The Economic History of Banking

Richard S. Grossman

June 10, 2011
The Economic History of Banking

Richard S. Grossman

June 10, 2011

Among the most important functions of a financial system is intermediation: channeling the aggregate savings of an economy toward productive use.¹ In the absence of intermediators, individuals, firms, and governments in need of finance and savers with surplus funds seeking investment opportunities would have to find each other and negotiate detailed contracts. These contracts would have to specify whether the funds are be loaned or to purchase a share of the enterprise. If loaned, the contract must stipulate the term and interest rate of the loan, as well as the amount and quality of collateral to be pledged as security. If the funds purchase an ownership share, the contract must stipulate the fraction of the profits and seats on the board of directors the new investor will be entitled to and, if the enterprise fails, how much liability will the shareholder bear. Without financial intermediators, negotiating such arrangements would be so costly and time-consuming that many worthwhile projects would go unfunded. The consequences for economic development would be severe.

The importance of finance for long term economic growth has been debated for years. Bagehot (1873), Schumpeter (1934), Gurley and Shaw (1955), Goldsmith (1969), Hicks (1969), McKinnon (1973), and Miller (1998) assert that finance plays a crucial role in fostering industrialization and long-term economic growth; Robinson (1952), Lucas (1988), and many development economists view it as being less essential. The weight of modern empirical analysis supports a finance-growth nexus, although the precise nature of the relationship and how financial systems and economies “co-evolve” is still not well understood (Levine (2005)).

Of the two major components of the financial system, securities markets and financial institutions, economic theory suggests that securities markets are more efficient as intermediators. Those who need funds can offer their wares (i.e., securities) for sale in the financial marketplace while savers can shop for the financial product with the attributes (e.g., maturity, risk, return, and collateral) that suit their preferences, and buy and sell those products as their personal circumstances change. If securities markets are active, transactions costs are low, and all parties have good information, both buyers and sellers will get the best price possible. Market prices will continuously adjust to reflect changes in supply and demand for securities which, in turn, reflect new information about underlying investments. If circumstances

¹ Financial systems provide other services as well: facilitating trade; monitoring managers and exerting corporate control; assisting in the trading, hedging, diversification, and pooling of risk; and allocating resources. Levine (1997: 691).
warrant, adjustments can take place through the market: firms and other demanders of funds can purchase their own debt or equity if the funds are no longer needed; savers can sell securities at any time in order to gain access to their funds.

Despite the many theoretical advantages of securities markets, highlighted in particular by Rajan and Zingales (2003), the conditions necessary to secure those advantages—an active market with many buyers and sellers, low transactions costs, and ample information—have been rare historically. Even with today’s active and sophisticated securities markets, the transactions and information costs for small firms (e.g., securing a bond rating, complying with regulations imposed by regulators and exchanges) and firms in less developed countries may render access to securities markets impractical. Historically, the obstacles were much higher: active securities markets were scarce, information flows were poor by contemporary standards, and high transactions costs limited market access to only the largest and most credit-worthy borrowers, typically governments. By contrast, far less infrastructure is required for a bank—or an individual banker—to intermediate between depositors and borrowers.

1. Banking prerequisites

Although incorporated commercial banking is only about 225 years old, it represents the culmination of centuries, even millennia, of financial and legal developments (Grossman (2010: 28ff.)). Among the oldest of these was the evolution of money, both as a medium of exchange and as a unit of account, and of money-changing. Other important ancient innovations include bookkeeping and mechanisms for making transfers among customers without the need for the exchange of cash, as well as oral and written contracts. A medieval development which combined elements of written contracts and cashless transfers was the negotiable bill of exchange. The demand for the bills of exchange was bolstered by the advent of medieval trade fairs, where the need for relatively complex credit transactions that spanned both time and geography was great.

An outgrowth of these developments was the ability to create credit, an essential function of commercial banks. Credit creation consists of issuing notes or creating deposits (or other forms of credit) beyond what the issuer has on deposit. Money changing (i.e., exchanging one
currency for another), a common activity of early bankers, does not create credit. Similarly, if banks merely provide depositors with safekeeping services, they do not create credit. Finally, if banks take in a quantity of money on deposit which they lend out, they do not create credit, since credit outstanding equals the total amount of deposits held. Put succinctly by Schumpeter (1934: 73), “It is always a question, not of transforming purchasing power which already exists in someone’s possession, but of the creation of purchasing power out of nothing…”

London’s goldsmith banks, which were active in the mid-seventeenth century, provide a good example of credit-creating banking. Suppose that gold is the only widely accepted money. Although the very wealthy might be able to afford sufficiently thick walls or armed guards to protect their treasure, those of more modest means might be prepared to pay a reputable third party with good security to keep their savings for them. Goldsmiths, since they dealt in large quantities of precious metals, would have had a safe or other means to secure their own gold and would have made ideal depositories. Thus, individuals could store their gold with a goldsmith for safekeeping, and in exchange receive a receipt indicating exactly how much gold was on deposit. Initially, depositors would have had to return to the goldsmith every time they wanted to make a transaction: the goldsmith paid out as much of the depositor’s gold as requested and altered the receipt accordingly. With the passage of time, instead of making frequent trips to the goldsmith in order to make withdrawals, receipts were passed from hand to hand and functioned as currency. The receipts retained value since they represented a quantity of gold that could be withdrawn at any time.

Because a substantial portion of receipts remained in circulation, goldsmiths eventually came to realize that they could issue more receipts than they had gold on deposit, and could lend these excess receipts at interest. These loans constituted the beginning of credit creation or, in the words of Schumpeter, creating “purchasing power out of nothing.” Credit creating loans (i.e., loans over and beyond the amount of gold reserves held) need not be receipts, but could take the form of deposits, as long as those deposits are somehow spendable, either via written order to a third party or via a transfer on the books of a banker.

2. Exchange banks, state banks, and private banks
The ability to transfer money without the use of cash was further developed by exchange banks, such as Venice’s Banco della Piazza di Rialto and Amsterdam’s Wisselbank. These banks, which were established in sixteenth and seventeenth centuries, were municipally owned and provided a mechanism for local merchants to make and receive payments from each other and from outside sources. An important consequence of these banks was to standardize the monetary unit on that used by the exchange bank. Because many different types of coins (often clipped or otherwise debased) circulated, it made sense for exchange banks to standardize all transactions on a single currency unit.

The exchange banks provided inspiration—and sometimes a model—for the subsequent establishment of state banks. In addition to their roles as the government’s banker, state banks were often established in order to bring order to a chaotic monetary situation. The state banks of Austria-Hungary (established 1816), Belgium (1850), Denmark (1818), and Norway (1816) were granted monopolies on the right to issue bank notes upon their establishment or soon thereafter, frequently because a predecessor institution had collapsed because it had over-issued currency. Although originally these institutions had very little in the way of public responsibility beyond stabilizing the monetary situation and lending to the government, during the course of the nineteenth century they evolved into the central banks that we know today, operating as bankers’ banks, lenders of last resort, and, sometimes, banking supervisors.

Another important predecessor institution was the private bank. Unlike state banks, which came into existence with a birth certificate in the form of a charter granted by a legislature or sovereign, the establishment of private banks required no special legislation. Instead, these institutions were set up and governed by the terms of their partnership agreements and were bound by no specific laws beyond those that applied to individuals. Because they operated without specific government sanction, they were not usually required to keep especially detailed records or submit reports to government authorities. Thus, our knowledge of specific private banking firms is limited to those firms with surviving records.

---

2 Johan Palmstruch’s original plan for the Swedish Riksbank envisioned it containing two departments, one of which was to be called the Wexelbank, a direct translation from the Dutch Wisselbank.
3 See, for example, Joslin (1954), Pressnell (1956), and Temin and Voth (2008).
Because private banks were unchartered, they were not restricted to specific types of business and hence could engage in a variety of banking and non-banking activities. These included money changing, pawn-brokering, and, most importantly, providing domestic and international trade credit. Frequently, private banks started out as purely commercial firms and, because of the need to extend credit for trade finance, evolved into institutions where banking functions soon overshadowed non-financial business. In discussing private banking in the Rhineland during the first half of the nineteenth century, Tilly (1966: 47) writes:

[T]here is great uncertainty as to when a man was a banker and when he was not. Many persons provided banking services—either as a sideline or on a small scale—long before they called themselves bankers… At the same time, the title ‘banker’ was sometimes given to businessman whose resources were employed mainly in non-banking lines.

3. Commercial banking

The first incorporated commercial banks emerged in the late eighteenth and early nineteenth centuries. Early examples include Australia’s Bank of New South Wales (established 1817), Belgium’s Société Générale (1822), Canada’s Bank of Montreal (1820), Denmark’s Fyens Discontokasse (1846), Germany’s A. Schaaffhausen’scher Bankverein (1848), Norway’s Christiana Bank og Kreditkasse (1848), and the U.S.’s Bank of North America (1782). Although commercial banks conducted many of the same operations as private banks, there were important differences.

Private banks were frequently family-run concerns, with resources limited to those of the extended family. Although some of these financial enterprises grew quite large (e.g., the Rothschilds), for the most part they were small by comparison with the commercial banks that would follow. Another constraint on the size of private banks was that they were not able to organize as corporations. Corporations are, in theory, immortal: when shareholders die, their ownership in the ongoing concern can be passed to others or sold, however the firm itself should be unaffected. The lifespan of a partnership is less certain: the death of a partner may necessitate the liquidation of the firm. The uncertainty over the lifespan of private banks may have presented an obstacle to their growth. Additionally, corporations are likely to grow larger than partnerships because their shares can be more easily bought and sold, allowing them to achieve a more diffuse—and larger—ownership base. The effect on firm size was magnified by the fact
that during the course of the nineteenth century limited liability became an option for incorporated banks. Limited-liability shares, which would not subject the shareholder to unlimited calls in case of failure, would have been more marketable to a wider audience than those of an unlimited liability partnership.

From its inception during the late eighteenth and early nineteenth centuries through the Great Depression, commercial banking in the industrialized countries followed a characteristic life-cycle pattern. Although the specifics differ considerably from country to country, the pattern is robust, and consists of an early phase of rapid growth, followed by slower growth and an eventual peak, followed by a more or less rapid decline. Throughout this lifecycle, the structure of the banking system is shaped by four key types of events: crises, bailouts, merger movements, and regulatory reforms (Grossman (2010)).

The earliest stages of the banking lifecycle are characterized by rapid expansion, during which both the number of banks and aggregate banking assets increase. The speed and strength of the initial growth depends positively upon the demand for banking services, derived from the current and expected future financing needs of commerce and industry, and negatively upon the severity of the regulatory constraints placed upon banks and competition from other intermediators. Commercial banking growth may accelerate taking over the functions of private banks and other preexisting intermediators (Lindgren (2002)). This early growth will be subject to a variety of economic and regulatory shocks, and will not in any sense be smooth or constant.

The upward trend continues until a turning point is reached. The turning points in the number and assets of banks may not occur at the same time in any given country, nor need they take place at the same time in any particular country’s economic development. The turning point in numbers usually occurs in close proximity to a banking crisis or a merger movement. Both of these events reduce the numbers of banks; however, crises typically lead to a reduction in aggregate banking assets, while mergers may be accompanied by a continued increase in banking assets.

Crises generate two reactions, both of which can have profound effects on banking structure. In the short-run, government or private actors may attempt to rescue some or all of the

---

A literature on product lifecycles posits a similar pattern in competitive industries. See Gort and Klepper (1982).
failed banks. In the slightly longer run, governments may impose more stringent regulations in an attempt to bolster banking stability. Mergers, if they result in a substantial increase in banking concentration, may also encourage regulators to intervene in order to reduce the likelihood of a monopolized banking sector. Government efforts to rein in banks via regulation—whether due to crises or heightened merger activity—will be met with increasingly inventive attempts to circumvent such regulation. These may lead to additional regulations and, in turn, even more ingenious efforts at avoiding such regulation. Kane (1977) aptly refers to these cycles of moves and counter-moves as the “regulatory dialectic.”

This lifecycle pattern was repeated over and over in the industrialized world from the beginnings of the nineteenth century through the Great Depression (Grossman (2010: 17-26). The first commercial banks in Australia and Canada, for example, were established in 1817. By 1837, the banking populations of Australia and Canada were 11 and 20, respectively. In Australia, the number of banks peaked at 31 in 1890; in Canada the number peaked at 45 in 1875. By the onset of the Great Depression, Australia had about a dozen banks; Canada had less than half a dozen. The banking populations of Denmark, Finland, Norway, and the United States all exhibited similar patterns, reaching peaks during the period between the end of World War I and the onset of the Great Depression.

As noted above, this lifecycle is neither regular, nor consistent. Although the early phase of banking growth usually accompanied rapid industrialization, in Britain, the first country to experience an industrial revolution, commercial banking growth lagged due to the influence of the Bank of England. The bank had been granted a charter by Parliament in 1694 in return for a substantial loan to the government; subsequent chartering acts prevented the establishment of banks with more than six partners. Thus, for the next century, English banking was the sole preserve of the Bank of England and a large number of small banking partnerships. By 1826, the instability of a banking system composed primarily of small banks, complaints from industry about the inadequate resources of those banks, and the diminishing importance of the Bank of England as a source of government revenue led to a repeal of the monopoly (Broz and Grossman (2004)). The Netherlands, an early commercial, if not industrial power was also slow to develop commercial banking. In the Dutch case, the slowness was due to the existence of a variety of alternative intermediators: kassiers (cash-keepers, a sort of financial agent), an active money market, and private merchant houses which provided ample intermediation.
As noted above, the lifecycle pattern was by no means smooth or regular and was frequently interrupted. Among the most important—and frequent—interruptions, were banking crises. Banking crises were ubiquitous during the nineteenth and early twentieth century: Grossman (2010: 297-313)catalogues more than 60 banking crises among the nations of western Europe, the United States, Canada, Japan, and Australia during the century following 1825. The majority of these resulted from “boom-bust” economic cycles. Irving Fisher (1932; 1933), one of the first modern economists to take an analytical approach to financial crises, argues that crises result from the cyclical nature of real economic activity. Economic expansion leads to a growth in the number and size of bank loans, as well as to an increase in the relative indebtedness of non-bank firms. As economic expansion proceeds and bankers seek more profitable investments, less worthwhile projects receive funding. Fisher (1932: 43) laments the excessive buildup of debt during cyclical upswings: “If only the (upward) movement would stop at equilibrium!” When the expansion ends, these marginal firms are the first to be unable to meet their debt service obligations. This leads to loan defaults and declines in the prices of outputs and securities. The debt-deflation spiral feeds upon itself: loan defaults lead to bank failures, exacerbating the macroeconomic downturn already underway.

4. The Great Depression and its aftermath

No matter where banking systems were within their evolutionary lifecycle, the Great Depression and World War II stopped the process dead in its tracks. The 1930s saw banking crises in Austria, Belgium, France, Germany, Italy, Norway, Switzerland, and the United States, as well as a number of other countries; in many countries, the crisis was of unprecedented severity. Given that banking crises also afflicted many other countries during the early 1920s, it is safe to say few industrialized countries emerged from the interwar period with their banking systems completely intact. In response to financial devastation and wartime needs, governments enacted strict, often anti-competitive rules and regulations aimed at stabilizing the banking system and directing credit towards favored sectors. These constraints—a sort of financial

---

5 For other catalogues of banking and financial crises, both historical and modern, see Kindleberger (1978), Bordo et al. (2001), Lindgren et al. (1996), and Caprio and Klingebiel (2003).
6 He was not the first to notice this phenomenon. Writing 150 years ago, Evans (1859 [1969]) argued that the pattern had already been well established during the previous 60 years.
“lockdown”--combined with low, stable interest rates and robust economic growth, led to the longest period of banking stability the industrialized world has ever known.

In the United States, where state and federal law already prescribed detailed banking regulations, including branching restrictions, capital requirements, and reserve requirements, the Banking (Glass-Steagall) Act of 1933 further curtailed the scope of permitted bank activities by prohibiting commercial and investment banking functions from being carried out by the same institution. The Banking Act also established the Federal Deposit Insurance Corporation (FDIC), which insured commercial bank deposits up to $2500 (soon doubled to $5000; today $250,000). This marked the first time that an explicit deposit insurance program had been enacted on a national level anywhere. The American deposit insurance legislation was anti-competitive in that it both limited entry into banking and made entry more discretionary, since federal banking authorities were required to consider capital adequacy, earnings prospects, managerial character, and community need before allowing prospective banks to commence operations with deposit insurance (Spong (2000: 24)).

In countries where there had been no specific banking law prior to World War I and enacted banking codes during the interwar period, the change in the regulatory environment was especially dramatic. In Belgium, Depression-era laws not only separated commercial banking from investment banking, but also forbade bankers from taking part in the operation of a non-financial company (Allen et al. (1938)). The laws also established capital and reporting requirements and established a banking commission with wide ranging powers over all deposit-taking institutions, including private banks. The commission was empowered to limit entry into banking, to set interest rates and liquidity ratios, and to establish reporting and auditing requirements. Switzerland also enacted its first-ever banking code during this period, which established detailed capital and liquidity regulations, specified certain rules of corporate governance, and held managers liable for intentional damage to creditors and shareholders. The law also established a banking commission, which was empowered to regulate entry, approve articles of association, certify auditing associations, and set required financial ratios, as well as giving the Swiss National Bank the authority to veto the conditions of certain bank undertakings.

The crisis also led to greater government ownership in—and direct control of--the banking sector. In Germany, virtually all important commercial banks received an infusion of
capital from the state in the aftermath of the 1931 banking crisis. It was estimated that in 1932
the state held 50 percent of the capital of the large commercial banks (Allen et al. 1938). The
1931 crisis in Italy led to substantial government participation in the banking industry through
the creation of the Istituto Mobiliari Italiano (IMI) in 1931 and the Istituto Ricostruzione
Industriale (IRI) in 1933 which took over the industrial participations, loan portfolios, and, in
some cases, the lending operations of distressed Italian banks (Gerbi (1956)).

Even in countries where banking systems remained essentially intact, important
Depression-era changes in the regulatory environment were enacted. In Canada, the Bank of
Canada was established, not because there existed any pressing economic need for a central
bank, but because of the government’s desire to be seen to be doing something in response to the
Great Depression (Bordo and Redish (1987: 417)). Sweden’s banking system also weathered
the 1930s without a serious crisis, despite the collapse of the Kreuger industrial empire in 1932,
although it enacted legislation in 1933 making it more difficult for banks to own industrial
shares. Sweden also established A/B Industrikredit which, like IMI and IRI in Italy, was to take
industrial shares off the hands of banks.

World War II led to the introduction of additional legislative and administrative measures
to control the uses and cost of bank credit, as governments harnessed the banking industry to
finance the war effort. Many of these reforms persisted long after the conflict ended. In
February 1942 the Commonwealth Bank of Australia was given powers to set maximum interest
rates as a wartime measure. At the war’s end, these powers were continued by the Banking Act,
1945 and lasted until the 1980s (Australia. Parliament. (1991: 20-21)). In Germany, interest
rates on deposits had been set during the war by the government and the bankers’ association;
following the war the job was taken over by West German state governments, which maintained
control until interest rate deregulation in 1967 (Irmler (1956)). Sweden introduced credit
controls at the beginning of World War II and interest rate controls at the war’s end, both of
which remained in force for nearly a half century (Jonung (1993)).

5. The era of deregulation—and instability—begins

The close of the 1960s marked the beginning of the end of tightly controlled domestic
banking and financial markets. Financial liberalization can be traced, at least in part, to the
exigencies of inflation-induced high and volatile market interest rates. Inflation contributed to, and was exacerbated by, the demise of the Bretton Woods regime in 1971-73 (Bordo and Eichengreen (1993: 465-466)). Important components of liberalization included the gradual elimination of interest rate controls on bank lending and deposit rates, the growth of market-oriented mechanisms for the allocation of capital, such as the development of money markets and auction techniques for the issue of government debt securities, and the elimination of direct controls on bank lending. Many countries made substantial strides towards liberalizing other aspects of bank regulation as well, including blurring the lines between securities firms, savings banks, commercial banks, and insurance companies, and relaxing rules that restricted branching and foreign entry into domestic banking markets.

Deregulation has been a double-edged sword for commercial banking. On the one hand, it has introduced dynamism into the financial industry that had been lacking since the Great Depression. Depositors are no longer at the mercy of government-imposed interest rate ceilings and banks have been mostly freed of the constraints that forced them to channel the accumulated savings of the economy toward government-favor objectives. Banks and other financial institutions are freer to develop new financial products to better suit the needs of their customers.

This freedom has come at a cost, however. Starting with the 1974 failures of Germany’s Bankhaus I.D. Herstatt and the U.S.’s Franklin National Bank the world banking system, which had been crisis-free since World War II, began to experience crises at a rate similar to the pre-World War I era (Bordo et al. (2001)). Notable episodes have included the US S&L crisis during late 1980s, the Nordic banking crisis during late 1980s and early 1990s, and Japan’s “lost decade” of the 1990s. In each of these cases the classic boom-bust cycle was aggravated by ill-timed deregulation and/or perverse government incentives, often combined with lax supervision.

The most severe post-World War II crisis to date has been the sub-prime meltdown, which originated in the United States in 2008, but soon spread worldwide. Many of the same macroeconomic and regulatory forces that contributed the other modern crises were at work in the run-up to the crisis. Expansionary fiscal and monetary policy, combined with unhelpful government intervention in the mortgage market and weak regulatory oversight fuelled a speculative boom in U.S. real estate. The boom was exacerbated by the growth--and misuse--new and complex financial instruments which, thanks to increasing globalization and
improvements in communications, facilitated worldwide participation in the boom—and the subsequent bust.

References


