrollers favored branching for national banks, but were opposed by the ever more numerous unit banks (White 1983). In 1900, a reduction in the minimum required capital for a national bank led to a surge in their numbers, further undercutting arguments for branch banking privileges. In 1927, Congress's McFadden Act eased the ban on branching by allowing national banks to have branches to the extent allowed by state banking laws in the states where they were located.

A system with tens of thousands of unit banks faced a major problem of network coordination. It was solved in the US case by the development of elaborate correspondent banking networks. These began early in the century when banks in cities and towns within a state or in other states began to maintain deposit balances in major economic and financial centers such as New York, Philadelphia and Boston so that their local clients could access funds in those centers to purchase both imported goods and domestic products. The National Bank Act furthered this development by allowing country banks to count as reserves deposits they maintained in a so-called

reserve-city bank, and reserve-city banks similarly could count as reserves deposits they maintained in central-reserve-city banks. Initially, New York was the only central reserve city, and its national banks were required to hold 25 percent legal reserves against their liabilities, a high reserve ratio. New York banks competed for these reserves by paying interest on them, which they could afford to do by lending out the funds as highly liquid call loans to participants in New York's securities markets, most prominently the New York Stock Exchange (James 1978). The high reserve ratio of New York City national banks helped substitute for a central bank by achieving limited reserve centralization, and the clearing house the banks organized in 1853 acted as a lender of last resort during the comparatively frequent bank panics the USA experienced in the absence of a central bank. The last of these panics, that of 1907, persuaded Congress to bring back a central bank in the form of the Federal Reserve System, in 1914.

### III

The 1920s were in general a prosperous decade for the USA. Nonetheless, by 1929 nearly 5,500 of its banks, about 18 percent of peak number in 1921, failed or suspended operations (see Table 1). The failures were not perceived as a national problem at the time. Most of them were small banks in agricultural regions, namely in 12 states between the Mississippi River and the Rocky Mountains, plus three states of the Old South (North Carolina, South Carolina and Georgia), and California (Board of Governors of the Federal Reserve System 1959). These states were home to 80 percent of the failures.

Despite the prosperity of the Roaring 20s, US farmers did not do well in the decade. Debts incurred for land and equipment in the period of inflated prices during and immediately after World War I became difficult to repay when commodity prices fell in the 1920s. When farmers could not repay their creditors – farm mortgage lenders and suppliers of agricultural inputs – the creditors could not repay their bank loans, and the small, undiversified unit banker therefore had a difficult time staying in business.

The problem was not unique to the USA. As economist W. Arthur Lewis (1949) explained in an underappreciated but insightful book long ago, the problem was worldwide and very much a consequence of World War I. Before the war, in the period 1870–1914, world trade grew at about 3 percent per year, and generated prosperity almost everywhere as the dynamic industrial centers of the world – Western Europe, the USA and Japan – imported foodstuffs and raw materials from the primary-producing peripheries of the rest of the world. Then the war resulted in a large negative population shock that reduced the growth of world trade to only 1 percent a year from 1914 to 1929, hitting primary producing economies harder than the industrialized centers.

The USA with its continent-wide economy was both an industrial center and a primary producer. Its industries thrived during the 1920s while its agricultural areas

experienced a depression and numerous small bank failures. Industrial states such as New York, New Jersey, Connecticut and Pennsylvania actually gained new banks during the decade, while four contiguous farming states of the upper Midwest lost hundreds of banks: Iowa, 515; Minnesota, 462; North Dakota, 421; and Nebraska, 339 (Board of Governors of the Federal Reserve System 1959).

#### IV

The collapse of US commercial banking from 1929 to 1933, which saw more than 10,000 banks disappear, is a different story. In it every state lost banks. It is also the well-known story of the Great Depression: collapsing prices and production, rising unemployment, runs on banks, and economic policy mismanagement. That story need not be repeated here. More relevant for present purposes are the lessons learned by policy makers from the debacle, and the measures they implemented to avoid another one. Those measures led to a very different outcome from the widespread bank failures of the 1920s and the Depression. For the next half century there would be very few bank failures, and the number of US banks would hardly change at all. But branch banking gradually became more popular. As indicated in Table 1, by the mid 1980s the number of banks with branches nearly equaled the number of unit banks, and by 2018 more than 80 percent of US banks branched. In contrast, from 1921 to 1933, only 2 to 4 percent of US commercial banks had branches.

The main lesson policy makers in early months of President Franklin Roosevelt's New Deal gleaned from the banking debacle of the Great Depression was that there was too much competition in US banking. That competition in their view had induced banks to take excessive risks. Shortly after the bank holiday of 1933 that temporarily shut down all US banks, Congress enacted the Banking Act of 1933, sometimes called the Glass–Steagall Act after its congressional sponsors. The Act called for a formal separation of commercial and investment banking, a measure favored by Senator Carter Glass. The reasoning was that large commercial banks entering the investment banking business before the Depression had generated too much competition, forcing down financial returns and inducing too much risk taking. Investment banks, of course, were relieved to have Congress eliminate the competition they were getting from large commercial banks.

Another provision of the 1933 Act insured most deposits in commercial banks via a newly created Federal Deposit Insurance Corporation. Both large banks and national leaders such as Senator Carter Glass and President Roosevelt disliked deposit insurance. They saw that most bank failures had been those of small, unit banks, and argued that relaxation of restrictions on branch banking would make for safer banking. But the numerous small unit bankers and their champion in Congress, Representative Henry Steagall, instead called for deposit insurance as a means of staving off the competition branch banking would pose for them. In a compromise, Glass supported deposit insurance in return for Steagall's support for the separation of commercial and investment banking. Restrictions on branch banking remained in place.



Still a third provision of the Act banned the payment of interest on demand deposits and enabled the Federal Reserve to regulate rates paid on time and savings deposits. The reasoning was that price competition for deposits drove down bank earnings and thereby induced greater risk taking. Large commercial banks, of course, liked this restriction on competition because it reduced their costs (Calomiris and White 1994).

Effectively, the measures introduced by the Banking Act of 1933 turned US commercial banking into something like a cartel that regulated price competition and controlled entry. It did so to such a degree that the country had about 14,000 banks for the next half century. State banking authorities went along with the changes at the federal level because it protected state as well as national banks from 'excessive' competition. But as that half century wore on, the New Deal regulatory regime, seemingly a success in preventing bank failures, began to unravel. Established banks found the regime too constraining of their ambitions. Potential entrants chafed at the hurdles they faced in trying to establish new banks. More tellingly, while competition within commercial banking was closely regulated and even suppressed, competition from outside the industry eroded commercial banks' market share of total finance. Thus the solutions to one set of problems in the 1930s became the problems the commercial banking industry would face in the 1960s and 1970s.

## V

Stability in commercial bank numbers from the mid 1930s to the mid 1980s masks some significant changes in those decades. Branch banking increased to the point where nearly half of the 14,000 banks had branches, up from 4 percent in the 1933 (Table 1). Unit banking, however, was still the rule in many states.

There were other changes as well. In the 1950s, large money-center banks began to experience a funding crisis after they transformed their balance sheets from holding mostly government debt, a legacy of World War II, to a more normal mix of private loans and investments. US non-financial corporations, traditional clients of money-center banks, could and did operate throughout the country and increasingly, throughout the world. In contrast, the large banks they dealt with often were confined by laws and regulations to one county, one city or even one office. There was a growing imbalance between the size and financial needs of US non-financial corporations and the size of US banks, constrained in their growth by banking laws and regulation. Indeed, the deep and innovative money and capital markets of the USA are rooted historically in an overregulated and fragmented commercial banking system that forced corporations to look for non-bank sources of finance.

To better meet the needs of their corporate clients in postwar America, large banks, especially those of New York City, made a series of innovative moves. They merged to expand their deposit bases in order to be able to make larger corporate or wholesale loans. In 1955, National City Bank merged with First National to form First National City (later Citibank), and Chase National merged with Bank of Manhattan, emerging



as Chase Manhattan. In 1960, J. P. Morgan merged with Guaranty Trust to form Morgan Guaranty.

The 1960s brought more funding innovations. The so-called federal funds market in which banks with surplus reserves traded them overnight to banks with reserve deficits had been dormant since the 1930s because most banks held excess reserves. Banks revived it in the 1950s. They also increased levels of 'repo' funding in which a bank would sell securities to a buyer with an agreement to repurchase them later, often a day later.

But the most important innovation of the 1960s was the large (\$100,000 or more), negotiable certificate of deposit, or CD, the key aspect of which was negotiability. That made CDs into a popular money-market instrument, attracting the surplus funds of non-financial corporations that might otherwise have been invested in Treasury bills. The innovative CD marked a transformation in banking, as it freed banks from the limits of traditional deposit funding. A fully 'loaned-up' bank wanting to make a new loan could fund it simply by selling a large CD in the money market. For banks' balance sheets, liability management became as important as asset management.

These domestic US banking innovations went along with a large-scale international expansion of money-center banks. By expanding overseas, they could better serve their American corporate clients which were also expanding internationally, they could pick up new foreign clients, they could access new sources of funding such as Eurodollars, and they could do so in a much less regulated banking climate than existed at home in the USA (Sylla 2002).

Even at home, regulations were loosening. President John F. Kennedy in 1961 appointed James Saxon to be Comptroller of the Currency, the regulator of national banks, as a part of his agenda to get the country moving again – growing faster economically – by fostering banking growth and consolidation. Saxon controversially ruffled many industry and regulatory feathers by routinely approving bank mergers and branch expansions, and by encouraging banks to enter new lines of business such as selling insurance and underwriting municipal securities issues. Saxon won some battles and lost others when opposed by unit banks, independent insurance agents, and investment banks, all of which wanted to protect their turf (Rose 2019). But his term in office, 1961–6, marked the start of a four-decade alliance of top national leaders and leading bankers that eventually would undo most of the New Deal's bank regulatory reforms. In the late 1960s, banks found a way to enter new lines of business by forming one-bank holding companies with the bank itself as the main subsidiary of the holding company. The holding company could then enter financial businesses off limits to the bank itself. And holding companies could fund their ventures by selling commercial paper, providing a new market source of funding for banks.

The Great Inflation of the late 1960s and 1970s – the US price level rose at a compound rate of 6 percent per year from 1966 to 1981 – did much to continue the impetus for weakening the old bank regulatory system by exposing its flaws. Persistent inflation led to persistent rises in market interest rates, while the interest

rates banks could pay their depositors were held down by regulation. The 1970s brought the innovation of money-market mutual funds that sold shares to investors and invested the money in short-term instruments such as Treasury bills and CDs. As the market yields of these instruments rose substantially above the rates banks could pay, depositors pulled their money from banks and bought higher-yielding money-market fund shares in what was termed 'disintermediation'. They also pulled their money from savings and loan associations (S&Ls), long the main source of US residential mortgage finance. The S&Ls found themselves in the unenviable position of having to pay more for funding than they earned on their mortgage portfolios. By 1990, most of them would fail, with their facilities often taken over by commercial banks in deals brokered by the Federal Deposit Insurance Corporation (FDIC).

Merrill Lynch, the largest US securities broker-dealer and an investment banker, further advanced bank disintermediation by introducing cash-management accounts that combined brokerage accounts with a money-market fund on which a customer could write checks. Since Merrill Lynch had offices throughout the USA, it thus gave the country a form of nationwide banking commercial banks on their own could not offer. Other broker-dealers followed Merrill's lead. More deposits left commercial banks, causing a steep drop in their market share of total American finance. Just after World War II in 1946, commercial banks held 57 percent of US financial assets. By 1979, their share fell to 38 percent (Rose 2019). That set the stage for the sweeping banking deregulations of the 1980s and 1990s.

## VI

The 1980s ushered in a series of crises for US commercial banks. First came the less-developed-country (LDC) debt crisis in 1982, which saddled large, money-center banks with defaulted loans previously made to less-developed nations. The loans had been funded by so-called petrodollars, dollars deposited with large, international banks by oil-producing nations after the OPEC-led increases in oil prices during the 1970s. In the early 1980s, energy deregulation in the USA combined with the disinflation to reduce oil, natural gas and other commodity prices so much that the borrowing nations could not service their debts.

Within the USA, itself a major producer of oil and gas, there was a similar effect of falling prices. Loans to domestic energy producers also defaulted, leading to the failure, among others, of Continental Illinois Bank, the seventh largest US bank, in 1984. The US government essentially bailed out the bank by taking it over because it was deemed too big to fail; its successor bank eventually was privatized and later folded into Bank of America.

Further domestic bank failures by the mid 1980s made feasible takeovers by strong out-of-state banks of troubled banks in a number of states. A prime example was the acquisition by Chemical Bank, a New York money-center bank, of troubled Texas Commerce Bank in 1987. The practice was soon formalized, and also limited, by regional banking compacts that authorized interstate banking within a region, but

excluded banks not within the region. The intent was to keep out the large, money-center banks, particularly those, such as Chemical, of New York City. In truth, the threat from the New York banks was not so great because several of them followed up their LDC debt problems with new ones related to bad domestic real estate lending during the mid and late 1980s. Well-managed Chemical Bank was an exception.

By the 1990s, widespread banking failures demanded regulatory change, as they had done in the 1930s. The need for change combined with a waning ability of small banks and non-bank protectors of financial turf such as insurance agents and securities broker-dealers to allow large commercial banks at last to achieve their deregulatory goals. In 1994, Congress allowed banks to operate nationwide with the Riegle–Neal Interstate Banking and Branching Efficiency Act. And in 1999, the Gramm–Leach–Bliley Financial Services Modernization Act essentially repealed Glass–Steagall by allowing commercial bank holding companies to combine with investment banks, securities broker-dealers, insurance companies and mortgage bankers. Institutions that did so were variously termed megabanks, universal banks and financial department stores.

Financial deregulation prompted large bank mergers, takeovers and other forms of consolidation. To follow a previous example, in 1991 Chemical Bank merged with Manufacturers Hanover, both of New York, combining the sixth and ninth largest US banks into its second largest. In 1996, Chemical took over Chase Manhattan, keeping the Chase Manhattan name and forming the largest US financial institution. In 2000, Chase Manhattan acquired JPMorgan and renamed itself JPMorgan Chase, solidifying its lead in total assets. Then in 2004, JPMorgan Chase acquired a Midwest superregional, Bank One, to further advance its scale and coverage of the USA. Prime motives for the consolidations were cost reductions and greater efficiency; in practice that often meant eliminating duplicative jobs.

Thus by the early twenty-first century four megabanks emerged: Citigroup, JPMorgan Chase, Bank of America and Wells Fargo. Deregulation came with some limits. A large bank was to have no more than 10 percent of US deposits or more than 30 percent of deposits in any state. As a group, the megabanks soon approached those limits. And, as [Table 1](#) shows, between 1984 and 2008, the USA lost more than half of its independent banks, while at the same time nearly doubling the number of branches.

Initially, deregulation and financial consolidation seemed to work. Megabanks along with US banks in general were relatively unscathed by the collapse during 2000–2 of a speculative boom on Wall Street that came to be called the dot com bubble. An associated short but sharp recession in 2001, however, led the Federal Reserve to reduce its policy rate from above 6 percent in 2000 to 1 percent by 2003. Cheap money triggered speculation in residential real estate and a housing boom that in retrospect featured extremely lax mortgage lending standards. The four megabanks along with large investment banks – Merrill Lynch, Goldman Sachs, Morgan Stanley, Lehman and Bear Stearns – fueled the housing bubble by

securitizing large amounts of mortgage debt including growing amounts of risky sub-prime loans. They sold mortgage-backed securities to investors around the world who were reaching for yield in a low-interest environment. More ominously, they retained large amounts on and off their balance sheets, seemingly coining profits by earning the spread between cheap money-market rates and higher mortgage yields. Excessive risk-taking was justified by the mantra that house prices never go down, so the collateral would be there even if borrowers on mortgages defaulted.

The housing bubble began to collapse in 2007 when defaults on recently made mortgage loans rose. House prices did go down. Mortgage-backed securities collapsed in value, threatening the solvency of the highly levered banks. Money-market lending, particularly overnight repos, froze up, triggering the global financial crisis of 2008–9, about which much has been written (Tooze 2018 offers an excellent, comprehensive global account).

In the crisis, some large banks failed. Others on the brink of failure had policymaker-arranged marriages or shotgun weddings with stronger institutions. Still others had to be bailed out by the US Treasury and the Federal Reserve. Two leading investment banks that survived, Goldman Sachs and Morgan Stanley, did so in part by taking out commercial banking charters that gave them access to the Fed discount window as well as the opportunity to develop the stable deposit base possessed by commercial banks. Bear Stearns failed and was merged into JPMorgan Chase, which at the height of the crisis also acquired a large and failing mortgage banker, Washington Mutual. Also at the height of the crisis, at the behest of US financial authorities Bank of America purchased Merrill Lynch. Lehman failed outright, its remnants scattered among a number of financial institutions.

## VII

Somewhat paradoxically, several of the largest US banks deemed too big to fail going into the crisis emerged from it even bigger than before. Congress concluded that they along with the entire US financial system had to be made safer by enhanced regulation. That was the goal of the Dodd–Frank Wall Street Reform and Consumer Protection Act of 2010. The Act was so complex and so vague in many ways that even now, a decade later, legislators, regulators, lawyers, banks and even the Supreme Court are still trying to determine what its hundreds of pages of provisions mean.

The complexity of Dodd–Frank, including its increases in bank reporting requirements, raised regulatory compliance costs, probably more, relatively, for small banks than large ones. That is one reason the USA lost a third of its remaining independent banks in the decade 2008–18 (Table 1). Not all of the disappearances, however, were small banks. Two large southeastern banks, BB&T and Sun Trust, merged in 2019 to form the renamed Truist Bank, which became the sixth largest US bank with nearly half a trillion dollars in total assets. Banks' urge to merge was slowed by the crisis, but it is still there. More mergers may be expected.

From 1933 to 2008, stable or declining bank numbers were accompanied by increases in total branches and offices. That changed in the last decade, as both branches and offices declined in number. Ubiquitous ATMs and the spread of online banking have reduced the need for customers to visit a bank branch to make deposits or withdraw cash, making branch elimination an efficient method of reducing costs. Closing branches is another trend that will continue.

For all its complexity, Dodd–Frank appears to have fostered safer banking in two traditional ways. Its demands for more bank liquidity do address the age-old incentive of banks to run down liquidity in the interest of making more profitable loans. And its demands for more bank capital do address the age-old incentive of banks to hold minimal capital in order to magnify returns on equity. Placing limits on those potentially perverse incentives always was, and will continue to be, among the leading rationales for regulating banks.

## VIII

The USA today, like most other developed nations, has a high degree of concentration in its commercial banking system. It is largely a result of a jump in concentration during the past quarter century, although concentration also increased in the late 1920s and 1930s when most large banks survived while thousands of smaller ones failed. A century ago when US banks reached their all-time peak in numbers, the five largest banks had less than 10 percent of bank assets; in recent years that increased to nearly 50 percent. The top 25 banks a century ago had less than 20 percent of assets; now they have about 70 percent (Fohlin and Jaremski 2020).

Increased banking concentration has its advocates and its skeptics. Before the 2007–9 global financial crisis, advocates argued that it made banking more competitive and efficient than its previous, highly fragmented and overly regulated system that protected banks from competition. It did so by increasing large-bank economies of scale and scope. After the crisis, concentration was defended for similar reasons, and bolstered by the Dodd–Frank reforms that made megabanks safer via enhanced regulatory scrutiny, as well as by periodic stress tests to monitor liquidity and capital levels.

Advocates also argue that worries about concentration reducing competition and raising costs to consumers are overwrought because oligopolistic banks do compete, and not just with each other and smaller banks. They also compete with shadow banks, meaning credit intermediaries existing outside the formal banking system. These include mutual funds, finance companies, hedge funds and private equity funds (Sargen 2020). Of late, moreover, they have to compete with fintech companies in such traditional banking areas as payments and lending.

In a twist from the usual lauding of competition, some defenders of concentrated banking claim that less competition can be a good thing if it promotes financial stability and avoids the economic costs of financial crises. Canada, with fewer banking

crises and bank failures than the neighboring USA, can be cited as an example (Kobrak and Martin 2018).

Skeptics of increased concentration point to the megabanks' role in the build-up to the 2007–9 crisis, as well as to conflicts of interest inherent in combining a wide range of financial functions in one organization. These critiques suggest that complex megabanks are difficult to manage, as the crisis seemingly demonstrated, and thus they may add to instead of ameliorating financial instability (Kaufman 2016). Some US political leaders go so far as to suggest breaking up the megabanks and reinstating the 1930s Glass–Steagall separation of commercial and investment banking. Barring another crisis such as 2007–9, that seems unlikely. But such a crisis cannot be ruled out, as banks in recent years appear to have gone into securitizing of loans into collateralized loan obligations with the same gusto that led to risky mortgage securitizations before 2008.

As a final point, a student of American and European banking history may notice an interesting transatlantic reversal of positions over the past century. At its start around 1920, the USA had its tens of thousands of independent unit banks fragmented across state boundaries, while Europe's leading nations had more concentrated systems featuring a relatively small number of large banks with extensive branch systems. This article illustrates how the USA over the course of the century became more like Europe then.

But it also seems that Europe now has become more like the USA then. The EU member states, each with its own state banking system and its banks protected to some extent from cross-border competition of banks from other European states, resembles the USA a century ago. The USA then had elements of unity across the union: the dollar common currency, a central bank with branches across the union, and a correspondent banking system to knit at least the leading banks in the fragmented state-centered banking systems together. Both US labor and capital were mobile across states, meeting one of the key criteria for an optimal currency area. By the 1930s, if not before, risk-sharing mechanisms for fiscal transfers to states and regions experiencing economic difficulties were also in place, meeting another criterion of an optimal currency area. Europe now has a mostly euro common currency, a central bank in the ECB with its branches of former national central banks, and in the EU's Single European Payments Area, perhaps something of a modern equivalent of the old US correspondent banking system to facilitate cross-border banking transactions. But the EU has more language barriers than the USA, which might limit labor mobility across member states. And Europe's sovereign debt crisis starting in 2009 was met more with resentment and calls for austerity than with fiscal transfers to share risks.

The USA spent much of the past century trying to form a more efficient banking union across its states, and eventually it succeeded. Today American consumers, businesses and the US economy benefit from the efficiencies of a single integrated and consolidated banking market across the union. The EU has been struggling to achieve such a banking union for some time, thus far without a great deal of success. One has to wonder how long it will take. On that, the US example is far from encouraging.

Submitted: 22 July 2020

Revised version submitted: 30 July 2020

Accepted: 10 September 2020

First published online: 3 November 2020

## References

- BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM (1959). *All-Bank Statistics: United States, 1896–1955*. Washington, DC: Board of Governors.
- CALOMIRIS, C. M. and HABER, S. H. (2014). *Fragile by Design: The Political Origins of Banking Crises and Scarce Credit*. Princeton, NJ: Princeton University Press.
- CALOMIRIS, C. M. and WHITE, E. N. (1994). The origins of Federal Deposit Insurance. In C. Goldin and G. D. Libecap (eds.), *The Regulated Economy: A Historical Approach to Political Economy*. Chicago: University of Chicago Press.
- CASSIS, Y. (2002). Before the storm: European banks in the 1950s. In S. Battilossi and Y. Cassis (eds.), *European Banks and the American Challenge: Competition and Cooperation in International Banking under Bretton Woods*. Oxford: Oxford University Press.
- FOHLIN, C. and JAREMSKI, M. (2020). US banking concentration, 1820–2019. *Economic Letters*, **190** (May), pp. 1–3.
- HAMMOND, B. (1957). *Banks and Politics in America, from the Revolution to the Civil War*. Princeton, NJ: Princeton University Press.
- Historical Statistics of the United States, Millennial Edition* (2006). Ed. S. B. Carter *et al.*, vol. 3. Cambridge: Cambridge University Press.
- JAMES, J. A. (1978). *Money and Capital Markets in Postbellum America*. Princeton, NJ: Princeton University Press.
- KAUFMAN, H. (2016). *Tectonic Shifts in Financial Markets*. Basingstoke: Palgrave Macmillan.
- KOBRAK, C. and MARTIN, J. (2018). *From Wall Street to Bay Street: The Origins and Evolution of American and Canadian Finance*. Toronto: University of Toronto Press.
- LEWIS, W. A. (1949). *Economic Survey, 1919–1939*. London: G. Allen and Unwin.
- LU, Q. (2017). *From Partisan Banking to Open Access: The Emergence of Free Banking in Early Nineteenth Century Massachusetts*. Basingstoke: Palgrave Macmillan.
- MICHIE, R. (2003). Banks and securities markets, 1870–1914. In D. J. Forsythe and D. Verdier (eds.), *The Origins of National Financial Systems: Alexander Gerschenkron Revisited*. London: Routledge.
- MITCHELL, B. R. (1998). *International Historical Statistics*, 4th edn. New York: Stockton Press.
- ROBERTSON, R. M. (1995). *The Comptroller and Bank Supervision: A Historical Appraisal*. Washington, DC: Office of Comptroller of the Currency.
- ROSE, M. H. (2019). *Market Rules: Bankers, Presidents, and the Origins of the Great Recession*. Philadelphia: University of Pennsylvania Press.
- SARGEN, N. (2020). *JPMorgan's Fall and Revival: How the Wave of Consolidation Changed America's Premier Bank*. New York: Palgrave Macmillan.
- SYLLA, R. (2002). United States banks and Europe: strategies and attitudes. In S. Battilossi and Y. Cassis (eds.), *European Banks and the American Challenge: Competition and Cooperation in International Banking under Bretton Woods*. Oxford: Oxford University Press.
- SYLLA, R. (2006). Schumpeter redux: a review of Raghuram G. Rajan and Luigi Zingales's *Saving Capitalism from the Capitalists*. *Journal of Economic Literature*, **45** (June), 393–406.
- SYLLA, R. (2011). Wall Street transitions, 1880–1920: from national to world financial centre. In L. Quennouëlle-Corre and Y. Cassis (eds.), *Financial Centres and International Capital Flows in the Nineteenth and Twentieth Centuries*. Oxford: Oxford University Press.
- TOOZE, A. (2018). *Crashed: How a Decade of Financial Crises Changed the World*. New York: Viking.
- WHITE, E. N. (1983). *The Regulation and Reform of the American Banking System, 1900–1929*. Princeton, NJ: Princeton University Press.