© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

1

If Stable and Efficient Banks Are Such a Good Idea, Why Are They So Rare?

> The majority of economists . . . tend to assume that financial institutions will grow more or less spontaneously as the need for their services arises—a case of demand creating its own supply. . . . Such an attitude disposes of a complex matter far too summarily.
>
> Rondo Cameron and Hugh Patrick, *Banking in the Early Stages of Industrialization* (1967)

Everyone knows that life isn't fair, that "politics matters." We say it when our favorite movie loses out at the Academy Awards. We say it when the dolt in the cubicle down the hall, who plays golf with the boss, gets the promotion we deserved. We say it when bridges to nowhere are built because a powerful senator brings federal infrastructure dollars to his home state. And we say it when well-connected entrepreneurs obtain billions in government subsidies to build factories that never stand a chance of becoming competitive enterprises.

We recognize that politics is everywhere, but somehow we believe that banking crises are apolitical, the result of unforeseen and extraordinary circumstances, like earthquakes and hailstorms. We believe this because it is the version of events told time and again by central bankers and treasury officials, which is then repeated by business journalists and television talking heads. In that story, well-intentioned and highly skilled people do the best they can to create effective financial institutions, allocate credit efficiently, and manage problems as they arise—but they are not omnipotent. Unable to foresee every possible contingency, they are sometimes subjected to strings of bad luck. "Economic shocks," which presumably could not possibly have been anticipated, destabilize an otherwise smoothly running system. Banking crises, according to this version of events, are much like Tolstoy's unhappy families: they are all unhappy in their own ways.

This book takes exception with that view and suggests instead that the politics that we see operating everywhere else around us also determines

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

whether societies suffer repeated banking crises (as in Argentina and the United States), or never suffer banking crises (as in Canada). By politics we do not mean temporary, idiosyncratic alliances among individuals of the type that get the dumbest guy in the company promoted to vice president for corporate strategy. We mean, instead, the way that the fundamental political institutions of a society structure the incentives of politicians, bankers, bank shareholders, depositors, debtors, and taxpayers to form coalitions in order to shape laws, policies, and regulations in their favor—often at the expense of everyone else. In this view, a country does not "choose" its banking system: rather it gets a banking system that is consistent with the institutions that govern its distribution of political power.

The Nonrandom Distribution of Banking Crises

Systemic bank insolvency crises like the U.S. subprime debacle of 2007–09 —a series of bank failures so catastrophic that the continued existence of the banking system itself is in doubt—do not happen without warning, like earthquakes or mountain lion attacks. Rather, they occur when banking systems are made vulnerable by construction, as the result of political choices. Banking systems are susceptible to collapse only when banks both expose themselves to high risk in making loans and other investments and have inadequate capital on their balance sheets to absorb the losses associated with those risky loans and investments. If a bank makes only solid loans to solid borrowers, there is little chance that its loan portfolio will suddenly become nonperforming. If a bank makes riskier loans to less solid borrowers but sets aside capital to cover the possibility that those loans will not be repaid, its shareholders will suffer a loss, but it will not become insolvent. These basic facts about banking crises are known to bankers or government regulators; they are as old as black thread.

By contrast, consider what occurs when bank capital is insufficient relative to bank risk. Bank losses can become so large that the negative net worth of banks totals a significant fraction of a country's gross domestic product (GDP). In this scenario, credit contracts, GDP falls, and the country sustains a recession driven by a banking crisis. Governments can prevent this outcome by propping up the banking system. They can make loans to the banks, purchase their nonperforming assets, buy their shares in order to provide them with adequate capital, or take them over entirely.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

If such catastrophes were random events, all countries would suffer them with equal frequency. The fact is, however, that some countries have had many, whereas others have few or none. The United States, for example, is highly crisis prone. It had major banking crises in 1837, 1839, 1857, 1861, 1873, 1884, 1890, 1893, 1896, 1907, the 1920s, 1930–33, the 1980s, and 2007–09.[1] That is to say, the United States has had 14 banking crises over the past 180 years! Canada, which shares not only a 2,000-mile border with the United States but also a common culture and language, had only two brief and mild bank illiquidity crises during the same period, in 1837 and 1839, neither of which involved significant bank failures. Since that time, some Canadian banks have failed, but the country has experienced no systemic banking crises. The Canadian banking system has been extraordinarily stable—so stable, in fact, that there has been little need for government intervention in support of the banks since Canada became an independent country in 1867.

The nonrandom pattern of banking crises is also apparent in their distribution around the world since 1970. Some countries appear immune to the disease, while others are unusually susceptible. Consider the pattern that emerges when we look at data on the frequency of banking crises in

[1] Throughout this book we regard banking crises as either systemic insolvency crises or systemic illiquidity crises. Some crises, like the subprime lending crisis in the United States, and the other U.S. crises in 1837, 1839, 1857, 1861, the 1920s, 1930–33, and the 1980s, have involved extensive bank insolvency, not just moments of illiquidity when banks experience severe withdrawal pressures. Thresholds of insolvency sufficient to constitute a crisis are defined differently by different scholars, but roughly speaking, bank insolvency crises are usefully defined as events during which the negative net worth of banks, or the costs of government interventions to prevent those insolvencies, exceed some critical percentage of GDP. This approach underlies the databases on banking crises for the recent era derived by researchers at the World Bank and International Monetary Fund (e.g., Caprio and Klingebiel [2003]; Laeven and Valencia [2012]). A second class of banking crises is those that entail systemic illiquidity disruptions (e.g., widespread bank runs) but do not involve significant bank insolvencies or costly government interventions to prevent those insolvencies. Calomiris and Gorton (1991), for example, categorize the U.S. banking panics of 1873, 1884, 1890, 1893, 1896, and 1907 as systemic and important liquidity shocks even though they did not produce a high degree of bank insolvency. Both of these definitions of crises are more restrictive than those that are sometimes employed in the "financial crisis" literature (e.g., Reinhart and Rogoff [2009]), where negative events, such as the failure of a single large bank, are considered to be evidence of a crisis. By those less restrictive standards, the world's banking systems would appear to be even more crisis prone.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

the 117 nations of the world that have populations in excess of 250,000, are not current or former communist countries, and have banking systems large enough to report data on private credit from commercial banks for at least 14 years between 1990 and 2010 in the World Bank's Financial Structure Database.[2] Only 34 of those 117 countries (29 percent) were crisis free from 1970 to 2010. Sixty-two countries had one crisis. Nineteen countries experienced two crises. One country underwent three crises, and another weathered no less than four. That is to say, countries that underwent banking crises outnumbered countries with stable banking systems by more than two to one, and 18 percent of the countries in the world appear to have been preternaturally crisis prone.

The country that experienced the most crises was Argentina, a nation so badly governed for so long that its political history is practically a synonym for mismanagement. The close runner-up (with three crises since 1970) was the Democratic Republic of the Congo, the nation whose brutal colonial experience served as the inspiration for Joseph Conrad's *Heart of Darkness,* which was governed after independence by one of the third world's longest-lived and most avaricious despots (Mobutu Sese Seko, who ruled from 1965 to 1997), and whose subsequent history is a template for tragedy.

The 19 countries that had two banking crises are also far from a random draw. The list includes Chad, the Central African Republic, Cameroon, Kenya, Nigeria, the Philippines, Thailand, Turkey, Bolivia, Ecuador, Brazil, Mexico, Colombia, Costa Rica, Chile, Uruguay, Spain, Sweden,

[2] We exclude former and current communist countries from this analysis because their state-run banking systems do not allocate credit but rather act as an accounting system for the state-controlled allocation of investment. The concept of a banking crisis has no real analytic meaning in such a system. Former communist countries have tended to be crisis prone. If we had included them in our data set, an even greater percentage of the countries of the world would be counted as crisis prone. We exclude countries that do not report at least 14 observations for the ratio of private credit by deposit money banks to GDP during the period 1990–2010. That is, in order to mitigate measurement error, we require observations for at least two-thirds of all possible observations for any country. We draw the credit data from the period 1990–2010 because the coverage of the World Bank Financial Structure Database tends to be less complete, especially for poorer countries, prior to 1990. We draw the data on banking crises from Laeven and Valencia (2012) and include both their "systemic" crises and their "borderline" crises in our definition of crises. We update their work by adding the case of Cyprus in 2013.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

and . . . the United States. One of the striking features of this list is the paucity of high-income, well-governed countries on it. Of the 117 countries in our data set, roughly one-third are categorized by the World Bank as high-income nations. But only three of the 21 crisis-prone countries, 14 percent, are in this group. This suggests that, for the most part, being crisis prone is connected to other undesirable traits and outcomes. But that raises another troubling question. Why is the United States on this list?

The Nonrandom Distribution of Under-Banked Economies

There is, of course, more to having a good banking system than simply avoiding crises. Equally problematic are banking systems that provide too little credit relative to the size of the economy—a phenomenon known as under-banking. This outcome, too, appears not to be randomly distributed. Consider the striking contrast between Canada and Mexico, the United States' partners in the North American Free Trade Agreement (NAFTA). From 1990 to 2010, private bank lending to firms and households averaged 95 percent of GDP in Canada, but in Mexico the ratio was only 19 percent. The dramatic difference in those ratios means that Mexican families have a much more difficult time financing the purchase of homes, automobiles, and consumer goods, and Mexican business enterprises have much more difficulty in obtaining working capital, than their Canadian counterparts. The result is slower economic growth. Little wonder, then, that over 500,000 Mexicans—roughly half of all new entrants to the Mexican labor market—illegally cross the border to the United States each year.

As figure 1.1 shows, the stark difference between Canada and Mexico is part of a recurring pattern. In the world's poorest countries (those on the far left-hand side of the figure), including for example, the Democratic Republic of the Congo, the ratio of bank credit to GDP averages only 11 percent. In the richest countries (shown on the far right-hand side of the figure), the ratio of bank credit to GDP averages 87 percent.

Crucially, there is also substantial variance across countries within each of the four income groups, which suggests that the amount of credit extended within countries is not solely a function of demand for credit but also reflects constraints on the supply of credit. In other words, the fact that some countries in each income group extend much more credit than

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

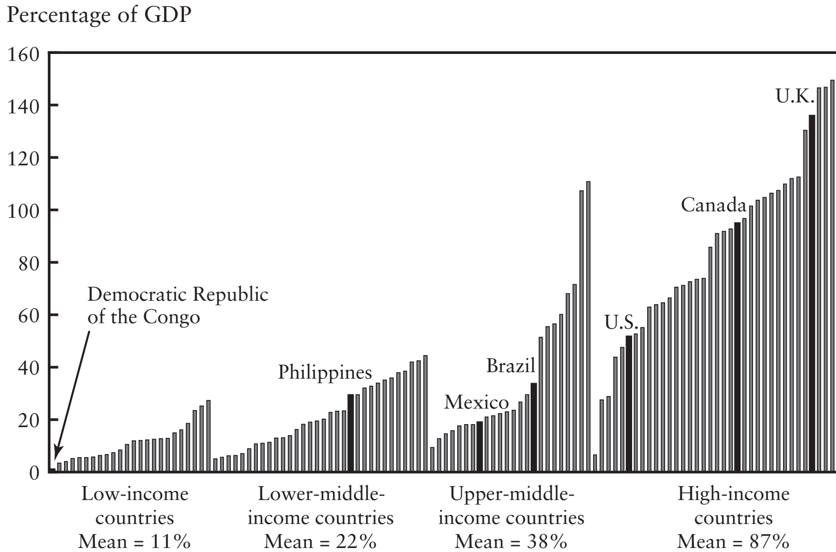

FIGURE 1.1 Average private credit from deposit money banks as a percentage of GDP, 1990–2010, by World Bank income classifications.

Source: World Bank (2012).

Note: For reasons of readability, only selected country names are shown on the *x* axis.

others in the same income group (or even the next-highest income group) suggests that many countries in all categories are under-banked. For example, Mexico appears to be under-banked relative to other countries in the same income group, and even has a lower ratio of credit to GDP than many countries in the next-lowest income group (e.g., the Philippines).

Being under-banked has huge social costs. A large and growing academic literature has shown that under-banked countries suffer lower economic growth than other countries. Economic historians have shown that Holland, Great Britain, and the United States experienced revolutions in financial intermediation and financial institutions *before* their rise to global economic hegemony in the eighteenth, nineteenth, and twentieth centuries, respectively. They also found that Russia, Germany, and Japan underwent similar financial revolutions before they narrowed the gap with the world's economic leaders in the late nineteenth and early twentieth centuries.[3] Financial economists using statistical methods to analyze the growth

[3] Gerschenkron (1962); Cameron et al. (1967); Sylla (1975; 2008); North and

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

of contemporary economies have reached similar conclusions. Whether they look at variance in outcomes across countries, across regions within countries, within countries over time, or across industries, their studies all demonstrate that higher levels of financial development produce faster rates of physical capital accumulation, faster economic growth, more rapid technological progress, faster job creation, and increased opportunities for social mobility.[4] Given the relationship between economic growth and the ability to project power internationally, under-banked countries are also at a disadvantage in defending their sovereignty and influencing events abroad. In short, being under-banked is a far more serious state of affairs than lacking capacity in the real sectors of the economy, such as textiles, sugar refining, or automobile manufacturing: finance facilitates the efficient operation of all other economic activities, including industrial sectors crucial to the defense of the state.

How Many Efficient and Stable Banking Systems Are There?

If very few countries have been free of banking crises since the 1970s, and if much of the world is under-banked, in how many countries is bank credit plentiful and the banking system stable? Answering this question requires us to draw a line between those economies where bank credit is abundant and those economies that are under-banked. If we define a country with abundant credit as one that has an average ratio of bank credit to GDP one standard deviation above the mean for the 117 countries in our data set (83 percent), which corresponds to the ratio in Australia, and

---

Weingast (1989); Neal (1990); de Vries and van der Woude (1997); Rousseau and Wachtel (1998); Rousseau (2003); Rousseau and Sylla (2003, 2004).

[4] King and Levine (1993), Levine and Zervos (1998), Taylor (1998), and Beck, Levine, and Loayza (2000) employed innovative statistical techniques to identify cross-country patterns. A later group of scholars—most notably Rajan and Zingales (1998), Wurgler (2000), Cetorelli and Gamberra (2001), Fisman and Love (2004), and Beck et al. (2008)—focused on the development of industries as well as countries, and they reached the same conclusion: finance leads growth. Research focusing on the growth of regions within countries by Jayaratne and Strahan (1996); Black and Strahan (2002); Guiso, Sapienza, and Zingales (2004); Cetorelli and Strahan (2006); Dehejia and Lleras-Muney (2007); and Correa (2008) produced broadly similar results. These studies built on the theoretical and narrative insights of Goldschmidt (1933); Gurley and Shaw (1960); Gurley, Patrick, and Shaw (1965); Goldsmith (1969); Shaw (1973); McKinnon (1973); and Fry (1988).

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

define a stable banking system as one that has been free of systemic crises since 1970, we arrive at a shocking answer: only six out of 117 countries—5 percent—meet those criteria.[5]

The Puzzling Pervasiveness of Dysfunctional Banking

The puzzle of why societies tolerate unstable and scarce bank credit deepens when one considers the costs imposed by unstable and under-banked systems on those societies. In addition to the slower long-term growth produced by under-banking, unstable banking systems entail other costs. Banking crises magnify recessions, resulting in greater job losses, and taxpayers are forced to pay the price of rescuing bankers from the consequences of their own mistakes. Why do citizens tolerate this? Worldwide, that tab has been huge. Over the period 1970–2011, the median direct fiscal cost of banking crisis resolution was 6.8 percent of GDP, and the median increase in country indebtedness during a crisis was 12.1 percent of GDP. The cost of banking crises in terms of lost GDP (due to the effects of credit contractions, heightened sovereign-debt risk, and currency collapse on economic activity) also tends to be enormous: from 1970 through 2009, the median lost output during a banking crisis amounted to 23 percent of GDP.[6]

In thinking about this puzzle, one shouldn't assume that taxpayers have always been willing to pay for bank bailouts. Taxpayer-funded bailouts of banks are a recent phenomenon. Until the mid-twentieth century, the costs of failure tended to be borne by the bankers themselves, along with bank shareholders and depositors. Since then, however, the costs have been progressively shifted to taxpayers. How did bankers, regulators, and politicians come to impose these costs on taxpayers, and why do taxpayers put up with bearing those costs?

[5] One might think that making the standard for being credit-abundant conditional on a country's World Bank income group might reveal a larger number of credit-abundant, low-crisis countries. In fact, the opposite is the case. If we define a credit-abundant country as one in which the ratio of private credit to GDP is at least one standard deviation above the mean for that country's World Bank income group, and then ask how many of those countries have not had a banking crisis since 1970, the answer is three.

[6] Laeven and Valencia (2012), 17.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

This shifting of losses onto taxpayers is especially puzzling because it tends to produce much larger losses and deeper recessions than a system in which shareholders and depositors bear the cost. A broad literature in financial economics has demonstrated that a system in which shareholders and depositors have money at risk imposes discipline on the behavior of bankers: at the first sign of trouble, stockholders start selling their shares, and depositors start moving their funds to more solvent banks.[7] As a result, some banks fail, some of the holders of bank liabilities (shareholders and depositors) are wiped out, credit contracts as bankers rush to reduce their exposure to risky classes of loans, and economic growth slows. The result is painful, but not tragic. Most important, bankers know the consequences of imprudent behavior and thus tend to maintain large buffers of capital and large portfolios of low-risk assets. As a consequence, systemic banking crises are rare. Contrast that outcome with the system that has come to be the norm since the mid-twentieth century. When losses are borne by taxpayers, the incentives of stockholders and depositors to discipline bankers are much weaker. Bankers are willing to take bigger risks, thereby increasing the probability of failure. As a result, after 1945 banks in the world's most developed economies became more highly leveraged and maintained smaller amounts of low-risk assets.[8]

[7] Calomiris and Powell (2001); Cull, Senbet, and Sorge (2005); Demirgüç-Kunt and Detragiache (2002); Demirgüç-Kunt and Huizinga (2004); Demirgüç-Kunt and Kane (2002); Calomiris and Wilson (2004); Barth, Caprio, and Levine (2006), chapter 4; Haber (2008a); Calomiris (2011a).

[8] Schularick and Taylor (2012). In theory, regulation can replace monitoring by shareholders and depositors, but the evidence is that regulators are subject to political pressures not to act (Brown and Dinc [2005]; Barth, Caprio, and Levine [2006], chapter 5). As a result, losses pile up as bankers throw good money after bad. Inevitably, the stock of unrecognized bad loans—known as "evergreened loans"—grows so large that the banking system is threatened with complete collapse, at which point the regulators are finally forced to step in. By then, however, the stock of bad loans is enormous. This means not only that the cost of cleaning up the failing banks is larger than it would be under a system in which shareholders and depositors disciplined bankers, but also that the recession that follows the banking crisis will be larger as well. It is not just that credit contracts; it contracts vertiginously. Moreover, the government has to find a way to reconcile the fiscal imbalance caused by the bank rescue, which means cutting spending, raising taxes, and increasing central bank interest rates in order to prevent a run on the currency. Not surprisingly, recessions associated with financial crises tend to be deeper than other recessions (Jordà, Schularick, and Taylor [2012]).

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

More puzzling still, costly crises and persistent under-banking occur even though banking systems are subject to close regulation and supervision by governments. In most countries, banks are regulated much like public utilities such as electricity generation: entry to the market is controlled by government agencies in order to assure that the firms providing the service remain profitable, and the government inspects their operations to make sure that they are providing efficient service to their customers while not taking imprudent risks. Why, then, do governments often look the other way when banks make loans to firms and households that have a high probability of default? In the same vein, why do some governments allow those same imprudent banks to deny service to customers who are good credit risks, to the point that in many countries, banks lend only to the enterprises controlled by their own board members?

Fragile by Design

If a stable banking system capable of providing stable access to credit to talented entrepreneurs and responsible households is such a good idea, then why are such systems so rare? How can it be that a sector of the economy that is highly regulated and closely supervised works so badly in so many countries? Our answer to this question is that the fragility of banks and the scarcity of bank credit reflect the structure of a country's fundamental political institutions. The crux of the problem is that all governments face inherent conflicts of interest when it comes to the operation of the banking system, but some types of government—particularly democracies whose political institutions limit the influence of populist coalitions—are better able to mitigate those conflicts of interest than others.

The next chapter examines these conflicts of interest more closely. For the moment let us simply say that they are of three basic types. First, governments simultaneously regulate banks and look to them as a source of finance. Second, governments enforce the credit contracts that discipline debtors on behalf of banks (and in the process assist in the seizing of debtor collateral), but they rely on those same debtors for political support. Third, governments allocate losses among creditors in the event of bank failures, but they may simultaneously look to the largest group of those creditors—bank depositors—for political support. The implication

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

is inescapable: the property-rights system that structures banking is not a passive response to some efficiency criterion but rather the product of political deals that determine which laws are passed and which groups of people have licenses to contract with whom, for what, and on what terms. These deals are guided by the logic of politics, not the logic of the market.

The fact that the property-rights system underpinning banking systems is an outcome of political deal making means that there are no fully "private" banking systems; rather, modern banking is best thought of as a partnership between the government and a group of bankers, a partnership that is shaped by the institutions that govern the distribution of power in the political system. Government policies toward banks reflect the deals that gave rise to those partnerships, as well as the power of the interest groups whose consent is crucial to the ability of the political group in control of the government to sustain those deals. Banks are regulated and supervised according to technical criteria, and banking contracts are enforced according to abstruse laws, but those criteria and laws are not created and enforced by robots programmed to maximize social welfare; they are the outcomes of a political process—a game, as it were—whose stakes are wealth and power.

We call this process of deal making the Game of Bank Bargains.[9] The players are those with a stake in the performance of the banking system: the group in control of the government, bankers, minority shareholders, debtors, and depositors. The rules, which are set by the society's political institutions, determine which other groups must be included in the government-banker partnership and which can be left out in the cold because the rules of the political system make them powerless. Coalitions among the players form as the game is played, and those coalitions determine the rules governing bank entry (and hence the competitive structure and size of the banking sector), the flow of credit and its terms, the permissible activities of banks, and the allocation of losses when banks fail. What is at stake in the Game of Bank Bargains is, therefore, the distribution of the

[9] Our approach builds on the classic work of Olson (1965), Stigler (1971), Krueger (1974), and many others who conceive of government policies as reflecting conflicts among vested interests. A distinguishing feature of our work is the focus on the formation of partnerships of interests that control banking policy. These partnerships often straddle ideological and partisan boundaries.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

benefits that come from a system of chartered banks. The group in control of the government always receives a share of those benefits, and the coalition that forges a partnership with the government splits the remainder.

Notice that our emphasis is not on the extent of regulation but rather on the goals that give rise to regulation and the way those goals are shaped by political bargains. In some countries, the institutions and coalitions are such that regulation improves market outcomes. In other countries, regulation is structured to achieve government objectives that also serve special interests, in spite of their disastrous consequences for society at large.

The struggle among political coalitions determines who gets to play what roles in the financial system; that is, who is granted what kind of banking charter, and which groups of borrowers get government-favored access to credit. A central aspect of the Game of Bank Bargains, therefore, is deciding the rules for entry into banking. It is rarely the case that government chooses a fully "open-access" chartering regime. Being selective about who gets to be a banker and deciding how much bankers are allowed to lend and to whom are crucial elements of the game.

Our goal in this book is to explain this game. We seek to understand the process by which different rules emerge in different countries and how the players operate within those differing sets of rules. We show how differences in fundamental political institutions across times and places produce differences in the rules of the game, and how those politically based rules sometimes result predictably in stable and plentiful bank credit, sometimes in unstable and scarce bank credit, and sometimes (as in the United States) in unstable and plentiful bank credit. Which players favor vertiginous increases in credit, and which players favor tight constraints on it? Under what circumstances can they forge durable political coalitions with other players that have an interest in the organization of the banking system? What are the terms of exchange among the members of these coalitions? Are there differences in the way the game is played under democratic and authoritarian political systems? Do those differences affect the durability of coalitions, the size and structure of banking systems, and the fragility of the banks?

In order to understand this inherently complex game, we trace the coevolution of politics and banking in detail for several countries, one country at a time, over long periods. We spell out how coalitions were

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

formed, why some endured whereas others were undermined, how they brought about specific and important changes in the policies governing banking, how those policies determined which groups could access credit and which could not, and how some of those policies produced disastrous banking crises.

We are not the first to trace the historical evolution of banking systems in various countries or to note the importance of politics for shaping the evolution of banking systems; most obviously, we are building on the seminal contributions to economic history by Bray Hammond (1957), Alexander Gerschenkron (1962), Rondo Cameron (1967), and Richard Sylla (1975), and more recent historical scholarship by Eugene White (1983), Howard Bodenhorn (2003), and Richard Grossman (2010), among others. We also build on the growing literature on the political economy of finance, such as recent work by James Barth, Gerald Caprio, and Ross Levine (2006) and Raghuram Rajan (2010). Indeed, we could not have written this book had we not stood on the shoulders of several generations of financial economists and economic historians. We see our unique contribution as conceptualizing a general framework for understanding how political factors shape banking-system outcomes. We illustrate that framework with detailed narratives that span hundreds of years and integrate evidence and insights from a broad range of academic disciplines, including political history, economic history, financial economics, and political science.

What We Do in This Book, and Why We Do It

A necessary first step toward any kind of productive reform of banking systems is to understand why and how banking-system outcomes are produced by political bargains. In showing how the Game of Bank Bargains is played, we seek to show readers how the variants of the game that they have been drawn into (wittingly or not) in their respective countries may be imposing costs on them while benefiting others.

To accomplish that goal, we have to engage in two quite distinct enterprises. First we lay bare the logic of the Game of Bank Bargains; second, we show how the game has been played, and is currently played, in real-world settings. As much as we might have liked to, we did not pick these settings purely on the basis of which countries have the best beaches: we

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

have chosen cases that allow us to demonstrate how variation in political institutions drives variation in the nature of government-banker partnerships and how variation in those partnerships then produces differences in the size, competitive structure, and stability of banking systems.

We begin by showing why there can be no banking systems without the police power of the state. In chapter 2 we focus, in particular, on the fundamental property-rights problems that societies have to solve in order to create a banking system. The idea that banking systems can exist outside a system of government regulation is simply a libertarian fairy tale.

Chapter 3 explains why governments need banks. We explore how and why the institution of the chartered bank emerged from the process by which Europe was reconfigured from a hodgepodge of duchies, principalities, city-states, and kingdoms into a set of modern nation-states beginning in the seventeenth century. We focus on the strong incentives for rulers to become aggressive proponents of many of the financial innovations that underpin all modern banking systems, such as chartered corporations, negotiable instruments, and sovereign-debt instruments. The rulers who encouraged innovation were able to create durable nation-states and global trading networks that dominated the rest of the globe. The rulers who failed to do so disappeared, and their territories were absorbed into those of some other sovereign. We then follow the evolution of government-banker partnerships, showing that as political systems became more complex in the nineteenth and twentieth centuries, those partnerships gave rise to a broad array of quasi-government, quasi-private entities, including central banks and special-purpose intermediaries (such as the mortgage repurchase giants Fannie Mae and Freddie Mac).

We then examine real countries over centuries of history to illustrate how the Game of Bank Bargains has been played in different political environments. We focus both on variation in political institutions across countries and, perhaps even more important, on variation in political institutions within countries over time.

We begin our case studies with England. In chapter 4 we show how and why the English government granted a monopoly charter to the country's first banking corporation, the Bank of England, after the Glorious Revolution of 1688—a political revolution that gave Parliament primacy over the crown. England no longer had an absolute monarch who could expropriate funds at will, but because it had an extremely limited elec-

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

toral franchise, a group of wealthy financiers was able to form a durable coalition with the parties in control of Parliament, giving the Bank of England a set of unique privileges in exchange for a series of loans to the government. The result was a monopoly banking system that allocated credit narrowly and was inherently unstable.

In chapter 5 we explore how the Pax Britannica permitted both an expansion of the franchise and a relaxation of the government's need to finance expensive wars, thereby giving rise to political coalitions that favored a greater openness in bank chartering. The stable, efficient, and competitive banking system that Britain had forged by the dawn of the twentieth century was then repressed by the government once again in order to fight the "Thirty Years' War of the 20th Century"—the combination of World War I, rearmament, and World War II. After 1945, political coalitions that favored the creation of a welfare state and nationalized industries made the banking system mostly irrelevant. In short, the banking system operating today in London—which is not just the center of British banking but a hub of global finance—is actually a very recent phenomenon.

In chapter 6 we turn to the United States, covering the period from the Revolutionary War in the 1770s until the repeal of restrictions on interstate banking in the 1980s and 1990s. We include the case study of the United States for a variety of reasons, foremost among which is the illustrative power of U.S. banking history. If there are any readers who doubt our claim that banking regulation is and always has been all about politics, some familiarity with the first two centuries of U.S. banking history should change their minds.

Today, institutions like the Bank of America seem to have a branch or ATM on every corner. Until very recently, however, the Bank of America, like all the other large banks that currently dominate the U.S. market, was legally enjoined from having branches in more than one state: the Bank of America was a California bank. Perhaps more surprising still, until the 1970s, most U.S. states had laws that prevented banks from opening multiple branches even within the state. The result was that the U.S. banking system was composed of tens of thousands of "unit banks" (individual banks, with no branches) operating in thousands of quasi-segmented local markets. No other country had a banking system anything like this, and for good reason: a system composed of tens of thousands of unit

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

banks is inherently unstable because banks can neither spread risks across regions nor move funds easily from one location to another to manage liquidity problems (like bank runs). Such a system is also operationally inefficient, because banks cannot take advantage of scale economies in administration. For all these reasons, Americans paid higher interest rates for loans (and received lower interest rates on their deposits) than they would have in a system of branching banks.

More surprising still, this was decidedly *not* the system that Alexander Hamilton had in mind when he crafted America's first banking institutions in the 1780s and 1790s. Hamilton's vision was undermined within a few decades by a very strange, and very determined, coalition of agrarian populists (who were opposed to corporations of any kind as well as to the elites who controlled them) and small bankers (who knew that they did not have a prayer of competing against big banks that could open branches as they pleased). One reason that this seemingly unlikely coalition was so successful was that it was able to exploit a fundamental institution of the American political system: federalism. Because banking legislation was largely the purview of states, not the central government, the populist-banker coalition could fight and win in the hallways of state capitols rather than face a national political debate. Successful state-level coalitions could then be used to influence the selection of locally elected congressmen to carry their cause to Washington. Every time a crisis wracked its inherently fragile system, this coalition managed to turn efforts at reform to its advantage—even outmaneuvering President Franklin Roosevelt in the writing of the Glass-Steagall Act. As we show in chapter 6, this coalition enjoyed a century and a half of dominance. Ultimately, it was undone by a combination of demographic, economic, and technological changes that undermined unit banks as a business model.

There is no escaping the Game of Bank Bargains: politics always intrudes into bank regulation. Chapter 7 drives that point home by examining how the U.S. banking system, freed of restrictions on branching and competition—a change that should have made the system more stable—became positioned during the 1990s for the spectacular banking crisis of 2007–09. The political coalition between unit bankers and small farmers was replaced by a new coalition between the rapidly growing megabanks and urban activist groups. Bankers had ambitious plans to merge and expand. Their plans were, however, subject to a political constraint: they

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

needed to be judged as good citizens of the communities they served in order for the Federal Reserve Board to approve the mergers. Activist groups wanted to be able to direct credit to their memberships and constituencies, and the "good citizenship" criterion gave them a powerful lever with which to negotiate with merging banks. The bankers and the activists forged a coalition that consolidated the American banking industry into a set of megabanks that were too big to fail. That coalition was formed, among other things, by the contractual commitments of merging banks to channel more than $850 billion in credit through activist groups in exchange for the political support of those activist groups for bank mergers between 1992 and 2007.

In addition to these explicit agreements between banks and activist groups, banks committed an additional $3.6 trillion over the same years in order to rate as good citizens. Many steps were required to make these arrangements work, and we return to them in detail, but one of the most crucial was that the government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac, which repurchased and securitized mortgages, were pressured by the Clinton administration to lower their underwriting standards dramatically so that these loans could become part of Fannie and Freddie's portfolios. Once they consented to doing so, their progressively weaker underwriting standards applied to *everyone*. We cannot stress this point strongly enough: the politics of regulatory approval for bank mergers set in motion a process whose ultimate outcome was that large swaths of the American middle class were able to take advantage of mortgage-underwriting rules that, compared to those of any other country in the world and of earlier periods of America's own history, were inconceivably lax. The result was the rapid growth of mortgages with high probabilities of default for all classes of Americans. To provide just a glimpse of the data, a 2006 survey by the National Association of Realtors found that 46 percent of first-time homebuyers made no down payment at all, and the median down payment for first-time buyers was only 2 percent of the purchase price.[10] In short, the subprime crisis of 2007–09 was the outcome of a series of spectacular political deals that distorted the incentives of both bankers and debtors.

[10] Pinto (2011), 25.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

As we have pointed out, severe bank-insolvency crises require a combination of imprudent lending and inadequate bank capital to back high-risk loans. In chapter 8 we focus on how, exactly, U.S. banks, like the GSEs, were allowed to back their portfolios of risky adjustable-rate mortgages, requiring low down payments and little or no documentation, with capital buffers that clearly were inadequate. Once again the devil is in the details, but one cannot escape the conclusion that the decisions made by regulatory agencies were driven by the logic of politics.

Our goal in examining this episode is not to blame either "Wall Street fat cats" or activist groups for America's 2007–09 banking crisis. Many books make those kinds of arguments—and, frankly, we think that they often miss the mark. The subprime lending crisis was simply the latest in a very long string of American banking crises. What must be explained is why the United States has a banking system that is so persistently crisis prone. What is it about American political institutions that generates incentives for bankers and populists to search one another out and forge such powerful coalitions?

Note, too, that our focus is on the institutions that shape the incentives of individuals and groups, not the individuals and groups per se. People everywhere (now and in the past) generally pursue their self-interest, and one of the ways they do that is by exercising their political rights. When political institutions encourage them to form coalitions, even with unlikely partners, in the pursuit of mutual advantage, they will do so. When political institutions offer only weak incentives for these parties to join forces, they will not do so. No individual or group of individuals is to blame for what happens. Indeed, placing blame is inherently unproductive because it distracts people from the important question: how could we change political institutions to reduce the incentives to form socially unproductive coalitions?

Chapter 9 highlights the crucial role played by politics in the organization of the U.S. banking system by contrasting it with that of Canada. Indeed, Canada presents something of a counterfactual experiment: it has a colonial and cultural heritage similar to that of the United States, but it does not have the particular set of political institutions and circumstances that created the United States' bizarre system of unit banking. Rather, Canada's political institutions were purposely constructed in such a way that almost all economic policies and regulations, including those pertain-

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

ing to banking, had to be decided by a national, bicameral legislature, one of whose houses was appointed, not elected.

Although Canada, too, certainly had its share of populists and would-be unit bankers, they could not succeed in controlling the banking system, because they had to win their political struggles all at once at the national level, within a parliamentary structure that was designed not to be easily controlled by populist factions. The structure of the Canadian banking system was therefore strikingly different: from its beginnings, it was characterized by a small number of very large banks with extensive national networks of branches. The owners of those banks were never drawn into coalitions with Canadian populists to create and share rents at the expense of everyone else. The result has been not just lower costs of credit in Canada but also a much more stable banking system. Since the 1920s, the United States has suffered three systemic banking crises—the widespread bank failures of the Great Depression, the savings and loan crisis of the 1980s, and the subprime crisis of 2007–09—while Canada has suffered none.

Our examination of Mexico in chapters 10 and 11 allows us to explore the differences between banking systems in authoritarian and democratic political systems. Unlike Britain, the United States, and Canada, where elections have been part of the political system for centuries, and in which the right to vote was gradually expanded during the nineteenth and early twentieth centuries, Mexicans were denied the right to effective suffrage until the late 1990s. During most of its history, Mexico has been governed by one type of authoritarian system or another, and on several occasions those governments have engaged in either partial or total expropriations of the banking system. Thus the case of Mexico not only allows us to understand how authoritarian political leaders form coalitions with bank insiders and minority shareholders to create a banking system but also to understand the conditions under which autocrats break those coalitions by seizing the wealth of those same insiders and minority shareholders. The Mexican case also raises the question of how governments managed to coax bankers into forming new coalitions to create banking systems despite their history of periodic expropriation. The answer is that throughout Mexican history, the government tightly regulated bank entry in order to drive up rates of return high enough to compensate bank insiders and shareholders for the risk of expropriation. In Mexico, as elsewhere, bank-

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

ing was all about politics. Mexican political outcomes oscillated between periods of chaos (like civil wars, during which banking systems collapsed) and periods of relative calm, during which crony banking systems comprising a small number of banks allocated scarce credit among politically influential insiders. This pattern has only been broken since 1997, after Mexico began to democratize and the government opened up the banking system to foreign entry.

When fundamental political institutions change, the Game of Bank Bargains changes. There is perhaps no better case to test this proposition than Brazil, the subject of chapters 12 and 13. For most of its history as an independent country, Brazil has been governed by one form of autocracy or another. Indeed, it was not until 1989 that Brazil staged its first direct election for the presidency under rules of universal adult suffrage. Since that time, Brazil has been a stable democracy, but one in which strong populist currents dominate. Indeed, Brazil's political institutions combine features that tilt politics heavily in favor of populist constituencies: a strong president and a weak legislature; centralized tax collection, but decentralized government spending; centralized political parties; universal suffrage for persons over the age of 16; and a constitution that specifies a long list of "positive rights."

Can we then point to any discernible differences in how banks were regulated and operated under autocracy and democracy? The differences are not just discernible; they are dramatic. For most of Brazilian history, autocrats were able to use the Brazilian banking system as little more than a mechanism to levy an inflation tax. There was nothing subtle about this practice: at its peak in the late 1980s, inflation ran at nearly 2,500 percent per year, and the banks and the government split nearly 8 percent of GDP between them in inflation-tax revenues. There have been two exceptions to this pattern. The first was during the period 1831–89, when local oligarchies were able to constrain the Brazilian emperor, and there was barely any banking system at all. The second has been the years since Brazil democratized in 1989. Brazil's democratically elected governments quickly brought an end to inflation taxation; by the mid-1990s, inflation had fallen nearly to U.S. levels, a shift that forced banks to get into the business of actually lending money rather than just earning income off the float on checking accounts (the profit a bank makes by not paying interest to depositors while a check clears).

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Populism creates its own set of demands on the banking system, however. Brazilian banks are no longer inflation-tax machines for the government, but some of them are now employment-generation machines. Two of the largest banks in Brazil are run by the government (although one of them has private shareholders). These banks allocate credit not according to market criteria but to help candidates win office by making sure that those candidates' business allies in politically contested districts have enough credit to maintain or increase levels of employment.

Chapter 14 summarizes the principal findings of the previous chapters and links the histories of our five case studies with the experiences of other countries to show that our approach has broad explanatory power.

In chapter 15 we consider what our findings have to say about the feasibility of libertarian policy advocacy, the limits of regulatory reform, the validity of various theoretical models of banking crises, and the durability of political institutions.

What We Don't Do in This Book, and Why

Some readers may wonder why we focus on banks rather than examining financial markets as well. We recognize, of course, that the history of financial markets and the institutions specific to them (for example, stock exchanges) also reflects important dimensions of financial-system development and can be the locus of financial instability (for example, sovereign-debt crises and stock-market crashes), and we explore some of those important connections in our historical narratives. We concentrate on banking systems, however, for a simple reason: financial markets have always been created and sustained by banks. No financial system has developed bond and stock markets without first developing a banking system. Brokers and dealers either are banks or rely on banks for the credit to manage their underwriting and trading activities, and the firms that access financial markets do so only after long periods in which they become seasoned prospects as the result of their interactions with banks. Conceptually, it is sometimes hard even to distinguish between bank-intermediation and financial-market instruments. Bills of exchange, for example, are bank IOUs, but they are also tradable financial instruments. Some early banks, such as the Bank of England, began as little more than sovereign-debt restructuring devices, and holding stock in those banks was akin to invest-

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

ing in sovereign debt. Throughout history and around the world, banking has been germinal and central to financial development of all kinds.

Furthermore, to the extent that financial markets can offer alternatives to banking finance, those alternatives tend to be constrained by the same factors that limit the supply of bank credit. In many countries the risk of expropriation by the government limits the ability of firms to use equity markets as a means of raising capital. This risk tends to be mitigated by managers' engaging in rent-seeking activities with members of the government itself, a practice that has the effect of reducing the amount of capital that can be mobilized by markets.[11] In short, securities markets also operate within a political context and do not allow an end run around the tight political constraints that limit the operation and performance of banks.

Other readers may wonder why we focus on chartered banks—the enterprises known as commercial banks, corporations that take deposits, make loans, and seem to place their ATMs just about anywhere they can rent six square feet of space—rather than private banks (also known as merchant banks or investment banks), such as the House of Rothschild or J. P. Morgan & Co. Private banks have existed since the invention of money in antiquity. The amount of capital that they could mobilize, however, was always constrained (see chapters 2 and 3). As chartered banks came into being, private banks deployed their skills and reputations to become coordinators of financial networks: they arranged lending syndicates, brokered international financial transactions, underwrote and distributed securities, and helped govern large nonfinancial corporations. The key credit suppliers of the modern era, however, have been chartered banks.

As we wrote this book, we were often asked by colleagues, referees, and students why we confined ourselves to a small number of case studies. Why five countries, instead of eight, ten, or twenty? As a practical matter,

[11] Haber, Razo, and Maurer (2003), chapters 2, 4, and 5, focus on the impact of rent seeking on the number of firms in any industry that can mobilize capital through securities markets. In their framework, rent seeking limits firm entry and exit. Stulz (2005) focuses on the impact of rent seeking on the amount of stock that any firm can offer to minority shareholders. In his framework, rent seeking requires a small number of powerful shareholders as a bulwark against government predation. These mechanisms are not mutually exclusive.

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

exploring how a country's institutions have changed over time is not an enterprise characterized by increasing returns to scale; there is an obvious tradeoff between the number of cases covered and the depth with which they can be discussed. Moreover, our purpose is not to provide an exhaustive history of the political economy of banking in every country around the globe; rather it is to develop and illustrate a general framework that we believe has wide explanatory power. We invite other researchers to test that framework against additional country cases as well as against large multicountry data sets.

Still other colleagues and students asked why we bothered with hundreds of years of history: why not just study the past couple of decades? For that matter, why not make things simpler yet and focus just on the last couple of decades in one country—a research strategy that has been employed to good effect by a number of scholars who have sought to unravel the causes of the U.S. subprime lending crisis?

Our answer is simple in principle, but the necessary methods are not simple in practice: when people want to understand the factors that shape complex systems and shift those systems from one equilibrium to another, they undertake multidimensional analyses—and one of the dimensions on which they focus intently is time. This is what a physician does, for example, when she is faced with a difficult diagnosis. She starts with fundamental, time-invariant factors, such as the patient's genetic predisposition toward particular diseases as deduced from a family medical history. She then looks at the patient's own medical history: what illnesses has the patient had in the past, what pathogens has she been exposed to, and what has been her health trajectory? She next gathers information on factors that affect health in the short run: diet, exercise, stress. The physician also gathers data from direct observation: tests of metabolic function, palpation of organs, and a review of symptoms. Finally, the physician compares the information from this patient to information from other patients she has treated and assesses all of it within a logical framework.

This approach allows her to rule out certain hypotheses because they are inconsistent with the timing of events (e.g., if A happened after B, then A could not have caused B); because they are inconsistent with comparative evidence (if A did not cause B in other cases, then A likely did not cause B in this case either, even if A preceded B in time); or because they are unlikely given the patient's underlying, time-invariant, characteristics

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

(if no one in a patient's family has ever had disease X and there is a strong genetic component to that disease, then disease X is likely not the cause of the patient's symptoms). What is true of medicine also holds in the natural sciences that rely on observational methods (e.g., epidemiology, astronomy, and evolutionary biology) and in which accurate causal statements are important. It should be no less true in the social sciences.

Indeed, the medical diagnostic analogy is particularly apt for studying the world's banking systems, where good health has been the exception and an almost pathological combination of unstable and scarce credit has been the rule. As is frequently the case in medicine, the causes of good health are fundamental: stable banking systems that allocate credit broadly are an outcome of political systems characterized both by broad suffrage and constraining institutions that limit the incentives for bankers to form coalitions with populists. Understanding why this is the case requires a deep historical exploration into the origins and consequences of those political institutions. To that exploration we now turn.