

Banking Fragility, United States, 1790–2009

Charles W. Calomiris

Columbia University, New York, NY, USA

OUTLINE

| | | | |
|---|----|---|----|
| Introduction | 15 | Summary of US Historical Experience | 25 |
| Banking Fragility in Theory and in Historical Reality | 15 | A Worldwide Tale of Two Banking Eras: 1875–1913 and 1978–2009 | 25 |
| US Banking Crises: 1790–1933 | 18 | Glossary | 27 |
| Central Banking and Bank Instability in US History | 21 | References | 27 |

INTRODUCTION

What are banking crises and when and why do they occur? Banking crises properly defined consist either of moments of widespread and panicked depositor withdrawal or waves of costly bank failures. A historical analysis of banking crises reveals that they are not random events, cannot be seen as the inevitable result of human nature or the liquidity transforming structure of bank balance sheets, and do not typically accompany business cycles or monetary policy errors. Rather, risk-inviting microeconomic rules of the banking game, or shocks to public finance that are relevant for banks, have always been the key contributors to a propensity for banking distress, whether in the form of banking panics or waves of bank failures.

Some risk-inviting rules take the form of visible subsidies for risk taking, as in the historical state-level deposit insurance systems in the United States, Argentina's government guarantees for mortgages in the 1880s, and the worldwide expansion of government-sponsored deposit insurance and other bank safety net programs in the last half of the twentieth century. Other risk-inviting rules historically have involved government-imposed structural constraints on banks, which include entry restrictions like 'unit banking' laws that constrain competition, prevent diversification of risk, and limit the ability of banks to deal with shocks. Another destabilizing rule of the banking game is the absence of a properly structured

central bank to act as a lender of last resort to reduce liquidity risk without spurring moral hazard.

Governments have also been sources of shocks to banks, which take the form of sudden changes in government policies toward banks, fiscal shocks that affect banks to the extent that banks are exposed to fiscal risks of governments, or monetary policy changes (e.g., the monetary collapse of 1929–33 in the United States).

BANKING FRAGILITY IN THEORY AND IN HISTORICAL REALITY

Pundits, policy makers, and economists often remind us that banking crises are nothing new, an observation sometimes used to argue that crises are inherent to the business cycle or perhaps to human nature itself. [Kindleberger \(1978\)](#) and [Minsky \(1975\)](#) were prominent and powerful advocates of the view that banking crises are part and parcel of the business cycle and result from the propensities of market participants for irrational reactions and myopic foresight.

Some banking theorists, starting with [Diamond and Dybvig \(1983\)](#), have argued in a somewhat parallel vein that the structure of bank balance sheets is itself to blame for the existence of panics; in their canonical model, banks structure themselves to provide liquidity services to the market and thus create large liquidity risks for themselves and also make themselves vulnerable to

self-fulfilling market concerns about the adequacy of bank liquidity. The theoretical modeling of banking theorists, like the myopia theory of Minsky, is meant to explain prevalent banking fragility – a phenomenon that any blogger can now trace at least as far back as AD 33, when Tacitus (Book VI) tells us that the Roman Empire suffered a major banking panic, which was quelled by a large 3-year interest-free loan to the banking system by Emperor Tiberius.

There is, however, at least one obvious thing wrong with arguments that purport to show how myopia, business cycles, and inherent bank liquidity transformation can explain the historical constancy of banking crises: in fact, the propensity for banking crises has not been at all constant over time or across countries. Banking crises, properly defined, have not regularly and consistently accompanied business cycles. In fact, banking crises have been much more frequent in some eras than in others and much more frequent in some countries than in others. The differences across countries and across time are dramatic.

This is, in fact, a central lesson of the history of banking crises, which economic historians emphasize: banking crises are not an historical constant, and therefore, the propensity for banking crises cannot possibly be said to be the result of factors that have been constant over time and across countries for hundreds of years, including business cycles, human nature, or the liquidity transformation inherent in bank balance sheets. Instead, it is the policy environment of a country that accounts for its propensity for banking crises, by which is meant both the structure of rules that gives rise to banks and regulates them, and the nature and extent of government policy shocks (regulatory, fiscal, and monetary) that buffet the financial system. Government policies have been at the heart of banking fragility throughout world history.

The first aspect, the structure of the rules governing the banking system within a country – defined by the rules that govern entry into banking as well as the location, powers, and the operations of each of the banks, including government subsidies or special rights granted to favored participants in the banking system and the incentive consequences of those subsidies and rights – encompasses a wide range of phenomena. In times and places where politically determined microeconomic rules of the banking game have encouraged risky practices or prevented effective private measures to limit banking crisis risk, the risk of banking crises is high; conversely, the absence of such adverse political rules of the game has resulted in stable banking systems.

Consider the differences between the history of banking instability in the United States and Canada. Canada's banking system has seen only one suspension of deposit convertibility, in 1837, which was quite early in Canadian banking history (30 years before the Dominion

Act that made Canada autonomous from Great Britain), and Canada has never suffered a large wave of costly bank failures, even during times of severe economic downturns (including the 1837 crisis and the Great Depression of the 1930s). In contrast, since 1830, the United States has suffered major banking crises – defined either as times of widespread sudden contraction of bank deposits or times of costly and widespread bank failures (and using a negative 1% of gross domestic product (GDP) threshold for the aggregate net worth of failed banks as a critical value to measure severity) – over and over again, including in 1837, 1839, 1857, 1861, 1873, 1884, 1890, 1893, 1896, 1907, the 1920s, 1930–33, the S and L Crisis of the 1980s, and the recent subprime crisis. The US propensity for banking crises was uniquely high in the nineteenth and twentieth centuries (Bordo, 1985). Only one of the many US banking crises (1861) is traceable to a government fiscal shock; as we shall see, the others reflect unusual aspects of the US regulatory regime – most importantly, the US decision to adopt a banking system that took the form of thousands of small 'unit' (single-office) banks.

The second government policy-related source of banking instability – fiscal or regulatory policy shocks – has also been important for explaining the timing and location of banking crises. Consider the first banking crisis for which any detailed information can be found, the Roman banking crisis of AD 33. Tacitus traces the crisis to government enforcement of a long-neglected usury law and subsequent additional rules that tried to limit lending to mortgages (as a way to boost the land market). These rules caused a sudden and massive contraction of the supply of credit, which resulted in a collapse of land values and the suspension of convertibility of banks, which were finally brought to an end by government intervention (a large, interest-free loan to the banks from the Roman treasury).

Government fiscal problems have been among the most important source of shocks that have provoked banking crises, and these shocks have been more frequent in some countries than in others. The Napoleonic War and its consequences for Britain's finances led the British government to require the suspension of convertibility by the Bank of England and other British banks, which permitted Britain to then use the Bank of England (freed from the constraint of convertibility of its currency into gold) as a source of inflationary finance for its war debt. The collapse of the value of US government debts in December 1861 (which had been sold to the major banks of New York, Boston, and Philadelphia a few weeks earlier) led to a suspension of convertibility of bank deposits, which was remedied by a government bailout of banks through the creation of legal tender currency: by redenominating deposits into the new debased currency unit and banks' liabilities were reduced in

value, thereby elevating the value of their net worth (Hammond, 1970). For Latin American banks, the fiscal risks of autocratic governments during the nineteenth and twentieth centuries, and the expropriations of banking resources that often accompanied fiscal problems, have been a dominant source of banking instability.

Monetary policy shocks have often set the stage for banking crises (see Bordo, 2007; Bordo and Wheelock, 2007, 2009), but monetary policy errors typically have not been enough to generate banking crises. With few exceptions (Norway's housing bubble and bust in 1899 is generally attributed to loose Norwegian monetary policy in the 1890s and the US monetary contraction of 1929–33, which was the primary source of the macroeconomic collapse in the United States during the Depression), destabilizing monetary policy generally has not been a primary contributor to banking crises (defined either as suspensions of convertibility or waves of costly bank failures). At the same time, global shifts in the supply and demand for money – related, for example, to discoveries of gold, or financial shocks in one or more major countries that increased the demand for gold – transmitted deflationary pressures or illiquidity problems from one country to another, especially when many countries operated on a common global gold monetary standard, as during much of the nineteenth and early twentieth centuries.

Another lesson from the history of banking instability is that the ability to derive useful lessons about banking crises depends on defining banking crises properly. Banking crises must be distinguished from the broader category of 'financial crises,' which include a variety of other phenomena (i.e., sovereign debt defaults, exchange rate depreciations, land price declines, and stock market declines), which may or may not be associated with banking distress. And banking problems, including the failure of one or two banks, or significant declines in deposits and loans for the system as a whole, which typically accompanied recession-related increases in bank losses, do not equate to a banking crisis.

That is not to say that contractions in lending were irrelevant; contractions in bank activity, manifested in significant loan and deposit shrinkage in the wake of loan losses, even when not associated with a banking crisis, typically impose substantial costs for the economy. For a cross-country analysis of those costs, see Bordo and Eichengreen (2003), who study the effects of historical banking distress (broadly defined) on business cycle severity. But these contractions historically typically did not result in true crises. What makes a banking problem into a banking crisis?

When defining banking crises, it is important to distinguish between two different aspects of banking crises – waves of bank insolvency (episodes in which bank losses result in many failed banks with large loan losses) and banking panics (moments in which the

banking system as a whole suffers from sudden, large withdrawals of deposits). Furthermore, it is important to define banking panics carefully. Panics are usefully defined as moments of confusion about the incidence of losses in banks that are sufficiently severe as to create systemic withdrawal pressure on a large number of banks that is sufficient to elicit collective action by the banks and/or the government (e.g., joint issuance of liabilities, like clearing house certificates or the undertaking of joint action, such as suspension of convertibility or other explicit attempts to coordinate behavior to alleviate the effects of panic). That definition creates an objective standard that distinguishes true panics from less severe moments of stress that are not truly systemic in scope.

Sometimes the two aspects of the definition of a banking crisis (panic, as opposed to many costly bank failures) have coincided (as during some episodes in the United States in the 1830s and the 1930s, and many recent episodes worldwide), but often, they have not coincided. The 1920s in the United States witnessed a severe wave of bank failures related to declines in agricultural income but not panics. The US experience between the Civil War and World War I witnessed several banking panics but no significant waves of bank failures. It is useful to recognize panic and insolvency as separate aspects of banking crises because, as we shall see, these different aspects reflect separate causal influences.

A final lesson from the history of banking crises concerns the circumstances that tend to produce effective learning in the policy responses to crises. Reforms have often followed in the wake of crises, but the record of reform is uneven. One successful historical reform occurred in Britain in the mid-nineteenth century, when the Bank of England changed the terms under which it gave financial institutions access to its discount window, and thereby reduced moral-hazard problems that had contributed to excessive risk taking. Britain had suffered a banking crisis about once a decade, but after 1866, it did not experience a banking crisis again until 2008. That example shows that meaningful structural reforms that reduce incentives to take on excessive risk can stabilize banking systems. On the other hand, policy responses sometimes make matters worse: the failure to recharter a central bank in the United States in the 1830s reflected, in part, mistaken views about the second bank of the United States (SBUS) during the crises of 1819 and 1825, and the bank regulatory changes in the United States in 1933 reflected political deal making rather than a proper response to the root causes of banking instability.

The remainder of this essay traces the peculiar history of US banking crises in detail. The banking history of the United States saw a uniquely high and persistent degree of bank instability, which is traceable to its unique

historical circumstances. We close with a broader overview of the global history of bank insolvency crises.

US BANKING CRISES: 1790–1933

As many scholars have recognized for many years, US banks were unusually vulnerable to systemic banking crises compared to banks in other countries (for reviews, see Bordo, 1985; Calomiris, 2000). The United States was especially unique in its vulnerability to panics in the years between the Civil War and World War I. Sprague (1910) and Calomiris and Gorton (1991) identify six episodes of particularly severe banking panics in the United States between the Civil War and World War I. Prior to the Civil War, there were other nationwide banking crises in 1819, 1837, 1839, and 1857, in which both system-wide panic and many bank failures occurred. In the 1920s, the United States experienced waves of bank failures in agricultural states, which have always been identified with fundamental shocks to banks and which did not give rise to national or regional panics.

The key difference between the United States and other countries historically lies in the structure of the US banking system. The US system was mainly based on unit banking – geographically isolated single-office banks. Unit banking meant that banks could not enjoy diversification economies by pooling loan risks from different regions. Unit banking, which resulted in thousands, and sometimes tens of thousands of banks, also limited the ability of banks to pursue collective action by pooling resources during periods of adverse shocks. A system with tens of thousands of geographically distant banks simply could not organize appropriate collective action to stem financial crises. Bank clearing houses or informal alliances among banks to make markets in each other's deposits during crises did exist in the United States, but these required that members adhere to guidelines and that they be able to monitor one another to ensure compliance. Not only did geography get in the way of such coordination, the sheer number of banks made collective action difficult. The benefits of one bank choosing to monitor another are shared, but the monitoring and enforcement costs are borne privately; coalitions with 30 members seemed able to motivate individual banks to bear the private costs of monitoring on behalf of the coalition, but coalitions of hundreds or thousands of banks unsurprisingly were not able to structure effective monitoring and enforcement.

Other countries did not choose the fragmented US approach to banking, and no other country experienced the US pattern of periodic banking panics prior to World War I or the waves of agricultural bank failures that gripped the United States in the 1920s.

During the antebellum period, a few southern branch banking states, and three Midwestern states that formed mutual guarantee systems with small numbers of members, were able to implement successful, stabilizing coalitions of banks at the state level for purposes of mutual protection during banking crises (Calomiris, 1989, 1990, 2000; Calomiris and Schweikart, 1991). But these were short-lived and isolated exceptions; nationwide branching was not permitted, and most states prohibited or limited within-state branching. US banks were numerous (e.g., numbering more than 29 000 in 1920), undiversified, insulated from competition, and geographically isolated from one another and thus were unable to diversify adequately or to coordinate their response to panics (US banks did establish clearing houses in cities, which facilitated local responses to panics beginning in the 1850s, as emphasized by Timberlake (1984) and Gorton (1985)).

The fragmented structure of US banking explains why the United States uniquely suffered banking panics in the years between the Civil War and World War I despite the fact that the vast majority of banks were healthy throughout this period and were consistently able to avoid failure. The absence of a lender of last resort in the United States was also an important contributor to bank instability, but the structure of unit banking appears to be the more important influence; Canada, which operated on a branching basis, avoided panics, although it did not charter a central bank until 1935. Empirical studies show that the major US banking panics of 1873, 1884, 1890, 1893, 1896, and 1907 were moments of heightened asymmetric information about bank risk but not times when bank failure risk was large for the country as a whole (Bruner and Carr, 2007; Calomiris and Gorton, 1991).

Calomiris and Gorton (1991) show that banking panics were uniquely predictable events that happened at business cycle peaks. In the pre-World War I period (1875–1913), every quarter in which the liabilities of failed businesses rose by more than 50% (seasonally adjusted) and the stock market fell by more than 8%, a panic happened in the following quarter. This happened five times, and the panic of 1907 was the last of those times. Significant national panics (i.e., events that gave rise to a collective response by the New York Clearing House) never happened otherwise during this period.

Bank failure rates in the years between the Civil War and World War I, even during these panic episodes, were low, and the losses to depositors associated with them were also small. In 1893, the panic with the highest failure rate and highest depositor loss rate, depositor losses were less than 0.1% of GDP. *Expected* depositor losses during the panics also appear to have been small. Sprague (1910, pp. 57–8, 423–424) reports that the discount applied to bankers' cashier checks of New York

City banks at the height of the panic of 1873 did not exceed 3.5%, and with the exception of an initial 10-day period, that discount remained below 1%. A similar pattern was visible in the panic of 1893. A 1% premium would be consistent with depositors in a New York City bank estimating a 10% chance of a bank's failing with a 10% depositor loss if it failed. Clearly, banking panics during this era were traceable to real shocks, but those shocks had small consequences for bank failures in the aggregate, and even at the height of the crisis, those consequences were expected to be small. Historical US panics teach us that even a small expected loss can lead depositors to demand their funds, so that they can sit on the sidelines until the incidence of loss within the banking system has been revealed (usually a process that took a matter of weeks).

Bank failure rates in the 1830s, 1850s, and 1920s were higher than those of the other pre-Depression systemic US banking crisis episodes. The 1830s, in particular, saw a major macroeconomic contraction that caused many banks to fail, which historians trace to large fundamental problems that had their sources in government-induced shocks to the money supply (Rousseau, 2002), unprofitable bank-financed infrastructure investments that went sour (Schweikart, 1987), and international balance of payments shocks (Temin, 1969).

The 1920s agricultural bank failures were also closely linked to fundamental problems, in this case, the collapses of agricultural prices at the end of World War I, which were manifested in local bank failures because of the lack of regional or national loan portfolio diversification (Alston et al., 1994; Calomiris, 1992).

In both the 1830s and the 1920s, some states suffered more than others from waves of bank distress. In the 1830s, states that had an active role in directing the credit of their banks fared particularly badly (Schweikart, 1987). Prior to both the bank failure waves of the 1830s and the 1920s, some states had enacted systems of deposit insurance in which neither entry nor risk taking was effectively constrained. These states experienced far worse banking system failure rates and insolvency severity of failed banks than did other states (Calomiris, 1989, 1990, 1992). Indeed, the basis for the substantial opposition to federal deposit insurance in the 1930s – an opposition that included President Franklin D. Roosevelt, his Treasury Secretary, and the Federal Reserve – was the disastrous experimentation with insurance in several US states during the early twentieth century, which resulted in banking collapses in all the states that adopted insurance and especially severe collapses in states that made deposit insurance compulsory.

In the 1920s, state-chartered banks that participated in deposit insurance fared much worse than either national banks in those states or state-chartered banks in neighboring states. The disastrous experience of those banks

reflected a combination of moral hazard and adverse selection. Moral hazard was reflected in the higher loan-to-asset ratios and lower capital-to-asset ratios of state-chartered banks in insured states. Furthermore, states that passed deposit insurance experienced substantial entry into banking by small operators in rural areas, who apparently overestimated the potential for agricultural prices (temporarily boosted by World War I) to remain high.

In contrast, in the 1920s, states that had enacted laws permitting branch banking tended to outperform unit banking states, both with respect to failure rates and failure severity (Calomiris, 1990, 1992). The evidence of the stabilizing effects of even limited branch banking in the United States (note that branching was not permitted across states and, in many cases, was constrained even when it was allowed within states) helped to produce significant relaxations of branch banking restrictions in many states and a merger wave of banks during the 1920s. From 1921 through 1931, more than 5000 banks were absorbed by acquirers. In 1910, for the United States as a whole, there were 292 branching banks operating 548 branches, with total loans and investments of \$1.3 billion, and in 1920, there were 530 branching banks operating 1281 branches, with total loans and investments of \$6.9 billion; by 1931, there were 723 branching banks operating 3467 branches, with total loans and investments of \$20.7 billion (Calomiris, 2000, p. 57).

The legacy of branch banking restrictions continued to destabilize banks during the Depression of the 1930s. Mitchener (2005) finds that states that prohibited branching had higher rates of bank failure, *ceteris paribus*. Despite these trends and evidence, the stabilizing trend toward bank consolidation and greater structural stability in the United States was derailed by the global macroeconomic policy disaster of the Great Depression and its adverse political consequences for continuing bank consolidation. Most importantly, Congressmen Henry Steagall of Alabama lobbied successfully on behalf of his state's unit bankers for federal deposit insurance, which was embraced by unit bankers as a political tool to prevent competition and continuing pressure for consolidation (Calomiris and White, 1994). Initially deposit insurance was passed as a temporary emergency measure limited to only cover small deposits (effectively a subsidy for small banks, for whom such deposits comprised a large fraction of their liabilities). Despite the opposition of Senator Carter Glass, the Federal Reserve System, the Treasury Department, and President Roosevelt – all of whom were aware of the disastrous consequences of deposit insurance in the states that had experimented with it in the early twentieth century – Steagall managed to succeed in passing deposit insurance, which was soon transformed from a

temporary to a permanent measure, and which now covers virtually all US deposits.

Beginning in the 1880s, there had been 150 attempts to introduce federal deposit insurance legislation in Congress (Calomiris and White, 1994). Opponents understood and espoused the theoretical arguments against deposit insurance that are familiar today – that deposit insurance removes depositors' incentives to monitor and discipline banks, that it frees bankers to take imprudent risks (especially when they have little or no remaining equity at stake and see an advantage in 'resurrection risk taking'), and that the absence of discipline promotes banker incompetence, which leads to unwitting risk taking. Deposit insurance won the day as legislation in 1933 for political, not ideological reasons, and ironically (given Roosevelt's opposition) remains the main surviving legacy of the banking legislation of the New Deal – a stark reminder of the power of crises to change the course of banking regulation.

Deposit insurance, which was very limited in coverage and became effective only in 1934, after the banking crises of 1930–33 had passed, had little role in stabilizing banks during the Depression of 1929–33. Bank failures and losses were high in the early 1930s by historical standards. Recent research on the Depression has investigated the extent to which those failures reflected extremely adverse macroeconomic shocks and their consequences for bank borrowers, as opposed to excessive, panicked responses to those shocks by depositors that may have forced many solvent banks into financial distress. Recent research shows that much if not all of the bank distress of the 1930s resulted from fundamental shocks to bank assets, much like the shocks that had buffeted agricultural banks in the 1920s.

The list of fundamental shocks that weakened banks during the Great Depression is a long and varied one. It includes declines in the value of bank loan portfolios produced by waves of rising default risk in the wake of regional, sectoral, or national macroeconomic shocks to bank borrowers as well as monetary policy-induced declines in the prices of the bonds held by banks.

Friedman and Schwartz (1963) argued that many bank failures resulted from unwarranted 'panic' and that failing banks were in large measure illiquid rather than insolvent. Friedman and Schwartz's emphasis on contagion posited that bank failures mainly reflected a problem of illiquidity rather than insolvency. Illiquid but solvent financial institutions, in their view, failed purely as the result of withdrawal demands by depositors, particularly during sudden moments of panic. In contrast, an insolvent institution fails to repay depositors as the result of fundamental losses in asset value, rather than the suddenness of depositor withdrawals.

Using a narrative approach similar to that of Friedman and Schwartz, but relying on data

disaggregated at the level of Federal Reserve districts, Wicker (1996) argues that it is incorrect to identify the banking crisis of 1930 and the first banking crisis of 1931 as national panics comparable to those of the pre-Fed era. According to Wicker, the proper way to understand the process of banking failure during the Depression is to disaggregate, both by region and by bank, because heterogeneity was very important in determining the incidence of bank failures.

Microeconomic studies of banking distress have provided some useful evidence on the reactions of individual banks to economic distress. White (1984) shows that the failures of banks in 1930 are best explained as a continuation of the agricultural distress of the 1920s and are traceable to fundamental disturbances in agricultural markets. Declines in railroad bonds were also significant in some cases (Meltzer, 2003, p. 346).

Calomiris and Mason (1997) study the Chicago banking panic of June 1932 (a locally isolated phenomenon). They find that the panic resulted in a temporary contraction of deposits that affected both solvent and insolvent banks. Fundamentals, however, determined which banks survived. Apparently, no solvent banks failed during that panic. Banks that failed during the panic were observably weaker *ex ante*, judging from their balance sheet and income statements, and from the default risk premia they paid on their debts. Furthermore, the rate of deposit contraction was not identical across banks; deposits declined more in failing banks than in surviving banks.

Calomiris and Wilson (2004) study the behavior of New York City banks during the interwar period and, in particular, analyze the contraction of their lending during the 1930s. They find that banking distress was an informed market response to observable weaknesses in particular banks, traceable to *ex ante* bank characteristics. It resulted in bank balance sheet contraction, but this varied greatly across banks; banks with higher default risk were disciplined more by the market (i.e., experienced greater deposit withdrawals), which encouraged them to target a low risk of default.

Calomiris and Mason (2003a) construct a survival duration model of Fed member banks throughout the country from 1929 to 1933. This model combines aggregate data at the national, state, and county level with bank-specific data on balance sheets and income statements to identify the key contributors to bank failure risk and to gauge the relative importance of fundamentals and panics as explanations of bank failure. Calomiris and Mason find that a fundamental-based model can explain most of the failure experience of banks in the United States prior to 1933. They identify a significant, but small, national panic effect around September of 1931 and some isolated regional effects that may have been panics, but prior to 1933, banking panics were not very important contributors to bank failures compared to fundamentals.

The fact that a consistent model based on fundamentals can explain the vast majority of US bank failures prior to 1933 has interesting implications. First, it indicates that the influence of banking panics as an independent source of shock to the economy was not important early in the Depression. Only in 1933, at the trough of the Depression, did failure risk become importantly delinked from local, regional, and national economic conditions and from fundamentals relating to individual bank structure and performance. Second, the timing of this observed rise in risk unrelated to indicators of credit risk is itself interesting. In late 1932 and early 1933, currency risk became increasingly important; depositors had reason to fear that President Roosevelt would leave the gold standard, which gave them a special reason to want to convert their deposits into (high-valued) dollars before devaluation of the dollar (Wigmore, 1987).

As part of their bank-level analysis of survival duration, Calomiris and Mason (2003a) also consider whether, outside the windows of ‘panics’ identified by Friedman and Schwartz, the occurrence of bank failures in close proximity to a bank affects the probability of survival of the bank, after taking into account the various fundamental determinants of failure. Calomiris and Mason consider this measure of ‘contagious failure’ an upper bound, since in part it measures unobserved cross-sectional heterogeneity common to banks located in the same area, in addition to true contagion. They find small, but statistically significant, effects associated with this measure. The omission of this variable from the analysis raises forecasted survival duration by an average of 0.2%. They also consider other regional dummy variables associated with Wicker’s (1996) instances of identified regional panics and again find effects on bank failure risk that are small in national importance.

The large number of bank failures in the United States during the Great Depression, a phenomenon that was largely confined to small banks, primarily reflected the combination of extremely large fundamental macroeconomic shocks and the vulnerable nature of the country’s unit banking system. Panic was not a significant contributor to banking distress on a nationwide basis until near the trough of the Depression, at the end of 1932. For these reasons, the Great Depression bank failure experience has more in common with the bank failures of the 1920s than the panics of the pre-World War I era.

Central Banking and Bank Instability in US History

Part of the microeconomic rules of the game in any banking system relate to the operations of the central bank, which include its policies for purchasing assets or lending against them, how it funds itself, and the

extent to which and the ways in which it competes with other banks.

The first central bank in the United States, the Bank of the United States (BUS), founded in 1791 and chartered for 20 years, was the only nationally chartered institution in the country and the only one to operate in more than one state. It operated as a for-profit banking enterprise (in which the government owned one-fifth of its \$10 million in initial capital stock), made loans, issued notes, and accepted deposits. It was not conceived as a tool for regulating other banks or acting as a lender of last resort to the financial system. The BUS’s most important role in the economy was as a lender to the government and as a fiscal agent for the government, managing the financial flows relating to taxes and debts.

The first panic in US history, in 1792, occurred just as the BUS was gearing up its operations. As Sylla et al. (2009) show, the early experience of the United States in dealing with the panic of 1792 illustrates that central bankers can be sources of banking system risk as well as mitigators of those risks. The nascent BUS actually fueled the panic through an overexpansion of credit in its first months of operation. But the Treasury Secretary, Alexander Hamilton, acted as an ad hoc central banker, inventing and applying ‘Bagehot’s (1873) rule’ of lending freely on good collateral at a high rate 8 decades before that rule would be penned. Thus, although the BUS itself was not a source of stability during the panic, Secretary Hamilton and the Treasury acted as an effective ad hoc lender of last resort.

Sylla et al. (2009) argue that Hamilton’s success in undoing the negative effects of the BUS’s destabilizing actions had significance beyond its immediate consequences for the financial system; the intervention avoided a political backlash against the Hamiltonian financial system, of which the BUS was a part. Such a backlash in response to failed financial policy innovations was more than a hypothetical possibility in 1792, given the experience of France and Britain decades earlier:

Earlier in the eighteenth century, John Law had attempted to modernize France’s financial system, but his efforts backfired when he failed to prevent the collapse of the Mississippi Bubble in 1720. At the same time, across the Channel, the collapse of the related South Sea Bubble also led to financial crisis. The British financial system, however, was more developed than that of France, as Britain had begun the modernization process in 1688, whereas France did not do so until 1715. A wounded but robust British financial system survived the shock, although legislation passed during the crisis stunted the development of Britain’s corporate sector for a century. (p. 63)

This discussion illustrates two broader points: (1) the actions of central banks are not always stabilizing for the banking system and (2) central banks’ privileges can be enacted and also withdrawn. Indeed, the charter of the BUS was allowed to lapse, largely as the result of

Jeffersonian objections to the concentration of financial power in a national bank.

Problems in managing fiscal affairs during the War of 1812, along with a desire to reestablish specie convertibility of state bank notes after the suspension of convertibility that had attended the War, led to the establishment of the SBUS, in which the government subscribed for one-fifth of its \$35 million in capital stock. The SBUS was charged with assisting the government in its financial affairs, reestablishing specie convertibility of other banks' notes, and operating a general banking business. Like its predecessor, it operated as the only nationally chartered bank until its charter's renewal was blocked by President Jackson in 1832. In 1836, the SBUS was granted a new charter by the state of Pennsylvania and operated as a state bank after that date.

As [Temin \(1969, p. 46\)](#) points out, like its predecessor in 1792, the SBUS became overextended almost immediately after it was chartered. Its western and southern branches did not coordinate their lending with the eastern branches. Notes issued by the SBUS branches in the periphery to fund loans were presented for payment in the east, where the SBUS branches accepted them at par, thereby encouraging further note issuance and lending by its branches in the west and south. The lack of discipline in 1817–18 extended to other banks as well. The Treasury asked the SBUS to delay the collection of balances owed to it by state banks in 1817 and 1818, which removed the SBUS as a source of interregional discipline over other banks' issuance and promoted increased leverage in the banking system. When as an act of self-preservation the SBUS finally cracked down on its branches in the west and south and on other banks, by demanding that they support their own note issues and pay their outstanding debts, a contraction of credit resulted, which according to [Catterall \(1902\)](#) 'precipitated the panic.' [Gouge \(1833\)](#) famously quipped that "The Bank was saved and the people were ruined." Public hostility toward the SBUS because of its role in the causing the panic of 1819 never disappeared ([Temin, 1969, p. 48](#)).

The hostility toward the SBUS was further fueled by perceptions of its behavior during the financial crisis in 1825–26, when once again it acted to limit its own credit and disciplined the state banks by demanding that they redeem their obligations. The public had expected the SBUS to prevent a financial contraction in 1825–26. According to [Hilt \(2009\)](#), the SBUS was a stabilizing force during the 1825–26 financial crisis. He argues that, despite widespread financial failures during the crisis, "there was no generalized banking crisis in the United States ... in part because of the Second Bank of the United States worked assiduously with its New York branch to provide credit to the banking community there." [Hilt \(2009, Footnote 36\)](#) also points to evidence

that Nicholas Biddle viewed this as an important part of the mission of the SBUS. Nevertheless, such lending was limited by the BUS's need to protect itself during the 1825–26 contraction.

The public hostility toward the SBUS that resulted from its failure to prevent financial crises was largely misplaced. The government itself had encouraged the excessive expansion of credit in the periphery in 1817–18 and had asked the SBUS to accommodate it. Furthermore, according to [Hilt \(2009\)](#), the SBUS had, in fact, been successful in preventing the financial collapse of 1825–26 from turning into a bona fide banking crisis as the result of the assistance it provided New York banks. The SBUS's decisions to contract in 1819 and 1825–26 were necessary to its own preservation; it was, after all, a privately owned bank and therefore responsible for its own survival and profitability. Most importantly, the upheavals of 1819 and 1825, like that of 1792, illustrated the limitations of the powers of the BUS and the SBUS. The BUS and SBUS lacked the full-fledged powers of a central bank to deal with crises. Indeed, some financial historians ([Temin, 1969, p. 45](#)) have questioned whether the BUS or SBUS qualify to be called central banks. Unlike the Bank of England, the BUS and SBUS did not have the power to issue an unlimited supply of their own bank notes with the implied backing of the sovereign. It is hard to fault the SBUS for failing to use powers that it did not possess.

That is not to say that the SBUS was entirely powerless or unsuccessful in reducing systemic risk in the banking system, as the successful interventions in New York by the SBUS in 1825–26 illustrate. Because of its special position as the only bank operating branches in various regions, the SBUS was large, had wide geographical reach, and played an especially important role in the bankers' acceptance market for financing commerce (intermediation via its various branches the financing of trade flows, as described in [Calomiris, 2000, Chapter 1](#)) and in the market for transporting and redeeming the notes of other issuing banks. In addition to its limited powers to assist banks during crises, it could act, and did act, to stabilize the system and help avoid the risk of panics in two ways: (1) as a source of discipline over other banks' note issuance, it limited the overextension of credit and bank leveraging during booms; and (2) as a unique interregional provider of trade credit, the SBUS reduced seasonal volatility in financial markets related to the planting and harvesting of crops. [Bernstein et al. \(2009\)](#) find evidence in support of increased average risk, and greater seasonality of risk, after the failure to recharter the SBUS. From 1816 to 1836, stock return volatility across the months of September and October (the harvesting season) averaged 2.45%, virtually identical to the 2.43% for the rest of the year. Following the SBUS's demise, from 1837 to 1860, stock return volatility rose

to 6.30% in September and October versus 5.02% during the rest of the year. These volatilities rose even higher during the National Banking Period, when they were 7.30 and 5.80%, respectively.

Despite its stabilizing role in the financial system, the inability of the SBUS to prevent financial crises, along with various other political and ideological battles in which the SBUS and its leader, Nicholas Biddle, became embroiled, ultimately resulted in a fierce battle over the future of the bank and Jackson's eventual veto of its rechartering.

The history of the SBUS illustrates a broader theme in the early history of central banking. Central banks, including the BUS, the SBUS, the Bank of England, and others, were chartered as for-profit companies with special privileges (in the case of the BUS and the SBUS, the special privileges had to do with their unique branching operations and their unique relationship with the government) that also gave rise to the expectation that they would undertake special responsibilities to the public in addition to maximizing their economic value for their stockholders. That dual mandate of profitability and social responsibility implied that central banks not only had to satisfy their stockholders that they were achieving a good return and acting prudently but also that they had to satisfy the public, through elected officials, that they were achieving their social mission. Because achieving social missions (like making markets in risky bank notes at par) tends to be costly, there has often been an inherent conflict between the private profitability of central banks and their public missions.

The US banking system operated without a central bank from 1836 until 1913 when the Federal Reserve System was established. The establishment of the Fed was a direct reaction to the panic of 1907, and the perception that private bankers acting as a coalition (organized to some extent through their local clearing houses and through ad hoc efforts like those undertaken in 1907 by J.P. Morgan) had insufficient ability to preserve systemic stability. In the wake of the 1907 panic, the National Monetary Commission was established, and it issued a voluminous and substantive report in 1910, which formed the factual and theoretical basis on which the Federal Reserve Act of 1913 was constructed, and to this day, the National Monetary Commission report still contains some of the most valuable information about the operations of banks of that era throughout the world.

The Federal Reserve System, like all central banks, was a creature of a political process and compromise that balanced various competing interests, and that compromise evolved over time. The structure of the system (12 regional Reserve Banks and a Board in Washington, with member bank ownership of the Reserve Banks) evolved into a system effectively owned by the taxpayers but still managed by a process of power sharing that gave weight

to local bank and business interests (who control the Reserve Banks' boards and executive appointments), political leaders in Washington (who appoint Federal Reserve Board members and to whom the board reports), and rural interests (who received special favors in the structuring of clearing arrangements and in the use of agriculture-related loans as collateral).

The philosophical foundations of the Fed are rather amorphous, and much of the logic that was embodied in its initial rules has been discredited by monetary economists, notably the 'real bills doctrine' that was supposed to govern its lending operations. Suffice it to say for our current purposes that the Fed obtained broad powers to lend to member banks against good collateral (initially construed as high-quality commercial bills and, later, also government securities) and to engage in open market operations to control the supply of reserve holdings by member banks at the central bank.

Importantly, the Fed's charter and its powers did not envision it as a crisis manager for failing banks or as a bailout agency, and the Fed's role in causing or averting banking crises primarily revolved around the way its policies affected *market* prices and flows, rather than the affairs of particular banks. The Fed was designed to use lending and other actions to regulate the aggregate supply of reserves, money, and credit in a way that would reduce seasonal and cyclical volatility of interest rates and increase the seasonal and cyclical elasticity of reserves and loans (initially, it was expected to accomplish those seasonal and cyclical objectives while remaining on the gold standard, a goal that was temporarily put aside several times and permanently abandoned in 1973).

The record of the Fed as a source of stability for the banking system is mixed. On the one hand, the Fed was sometimes a source of great instability in the system because the policy rules it followed for targeting monetary policy were often ill-conceived. [Friedman and Schwartz \(1963\)](#), and many others since, have shown that Fed monetary policy errors produced the monetary collapse that caused the economic and banking crises of the 1930s, and [Calomiris and Mason \(2003a\)](#) document measurable connections between the deteriorating macroeconomic and local economic environments in which banks operated and the resulting bank failures.

[Wicker \(1966\)](#), [Brunner and Meltzer \(1968\)](#), and [Wheelock \(1991\)](#) trace the Fed's policy errors in the 1930s and at other times to the misuse of interest rates and borrowed reserves as short-term monetary policy instruments. Ideologically, this view was related to the 'real bills doctrine,' the notion that the Fed's role was to accommodate commercial demand ([Meltzer, 2003](#), pp. 273–74). [Wheelock \(1991\)](#), in particular, argues that it was a consistently faulty monetary policy methodology, rather than a lack of leadership at the Fed following

Benjamin Strong's death (which [Friedman and Schwartz, 1963](#) posit to explain Fed failures after 1929), that explains the policy errors that gave rise to the Great Depression. The monetary policy errors that caused the Great Depression show that vesting authority in a central bank can be risky; although the central bank may intend to stabilize the system, it may, in fact, have the opposite effect.

On the other hand, there is substantial evidence ([Bernstein et al. 2010](#); [Miron, 1986](#); [Richardson and Troost, 2006](#)) that the founding of the Fed reduced liquidity risk in the banking system, which in turn reduced the propensity for bank panics. [Miron \(1986\)](#) showed that the founding of the Fed was associated with reduced seasonal variability of interest rates and increased seasonal variability of lending. Miron, however, did not explain how the Fed achieved this result. Why, exactly, did Fed lending practices make the loan supply function more elastic?

[Miron's \(1986\)](#) findings can be explained by a variant of the deposit risk targeting model in [Calomiris and Wilson \(2004\)](#). In that model, the riskiness of deposits is a function of bank asset risk and bank leverage. Because total bank capital and total cash assets in the economy do not vary much over the year, a seasonal increase in bank lending (especially to finance crop harvesting and transport in the fall, which [Davis et al. \(2007\)](#) show was largely driven by the cotton cycle) implies a commensurate increase in bank asset risk and in bank leverage, which unambiguously means an increase in the riskiness of deposits (the actuarially fair default risk premium). This is a source of seasonal variation in the risk of deposit withdrawals, since market discipline makes the risk of withdrawal in the deposit market sensitive to increases in default risk (i.e., some depositors are intolerant of risk and will withdraw when risk increases). A bank that increases its lending, *ceteris paribus*, faces increased deposit withdrawal risk, particularly if an adverse cyclical shock hits during a seasonal lending spike. All six of the major banking panics of the pre-World War I era happened at cyclical peaks; they were clearly responses to adverse economic shocks to banks' balance sheets ([Calomiris and Gorton, 1991](#)). Furthermore, these panics all occurred either during the spring planting season or the fall harvest, at times when lending (and bank liquidity risk) was at a seasonal peak.

From the perspective of this model, the founding of the Fed provided a means of reducing liquidity risk to banks by giving them a source of liquidity to stem deposit withdrawals (making them less vulnerable to withdrawal risk at times when seasonal lending peaks coincided with cyclical downturns). The founding of the Fed thus flattened the bank loan supply function, making loans vary more over the cycle, and interest rates vary less.

[Bernstein et al. \(2010\)](#) provide additional evidence consistent with that interpretation. They compare the

standard deviations of stock returns and short-term interest rates over time in the months of September and October (the 2 months of the year when markets were most vulnerable to a crash because of financial stringency from the harvest season) with the rest of the year before and after the establishment of the Fed. Stock volatility in those two months fell more than 40%, and interest rate volatility more than 70%, after the founding of the Fed. Like the SBUS before it (discussed above), the Fed succeeded in reducing seasonal variations in liquidity. They also show that this result is driven by years in which business cycles peaked. In other words, the main risk that the Fed eliminated was associated with combined cyclical peaks in economic activity and seasonal peaks in lending.

Many commentators have faulted the Federal Reserve for failing to prevent bank failures during the Great Depression with more aggressive discount window lending. While it is certainly true that expansionary monetary policy, particularly in 1929–31, could have made an enormous difference in preventing bank distress (through its effects on macroeconomic fundamentals), that is not the same as saying that more generous terms at the discount window (holding constant the overall monetary policy stance) would have made much of a difference. Discount window lending only helps preserve banks that are suffering from illiquidity, which was not the primary problem underlying large depositor withdrawals.

Indeed, in 1932, President Hoover created the Reconstruction Finance Corporation (RFC), to enlarge the potential availability of liquidity, but this additional source of liquidity assistance made no difference in helping borrowing banks avoid failure ([Mason, 2001](#)). As commentators at the time noted, because collateralized RFC and Fed loans were senior to deposits, and because deposit withdrawals from weak banks reflected real concerns about bank insolvency, loans from the Fed and the RFC to banks experiencing withdrawals did not help much, and actually could harm banks, since those senior loans from the Fed and the RFC reduced the amount of high-quality assets available to back deposits, which actually increased the riskiness of deposits and created new incentives for deposit withdrawals. In 1933, however, once the RFC was permitted to purchase banks' preferred stock (which was junior to deposits), RFC assistance to troubled banks was effective in reducing the risk of failure ([Mason, 2001](#)).

Despite the limitations inherent in the ability of collateralized lending to prevent bank failure, there is some evidence that greater Fed assistance to banks early in the Depression could have been helpful in avoiding some bank failures. [Richardson and Troost \(2006\)](#) show that, despite the limited ability of Fed discount window lending to absorb credit risk, Fed provision of liquidity

to member banks mitigated bank failure risk associated with illiquidity somewhat in 1930 and could have played a greater role in stemming illiquidity-induced failures if the Fed had been more willing to relax lending standards to member banks. They study the failure propensities of Mississippi banks. The Federal Reserve Act of 1913 divided Mississippi between the sixth (Atlanta) and eighth (St. Louis) Federal Reserve Districts. The Atlanta Fed championed a more activist role in providing loans to member banks experiencing troubles, while the St. Louis Fed rigidly adhered to the real bills doctrine and eschewed the extension of credit to troubled banks. Mississippi banks in the sixth District failed at lower rates than in the eighth District, particularly during the banking panic in the fall of 1930, suggesting that more aggressive discount window lending reduced failure rates during periods of panic.

Summary of US Historical Experience

The unusually unstable US historical experience of frequent nationwide banking panics and a propensity for unusually severe and widespread waves of bank failures (the 1830s, the 1920s, and the 1930s) reflected a unique feature of the microeconomic structure of US banking – namely, the fragmented banking structure of unit banking – which made it harder to diversify lending risk *ex ante* and coordinate the management of banking system risk *ex post*.

Comparisons across regions and across states within the United States also reveal important cross-sectional differences in banking stability that are similarly traceable to structural features. The presence of branch banking, clearing houses, or other local institutional arrangements for collective action were stabilizing forces, but these stabilizing mechanisms were only permitted on a local or statewide basis. The presence of deposit insurance, which was advocated by unit bankers as a means of protecting them from debt market discipline, resulted in adverse selection in bank entry and moral hazard in bank risk taking, and was a destabilizing force that produced the worst localized bank failure experiences of the 1830s and the 1920s.

Early experiments with limited central banking in the United States resulted in the failure to recharter central banks twice in the early nineteenth century, which reflected, in part, a difficulty in reconciling the financial limitations of a private bank of limited means with the public pressures on that bank to ‘pay for’ its privileges by performing unprofitable services in the public interest. Although some observers accused the SBUS of contributing to financial instability through contractionary policies prior to and during both the panic of 1819 and the financial crisis of 1825–26, those accusations say

more about unrealistic public expectations of the power of the SBUS to prevent systemic problems than they do about the desirability of rechartering the SBUS. Although neither the BUS nor SBUS was equipped to act as lenders of last resort during crises, the SBUS succeeded in reducing systemic financial risk on average and over the seasonal cycle, foreshadowing the stabilizing effect of the Fed after 1913.

After the demise of the SBUS, the United States functioned without a central bank until the founding of the Fed in 1913. The record of the Fed *vis-à-vis* banking crises is mixed. On the one hand, the Fed (like the BUS in its first year of operation) could be a source of substantial risk to the system, resulting from inappropriate policy responses. The mistaken use of borrowed reserves and interest rates as monetary instruments created false impressions in 1929–32 that encouraged monetary contraction, which precipitated the Great Depression, the real effects of which produced massive bank failures in the 1930s. On the other hand, the existence of the discount window substantially reduced systemic liquidity risk, especially the risk that banks would be caught in an illiquid position at times of seasonal peaks in lending that coincided with cyclical peaks in economic activity. Although the ability to employ the discount window to stem bank failures during the Depression was limited – since shocks buffeting banks were primarily related to solvency rather than illiquidity – there is evidence that relatively aggressive discount window lending by the Atlanta Fed during 1930 did help to prevent some bank failures.

In summary, the microeconomic rules of the banking game – the unit banking structure of the industry, the occasional reliance on destabilizing deposit insurance, and the lack of an effective lender of last resort for the pre-World War I era – all contributed to the peculiar historical instability of the US banking system. The key destabilizing elements of the US system – a fragmented industrial structure, the absence of an effective lender of last resort, and the occasional presence of a destabilizing deposit insurance regime – compounded one another. Canada, which avoided chartering a central bank until 1935, managed to avoid banking crises due to the stabilizing role of its branch banking system, despite the absence of a central bank. In the United States, the fragility of the banking structure made the absence of a central bank more harmful than it otherwise would have been; likewise, the absence of an effective central bank magnified the destabilizing effects of unit banking.

A WORLDWIDE TALE OF TWO BANKING ERAS: 1875–1913 AND 1978–2009

Although the United States was unique in its high propensity for panics (reflecting its peculiar banking

structure), and it occasionally experienced high rates of banking loss, other countries sometimes experienced loss rates that exceeded that of the United States. In the pre-World War I period (1875–1913), the highest nationwide banking system loss rate (i.e., the negative net worth of failed banks relative to GDP) for the United States was roughly 0.1%, which was the loss rate for bank failures in the panic of 1893. Other countries generally experienced even lower bank failure rates, but there were a handful of episodes (probably only four) in the world during this period in which the negative net worth of failed banks exceeded 1% of GDP (a minimal severity standard used by Caprio and Klingebiel to gauge banking crises today).

During the pre-World War I era, Argentina in 1890 and Australia in 1893 were the exceptional cases; they each suffered banking system losses of roughly 10% of GDP in the wake of real estate market collapses in those countries. The negative net worth of failed banks in Norway in 1900 was roughly 3% and in Italy in 1893 roughly 1% of GDP, but with the possible exception of Brazil (for which data do not exist to measure losses), there seem to be no other cases in 1875–1913 in which banking losses in a country exceeded 1% of GDP.

By recent standards, this record for the pre-World War I period is one of impressive banking stability, especially considering the high volatility of the macroeconomic environment during that period. In contrast, since 1980, about 140 episodes have been documented in which banking systems experienced losses in excess of 1% of GDP and more than 20 episodes resulted in losses in excess of 10% of GDP, more than half of which resulted in losses in excess of 20% of GDP (these extreme cases include roughly 25–30% of GDP losses in Chile in 1982–82, Mexico in 1994–95, Korea in 1997, and Thailand in 1997, and a greater than 50% loss in Indonesia in 1997).

Loss rates in the pre-World War I period tended to be low because banks structured themselves to limit their risk of loss by maintaining adequate equity-to-assets ratios, sufficiently low asset risk, and adequate liquidity. Market discipline (the potential for depositors fearful of bank default to withdraw their funds) provided incentives for banks to behave prudently (for a theoretical framework, see Calomiris and Kahn, 1991). The picture of small depositors lining up around the block to withdraw funds has received much attention by journalists and banking theorists, but perhaps the more important source of market discipline was the threat of an informed ('silent') run by large depositors (often other banks). Banks maintained relationships with each other through interbank deposits and the clearing of deposits, notes, and bankers' bills. Banks often belonged to clearing houses that set regulations and monitored members' behavior. A bank that lost the trust of its fellow bankers could not long survive.

Recent research attempting to explain the unprecedented systemic bank failures worldwide over the past 3 decades has emphasized the destabilizing effects of bank safety nets. This has been informed by the experience of the US Savings and Loan industry debacle of the 1980s, the banking collapses in Japan and Scandinavia during the 1990s, and similar banking system debacles throughout the world. Empirical studies of this era of unprecedented frequency and severity of banking system losses have concluded uniformly that deposit insurance and other policies that protect banks from market discipline, intended as a cure for instability, have instead become the single greatest source of banking instability (see, e.g., Barth et al., 2006; Boyd et al., 2000; Caprio and Klingebiel, 1996; Demirguc-Kunt and Detragiache, 2000; Demirguc-Kunt et al., 2009).

It is also significant that the two countries that suffered the most severe bank failure episodes of the pre-World War I era – Argentina and Australia – had two things in common: (1) both of them suffered real estate boom and busts that exposed their financial systems to large losses, and (2) prior to these crises, both of them had employed unusually large government subsidies for real estate risk taking. In Argentina, that subsidy took the form of special mortgage guarantees issued by the government, which guaranteed holders of the mortgages repayment. Banks were licensed to originate these guaranteed mortgages, and then resold them as guaranteed liabilities in the London market, where they were traded as Argentine sovereign debts. This is akin to deposit insurance in that it makes the financing cost of the mortgage invariant to its risk, which entails the same moral hazard as deposit insurance: the guarantee makes the profitability of mortgage lending increasing in the riskiness of the mortgage portfolio and thus encourages originators to lend to risky borrowers.

The Australian case was a bit different; financial market policies toward the private sector were not the primary means through which the government promoted the land boom that preceded the bust of 1893. The pre-1890 Australian economic expansion was largely an investment boom in which the government played a direct role in investing in land and financing farmers' investments. Government investments in railroads, telegraphs, irrigation, and farms were financed by government debt floated in the British capital market and by government-owned savings banks and postal savings banks (Butlin, 1961, 1964, 1987; Davis and Gallman, 2001).

The theory behind the problem of destabilizing subsidization of risk taking has been well known for well over a century, and we have already noted that it was the basis for opposition to deposit insurance in the United States in 1933. Deposit insurance was seen by opponents as undesirable special interest legislation designed to

benefit small banks (Calomiris and White, 1994). Roosevelt, Glass, and others acquiesced for practical political reasons, to get other legislation passed, not because they wanted deposit insurance, per se. Bad economics is sometimes good politics. Similarly, Argentine mortgage subsidies were transparently intended to benefit land-owners in the pampas, just as the real estate risk subsidies in Australia, Rome, and Norway were conscious attempts to support constituencies that favored real estate development.

It is worth emphasizing that all of these risk subsidizing government interventions (mortgage guarantees, liability insurance, government lending on land) were intended to overcome market discipline that had been limiting risk taking. Whatever their merits, these interventions served powerful special interests by subsidizing real estate risk, destabilized their country's banking systems, and produced substantial losses. Bank insolvency crises in the pre-World War I era fundamentally were about imprudent government policies.

Research on the banking collapses of the last three decades offers a similar message. Empirical findings uniformly show that the greater the role of government in directing credit or in providing protection to private banks through the government safety net (e.g., deposit insurance), the greater the risk of a banking collapse (Barth et al., 2006; Boyd et al., 2000; Caprio and Klingebiel, 1996; Demirguc-Kunt and Detragiache, 2000; Demirguc-Kunt et al., 2009). Some of this research has identified the political economy of subsidizing risk taking as a core problem, and one that would-be reformers and financial regulators have had a difficult time overcoming. Empirical research on prudential bank regulation emphasizes the inefficacy of government regulations in preventing risk taking (since they are subject to the same political forces that purposely subsidize risk) and the importance of subjecting some bank liabilities to the risk of loss to promote discipline of risk taking as the primary means of reining in excessive risk taking (Barth et al., 2006; Board of Governors, 1999; Calomiris and Powell, 2001; Mishkin, 2001) – in other words, finding a means to use markets to constrain risk taking.

These studies of recent experience echo the conclusions of the studies of historical deposit insurance discussed above (Calomiris, 1990, 1992). The difference is that what used to be the exception – moral hazard and adverse selection resulting from government protection that give rise to excessive risk taking – has become the rule.

This evidence stands in sharp contrast to the theoretical approaches of Minsky (1975), Kindleberger (1978), and Diamond and Dybvig (1983) for explaining bank fragility. Rather than seeing market behavior, human nature, and the market-determined structure of bank balance sheets as the root cause of banking crises, this

new literature argues that the solution to banking crises lies in empowering markets to rein in the risk taking that is otherwise subsidized by the government.

Glossary

- Banking panic** A time when many depositors suddenly try to convert large amounts of their demand deposits into cash; this typically occurred either in reaction to perceived weaknesses in banks' risky asset portfolios or, less frequently, out of fear that an exchange rate devaluation would lower the real value of bank deposits.
- Liquidity transformation of banking** When banks hold large amounts of illiquid assets (nontraded loans) with maturities greater than the maturities of their debts (which typically consist of liquid, demandable deposits), the bank is said to engage in a process of liquidity production, which is sometimes called the liquidity transformation of banking.
- Real bills doctrine** The discredited view that banking systems will remain sound, and also be optimally responsive to legitimate credit needs, if they confine their lending to the discounting (lending against) real bills, which are effectively loans used to finance the movement of traded goods.
- Resurrection risk taking** When banks have large, hidden losses that threaten their solvency, bankers may opt to increase the riskiness of their investments, hoping for a positive outcome that will restore their solvency.
- Suspension of deposit convertibility** A bank or banks may suspend the rights of depositors to convert their demandable deposits into cash on demand; this was sometimes done en masse, usually with legislative or judicial approval, in response to sudden large withdrawal pressures.
- Unit banking** A banking system where chartered commercial banks are only permitted to operate from a single office, in contrast to a branch banking system, where banks have multiple offices in different locations.

References

- Alston, L.J., Grove, W.A., Wheelock, D.C., 1994. Why do banks fail? Evidence from the 1920s. *Explorations in Economic History* 30, 409–431.
- Bagehot, W., 1873. *Lombard Street: A Description of the Money Market*. Richard D. Irwin, Homewood, IL (reprinted, 1962).
- Barth, J., Caprio Jr., G., Levine, R., 2006. *Rethinking Bank Regulation: Till Angels Govern*. Cambridge University Press, Cambridge.
- Bernstein, A., Hughson, E., Weidenmier, M.D., 2009. Quasi-central banking and liquidity risk: lessons from the Second Bank of the US. Working Paper.
- Bernstein, A., Hughson, E., Weidenmier, M.D., 2010. Identifying the effects of a lender of last resort on financial markets: lessons from the founding of the Fed. *Journal of Financial Economics* 98 (1), 40–53.
- Board of Governors of the Federal Reserve System, 1999. Using subordinated debt as an Instrument of Market Discipline. Staff Study 172, December.
- Bordo, M., 1985. The impact and international transmission of financial crises: some historical evidence, 1870–1933. *Revista di Storia Economica* 2, 41–78 2d ser.
- Bordo, M.D., 2007. The crisis of 2007: the same old story, only the players have changed. Rutgers University Working Paper.
- Bordo, M.D., Eichengreen, B., 2003. Crises now and then: what lessons from the last era of financial globalization? In: *Monetary History, Exchange Rates and Financial Markets: Essays in Honour of Charles Goodhart*, vol. 2. Edward Elgar, Cheltenham, UK, pp. 53–84.

- Bordo, M.D., Wheelock, D., 2007. Stock market booms and monetary policy in the twentieth century. *Federal Reserve Bank of St. Louis Review* 89, 90–122.
- Bordo, M.D., Wheelock, D., 2009. When do stock market booms occur: the macroeconomic and policy environments of 20th century booms. In: Atack, J., Neal, L. (Eds.), *The Origins and Development of Financial Markets and Institutions*. Cambridge University Press, Cambridge, pp. 416–449.
- Boyd, J., Gomis, P., Kwak, S., Smith, B., 2000. A user's guide to banking crises. *The World Bank Conference Paper*.
- Bruner, R.F., Carr, S.D., 2007. *The Panic of 1907: Lessons Learned from the Market's Perfect Storm*. Wiley, Hoboken, NJ.
- Brunner, K., Meltzer, A.H., 1968. What did we learn from the monetary experience of the United States in the Great Depression? *Canadian Journal of Economics* 1, 334–348.
- Butlin, S.J., 1961. *Australia and New Zealand Bank: The Bank of Australasia and the Union Bank of Australia Limited, 1828–1951*. Longmans, London.
- Butlin, N.G., 1964. *Investment in Australian Economic Development, 1861–1914*. Cambridge University Press, Cambridge.
- Butlin, M.W., 1987. Capital markets. In: Maddock, R., Mclean, I.W. (Eds.), *The Australian Economy in the Long Run*. Cambridge University Press, Cambridge.
- Calomiris, C.W., 1989. Deposit insurance: lessons from the record. *Federal Reserve Bank of Chicago Economic Perspectives* 13, 10–30.
- Calomiris, C.W., 1990. Is deposit insurance necessary? A historical perspective. *Journal of Economic History* 50, 283–295.
- Calomiris, C.W., 1992. Do vulnerable economies need deposit insurance? Lessons from U.S. agriculture in the 1920s. In: Brock, P.L. (Ed.), *If Texas Were Chile: A Primer on Bank Regulation*. The Sequoia Institute, San Francisco, CA, pp. 237–349 450–458.
- Calomiris, C.W., 2000. *U.S. Bank Deregulation in Historical Perspective*. Cambridge University Press, Cambridge.
- Calomiris, C.W., Gorton, G., 1991. The origins of banking panics: models, facts, and bank regulation. In: Glenn Hubbard, R. (Ed.), *Financial Markets and Financial Crises*. University of Chicago, Chicago, pp. 107–173.
- Calomiris, C.W., Kahn, C.M., 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81, 497–513.
- Calomiris, C.W., Mason, J.R., 1997. Contagion and bank failures during the Great Depression: the June 1932 Chicago banking panic. *American Economic Review* 87, 863–883.
- Calomiris, C.W., Mason, J.R., 2003a. Fundamentals, panics and bank distress during the depression. *American Economic Review* 93, 1615–1647.
- Calomiris, C.W., Powell, A., 2001. Can emerging market bank regulators establish credible discipline: the case of Argentina, 1992–99. In: Mishkin, F.S. (Ed.), *Prudential Supervision: What Works and What Doesn't*. University of Chicago Press, Chicago, pp. 147–196.
- Calomiris, C.W., Schweikart, L., 1991. The panic of 1857: origins, transmission, and containment. *Journal of Economic History* 51, 807–834.
- Calomiris, C.W., White, E.N., 1994. The origins of federal deposit insurance. In: Goldin, C., Libecap, G. (Eds.), *The Regulated Economy: A Historical Approach to Political Economy*. University of Chicago Press, Chicago, pp. 145–188.
- Calomiris, C.W., Wilson, B., 2004. Bank capital and portfolio management: the 1930s 'capital crunch' and scramble to shed risk. *Journal of Business* 77, 421–455.
- Caprio, G., Klingebiel, D., 1996. *Bank insolvencies: cross country experience*. The World Bank, Washington, DC Working Paper No. 1620.
- Catterall, R.C.H., 1902. *The Second Bank of the United States*. University of Chicago Press, Chicago.
- Davis, L.E., Gallman, R.E., 2001. *Evolving Financial Markets and International Capital Flows: Britain, the Americas, and Australia, 1865–1914*. Cambridge University Press, Cambridge.
- Davis, J., Hanes, C., Rhode, P., 2007. *Harvests and business cycles in nineteenth century America*. Working Paper.
- Demirguc-Kunt, A., Detragiache, E., 2000. Does deposit insurance increase banking system stability? Conference Paper, The World Bank, Washington, DC.
- Demirguc-Kunt, A., Kane, E., Laeven, L. (Eds.), 2009. *Deposit Insurance Around the World*. MIT Press, Cambridge, MA.
- Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91, 401–419.
- Friedman, M., Schwartz, A.J., 1963. *A Monetary History of the United States, 1867–1960*. Princeton University Press, Princeton.
- Gorton, G., 1985. Clearing houses and the origin of central banking in the United States. *Journal of Economic History* 45, 277–283.
- Gouge, W.A., 1833. *An inquiry into the principles of the American banking system*. Philadelphia.
- Hammond, B., 1970. *Sovereignty and an Empty Purse: Banks and Politics in the Civil War*. Princeton University Press, Princeton.
- Hilt, E., 2009. Wall Street's first corporate governance crisis: the panic of 1826. NBER Working Paper No. 14892, April.
- Kindleberger, C.P., 1978. *Manias, Panics, and Crashes: A History of Financial Crises*. Basic Books, New York.
- Mason, J.R., 2001. Do lender of last resort policies matter? The effects of Reconstruction Finance Corporation assistance to banks during the Great Depression. *Journal of Financial Services Research* 20, 77–95.
- Meltzer, A.H., 2003. *A History of the Federal Reserve, 1913–1951*. University of Chicago Press, Chicago.
- Minsky, H.P., 1975. *John Maynard Keynes*. Columbia University Press, New York.
- Miron, J., 1986. Financial panics, the seasonality of the nominal interest rate, and the founding of the Fed. *American Economic Review* 76, 125–140.
- Mishkin, F.S. (Ed.), 2001. *Prudential Supervision: What Works and What Doesn't*. University of Chicago Press, Chicago.
- Mitchener, K.J., 2005. Bank supervision, regulation, and instability during the great depression. *Journal of Economic History* 65, 152–185.
- Richardson, G., William, T., 2006. Monetary intervention mitigated banking panics during the great depression? Quasi-experimental evidence from the Federal Reserve District Border in Mississippi, 1929 to 1933. NBER Working Paper No. 12591.
- Rousseau, P., 2002. Jacksonian monetary policy, specie flows, and the Panic of 1837. *Journal of Economic History* 62, 457–488.
- Schweikart, L., 1987. *Banking in the American South from the Age of Jackson to Reconstruction*. Louisiana State University Press, Baton Rouge.
- Sprague, O.M.W., 1910. *History of Crises Under the National Banking System*. National Monetary Commission, Government Printing Office, Washington, DC.
- Sylla, R., Wright, R.E., Cowen, D.J., 2009. Alexander Hamilton, Central Banker: crisis management during the U.S. financial panic of 1792. *Business History Review* 83, 61–86.
- Temin, P., 1969. *The Jacksonian Economy*. W.W. Norton, New York.
- Timberlake, R.H., 1984. The central banking role of Clearinghouse Associations. *Journal of Money, Credit and Banking* 16, 1–15.
- Wheelock, D.C., 1991. *The Strategy and Consistency of Federal Reserve Monetary Policy, 1924–1933*. Cambridge University Press, Cambridge.
- White, E.N., 1984. A Reinterpretation of the Banking Crisis of 1930. *Journal of Economic History* 44, 119–138.
- Wicker, E., 1966. *Federal Reserve Monetary Policy, 1917–1933*. Random House, New York.
- Wicker, E., 1996. *The Banking Panics of the Great Depression*. Cambridge University Press, New York.

Wigmore, B.A., 1987. Was the bank holiday of 1933 a run on the dollar rather than the banks? *Journal of Economic History* 47, 739–756.

Further Reading

- Bernanke, B.S., 1983. Nonmonetary effects of the financial crisis in the propagation of the great depression. *American Economic Review* 73, 257–276.
- Bernanke, B.S., James, H., 1991. The gold standard, deflation, and financial crisis in the Great Depression: an international comparison. In: Glenn Hubbard, R. (Ed.), *Financial Markets and Financial Crises*. University of Chicago Press, Chicago.
- Calomiris, C.W., Glenn Hubbard, R., 1989. Price flexibility, credit availability, and economic fluctuations: evidence from the U.S. 1894–1909. *Quarterly Journal of Economics* 429–452.
- Calomiris, C.W., Mason, J.R., 2003b. Consequences of bank distress during the Great Depression. *American Economic Review* 93, 937–947.
- Canovai, T., 1911. *The Banks of Issue in Italy*. National Monetary Commission, Washington, DC US Senate, 61st Congress, 2d Session, Document No. 575.
- Conant, C.A., 1910. *The Banking System of Mexico*. National Monetary Commission, Washington, DC 61st Congress, 2nd Session.
- Davis, W.S., 1910. *The Influence of Wealth in Imperial Rome*. Macmillan, New York.
- Goodhart, C.A.E., 1995. *The Evolution of Central Banks*. MIT Press, Cambridge, MA.
- Grossman, R.S., 1994. The shoe that didn't drop: explaining banking stability during the Great Depression. *Journal of Economic History* 54, 654–682.
- Lucia, J.L., 1985. The failure of the bank of United States: a reappraisal. *Explorations in Economic History* 22, 402–416.
- Schweikart, L., 1988. Jacksonian ideology, currency control, and central banking: a reappraisal. *The Historian* 78–102.
- Tacitus, C., 1996. *The Annals of Imperial Rome*. Penguin, New York.
- Temin, P., 1976. *Did Monetary Forces Cause the Great Depression?* W.W. Norton, New York.
- Temin, P., 1989. *Lessons from the Great Depression*. MIT Press, Cambridge, MA.
- White, E.N., 2009. *The Banque de France as a lender of last resort in the nineteenth century*. Rutgers University Working Paper.
- Wicker, E., 1980. A reconsideration of the causes of the banking panic of 1930. *Journal of Economic History* 40, 571–583.