Cambridge University Press
978-1-107-03094-7 - Banking in Crisis: The Rise and Fall of British Banking Stability, 1800 to the Present
John D. Turner
Excerpt
More information

1 Introduction: Holding shareholders to account

> But we know that generations do not always act upon the experience of their predecessors. There are periods of confidence in which all ordinary maxims of prudence are neglected . . . and all banking is in its very nature liable to abuse.[1]
>
> Thomas Tooke

Looking to the past

My first introduction to banking was playing *Monopoly*, the popular board game, with my siblings on rainy Sunday afternoons in the early 1980s. I learned two things from *Monopoly*. First, if one wished to mortgage property, the bank would advance no more than 50 per cent of the property's value. In other words, the loan-to-value ratio was 50 per cent. Second, the banker, usually my brother, had to be constrained from cheating via a combination of monitoring, punishment and appropriate incentives. Fast-forward several decades and real British banks were granting mortgages with loan-to-value ratios of up to 125 per cent and British bankers, instead of being constrained to behave prudently and cautiously, were incentivised to increase bank leverage and take imprudent risks with other people's money. The lessons of my youth suggested that such a system was doomed to implode, which it duly did in spectacular fashion in the autumn of 2008.

The portents of the collapse of the British banking system, as well as the breakdown of the banking system in the United States and in European economies, appeared in the summer of 2007, when banks ceased lending to one another. By September 2007, Northern Rock was receiving emergency loans from the Bank of England and facing depositor runs, with long queues of depositors outside many of its branches shown on BBC news broadcasts. It took an announcement by Alastair Darling, the Chancellor of the Exchequer at the time, of a taxpayer guarantee for

[1] *Committee of Secrecy on the Bank of England Charter*, P.P. 1831–32 VI, Evidence of Thomas Tooke, q. 3918.

all of Northern Rock's deposits and various wholesale liabilities to bring the run on the Rock to an end.[2] The financial condition of Northern Rock was so poor that it was eventually nationalised in February 2008. Then, following the failure of Lehman Brothers in the United States on 15 September 2008, banking and financial systems across much of the developed world experienced a collapse, which resulted in taxpayer-funded bailouts and emergency loans unprecedented in their scale and scope. The United Kingdom was at the epicentre of the crisis, with the Royal Bank of Scotland, HBOS (the result of the Halifax and Bank of Scotland merger), Lloyds-TSB and Bradford and Bingley, as well as Northern Rock, all requiring taxpayer support to prevent their collapse.

The 2007–8 crisis has resulted in economists, policy makers and ordinary citizens now looking to past financial crises to understand more about the anatomy of banking crises and the appropriate policy responses of governments, monetary authorities and financial regulators. As a result, there is renewed interest in economic history and, in particular, financial history and historical banking crises.[3] For example, in its final report, the Parliamentary Commission on Banking Standards suggests that the 2007–8 crisis might not have happened had the lessons of past failures been heeded.[4] This book ties into this appreciation of the importance of historical research by analysing the stability of the British banking system in the past two centuries, from immediately before the point at which modern joint-stock banking emerges until the Great Crash of 2007–8.

Economists have been lambasted for the inability of the profession to predict the Great Crash of 2007–8.[5] One possibility is that the economics profession came to be dominated by the wrong ideology – that is, a blind faith in competition and the free market. Over time, those who disagreed with the new ideology were excluded from the profession so that there were few dissenting voices and the free market became the 'new policy metaphysics'.[6] Another more worrying possibility is that the profession

2 House of Commons Treasury Committee, *Banking Crisis*, p. 45.
3 Notable examples include Reinhart and Rogoff, *This Time Is Different*; Schularick and Taylor, 'Credit booms gone bust'; and Gorton, *Misunderstanding Financial Crises*.
4 Parliamentary Commission on Banking Standards, *Changing Banking for Good*, vol. 1, pp. 15–16. This Commission recommended that the Bank of England's Financial Policy Committee should have an external member, 'with particular responsibility for taking a historical view of financial stability and systemic risk' (vol. 1, p. 62).
5 See, for example, Buiter, 'The unfortunate uselessness'; Gorton, *Misunderstanding Financial Crises*, vii–xii; and Hodgson, 'The great crash of 2008'. According to Frydman and Goldberg in *Beyond Mechanical Markets*, the mechanical and mathematical models of macroeconomists and financial economists were particularly to blame.
6 Offer, 'Narrow banking', p. 15.

simply supplied the ideology that was demanded by the economic and political elite. In other words, economics 'sold its soul'.

We must ask, however, whether economic historians fared any better. Notable economic historians or economists with a good knowledge of banking history did not predict the crisis. One possible explanation for this is that economic history has picked up the bad as well as the ideological habits of economics.[7] However, it is perhaps unfair to expect predictions from economic historians. To use a medical analogy, economists diagnose problems and prescribe preventive medicine, whereas economic historians are pathologists because they try to uncover what happened in the past.[8] Nevertheless, modern medicine started with pathology and medical students begin their training by examining cadavers. Similarly, this book examines the 'cadavers' of past banking crises to learn about the anatomy of banking crises and, in the process, perhaps learn something about preventive measures.

Three questions are explicitly addressed in this book. First, how often did banking crises occur in the past two centuries and how severe were they? Second, why did banking crises occur? Third, what role did the government and the Bank of England play in crises: Did they alleviate or exacerbate matters? Another question is implicitly addressed throughout the book: What insight does the history of banking stability in Britain provide about the reasons for the Great Crash of 2007–8? Although the book attempts to explain why the Great Crash occurred, it does so only in the context of two centuries of banking history. Those looking for a detailed narrative of the Great Crash should turn to the voluminous literature that it has already generated.[9]

Why should one care about banking stability?[10] As a society, we care because of the important roles of banks in the economy. First, banks provide most of an economy's money supply in the form of transaction deposits, which greatly reduces the costs of engaging in trade and exchange. Second, banks provide intermediation of funds between

7 Solow, 'Economic history and economics'.
8 I thank Cormac Ó Gráda for this analogy.
9 See, for example, Booth, *Verdict on the Crash*; Brunnermeier, 'Deciphering the liquidity and credit crunch'; Diamond and Rajan, 'The credit crisis'; Dowd and Hutchinson, *The Alchemists of Loss*; French et al., *The Squam Lake Report*; Gorton, *Slapped by the Invisible Hand*; Johnson and Kwak, *13 Bankers*; Mian and Sufi, 'House prices'; Mishkin, 'Over the cliff'; Peston and Knight, *How Do We Fix This Mess?*; Rajan, *Fault Lines*; Schwartz, 'Origins of the financial market crisis'; Shiller, *The Subprime Solution*; and Sorkin, *Too Big to Fail*.
10 Most scholars use the terms *banking crisis* and *financial crisis* synonymously, but some economists would also regard a currency crisis as a financial crisis. See Kaminsky and Reinhart, 'The twin crises'.

borrowers and savers. This credit intermediation ultimately facilitates investment by businesses and enables individuals to provide for their future consumption needs. Banking instability implies that these important services provided by banks are detrimentally affected, with potentially catastrophic consequences for both ordinary citizens and businesses.

A study of banking stability in the past two centuries is – in one sense – of purely historical interest. However, an historical examination of banking stability sheds light on the Great Crash because by studying the past, we understand how the banking system evolved and the origins of vulnerabilities in the banking ecosystem. Another reason that a long-run perspective is useful is that banking crises are low-frequency events. Hence, past crises are additional observations that are useful in understanding the dynamics and commonalities, as well as the basic anatomy, of banking crises. However, in looking at historical crises, one must be careful not to 'see history as a homogeneous data pool with which to test modern theories'.[11]

A benefit of focusing on only one country rather than conducting a comparative study is that a higher level of institutional detail is obtained. Furthermore, the unique methods used to measure banking stability throughout a two-century window in Britain would be extremely difficult to replicate for other economies. Of course, the downside of the single-country study is that the cross-sectional correlations and insights provided by a comparative analysis are lost. To compensate for this, a comparative analysis is utilised – whenever it is warranted – throughout the book. Nevertheless, the British case is informative about the global banking system for a number of reasons. First, for more than two centuries, Britain has been a – if not the – major player in world finance. For most if not all of the past two centuries, London has been the world's leading financial centre. Indeed, Britain has had a sophisticated and highly developed financial system longer than any other economy.

Second, it is traditionally believed that the British banking system was one of the most stable in the twentieth century. This was borne out, in particular, by the relative stability of the British banking system during the Great Depression.[12] Whereas other banking systems were collapsing and suffering panics, Britain's banking system was relatively calm, despite the substantial contraction in the wider British economy.

[11] Dow and Dow, 'Economic history', p. 3.
[12] Grossman, 'The shoe that didn't drop'; Capie and Wood, *Money over Two Centuries*, p. 333.

Third, central-banking practice and theory were developed mainly in Britain during the nineteenth century, notably with regard to the function of 'lender of last resort', which is believed to underpin banking stability in modern economies.[13] Britain is therefore interesting because it developed the prototype central bank and lender of last resort.

Fourth, unlike nearly all other developed nations, the United Kingdom had minimal statutory regulation of banking until 1979. Apart from Peel's Bank Charter Act (1844), which restrained bank-note issuance, no other major statutory attempts had been made to control or regulate banking. Thus, on the surface, it appears that Britain was unusual on two counts in the twentieth century: the stability of its banking system and the absence of statutory regulation for banks. Could this imply that the statutory regulation of banks and banking stability are mutually exclusive?

Fifth, as highlighted throughout this book, Britain's experience with banking stability in many cases mirrors what happened in other major economies.[14] In other words, some reasons for the rise and fall of British banking stability in the past two centuries have many parallels in other nations, particularly in the case of the Great Crash of 2007–8. As a result, the lessons and insights of this book stretch beyond the shores of Britannia.

There are, of course, risks to be avoided in a study of banking stability over the long run. One prominent risk is that of nostalgia, wherein the study of history enables one to look back fondly on the halcyon days when banking was stable. As a result of such nostalgia, one could ultimately recommend that banking should return to the way it was in the 'good old days' and that all subsequent socially optimal banking innovations that occurred during the intervening decades should be removed.[15]

Another risk of looking at the long run is that the nature of banking may have changed during the period so that banking in the nineteenth century has no similarities with banking (and, consequently, with banking crises) in the twenty-first century. One way of minimising this danger, which is adopted in this book, is to focus on commercial banks – that is, those banks that take in deposits and lend to businesses, governments and individuals. In the past two centuries, British commercial banks often

[13] Smith, *The Rationale of Central Banking*, pp. 8–24, 71–80.
[14] For example, the 2007–8 crisis resulted in bank failures or bailouts of major banks in Belgium, France, Germany, Iceland, Ireland, Netherlands, Spain, Switzerland and the United States.
[15] See Bhidé, *A Call for Judgment*, who, in his insightful critique of financial innovation, perhaps goes too far in recommending a return to banking as practised a half-century ago.

have changed substantially: modern banks offer a wider range of services than their ancestors, they are considerably larger in terms of scale and scope, and many of them operate across the globe. However, the two key economic functions of commercial banks in the past two centuries have not changed. First, British commercial banks have always provided a means of payment to their customers, whether in the form of bank notes or transaction deposits. Second, British commercial banks have always intermediated funds between savers, who are typically individuals, and borrowers, who are typically but not exclusively businesses.[16]

The basic argument

Banking is an intrinsically risky business, and the reason is simple: bankers lend other people's money, not their own. This creates an incentive problem because bankers get most of the benefit if the risky loans they make do well, whereas depositors, not bankers, incur most of the costs if loans go bad. Unless it is addressed, this incentive problem eventually results in unstable banking. The basic argument advanced in this book is that banking is at its most stable when one of two conditions exists, both of which address this intrinsic incentive problem at the heart of banking.

The first condition is that bank shareholders are held to account for bank failures. What does this mean? The basic idea is that when bank shareholders stand to lose substantially from a bank failure, they will ensure that their bank is properly and prudently run, thereby greatly reducing the probability of it failing in the first instance. As a result, bank depositors are assured that their deposits are safe because bankers have an incentive to ensure that they are judicious in their treatment of depositors' funds.

What can shareholders lose when their bank fails? First, shareholders can lose all of the capital they invested in the bank. Second, if their liability was not limited, shareholders could also face a call on their personal wealth in the event of bank failure. For example, bank shareholders could have unlimited liability, whereby they are liable to make good the deficit between their bank's assets and liabilities whenever it fails, down to their last penny or – in the quaint terminology of the nineteenth century – to their 'last acre and sixpence'. Alternatively, bank shareholders could

[16] Offer, in 'Narrow banking', argues that Victorian commercial banks, unlike modern banks, were mainly providers of liquidity to businesses and were not involved in much maturity transformation. However, although lending in this era was short term in duration, much of it was rolled over.

establish a type of 'halfway house' between pure limited liability and unlimited liability; for example, shareholders could be held liable for a defined multiple of their paid-up capital.

The second condition under which banking is at its most stable is when banks are constrained by onerous government controls. For example, banks could be required to hold a significant amount of low-risk government debt and be restricted from lending to risky or speculative sectors of the economy. Such onerous restrictions place bankers in a figurative straitjacket, which severely constrains their proclivity to take excessive risk and thereby keeps banking stable.

This book attempts to explain the stability (or otherwise) of the British banking system since 1800. To measure banking stability over the long run, an innovative approach is used. With its detailed study of bank-share prices and failure rates during a two-century period, this approach produces different results from the standard narrative accounts of British banking history, which classifies many more episodes as crises and fails to distinguish between the seriousness of various episodes of banking instability.[17]

Using this innovative approach suggests that there have been only two major banking-system crises in Britain in the past two centuries. The first major crisis was in 1825–6; the second was the Great Crash of 2007–8. In the interim, there were periods when the banking system was under stress and weak banks failed, but at no time was there a major crisis or a threat to the overall stability of the banking system. Notably, these minor crises or episodes of instability had a limited real economic effect, compared to the decreases in economic output associated with the two major crises. Indeed, some of the minor crises – in particular, those in the nineteenth century – may have had a role in strengthening banking systems because they eliminated weak and risk-loving banks.

This long-run perspective on banking stability also reveals that the severity and the scale of the 2007–8 banking crisis are unprecedented in British banking history. No previous crisis witnessed the collapse of such a large proportion of the banking system. Neither did any previous crisis necessitate such large-scale intervention by the taxpayers and the monetary authorities to save the system. No previous crisis was followed by such a steep decline in economic output, such a prolonged economic malaise, and such a large increase in public indebtedness. In other words, to quote an overused phrase, this time really does differ.

[17] Collins, *Money and Banking in the UK*; Baker and Collins, 'Financial crises'; Reinhart and Rogoff, *This Time Is Different*; Grossman, *Unsettled Account*; Capie and Wood, *Money over Two Centuries*.

Having identified that the UK banking system was relatively stable between 1826 and 2007, the remainder of the book addresses two principal questions: (1) Why was the British banking system crisis-free for such a long period? and (2) Why did it crash in 2007–8? A subsidiary question that the book addresses is: Why did the 1825–6 crisis happen?

Prior to 1826, the English banking system experienced frequent bouts of instability, but the crisis of 1825–6 was by far the most severe of the era. This crisis, which was purely English, occurred because banks were constrained to the partnership organisational form, which meant that they were small and therefore had inadequate capital. Scotland was able to escape the 1825–6 crisis unscathed largely because Scottish partnership law was highly flexible compared to English and Irish law. The result was that Scottish banks were more like joint-stock companies, making them more robust to economic shocks. Notably, the post-crisis reforms introduced into the English banking system in 1826 allowed banks to become more like Scottish banks in that they could be formed as joint-stock companies.

Although experiencing periodic bouts of nervousness, money-market strains and episodic bank failures, the UK banking system remained relatively stable throughout the *c.* 175 years since 1826 in that it did not experience any systemic or major crises. What explains this remarkably long period of relative stability? Briefly, the two conditions for stable banking (outlined previously) held during most of this era, with the result that banks did not take excessive risks and that the banking system was stable.

When banking incorporation law was liberalised in the mid 1820s, banks were required to have unlimited shareholder liability. This meant that when a bank failed, shareholders were liable to their last penny to repay depositors for any losses incurred as a result of the collapse. Because the unlimited liability was joint and several, the inability of some shareholders to meet their calls simply meant that wealthier and still-solvent shareholders subsequently faced larger calls. Consequently, one might expect that unlimited-liability banks would not have many wealthy shareholders. However, the voluminous evidence presented in this book suggests otherwise, and depositors typically had all of their deposits returned even when their bank failed. In addition, bank failures in this era stood as constant reminders that owners were held to account because shareholders faced calls to make good the deficit between their bank's liabilities and assets.

The incentives arising from the existence of unlimited liability constrained banks from excessive risk taking because shareholders and, more important, bank directors and managers stood to lose all of their wealth in the event of bank failure. Thus, because banks were not overextended, the banking system could withstand periodic shocks, and there were no endogenous bank-credit-fuelled asset booms followed by a crisis.

The failure of the City of Glasgow Bank in 1878, which resulted in the personal bankruptcy of most of its shareholders, was too much to bear for shareholders in other banks. Consequently, the shareholder-liability regime was diluted so that banks could adopt a halfway house between pure limited liability and unlimited liability. All British banks quickly converted to this new liability regime, under which they could choose and define exactly the extent to which shareholders were liable in the event of bank failure. In the 1880s median British bank, shareholders were liable for up to £2 for every £1 of capital held if their bank failed. Because the median bank also had a high ratio of total capital resources to deposits, this new regime still provided shareholders and managers with adequate incentives to avoid taking excessive risks. Thus, even after the demise of unlimited liability, bank shareholders continued to be held to account.

The extended shareholder liability described previously persisted in British banking until the 1950s, when there was a coordinated removal of it. However, this removal was largely symbolic because the average ratio of what shareholders were potentially liable to pay in the event of failure to total deposits in 1950 was about 3 per cent, having fallen from 33 per cent in 1900. At the same time that this decrease occurred, the ratio of capital to deposits also fell to very low levels, from 18 per cent in 1900 to 4 per cent in 1950. Both of these declines were largely a result of high inflation during the two world wars, during which deposits increased substantially without any commensurate increase in extended liability or banks' capital resources. Essentially, by the 1940s, shareholders were no longer being held to account – indeed, in the event of bank failure, they stood to lose very little. Why, then, did banks not take excessive risks? Why did banks remain stable? The answer provided in this book is that banks did not take excessive risks and remained stable because of substantial constraints placed on them by the Bank of England and the Treasury.

From 1939 until the 1970s, the Treasury adopted financial-repression policies partly to fund its high debt issuance, which had arisen as a result of fighting World War II and the cost of postwar reconstruction, and partly to guide lending towards strategic sectors and industries. These policies meant that banks were constrained, facing onerous 'requests' with regard to their liquidity ratios and their lending, which precluded them from excessive risk taking. Ultimately, financial-repression policies constrained banks from risk taking, with the result that their depositors and the financial authorities were totally unconcerned about the low levels of bank capital. One could therefore view financial-repression policies as a substitute for shareholder capital.

Financial-repression policies in the United Kingdom were not enforced on banks in a formal sense through statutory law; rather, British banks participated in an informal supervisory regime that had the Bank of England at the centre. The Treasury relayed its needs and wishes to the Bank, which in turn relayed them to the clearing banks – that is, the principal commercial banks.[18] The clearing banks were always sure to align their policies to the Bank's and, by extension, the Treasury's wishes. This informal relationship between the Bank of England and the clearing banks had developed in the interwar period under the long suzerainty of Governor Montagu Norman. It was held together in part by the fact that both the banks and the Bank of England were increasingly aware of the threat of nationalisation. The clearing banks met all of the requests made of them either as a quid pro quo for their being allowed to operate a cartel or because of implicit threats from the Bank or the Treasury.

Thus, the long period of banking stability from 1826 until the interwar period was mainly due to shareholders being held to account; from the end of the interwar period until the 1970s, it was due to austere financial-repression policies, which meant that banks had no capacity to engage in risk shifting. Why then did the Great Crash of 2007–8 happen? The simple answer is that with the end of financial repression, constraints were gradually removed from banks and there was no attempt to return to the pre-1939 world in which shareholders were held to account. Add to this the perception that banks would ultimately be bailed out by the taxpayers if they collapsed – a perception that a century of rescues of minor banking institutions had done nothing to assuage – and one can begin to see the malincentives facing bankers in an era when restraints on their business activities had been removed. Although attempts were made to constrain excessive risk taking via supervision and risk-weighted capital-adequacy ratios, those attempts were ultimately fruitless at best and counterproductive at worst because they may have actually created perverse risk-taking incentives for banks.

Although Britain experienced a severe downturn in economic output during the Great Depression of the 1930s, it is remarkable that it did not experience a banking crisis unlike many other economies at the time. At least three reasons are highlighted in this book that saved the British system from experiencing a crisis during the Great Depression: (1) the

[18] Clearing banks were so called because they controlled and were members of the London Clearing House, where cheques and other payment claims against banks were cleared.