CHAPTER 14

Financial Crises

For financial market participants, financial crises are very difficult to anticipate and plan for, yet extremely destructive when they occur. From the standpoint of public policy, there is insufficient agreement on how to prevent or cope with crises. For researchers, financial crises remain a diffuse area, with poor problem definition, yet intense controversy. Crisis episodes present the most dramatic departures from the standard model of asset price behavior. The subprime crisis, by some measures, is the most severe financial crisis since the Great Depression of the 1930s, and will drastically change the financial system in ways not yet known.

Financial crises can be summarized as episodes displaying some or all of these symptoms or observable hallmarks:

Asset prices change sharply. The changes in asset prices are often a large multiple of the recently observed volatility, and thus would be judged extremely low-probability by the standard model.

Return moments change sharply. Volatility increases for almost all assets. Correlations "break down," that is, change from previous values, drawing much closer to ± 1, or to 0, or changing sign.

Aggregate credit shrinks drastically. Balance sheets shrink rapidly as assets lose value, lenders seek to reduce leverage, the credit intermediation mechanism is impaired, and borrowers are forced to curtail activity.

Market liquidity conditions deteriorate and liquidity impasses arise. Transaction volumes for many assets spike briefly, and then decline and remain low.

Payments systems integrity may be impaired. The "plumbing" of financial markets is intertwined with credit intermediation, as seen

in Chapter 12, and may be dependent on the financial viability and creditworthiness of a small number of firms. Payment systems can therefore be affected by credit and liquidity problems.

Systemic risk increases; while ill-defined, it is part and parcel of a crisis.

Contagion: Dysfunction can spread to related and unrelated markets, via sentiment, hedging, and sharing the same or similar fundamentals.

Market functioning is impaired, that is, the price mechanism—security prices, interest rates, and credit spreads—cease to be the chief mechanism by which capital is allocated. Rather, various types of credit rationing come to the fore.

Economic activity, output and employment decline sharply.

Duration: These shocks and the credit contraction can last a long time. The depression following the initial financial panic of 1873, for example, lasted nearly a quarter-century.

Financial crises commonly break out openly upon a severe decline in some asset prices, a major default by a government or large financial intermediary, disruption of currency markets, or all three. But crises are the manifestation of longer-standing and less-visible problems. The causes of credit and currency market disruptions go to some of the oldest and deepest controversies in economics. In this chapter, we will discuss crises in the context of actual episodes, with special attention to the still ongoing subprime crisis.

Financial crisis is not a precisely defined technical term, but rather loosely describes certain historical episodes. Crises vary widely in severity and in the nature of the disruptions. One difficulty in studying crises is that the most severe crises are quite rare. There is, of course, disagreement as to which episodes of the past few centuries qualify as truly extreme, but, prior to the subprime crisis, the crises and associated depressions that began in 1793, 1825, 1873, and 1929 are often cited. Data on all but the last are far sketchier than that available for recent decades, making generalizations more difficult.

The paradigmatic financial crisis, at least prior to the subprime crisis, was the Great Depression, which lasted several years beginning in 1929—by some accounts over a decade—and led to tremendous suffering over large parts of the world. But the term "crisis" is also applied to events such as the breaking of the European Monetary System in 1992–1993, which had relatively limited consequences. The term "extreme events" is often used

in the context of financial crises. We reserve the term "crisis" for more general and widespread episodes of distress, and describe episodes that were contained to just a few markets or a short period of time, such as the stock market crash of October 1987, as extreme events.

Our discussion starts with some "stylized facts" about the behavior of financial markets during crises, drawn from episodes of the past few decades. We then discuss theories about the causes and prevention of crises. We also see how researchers have defined and classified crises, and distinguished them from other extreme events in financial markets. We then turn to the issue most directly relevant for risk managers: Can crises be anticipated?

14.1 PANICS, RUNS, AND CRASHES

Two of the most characteristic features of financial crises are the

Credit crunch, a pervasive withdrawal of credit and decline in willingness to lend. Wojnilower (1985) defines a credit crunch as a "blockage in the supply of credit—a sudden and unanticipated intensification of nonprice rationing" (p. 351). At the time the term was introduced, it referred largely to bank lending, but fairly describes the behavior of nonbank intermediation as well.

Liquidity crunch, an intense increase in liquidity preference and the desire to exchange other financial assets for money. In the most extreme cases, all but the closest money substitutes are rejected in favor of cash or precious metals.

In this section, we describe these and other typical overt manifestations of crises. In the next section, we focus on specific self-propagating mechanisms by which financial crises take hold and worsen.

14.1.1 Monetary and Credit Contraction

A sudden decline or even a collapse in measures of financial activity is a universal hallmark of severe crises. But from a welfare point of view, the decline in real economic activity is the most important manifestation of a crisis. Unemployment rises, economic growth slows or becomes negative, and output declines, causing widespread suffering. The deterioration in business conditions tends to be worse and longer-lasting, the more badly the financial system is impacted.

Credit Expansion and Contraction A severe contraction in monetary and credit aggregates is often the most salient characteristic of a crisis. The period before the onset of a crisis typically sees strong credit expansion. In a commercial bank dominated financial system, that meant an expansion of bank lending and balance sheets. In a financial system in which capital markets and securitization play a large role in addition to that of banks, much of the credit expansion takes place through growth in the outstanding volume of securities. A crisis may occur when credit expansion abruptly slows, or it may deepen when a crisis that has begun for other reasons is exacerbated by a sharp reduction in the pace of credit expansion.

Credit growth during expansions is generally smooth and gradual, apart from recoveries immediately following credit crunches. Credit contractions are quite abrupt. The subprime crisis led to an extraordinarily sharp contraction in credit. Figure 14.1 displays net borrowing in credit markets, a measure of changes in the total volume of credit extension in the United States. It turned negative in 2009 for the first time since the end of 1946, but this time because of a sharp decline in the volume of intermediation by the

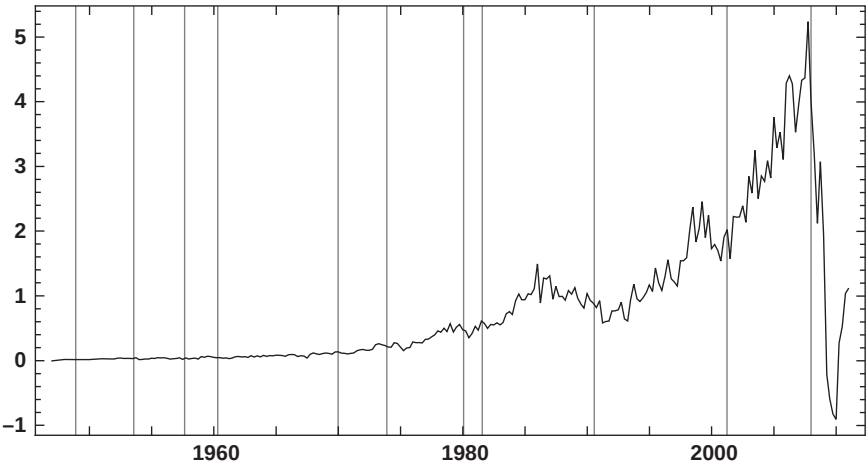

FIGURE 14.1 Net Borrowing in U.S. Credit Markets 1946–2010
All sectors, seasonally adjusted, quarterly, Q4 1946 to Q4 2010, trillions of dollars. This time series is a flow concept summarizing the net creation of credit in the United States. The vertical grid lines represent the dates of business cycle peaks, when contractions begin, as determined by the National Bureau of Economic Research (NBER).
Source: Federal Reserve Board, Flow of Funds Accounts of the United States (Z.1), Table F.1, line 1.

financial sector, rather than because of the cessation of federal government borrowing to finance World War II.

Contraction of Bank Lending Regulatory restrictions aside, bank lending is profitable as long as the interest on loans, adjusted for credit risk, exceeds the cost of capital. At times, however, banks face additional balance-sheet constraints on lending. Funding constraints occur when banks have difficulty borrowing from depositors or longer-term lenders. They may be unable to borrow at all, or the interest rates at which they can borrow rise suddenly and drastically. Balance-sheet constraints can become binding when banks are eager to preserve capital and decrease leverage, because they anticipate higher default or mark-to-market losses, or feel insecure about maintaining short-term funding as it rolls off. Regardless of net interest margin or prospective return on capital, at such times, they prefer at the margin to reduce lending and husband equity. These "non-price," i.e. non-interest rate constraints on lending distinguish credit crunches.

Figure 14.2 illustrates the behavior of bank lending, including business lending as well as mortgages and consumer loans, during economic downturns. Over the course of a credit cycle, the annual rate of growth of lending fluctuates widely, with high growth rates well in excess of 10 percent

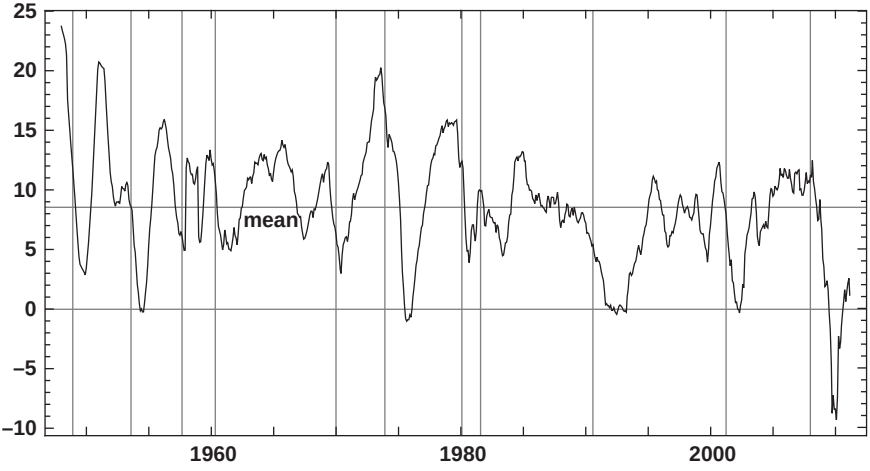

FIGURE 14.2 Growth in U.S. Bank Credit 1947–2011
Annual growth rate of loans and leases in bank credit of all commercial banks, seasonally adjusted, monthly data. Each point represents the logarithmic change, in percent, over the prior year. Vertical grid lines represent the dates of NBER business cycle peaks.
Data source: Federal Reserve Board, H.8 data release.

and an average just over $7\frac{1}{2}$ percent. In recent years, the growth rate has been somewhat lower, as capital markets and securitization accounted for much of the incremental growth in credit in the financial system as a whole.

Contractions in bank credit extension took place on five occasions during the last 65 years. By far the most severe—a decline of nearly 10 percent—occurred as the subprime crisis entered its intense phase following the Lehman bankruptcy. The overall contraction in credit was even more severe than Figure 14.2 indicates, since there was also a complete halt in credit intermediation through securitization from the fall of 2008. As we will see in a moment, the recorded decline in bank credit was somewhat attenuated by an "involuntary," but transitory, return of intermediation to the banking system.

A credit crunch can be amplified by liquidity concerns similar to those provoking a bank run. During a crisis, even financial intermediaries that are in no immediate danger of collapse become concerned about the potential demands on them for liquidity and about how their creditworthiness will be perceived. They therefore become eager to accumulate liquidity reserves of cash and highly marketable securities such as Treasury bills and notes, and reluctant to put money to work by lending.

Credit crunches are often "capital crunches," a term that describes one motivation for a credit crunch. Banks become more reluctant to extend credit, not in the first instance because they have become more concerned about the creditworthiness of the borrower or the use of the funds, but because they are concerned that their equity capital is insufficient. This was the case following large bank loan losses in the wake of the S&L crisis of the late 1980s, which led to the large recession of 1990–1991. It was very much the case during the subprime crisis, driven by the possibility of large future losses, over and above any losses that had already been realized. Uncertainty about the value of investment-grade securitized credit products, of which some banks had large holdings, was one major factor. Another was uncertainty about how long and severe the economic downturn would be and how severe the consequent loan losses would be.

Apart from a sharp decline in the volume of credit extended, a credit crunch is also characterized by a change in the terms on which credit is extended, which include not only the credit spread charged on a loan, but also the underwriting standards. In the United States, data on these terms is collected via the Federal Reserve's Senior Loan Officer Opinion Survey on Bank Lending Practices, which polls loan officers at large U.S. and foreign banks located in the United States. As seen in Figure 14.3, the reduction in credit volume was associated with a dramatic tightening of credit terms, both with respect to underwriting standards and pricing: By the end of 2008, three-quarters of respondents were tightening standards, and nearly all were

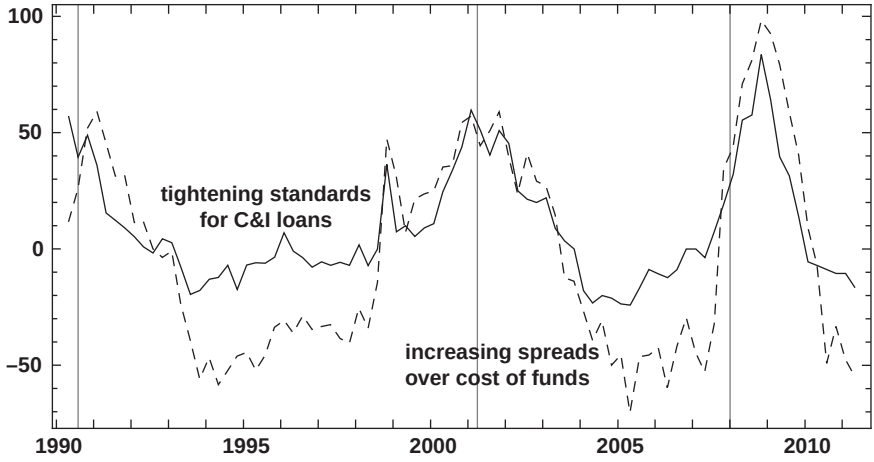

FIGURE 14.3 Tightening of Credit Terms 1990–2011
Quarterly data through Q2 2011. The solid line plots the net percentage of respondents tightening standards for C&I loans to large and middle-market corporate borrowers. The dotted line plots the net percentage of respondents increasing spreads of loan rates over banks' cost of funds for such loans. "Net percentages" equal the percentage of banks that reported tightening standards minus the percentage of banks that reported easing standards. The vertical grid lines represent the dates of NBER business cycle peaks.
Source: Federal Reserve Board, Senior Loan Officer Opinion Survey on Bank Lending Practices, Table 1.

increasing credit spreads. As with all market-clearing prices and quantities, it is difficult to establish empirically whether a contraction in the volume of credit during a downturn is a credit crunch, that is, the result of abrupt restriction by banks, or of lack of demand for credit brought about by bad business conditions. The tightening of terms is evidence that the reduction is induced by a credit crunch, and not only by lack of demand.

Figure 14.3 also illustrates the positive correlation between the pricing of loans and underwriting standards. When funding is easy, underwriting standards are loosened, and spreads are reduced.

The most severe form of credit contraction is caused by widespread bank runs, which we've discussed in Chapter 12 and return to just below. A bank run is a type of funding constraint. It occurs when all or a sizable fraction of depositors simultaneously demand the return of the funds they've lent a bank. Bank runs occur sporadically as isolated events even in quiet times. But in the most serious financial crises, they can become widespread.

A bank run often leads to the failure of the impacted bank. In Chapters 1 and 12, we discussed the fragility of banks, which carry out credit

intermediation via maturity transformation, that is, by using their balance sheets to turn short-term deposits into long-term loans to firms. They also carry out a liquidity transformation, turning illiquid assets such as loans into liquid deposits. Under fractional-reserve banking, liquidity reserves may be quite small compared to a bank's deposit base, and can be quickly depleted during a run. Once this happens, the bank becomes insolvent.

Fear of banks runs, even if they do not actually occur, can also have an impact during a crisis. These fears contributed to the widening of interbank lending spreads during the subprime crisis, as seen in Figures 14.9 and 14.10.

While bank runs are a typical phenomenon of financial crises, crises can also lead to a consolidation of intermediation within the commercial banking system. Banks are fragile, but the rise of market-based lending has placed some of the most fragile parts of the financial system outside of banks. Highly leveraged transactions in structured credit products financed via short-term borrowing, for example, are even more readily disrupted than bank lending.

In addition, banks function as "lenders of second resort," a phenomenon also called "involuntary reintermediation" that describes the increase in bank lending resulting from credit implicitly or explicitly granted prior to a credit crunch. The phenomenon was quite pronounced during the subprime crisis. One visible sign was the increase in drawnings on banks' lending commitments. Banks also felt obliged in many cases to assume onto their balance sheet securitization vehicles that could no longer be funded in the asset-backed commercial paper (ABCP) market, such as the ABCP conduits and structured investment vehicles (SIVs), described in Chapter 12; many of those without bank-provided backstops simply collapsed. Figure 14.4 displays the brief but sharp increase in non-mortgage bank lending to businesses in the immediate aftermath of the Lehman bankruptcy that preceded the protracted decline in lending.

Bank lending grew steadily until the panic phase of the crisis at the end of 2007. It then leveled off for several quarters. One would have expected, following the Lehman bankruptcy and the beginning of the most intense phase of the crisis, that bank lending would immediately contract sharply. The contraction did indeed occur, but not for another quarter. In fact, the immediate effect was a sharp increase in bank lending in October 2008, as borrowers dependent on access to securities markets drew on bank-based sources of financing such as credit lines and revolving credit agreements, and banks provided on-balance-sheet funding for off-balance-sheet vehicles that could no longer be financed in the capital markets. Only once these explicit and implicit guarantees by banks to extend credit had been fully met, could banks begin to effect a decrease in their lending assets.

Another example of the "lender of second resort" phenomenon that appeared during the subprime crisis is displayed in Figure 1.6 of Chapter 1.

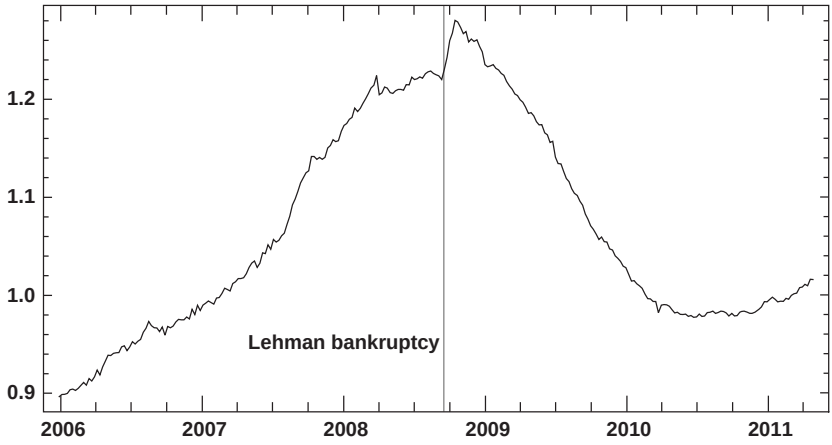

FIGURE 14.4 U.S. Bank Lending during the Subprime Crisis 2006–2011
Commercial and industrial loans, all commercial banks, seasonally adjusted, weekly, trillions of dollars.
Source: Federal Reserve Board, H.8 data release.

With market-based intermediation paths such as securitization collapsing, as we will see in a moment, banks took on a larger share of the remaining intermediation activity for some types of credit. As the share of securitization via commercial mortgage-backed securities (CMBS) in commercial real estate lending declined after its peak in 2007, for example, banks' share grew by about the same fraction of the total.

Contraction of Securities Markets Not only bank lending, but also other forms of credit intermediation decline rapidly during crises. We have seen one important example, the reduction in the volume and shortening of the maturities of commercial paper issued by financial firms following the Lehman bankruptcy, in Chapter 12.

Access to credit for ABCP conduits and SIVs, off-balance-sheet vehicles set up to finance loans and structured credit products via short-term funding, contracted even more rapidly than for financial firms during the subprime crisis. Short-term lenders to intermediaries and ABCP conduits through the commercial paper market responded to concerns about the creditworthiness and liquidity of the borrowers by stepping back from the market as their holdings rolled off. Commercial paper issuance by financial firms and ABCP conduits together accounted at the mid-2007 peak for over 90 percent of total U.S. commercial paper issuance and declined by over 50 percent over the subsequent three years, as seen in Figure 14.5. Part of this decline is mirrored in the post-Lehman "involuntary reintermediation" seen in Figure 14.4.

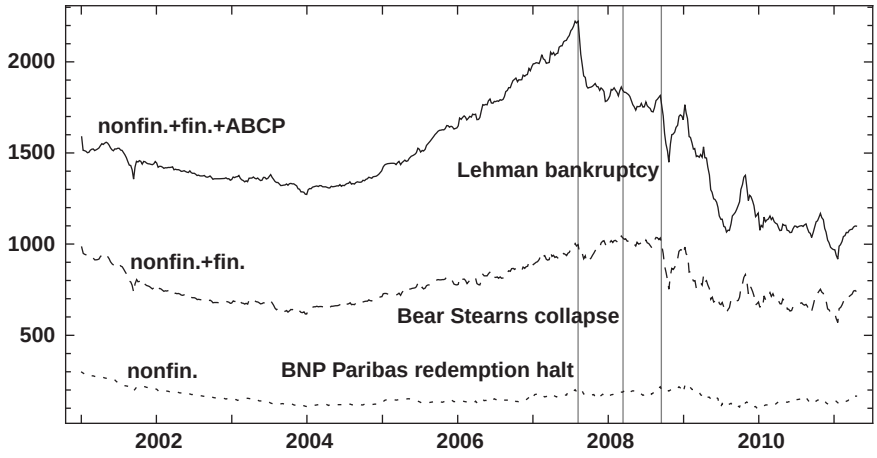

FIGURE 14.5 Outstanding Volume of Commercial Paper 2001–2011
Seasonally adjusted, billions of dollars, weekly, Jan. 10, 2001 to Apr. 20, 2011.
Source: Federal Reserve Board

Among the key investors in financial and ABCP were non–government-backed or "prime" money market mutual funds (MMMF). Institutional investors fled such funds following the Lehman bankruptcy in the MMMF equivalent of a bank run, as illustrated in Figure 14.7. Commercial paper credit spreads also rose sharply during the crisis.

Bond market issuance also declines during a crisis. This has a potentially much larger impact on real economic activity today in view of the greater importance of capital markets relative to bank lending in aggregate credit intermediation. Figure 14.6 displays bond issuance in three important sectors. The top panel displays issuance of ABS outside the real-estate sector, the center panel, issuance of residential and commercial real-estate related ABS, and the bottom panel issuance of high-yield corporate bonds. All show a drastic decline in 2008 and 2009.

The real-estate sector shows the most dramatic decline, from about $1.4 trillion in each of 2006 and 2007 to near-zero since. ABS not related to real estate and high-yield bond issuance each fell more than 50 percent in 2008. These declines mirrored a credit crunch in the nonbank sector. As the investor base for securitized loans collapsed, the capacity to originate underlying loans collapsed with it. High-yield bond issuance, however, not only recovered, but even reached a record pace in 2010. This is in part due to firms, including nonbank financial intermediaries, issuing unsecured debt rather than ABS. The underlying assets are then financed on the balance sheet rather than being moved off-balance-sheet to serve as collateral for specific

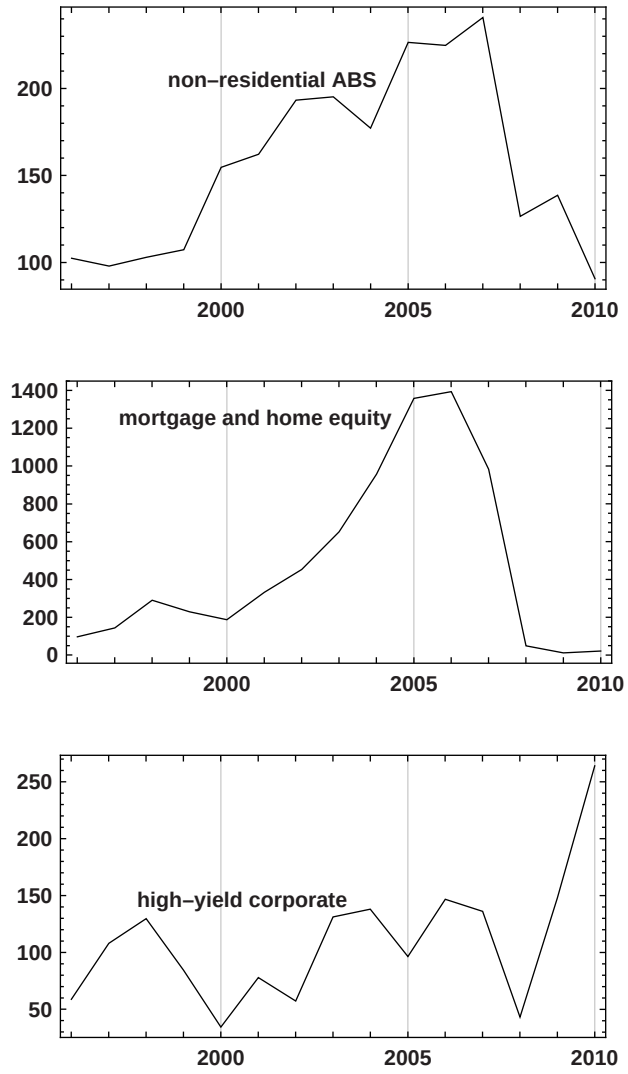

FIGURE 14.6 U.S. Bond Issuance 1996–2010
Bond issuance volume in the United States, billions of U.S. dollars, annual.
Upper panel: Issuance of ABS backed by auto loans, credit card receivables, equipment, and student loans.
Center panel: Issuance of private-label RMBS and CMBS, and of ABS backed by home-equity loans and manufactured housing.
Lower panel: Issuance of high-yield corporate bonds.
Data source: Securities Industry and Financial Markets Association, available at http://www.sifma.org/ research/statistics/statistics.aspx.

bonds. But it also reflected the weakness of banks during the subprime crisis and the ability of larger firms with direct access to capital markets to find alternatives to bank lending.

14.1.2 Panics

Banks runs have been a frequent marker of financial crises for centuries since the development of fractional reserve banking. We have discussed runs as an extreme form of the contraction of credit volumes. Now we look at runs from the standpoint of the intermediaries from which funding is withdrawn. The classic form of a run is the withdrawal of deposits from commercial banks. In the modern financial system, as noted, runs also occur in other forms: Nonbanks that are relatively dependent on wholesale funding can also experience runs in the less-visible form of a withdrawal of short-term credit.

Bank Runs and Run-Like Behavior An instance of a run that figured importantly in the early history of the subprime crisis and its spread beyond the United States was the run on Northern Rock, a British building society, or mortgage lender, which began after the lender sought emergency liquidity support from the Bank of England on September 13, 2007. The next day, the classic overt symptom of a run appeared, as depositors lined up outside the bank's branches to withdraw funds. However, what amounted essentially to a run of its wholesale funding sources had already begun. Northern Rock was exceptionally vulnerable to a run, as its leverage, measured as the ratio of assets to common stock, was close to 60 by mid-2007. A loss in the value of its assets, composed largely of mortgage loans, of even $1\frac{2}{3}$ percent would therefore have sufficed to wipe out its equity and begin to impair lenders. Under these circumstances, as soon as doubts arose as to the true value of its assets, Northern Rock was unable to roll over its short-term debt.

During the subprime crisis, runs took unusual forms not seen in the past that reflected changes in the financial system. We have seen, in Figure 12.5, the drastic reduction, over just a few weeks beginning in August 2007, in the outstanding volume of ABCP. Conduits and SIVs experienced a rollover risk event and were unable to replace ABCP borrowing as it matured. The sharp rise in draws on existing bank lines is another example of an unusual type of bank "run" by obligors. In Chapter 12, we saw that banks' and broker-dealers' prime brokerage business with hedge funds was an important source of cash funding, and of collateral that could be rehypothecated and thus used as a funding source. As doubts about Bear's and Lehman's liquidity and solvency grew in the course of 2008, they rapidly lost their hedge fund customers, and with them, access to the cash and collateral the hedge funds deposited with them.

Another unusual run-like phenomenon occurred in the credit default swap (CDS) markets. Counterparties became eager to novate outstanding CDS away from Bear and Lehman. This sent an adverse signal about market perceptions of these broker-dealers' strength, but also, to the extent novation could be achieved, deprived Bear and Lehman of the cash collateral associated with the CDS: Prior to the subprime crisis, CDS margin would have been disproportionally in the hands of large broker-dealers, rather than symmetrically in the hands of counterparties with positive net present value (NPV) positions.

Other run-like phenomena affected nontraditional intermediaries. We noted in Chapter 12 that money market mutual funds are exposed to liquidity risks similar to those of depository institutions. Immediately following the Lehman bankruptcy, a run on some MMMFs took place. It was triggered by reports that Reserve Primary Fund Class Institutional, a large and prominent MMMF, had invested a significant proportion of its assets in Lehman Brothers Holdings' commercial paper and other debt securities. There was, as noted earlier, also concern about MMMF holdings of other financial and asset-backed commercial paper.

Reserve Primary suffered massive redemptions and "broke the buck," reporting an NAV of $0.97 on September 17, 2008. Other MMMFs avoided breaking the buck, but had large redemptions, mainly by institutional investors, and reallocations by investors away from MMMFs investing mainly in commercial paper and other corporate, financial, and asset-backed issues, and into funds investing primarily in short-term government debt (see Figure 14.7). The institutional investor rejection of these assets also had important consequences for other parts of the financial system. ABCP was an important source of funding for both structured credit products, via structured investment vehicles (SIVs), and for the underlying loans, via ABCP conduits. The disappearance of the ABCP investor base thereby contributed to the shutdown of securitization markets in the fall of 2008. Part of that investor flight was indirect, through withdrawals from MMMFs.

Run-like behavior was also observed for other nonbank market participants. For example, hedge funds experienced large redemptions by investors, as discussed in Chapter 12. Some of these withdrawals were motivated by fear of losses. But in many instances, withdrawals were especially "run-like" in that they were motivated by investors' desire for liquidity and by their awareness of the sequential-service constraint on the funds. That is, they wished to retrieve liquidity from otherwise sound hedge funds before it was exhausted by other investors' redemptions.

Dealing with bank runs once they have begun is exceptionally difficult. Two drastic approaches are closing banks, as done for example during the "bank holiday" of 1933, and by guaranteeing deposits, or even banks'

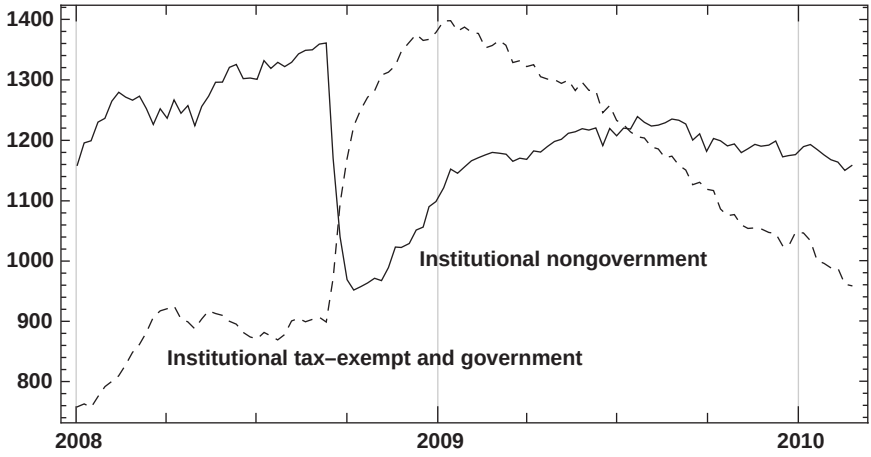

FIGURE 14.7 Institutional Investor Assets in MMMFs 2008–2010
Weekly, billions of dollars.
Source: Investment Company Institute.

debt generally. Early in the Swedish banking crisis of 1992, for example, the Swedish government guaranteed all bank debt. Ireland issued a similar guarantee in September 2008.

Liquidity Impasses and Hoarding The Northern Rock experience of a retail depositor run was not, in fact, typical of the subprime crisis. Widespread depositor runs, the most-feared form of financial crisis historically, in which depositors, particularly retail ones, seek the immediate return of their funds, did not occur; instances, such as the run on mortgage-focused IndyMac Bancorp in late June 2008, remained isolated events. But Northern Rock's demise combined this feature of classic bank runs with a feature more typical of the subprime crisis, namely, the withdrawal of wholesale funding to the financial sector, as lenders sought to preserve their own cash balances and became extremely wary of the financial condition of intermediaries. Even where the reluctance to provide funding did not lead to the collapse of an intermediary, it led to an enormous increase in their cost of funding.

Bank runs and run-like phenomena are, in essence, the result of many, if not all, market participants' desire to become more liquid simultaneously. They endeavor to keep their assets in the most liquid form possible, and to "term out," that is, to extend the maturity, of their short-term borrowings, in order to reduce rollover risk. They therefore become extremely reluctant to lend out their liquid assets; that is, in the language of the 1930s debates, they "hoard" them. The shift can be very rapid. From a situation in which

most market participants are content to borrow short-term and invest at longer terms or in riskier assets, the markets can within days move to one in which market participants strive to borrow at longer terms and seek to drastically reduce the term and risk of their assets.

This was illustrated when the subprime crisis entered a new and more intense phase on August 9, 2007, following the announcement by BNP Paribas that it was suspending redemptions from three residential-mortgage hedge funds. This was the latest blow to market confidence, following a string of residential mortgage originator bankruptcy filings since the end of 2006, rising mortgage delinquencies on late-vintage loans, and announcements of large losses on subprime loans and securitizations by banks and by hedge funds managed by UBS and Bear Stearns. Concern about potential bank losses became more acute. The next day, the interbank money market began to clearly reflect the disturbance: Widening spreads on unsecured short-term loans to banks relative to other money market rates signaled reluctance to lend to banks. Run-like behavior affected all cash operations of intermediaries and could take unusual forms, such as the hedge fund run on prime brokers and investors' run on viable hedge funds.

In the subprime crisis, as is typical in a financial crisis, the focus of concern was the creditworthiness of banks. One source of particular worry for potential lenders to banks was the disposition of off-balance-sheet vehicles such as SIVs, the liquidity risks of which we described in Chapter 12. A number of SIVs, though by no means all, were sponsored by large banks. The vehicles invested in subprime residential securities and other structured credit products, and financed these investments by issuing a variety of debt instruments, including short-term debt. When concern emerged about the creditworthiness of the bonds the SIVs held as assets, investors grew reluctant to roll over the short-term funding they had provided. This initiated the typical "fire sale" mechanism of an asset price decline, as the SIVs were then obliged to raise capital by selling assets, compounding the price declines and accelerating withdrawals of funding. It also had features of a bank run, in that investors were eager to withdraw funds before the SIVs liquidity or access to liquidity was exhausted by other investors' withdrawals.

While banks were in general not contractually obliged to provide backstop liquidity to the SIVs they sponsored, most felt compelled to do so in order to protect their reputations and avert long-term harm to their business as a whole. This added to the considerable strain on their balance sheets. The SIVs were, even before the crisis, highly levered, so overall leverage of the bank assuming the SIV increased. Bank leverage was also increasing because the SIVs were experiencing losses, eroding the banks' capital.

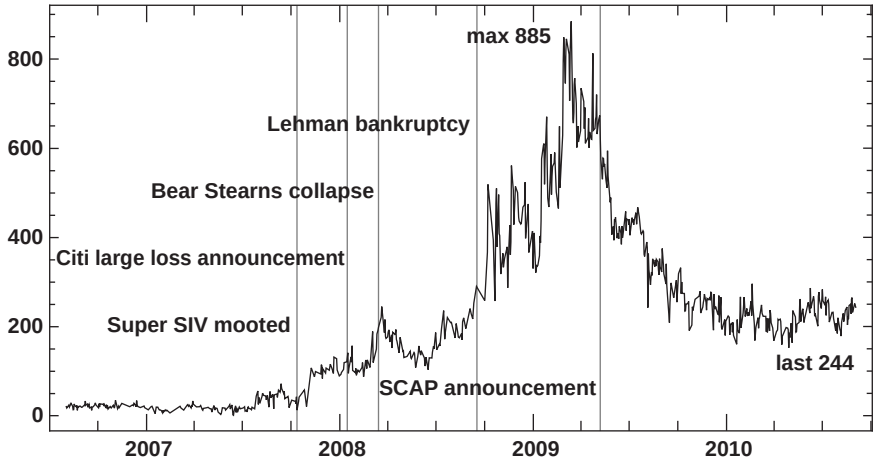

FIGURE 14.8 Citigroup Credit Spreads during the Subprime Crisis 2006–2010
Spread over Libor (*z*-spread) of Citigroup Inc. senior unsecured bonds with maturities of about 10 years, daily. The construction of the data is described in the caption to Figure 13.4.
Source: Bloomberg Financial L.P.

Also typical of a bank run was the difficulty potential lenders to banks had in distinguishing which intermediaries had losses great enough to cast doubt on their ability to repay loans. The initial shock, losses on subprime residential mortgage debt, was known to be large enough to affect a number of firms. While it was widely and firmly believed that at least some banks had experienced serious losses that had not yet been reported, there was considerable uncertainty as to which banks were worst affected. Potential counterparties also took into account not only the correlation of losses brought about by the direct effect of the shock, but also correlation indirectly induced by counterparty withdrawal even from firms less affected by subprime losses, since these could not be accurately ascertained in real time. Rumors and adverse information on the financial health of intermediaries play an important role in crises.

Figure 14.8 illustrates the extent and sudden onset of distrust of banks as borrowers. It displays the senior unsecured funding spread of Citigroup, one of the largest banks. Prior to the onset of the crisis, it paid only a narrow spread, often less than 20 basis points, for long-term unsecured funding. Its debt spread began to widen in late 2007, as awareness grew of the extent of off-balance-sheet funding it had carried out via SIVs, among other asset-quality issues. In late 2007, the U.S. Treasury abandoned a plan to address the problem for Citi as well as other large banks. Citi's cost of

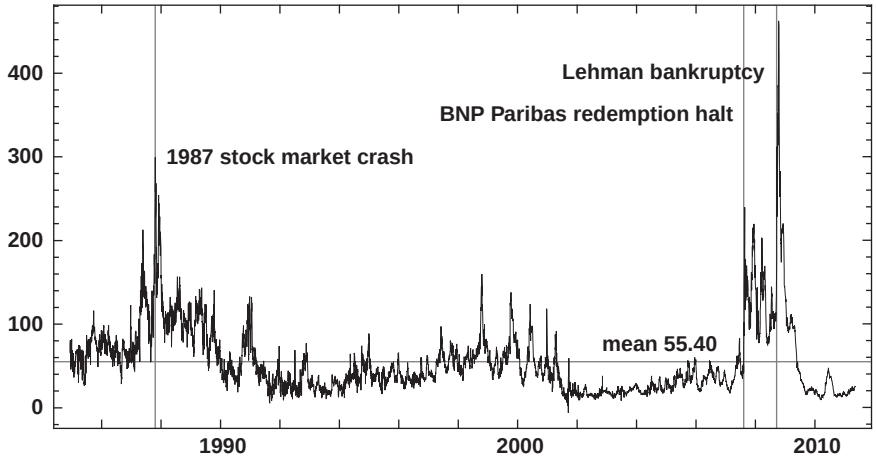

FIGURE 14.9 Three-Month TED Spread 1985–2011
Spread of three month USD BBA Libor minus the yield of the three month Treasury bill, basis points, daily, December 6, 1984 to May 11, 2011.
Source: Bloomberg Financial L.P.

funding spiked briefly following its first announcement of extremely large losses of over $18 billion on January 15, 2008, and then again amid the Bear Stearns collapse. But the funding spread climbed to extreme levels in the months following the Lehman bankruptcy, reaching a maximum near 900 basis points in mid-March 2009.

Financial crises as well as periods of less intense stress also lead to widening of money market spreads, due to increased anxiety about credit and liquidity risk and concern about the counterparty risk of dealing with banks. Figure 14.9 displays one measure of this anxiety, the spread between U.S. dollar Libor money market rates and yields on U.S. Treasury bills. This spread is one of a class called the *TED spread*, a term originally applied to spreads between eurodollar and Treasury bill futures. As can be seen, this spread tends to widen sharply during stress periods, for example, following the October 1987 stock market crash and during the subprime crisis.

As banks grew more anxious about the creditworthiness of their counterparties in mid-2007, term Libor rates increased dramatically compared to those for short-term interbank loans. This hindered the interbank money market in its function of distributing funds, at a price, from banks with a net surplus liquidity position to those needing short-term funds. Figure 14.10 displays the spread between Libor and *overnight interest rate swap* (OIS) rates, for one and three month month terms. OIS are OTC contracts in which one counterparty pays the other the difference between a predetermined level

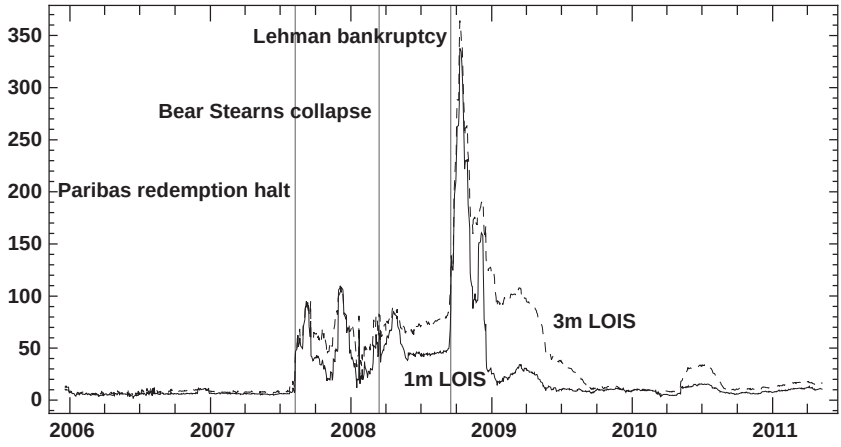

FIGURE 14.10 Libor-OIS Spread 2006–2011
Spread between three month USD BBA Libor and OIS, daily, basis points, January 2006 to May 2011. The solid (dashed) plot represents three month (one month) BBA Libor minus the price of a three month (one month) OIS swap.
Source: Bloomberg Financial L.P.

and the daily realization of the overnight money market benchmark rate over the term of the contract. The predetermined level is the quoted price, and reflects market expectations of where the overnight rate will be over the term of the OIS.

A wide spread between term Libor and OIS is unusual, since if it is costly to borrow at term, banks will prefer the overnight market, and the spread will diminish. The extreme widening of this spread in 2007 reflected the eagerness of banks with a funding need to lock up funds for longer than overnight or a few days, even if it was expensive. But it also reflected an increased desire of banks for surplus liquidity and the deep aversion of banks in a surplus liquidity position to lend at term to other banks. At the worst point of the crisis, the one-month Libor-OIS spread was nearly as wide as the three month spread. This reflected the extremity of banks' desire to shake their dependence on overnight funding. One-month funding had become almost as precious as three-month in view of the precariousness to which dependence on overnight funding had suddenly exposed them; the difference between the one- and three-month spreads was very small compared to the spread between either and the expected overnight rate.

Not only were there doubts about any particular counterparty's solvency; potential lenders also knew that the counterparty would have to find many other lenders willing to provide funding in order to remain liquid. In

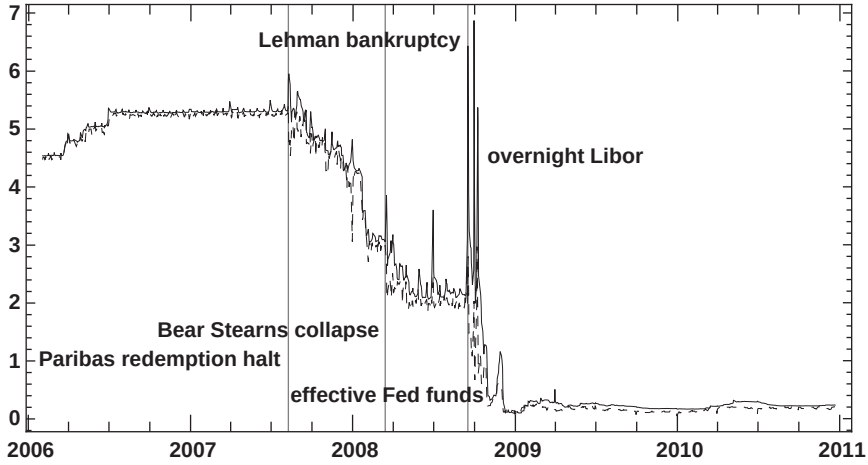

FIGURE 14.11 U.S. Dollar Overnight Rates 2006–2011
Effective Fed funds rate (dashed plot) and overnight U.S. dollar BBA Libor rates (solid plot), percent, daily. The effective Fed funds rate is published by the New York Fed.
Source: Bloomberg Financial L.P.

the terminology often used in the finance literature, potential lenders were conscious of *endogenous liquidity risk*.

Another symptom of the disruption of money markets was the spread between interbank money market rates and the federal funds market, in which balances at Federal Reserve banks are traded. Figure 14.11 displays the realized or "effective" Fed funds rate and the overnight Libor rate. Both became extremely volatile, as it became more difficult for ordinary monetary operations to bring the rate close to the target. The spread between the Fed funds and Libor rates, although for the same term, also widened out sharply.

14.1.3 Rising Insolvencies

Financial crises are associated with an increasing incidence of insolvency. Figure 6.1 displays corporate defaults for the past 90 years, measured by the fraction of bond issuers defaulting in a year. Defaults by issuers of high-yield bonds in 2009 were close to their highest level in the nearly century-long time series, exceeded only during the Great Depression. Both high-yield and investment grade defaults were at postwar highs in 2009.

Another measure of insolvency is the rate at which loans default. *Delinquency* occurs when borrowers stop paying interest and other debt service.

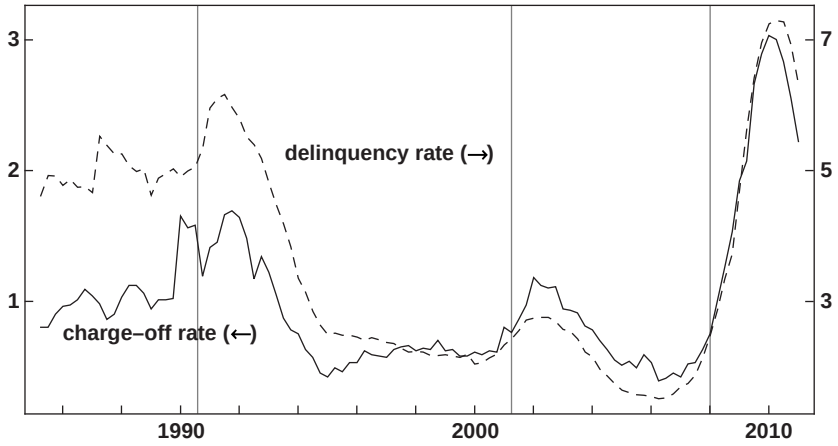

FIGURE 14.12 U.S. Commercial Bank Charge-Off and Delinquency Rates 1985–2010
Quarterly, as a percent of aggregate loan balances. Charge-offs (solid line, left axis) are the value of loans removed from the books and charged against loss reserves, net of recoveries. Delinquent loans (dashed line, right axis) are those past due 30 days or more and still accruing interest as well as those in nonaccrual status. Vertical grid lines represent the dates of NBER business cycle peaks.
Source: Federal Reserve Board.

Borrowers may resume paying later, or eventually default. Once the loan defaults or it is determined that default is highly likely, the loan is *charged off*, that is, a loss is taken equal to the unrecovered portion.[1] Figure 14.12 displays these measure for loans on the books of U.S. commercial banks over the past quarter-century. The data are obtained from regular *call reports* banks submit on their financial condition. The charge-off rate was twice as high at the end of 2009 as at its previous peak, in the aftermath of the 1990–1991 recession. Delinquency rates were also at their record high within the data set.

During a financial crisis, insolvencies are more difficult to address via restructuring, so that not only do default and bankruptcy become more frequent, but liquidation becomes a more frequent result of insolvency.

[1]*Discharge* is a legal term for extinguishing a debt obligation. *Charge-off* is an accounting term for reducing the value of the obligation on the balance sheet of its owner.

14.1.4 Impairment of Market Functioning

Market functioning is a somewhat elusive concept that describes how smoothly markets are working to equilibrate asset supplies and demands, and to intermediate credit and liquidity. "Broken markets" are frequently observed in crises. Some symptoms of poor financial market functioning are

- Observable asset prices that fluctuate rapidly, or visibly violate no-arbitrage conditions, and at which market participants don't believe they can execute transactions in size. An example of an apparent violation of arbitrage constraints on prices took place at the auctions to determine recovery values for Fannie Mae and Freddie Mac CDS after they were placed in conservatorship by the Federal Housing Finance Agency (FHFA). The recovery rates for the subordinate debt ended higher than those for the senior debt, 99.9 percent versus 91.51 for Fannie Mae and 94 percent versus 98 for Freddie Mac. Senior Supervisors Group (2009a) attributes the result to hedging by protection sellers. A similar example is the emergence of extremely low and eventually negative swap spreads after the Lehman bankruptcy (Figure 14.15).
- Credit is always allocated in part via rationing rather than purely via prices, that is, loan rates. But in a poorly functioning credit market, rationing can become predominant.

A dramatic example of poor market functioning is the phenomenon of *settlement fails* in repo markets. A settlement fail typically occurs when a repo is not closed out as contracted via the delivery of a security. The lender of cash is the seller of the security in this closing leg. When a fail takes place, the cash collateral is also not returned to the counterparty of the failing party. The convention is not to consider the fail a default, but rather to leave the repurchase price unchanged. The securities borrower (lender of cash) is then making the economic equivalent of a zero-interest loan to the lender of the security until he delivers the security, curing the fail.

Fails can occur for a number of reasons, the last two of which were especially relevant during the subprime crisis:

1. *Operational problems and miscommunication* can arise for individual market participants. But they can also be system-wide, as occurred on September 11, 2001, in which broker offices, their personnel, records, and infrastructure were destroyed.
2. *Concern about scarcity of collateral* can motivate borrowers of securities to keep the securities and lenders of securities to step back from the market. These concerns can focus on the potential for them to be "failed to" once they enter into a new repo. When one counterparty fails to

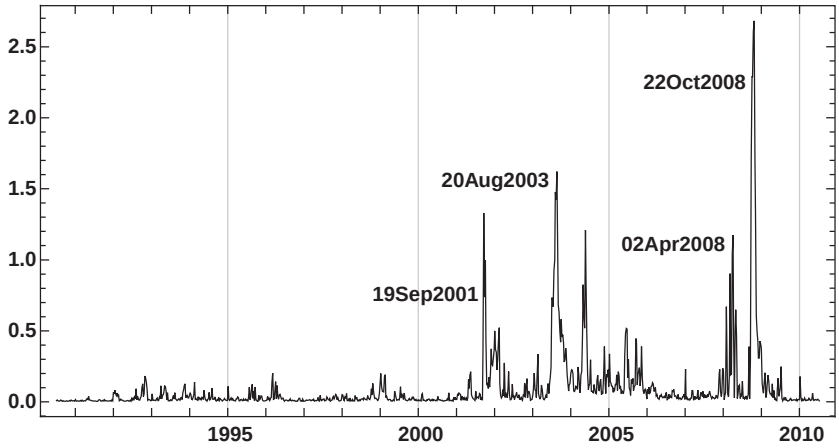

FIGURE 14.13 Settlement Fails in the Treasury Market 1990–2010
Cumulative fails to receive, weekly, trillions of dollars. Fails to receive should equal fails to deliver, but may be reported more accurately.
Source: Federal Reserve Bank of New York.

deliver to another, the latter is more likely to fail on repos in which he is obligated to deliver. This can lead to a cascade of connected fails, so that a small surge in fails can easily turn into a large one.

It can also result from hoarding of high-quality collateral, particularly in an environment in which there is concern about fails, and thus whether one will get the security back when expected, and about counterparty risk generally.

3. *A low interest cost of failing* reduces the opportunity cost of failing. If the interest earned by lending cash and borrowing a security is high, there is little incentive to fail, since the securities borrower forgoes interest for as long as a fail persists. This can happen in two ways. Specific securities much in demand as collateral can have lending rates close to zero. Or the money market rate itself can be low, as occurred during the subprime crisis. Low rates reduce the incentive to avoid failing, since foregone interest is lower.

Figure 14.13 displays data on fails reported by *primary dealers*, broker-dealers that are authorized to conduct bond transactions with the Federal Reserve. Four large spikes in fails are identified. The first occured after September 11, 2001, and was induced primarily by operational issues in the wake of the terror attack. The next, in August 2003, is related to a sharp runup in Treasury yields during the preceding two months. We discuss this episode in the next section as an example of "positive feedback" in financial

markets. The increase in fails was due to market participants' concern about being able to borrow Treasury securities to sell them short. Traders who had a borrow did not want to give it up.

The last two spikes occured during two acute phases of the subprime crisis, the Bear Stearns collapse and the Lehman bankruptcy. Both were due to the scarcity of collateral as well as low interest rates. One consequence was a further impairment of the interbank market. In addition to the drain of liquidity that is a typical feature of financial crises, the surge in fails showed that liquidity was not being appropriately distributed among intermediaries. Rather, intermediaries with cash or collateral that could reliably be used to raise cash were reluctant to part with it, while those short of cash and collateral found it extremely difficult to obtain it.

The drastic increase in fails during the subprime crisis is an illustration, not only of the impairment of market functioning, but also of a number of other recurrent features of financial crises, such as the operation of self-propagating mechanisms that deepen financial stress, the impact of "interconnectedness" in financial markets, and the stresses on interbank lending.

Scarcity of Collateral During a crisis, collateral that can be used to support borrowing becomes more scarce, and high-quality collateral becomes particularly scarce via a reduction in pledging of high-quality collateral and increasing haircuts on all types. Another manifestation of scarcity of collateral is the increase in fails discussed above. Banks and other street firms reduce their own counterparty risk, a stabilizing influence, but also abruptly reduce the volume of flow of collateral. Hedge funds and other clients move their collateral from margin to cash accounts.

Scarcity of collateral is another form the contraction of securities markets; the impact is not on the volume of intermediation achieved directly through the bonds themselves, but rather on liquidity, as fewer high-quality bonds are available for rehypothecation. There is strong evidence, from the rise in fails and balance sheet reports of financial intermediaries, that collateral, like cash, was "hoarded" during the financial crisis.

The phenomenon can be seen as accentuating the demand for high-quality assets that drove yields lower prior to the crisis. But also it also led to the rejection of assets such as structured credit products that earlier had satisfied that demand.

14.2 SELF-REINFORCING MECHANISMS

In a panic, asset owners may be eager to sell, even at much lower prices than have recently been observed. This can occur for a number of reasons. They may view even the lower current price as still well above fundamental value.

They may have bought the asset during a period of rising prices, planning to sell at a higher price that now seems unattainable. They may no longer be able to fund the asset, as, for example, in the case of ABCP conduits' and SIVs' holdings of structured credit products. Or their risk aversion may have dramatically risen.

Regardless of initial motivation, the desire to sell assets becomes self-perpetuating during crises. The mechanisms by which the desire to sell reinforces and intensifies itself are sometimes called "positive feedback loops," although of course there is nothing positive about them. The result is sometimes called an asset "fire sale." Certain common mechanisms make these sales self-perpetuating.

A related concept is *contagion*, the observation that localized financial stresses very often "spread" to neighboring geographical regions or to other markets and financial institutions. A run on one bank may readily trigger runs on others. A speculative attack on one currency may lead to attacks on others. In our discussion of Citigroup's SIV-induced losses, we alluded to the distinction between contagion induced by the vulnerability of many financial firms to a common shock, that is, direct correlation of losses, and *pure contagion* induced via fear and panic, a similar concept to that of endogenous liquidity risk. A similar ecologically tinged concept is that of *herding behavior*, the observation that market participants may have strong incentives to mimic or preempt actions by others.

In this section, we discuss some of these mechanisms. A broader discussion of whether or not the financial system is inherently stable or cyclical follows later in this chapter. A closely related question is whether the financial system tends to *procyclicality*; that is, a tendency for a positive credit environment, in which both the demand for credit and eagerness to lend are growing, to create an even more positive credit environment, and vice versa, leading to booms and busts. Some of these forces, such as procyclical capital standards, may be regulatory, and are discussed in Chapter 15.

14.2.1 Net Worth and Asset Price Declines

At the start of a crisis, a firm may suffer market or credit losses that reduce its equity capital. As its net worth falls, it finds itself more highly leveraged than it was before. A financial institution that is leveraged 20-to-1 might have credit losses of 1 percent of the value of its assets, reducing its equity capital by 20 percent. Suppose, for example, the firm's starting position is:

Assets	Liabilities
$A = 20$	Equity $E = 1$
	Debt $D = 19$

If it experiences a loss of 0.2 and there is no other change in its assets, and it doesn't raise additional equity, its leverage will rise to

$$\frac{20 - 0.2}{1 - 0.2} = 24.75$$

so the loss, although accounting for only 1 percent of its assets, has increased its leverage by nearly 25 percent:

Assets	Liabilities
$A = 19.8$	Equity $E = 0.8$
	Debt $D = 19$

The firm, even if it only wishes to restore its initial leverage ratio, must either raise capital equal to 25 percent of its now-diminished equity, or sell an additional 19.2 percent of its initial assets. In its more highly leveraged state, it will find it more difficult to raise additional or even maintain its existing funding; in particular, it becomes more vulnerable to rollover risk. If the defaults are severe enough, even if the firm's equity is not wiped out by the initial losses, the induced illiquidity may eventually lead to insolvency.

The asset owner's need for cash may arise because loans against other assets are coming due or their terms are tightening. Unencumbered assets purchased for cash, without leverage, may then need to be sold to raise additional cash to continue financing those other assets.

Losses sufficient to raise questions about solvency among market participants began to manifest themselves among large banks by the end of 2007, as the subprime crisis began to impact their balance sheets. For example, Citigroup announced on January 15, 2008, that it had lost $9.83 billion during the fourth quarter of 2007. This was a loss of about 0.5 percent of its assets, but at a leverage ratio of 20-to-1, it represented about a 10 percent reduction in equity. As noted in Chapter 12, the extent of this leverage is hard to discern from firms' financial reporting, and some critical information must be estimated or painstakingly extracted from various text elements rather than simply read off the tabular balance sheet presentation. Citi's counterparties may well have attributed far higher leverage to it, and thus considered the January 15 loss as potentially jeapordizing solvency.

We can now better understand the subsequent deterioration in lending terms it faced (see Figure 14.8). The widening of spreads reflects the "capital crunch" effect of these losses. Losses cause capital to decline dollar-for-dollar with assets, reducing the capital and increasing the leverage ratio. At the same time, Citigroup, like other large banks, had felt obliged to assume the assets of off-balance-sheet vehicles such as SIVs onto its balance

sheet. This increased leverage further, since no equity capital came back along with the SIVs. Moreover, the SIV assets were expected to continue to decline in value. Once back on the balance sheet, these losses, too, would increase leverage. Under these circumstances, the cost of fresh equity capital increased appreciably. Banks' willingness to extend credit plummeted in response. Rather, they sought all opportunities to reduce assets.

14.2.2 Collateral Devaluation

As noted in Chapter 12, a market participant is liquid if he holds sufficient unencumbered, high-quality assets. As asset prices and thus collateral values rise, market participants are able to borrow more and finance larger stocks of assets with a given amount of collateral. This applies generally, but is particularly important for financial intermediaries, for whom the capacity to put on leverage and to effect a liquidity transformation are the primary means of production of financial services.

The process works in reverse, but much more rapidly and disruptively, during a downturn. As intermediaries' balance sheets weaken, they must act to preserve liquidity and capital, and begin to shed assets. For market participants generally, if an asset purchase has been financed by using the asset itself as collateral, the loan may be terminated by the lender, or the loan terms may be changed adversely. The asset owner will then require cash or other still-liquid assets, either to finance the asset entirely, or to meet the changed loan terms.

One of the key loan terms that can change abruptly is the haircut. While data on haircuts, especially for lower-quality collateral, is not readily available, evidence is strong that haircuts rose drastically during the subprime crisis; Gorton and Metrick (2010) compute an index of repo haircuts that shows an increase in haircuts from near-zero in early 2007 to about 45 percent by mid-2009. Increasing haircuts are a key feature of episodes of collateral devaluation, so this self-reinforcing mechanismn is sometimes called a "haircut spiral."

The amount of top-quality, "unimpeachable" collateral, such as Treasury bonds, is limited. However, with higher haircuts, or at least greater acceptance by the market, other types of financial assets can also serve as collateral. The devaluation of collateral played a crucial role in the subprime crisis, primarily because the range of collateral had expanded so greatly during the preceding years.

By using collateral markets, banks and securities firms were able to finance purchases or borrowing of securitized credit products by investors, and then refinance themselves, that is, fund the loans to investors, by repledging the collateral obtained. For example, an insurance company or a

pension fund might pledge U.S. Treasuries to a dealer, and use the cash collateral thus obtained to invest in higher-yielding, but riskier securities. The dealer can refinance itself using the pledged Treasuries. In this respect, securities firms carried out a bank-like function and increased the amount of leverage in the financial system as a whole.

Collateral devaluation also has an impact on capital and credit risk calculations by banks. Unless the dependence of recovery on the state of the economy is incorporated into the models with which banks estimate the value of loan collateral, they may overestimate recovery in stress scenarios. They may thus underestimate the extent of systematic risk to which they are exposed and the amount of risk capital required.[2]

14.2.3 Risk Triggers

Risk triggers are mechanisms by which traders are forced to unwind positions in response to rising volatility or other metrics of risk. The unwinding of positions in turn causes large changes in asset prices, amplifying the increase in volatility, and intensifying the unwinding pressure imposed by the initial risk trigger. If volatility is high enough, it may render a leveraged market participant insolvent: in the example at the beginning of this section, a drop in asset value of 5 percent would trigger insolvency, forcing a bankruptcy sale or other resolution of the firm and its assets. The phenomenon of increased measured risk inducing behavior that further increases risk is sometimes called *endogenous risk*, in analogy to endogenous liquidity.

Risk triggers can take various forms. We will look at some that are defined quantitatively. But risk triggers can also be set informally. Portfolio or risk managers may simply require traders to "lighten up" in response to a downturn in the markets, but the effect on asset prices will be the same.

VaR Triggers VaR triggers require traders to unwind positions when VaR reaches certain levels. Even in the absence of formal triggers, rising VaR may induce firms to trim positions. VaR can increase in response to higher volatility or changes in correlations, even with unchanged positions. So traders may be forced to reduce portfolios that had previously been approved as appropriately sized. Unwinding positions can add to downward pressure on asset prices and upward pressure on the prices of hedging instruments.

The impact of VaR on positions during periods of stress is the converse of its impact during periods of low volatility. As we see later in this chapter,

[2]See Frye (2000, 2001).

volatility can be very low for long periods but suddenly increase during a financial crisis. During the low-volatility period, VaR will naturally be quite low, regardless of how it is computed. To the extent that market participants rely on VaR either for setting formal risk limits or as an overall metric of risk, or as a basis for setting capital requirements, they will be encouraged to put on larger positions. VaR can thus contribute to greater risk taking and to the procyclical behavior of the financial system.

Potential VaR triggers include VaR-based regulatory risk capital. As discussed in Chapter 15, under the Basel capital standards, the capital required to protect against the market risk of banks' trading book portfolios may be estimated using VaR. Capital standards then rise with volatility, even with unchanged risk-weighted assets, a phenomenon called *procyclicality of capital requirements*. In Chapter 11, we reviewed several critiques of VaR. The claim that it has a potentially destabilizing influence is among the more biting because of VaR's regulatory role, akin to finding that mandatory airbags cause an increase in auto accidents.

Stop-Loss Orders *Stop-loss orders* are buy or sell orders that are set at the time a trade is initiated, and are intended to unwind the trade in the event that prices move far enough in an adverse direction. They are attractive to portfolio manager because they appear to be automatic. Stop-loss orders are also said to align incentives appropriately, by countering the traders' typical insistence on "doubling down," that is, adding to a position rather than reducing it in response to adverse price moves. But they may prove ineffective if the orders cannot be filled in a volatile market; stop-loss orders are exposed to significant execution risk.

Stop-loss orders can add to selling pressure in declining markets. Frequently, stop-losses for an asset are set at roughly the same level, or within a narrow range, by many different traders. These bunched levels may be determined, for example, by technical trading indicators such as a previous low price. When the asset price approaches this level or range, it may quickly drop further as the stop-loss orders come into effect.

Dynamic Hedging Dynamic hedging can amplify the momentum of asset prices. One example is the dynamic hedging of options. As the underlying price declines, the delta of a long call also falls. Sellers of calls who hedge their positions by taking a long position in the underlying asset must then sell the underlying asset, amplifying the downward pressure. Sellers of puts must increase short positions in the underlying as prices fall, increasing put deltas, also amplifying the fall. Similar dynamics hold for increases in the underlying price. Option dealers are typically short options and hedge their

positions. Their hedging behavior is therefore sometimes held responsible for exacerbating price volatility.

A similar phenomenon can occur with dynamic strategies other than option hedging. Two examples are particularly noteworthy. The first is *portfolio insurance*, an approach to asset management in which the investor replicates the impact of an option hedge by selling the asset and increasing cash holdings on price declines, and vice versa. The stock market crash of 1987 has been attributed by some observers to the use of portfolio insurance.

Another example of option-induced price dynamics is the hedging of mortgage-based securities (MBS) portfolios. The interest rate risk of a long MBS position can be hedged by shorting U.S. Treasury bonds. But as discussed in Chapters 1 and 9, MBS are also exposed to prepayment risk. When interest rates decline, creditworthy homeowners will refinance their mortgages at lower rates, causing the expected duration of higher-coupon MBS to decline. The MBS therefore doesn't fully participate in the bond rally, causing a mismatch between the MBS and the Treasury hedge. The mismatch must be addressed by buying back part of the Treasury bond short position, further depressing yields.

Usually the Treasury market, one of the deepest and most liquid of asset markets, can readily absorb the fluctuation in hedging demand by MBS investors. Occasionally, however, long-term interest rates can be persistently dislodged, and some market participants can take large losses. One such episode occured in the summer of 2003, following a period in which Treasury yields had sharply declined on concern that the supply of bonds issued by the Treasury would decline. We describe this background in more detail in our discussion of swap spreads below. The end of the bond rally triggered a sharp reversal in yields. MBS durations had become very short, so Treasury hedges were small. As bond yields rose, expected MBS prepayments fell, and MBS durations lengthened. MBS investors then had to sell more Treasury bonds as a hedge, amplifying the rise in yields. The 10-year yield rose 140 basis points in the 10 weeks beginning in mid-June 2003. To get a sense of how large this move in yield was, we can apply the price-yield calculation we developed in Chapter 4; at an annual yield volatility of 15 percent, and with yields about 4 percent, this represented about a 5-standard deviation move. As noted earlier, repo fails also spiked (Figure 14.13).

Option Exercise Prices Concentration of option exercise prices can have a similar effect to stop-loss orders. A long position in a put option with a strike equal to the stop-loss level is economically similar to a stop-loss order. As the underlying asset price approaches the exercise price, delta hedging by traders tends to press the asset price lower. Like stop-loss levels, the exercise prices may be bunched at certain key asset price levels, compounding the effect.

In some markets, such as foreign exchange, *barrier options* are relatively common. These are an exotic option type that comes into existence (*knock-ins*) or is canceled (*knock-outs*) when the underlying price touches a stated level. Like exercise prices, barrier levels may be concentrated in the set of option contracts outstanding at any point in time. Barrier options can have extremely high deltas, well outside the range $(-1, 1)$, and correspondingly high gammas. Barrier options can thereby have an amplifying effect on underlying price behavior.

Credit Ratings Credit ratings are ubiquitous in credit portfolio management. Minimum or minimum average credit ratings are legally mandated for many institutional investors' portfolios, though, as we see in Chapter 15, the Dodd-Frank Act aims to diminish that role. They also play a critical role in regulatory capital standards. Ratings downgrades can force many investors to shed securities simultaneously. Credit ratings can therefore amplify credit cycles. Credit ratings affect the demand for bonds and credit spreads in at least three ways:

1. To the extent that ratings are perceived to provide accurate information, ratings upgrades will increase demand for an issue.
2. Some investors are restricted by law or regulation to meet ratings criteria. Downgrades of investment-grade bonds can materially decrease the audience for a bond.
3. High ratings increase the ability of a security to circulate as collateral for securitized loans, so a downgrade decreases a bond's liquidity as well as value.

All these factors tend to tighten spreads during a credit expansion. If ratings are procyclical, that is, if there is a tendency to award higher ratings during expansions, when balance sheets appear stronger and potential credit problems are less readily apparent, these influences will be intensified further.

A drastic example of the impact of ratings triggers manifested itself in mid-2007. Rating agencies began to downgrade the highest-rated tranches of subprime residential mortgage-based securities (RMBS), the largest bonds by original par value. Many banks had invested in these bonds because of their low regulatory capital requirements, and a few large banks and securities dealers had large concentrations of such paper. The downgrades forced some sales and triggered price declines, obliging some institutions to realize losses or mark their books lower, increasing the downward pressure on the bonds' prices. Apart from the self-propagating impact on asset prices,

the downgrades added to uncertainty and anxiety about the extent and distribution of capital attrition among large intermediaries.

14.2.4 Accounting Triggers

Accounting rules may also reinforce asset price declines. Some market participants are obliged to publicly report the *fair market value* (FMV) of at least some positions at regular intervals. For example, as discussed in Chapter 15, banks in most jurisdictions are required by regulation to mark-to-market positions in their trading books daily. Hedge funds are required to mark their books at least monthly in order to arrive at a fund NAV. Other market participants with similar positions may not be subject to such requirements, or only at less frequent intervals. Rule 157 ("FAS 157") issued by the U.S. Financial Accounting Standards Board (FASB), which went into effect after November 15, 2007, expanded the requirements for reporting estimates of FMV even for hard-to-price assets.

Some observers have taken the view that mark-to-market requirements can amplify asset price declines, generally create additional volatility, and, at the extreme, exacerbate financial instability. First, they can abruptly introduce a large gap between market and fundamental value during a period of stress. In this way, mark-to-market requirements make it more difficult for financial intermediaries to avoid "fire sales," selling assets that have experienced or are expected to experience sharp price declines because liquidity and risk premiums are rising, even though their fundamental values, that is, the expected present value of their future cash flows, may not have changed.

Second, by forcing intermediaries to realize losses and making those losses public, their capital bases and confidence in their solvency is reduced, exposing them to greater funding liquidity problems. The negative effects are particularly harsh when transactions liquidity is poor, and only a few or no transactions are being effected, and these only at distressed or hard-to-observe prices. FAS 157 in particular has been held responsible for forcing banks to report large losses early in the subprime crisis, damaging the perception of intermediaries' solvency at a critical moment. Both effects potentially induce contagion via the impact on asset prices and market confidence in the liquidity and solvency of banks.

There are several counterarguments to these concerns. It is not clear, first of all, whether banks have in fact been required under FAS 157 to report unrealistically depressed market values. Moreover, transparency can itself play an important positive role in enhancing financial stability. Finally, as we discuss in Chapter 15, capital and other regulatory standards can be set so as to mitigate any impact of mark-to-market rules in amplifying volatility,

for example by requiring higher capital ratios, and are a more appropriate tool for doing so than suspension of mark-to-market accounting.

14.3 BEHAVIOR OF ASSET PRICES DURING CRISES

In this section, we explore the impact of financial crises on asset return behavior. There are, as we've already seen, myriad examples of dramatic changes in returns around crises. But low-frequency changes in prices and spreads are also important in understanding return behavior in crises.

The most immediately visible characteristic of financial crises is an increase in return volatility. Not only do asset price fluctuations increase drastically, but the character of volatility changes. In particular, the volatility of volatility, or "vol of vol," increases unpredictably and fluctuates widely. New information becomes much more important, but also much harder to interpret. Traders generally like volatility, since they can't profit without changes in asset prices, but in crises, the increase in vol of vol makes volatility hard to profit from and induces traders to "pull in their horns." Reluctance to put on positions and to trade contributes to asset liquidity difficulties.

High realized volatility, and the other alarming phenomena seen in crises, also affect the level of implied volatility and the shape of the implied volatility surface. Crises, finally, see sharp changes in some correlations among asset returns.

Increased volatility during crises is preceded, by definition, by lower volatility during the precrisis period. But volatility also displays important low-frequency fluctuations; pre-crisis volatility tends to be lower not only by comparison with the crisis, but also with long-term average volatility. Lower volatility before the onset of crisis is also associated with rising asset prices and an attendant reduction in *ex ante* risk premiums. Real estate returns are particularly prone to large low-frequency fluctuations.

The behavior of asset prices during crises is closely connected to the credit contraction typical of crises. During a crisis, market participants are obliged to repay debts they expected to extend or roll over. They will not be able to use some assets as collateral, or will be able to borrow much less against the collateral as haircuts and margin requirements are increased. They may also be forced to sell unpledged assets.

Financial crises also bring about a *flight to quality*, that is, an increase in the values of less risky asset relative to more risky ones. Investors also favor more liquid assets in a flight to quality. A flight to quality may also take the form of *safe haven* buying, in which currencies perceived as less

geopolitically risky, such as the Swiss franc and U.S. dollar, appreciate relative to others.

In an inflation crisis, the flight may take the form of a *flight into real assets*, that is, assets such as commodities, real estate, and art works. The nominal prices of these assets tend to rise at the same pace as the general level of prices in normal times and to outpace the general price level during more extreme inflations, providing an inflation hedge.

14.3.1 Credit Spreads

As concern about rising defaults increases in a crisis, credit spreads move wider. Figure 14.14 displays credit spreads for three types of bonds, U.S. investment-grade and high-yield corporate bonds, and highly rated credit card asset-based securities (ABS). They move wider during recessions and periods of increased concern about credit quality, such as that following the Enron bankruptcy in late 2001 and following the downgrades of the major U.S. automobile manufacturers in 2005.

Spreads declined steadily after the credit scare of 2005, as did virtually all other risk spreads. Higher-rated bonds reached their tights in mid-February

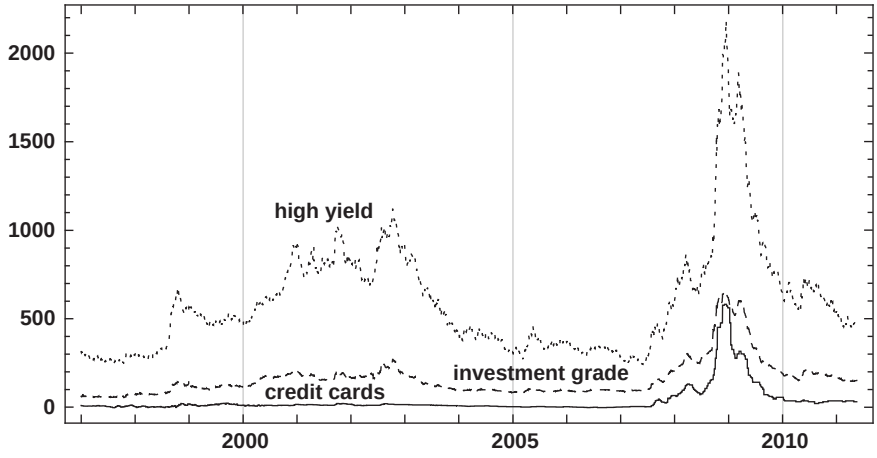

FIGURE 14.14 U.S. Credit Spreads 1997–2011
Spreads over swaps, basis points, daily, December 31, 1996, to May 19, 2011. Investment grade: BofA Merrill Lynch U.S. Corporate Index (C0A0); high yield: BofA Merrill Lynch U.S. High Yield Index (H0A0); credit cards: five year AAA U.S. credit-card ABS.
Source: Bloomberg Financial LP, JPMorgan Chase.

2007, and high-yield somewhat later, in May 2007. But during the subprime crisis, all spreads quickly widened to record levels.

The tightness of credit spreads just before the crisis began is a similar phenomenon to low precrisis volatility. It reflects low risk aversion, "search for yield," and low risk premiums. An open question regarding the origin of crises is: Do low spreads and volatility merely contrast sharply with asset price behavior during the crisis? Or do they also play some role in causing the crisis?

All three bond categories reached their wides at the end of 2008. Among the worst-hit bonds were structured credit products such as asset-backed securities (ABS) backed by auto loans and credit-card receivables. They were affected not only by the increased fear about the economy and risk aversion that caused all risk assets to sell off. They were also affected by aversion to securitized credit markets as the mechanisms and off-balance sheet vehicles built up to invest in them collapsed. Highly rated bonds backed by credit cards were by no means the hardest-hit among ABS; residential MBS prices fell even more sharply. Moreover, even after it became clear that not only financial markets, but also the real economy would be severely impact by the crisis, credit losses on senior nonmortgage ABS were not expected to be large. Nonetheless, they sold off much more severely than corporate bonds, relative to their tightest spreads before the crisis. This table displays values on selected dates from Figure 14.14:

Date	C0A0	H0A0	AAA Cards
15Feb2007	87	259	1
01Jun2007	94	241	3
15Dec2008	651	2,182	580

The wildly disproportionate widening of ABS spreads relative to corporate bonds with arguably comparable credit risk represents a large premium, with components of general risk aversion, specific aversion to the ABS asset class, and a substantial liquidity risk component. The liquidity risk premium, in turn, was generated by the disappearance of both funding and market liquidity for senior nonmortgage ABS.

Another type of credit spread, the swap spread, also widens dramatically during crises. The swap spread is the spread between U.S. Treasury notes and plain-vanilla swap rates with the same maturity, and are a widely used barometer of market anxiety. Both swaps and Treasuries trade daily in highly liquid markets. U.S. Treasury yields drop in times of heightened risk aversion, especially during more extreme episodes of flight to quality

and liquidity by investors. Swaps, in contrast, are not free of credit risk, and widen relative to Treasuries when concerns about the fragility of the banking system grow more pronounced. Figure 14.15 displays 10-year swap spreads over the past two decades. Swap spreads were above-average during stress periods such as the currency crisis period of the early 1990s, and the dot-com bust.

Swap spreads, however, are also driven by supply and demand factors that are not closely related to financial distress. The highest swap spreads observed during the past two decades were driven by the declining supply of U.S. Treasuries at a time of rapidly shrinking U.S. budget deficits in the late 1990s. The U.S. Treasury conducted a series of reverse auctions to repurchase outstanding debt between 2000 and 2002, following an announcement on January 13, 2000, and announced the suspension of 30-year bond issuance on October 31, 2001. No auctions of 30-year bonds took place between August 2001 and February 2006. Anxiety on the part of institutional investors about the possible disappearance of a security type, U.S. government bonds, to which a significant fraction of their portfolios are allocated, depressed Treasury yields. The decline in U.S. federal deficits thus contributed to the general decline in yields and risk premiums attributed to the so-called "savings glut" of the late 1990s and early 2000s, which we will discuss further below in the context of the causes of financial crises.

As can be seen in Figure 14.15, swap spreads widened more sharply in response to these market-specific pressures than to financial crises. Swap spreads initially widened during the earliest phase of the subprime crisis, but then, following the Lehman bankruptcy, collapsed and ultimately turned negative for the first time in 30-odd years of swap history. Among the reasons were expectations of large-scale Treasury issuance and institutional investors' desire to receive fixed, locking in yields as interest rates dropped rapidly. Hedging of swaptions may also have played a part. But lack of risk capital was likely the key factor, as in the similar widening of cash-CDS spreads described at the end of Chapter 13 and the increase in convertible bond cheapness displayed in Figure 12.2. This mechanism must be taken into account in using the swap spread as a gauge of overall market sentiment. As we have seen, the focus of the panic from 2007 to 2009 was on short-maturity borrowing, and the widening of credit and term spreads was greatest at maturities of a few months.

14.3.2 Extreme Volatility

To begin our discussion of extreme volatility, let's look more closely at S&P 500 return volatility, as illustrated in Figure 10.2. It is useful to look also at Figure 14.16, which displays logarithmic prices for the same period, 1927

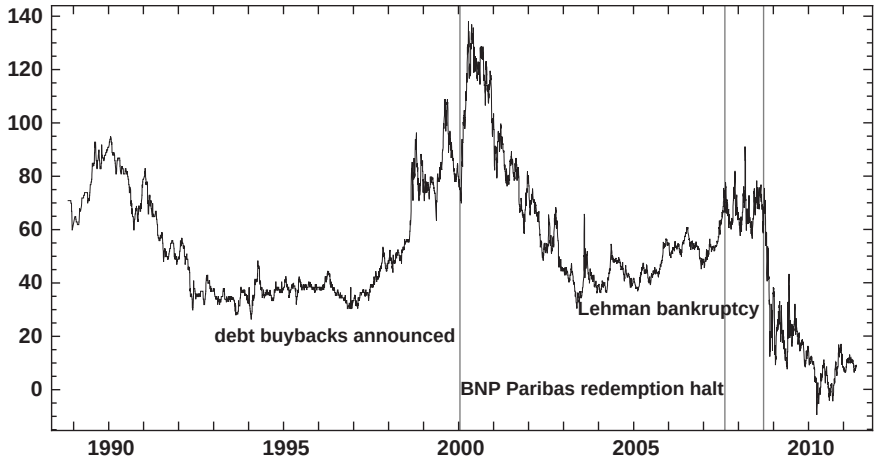

FIGURE 14.15 U.S. Dollar Swap Spreads 1988–2011
Spread between 10-year plain vanilla interest-rate swap spreads and the yield to maturity of the on-the-run 10-year U.S. Treasury note, daily, November 1, 1988 to May 11, 2011.
Source: Bloomberg Financial L.P.

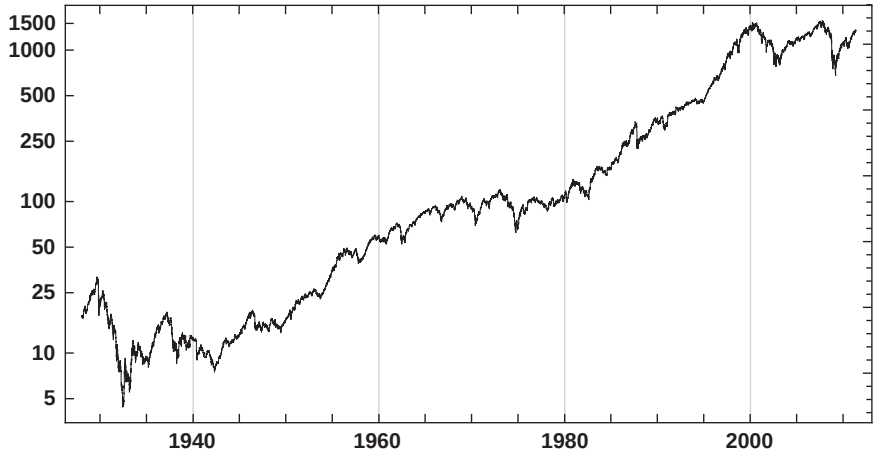

FIGURE 14.16 S&P 500 Prices 1927–2011
Logarithm of the S&P 500 index closing prices, daily, December 30, 1927, to April 14, 2011. The level of the index rather than its log is indicated on the *y*-axis.
Source: Bloomberg Financial L.P.

to date. The solid plots in Figure 10.2 indicate the 99.8 percent confidence interval for the next day's return, based on recent returns, using the EWMA approach (see the caption of Figure 10.2 for details). The horizontal grid lines mark the 99.8 percent confidence interval, using the standard deviation of daily returns of 1.2 percent over the entire period. That return volatility is just about 19 percent annually and corresponds to a 99.8 percent confidence interval for daily returns of ±3.7 percent.

Before the onset of the subprime crisis, there had only been one sustained, long-lasting episode of volatility far in excess of the long-term average: subsequent to the Great Crash of 1929. It lasted until the beginning of U.S. participation in World War II, or well over a decade.

Another episode of extreme volatility, but much shorter in duration, took place in October 1987. The S&P 500 fell over 20 percent on October 19, about double the magnitude of the next-largest daily return in either direction. Relative to the conditional return forecast, it was an even larger outlier, as we see later in this chapter. But this increase in volatility remained an isolated episode. Volatilities may go up or correlations change in the absence of a crisis, due to events in local markets. But isolated increases in volatility tend not to endure long.

The increased volatility since the onset of the subprime crisis, in contrast, has been sustained, though as of early 2011, it had not been as extreme as either the 1929 or 1987 crashes, and its duration is yet to be determined. Since conditional volatility forecasts increased gradually, it displayed fewer extreme outliers; the S&P 500 index "eased itself in" to extreme volatility, in contrast to 1929 and after. Figure 10.2 shows that the amplitude of the price swings has also thus far in the subprime crisis been smaller than during the Great Crash.

Periods of low volatility are also important. The most pronounced occurred between the recovery from the 1987 crash and the onset of the subprime crisis. It is interrupted by higher, though not extreme, volatility between 1997 and 2002, a period that covers the Asian and Russian crises, the end of the NASDAQ bubble, September 11, 2001, and the Enron bankruptcy. The period from the late 1980s to early 2007 is otherwise the longest in the historical record of sustained low volatility.

Pegged currencies show particularly dramatic increases in volatility. On the one hand, their volatility will have been dampened by monetary and exchange-rate policy, as long as the peg is maintained. On the other, once the peg is broken, large capital flows often lead to extreme changes in the exchange rate. We will look at such episodes in more detail later in this chapter.

Implied Volatility At least for the past few decades, with the growth of options markets, it has become possible to observe not only realized volatility, but also expectations of future volatility via option implied volatility. As discussed in Chapters 2 and 10, option prices contain information about risk-neutral probabilities, a probability measure that contains a mix of information about the probabilities market participants assign to different future price outcomes, and which future price outcomes they most want to protect themselves against. Implied volatility, like any market price, may contain a risk premium; the expectations expressed in it are risk-neutral, not the subjective but unobservable expectations of market participants. Nonetheless, they are informative about market participants' view of future risks.

Implied volatilities rise very quickly with the onset of a crisis or market disturbance. The term structure of implied volatility inverts, with longer-term vols rising, but not as sharply as shorter-term vols. As we see later, there is some evidence that implied vol can anticipate disturbances to some extent. Figure 14.17 displays the VIX volatility index, an important measure of stock market implied volatility, over the past two decades, covering a number of periods of significant stress, including the subprime crisis. The VIX is a composite of implied volatilities of options on the constituents of the S&P 500 with different strike prices and tenors.

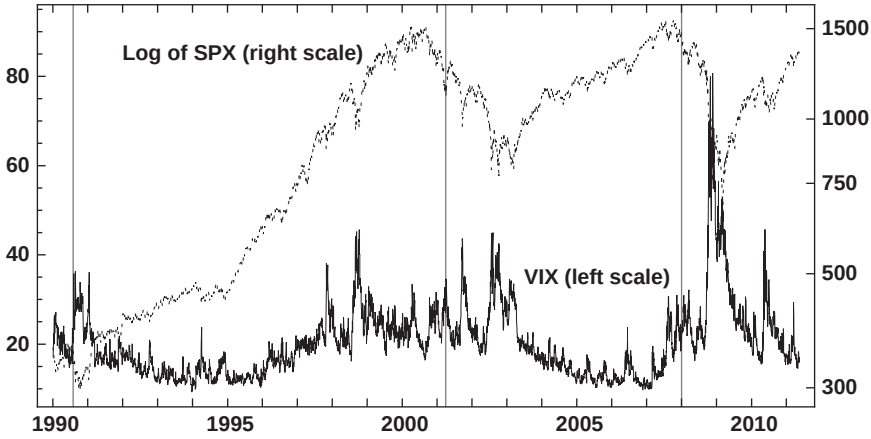

FIGURE 14.17 U.S. Equity Implied Volatility 1990–2011
VIX volatility index, percent per annum, left axis, and the S&P index, plotted on a logarithmic scale. Vertical right-axis tick labels are expressed as levels. Daily data, January 2, 1990 to May 13, 2011. Vertical grid lines represent the dates of NBER business cycle peaks.
Source: Bloomberg Financial L.P.

The spikes in the VIX are the most evident feature of the plot. There are five episodes of a VIX reading above 40, coinciding with the episodes of greatest financial market turmoil over the past two decades. The triggers were

1. The second phase of the 1997–1999 Asian crisis, during which Russia defaulted on its sovereign debt and the Long Term Capital Management hedge fund collapsed.
2. The 9/11 terrorist attack.
3. The post-Enron accounting scandals and final phase of the technology stock bust in 2002. Geopolitical stress and the rise in oil prices also played a role.
4. The Lehman bankruptcy filing, which caused by far the largest spike, to an unprecedented 80 percent.
5. The European debt crisis, which became acute in April 2010.

These spikes all coincide with sharp declines in the S&P 500 index and with some degree of market turmoil. Rising markets are associated with a low VIX. The only occasions on which the VIX recorded a closing value below 10 percent were at the end of 2006 and in January 2007, just before the subprime crisis began.

Volatility of Volatility Not only the level of volatility, but its variability, increases during crises. Since a great deal of risk management depends on accurately measuring volatility, this presents market participants with tremendous difficulties and contributes to risk aversion and the desire to hold smaller positions, that is, to delever, during a crisis.

Extreme fluctuations affect implied as well as actual volatility, signaling, unsettled expectations, or a period in which market participants expect important news to arrive at an accelerated pace. As seen in Figure 14.17, implied volatility can double in a matter of a few trading sessions, as one would expect when surprising news arrives and more is expected. For example, the VIX index doubled in one month, August 1998, as the Russian debt crisis unfolded, from a level of just over 22 to over 44. It more than tripled from its level of about 26 the Friday before the Lehman bankruptcy (September 12, 2008), to its all-time high, over 80, on October 27, 2008.

In addition to the volatility of implied volatility over time, there is typically a cross-sectional rise in the dispersion of implied volatilities of the individual assets within an asset class. Figure 14.18 illustrates for the U.S. stock market. It displays the cross-section variance of the implied volatilities largest 193 constituents of the S&P 500 index. This measure is closely related to the equity implied correlation, discussed in Chapter 10, and measures

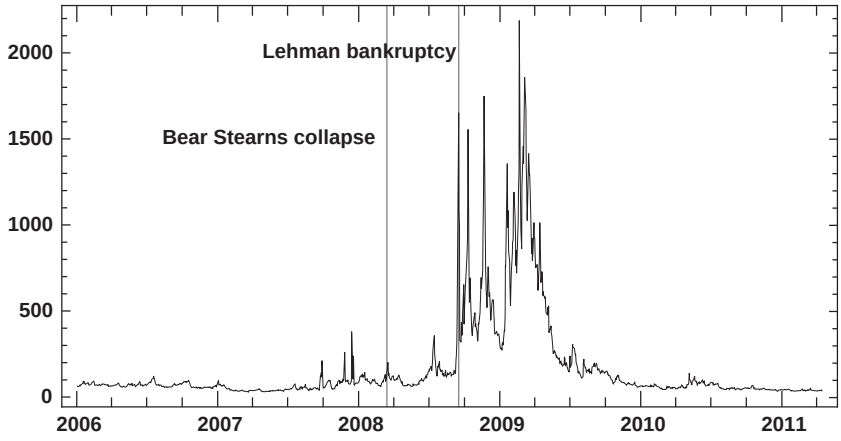

FIGURE 14.18 Equity Volatility Dispersion
Standard deviation of the implied volatilities of the largest 193 constituents of the S&P 500 index, using market-capitalization weights. The included stocks accounted for about 80 percent of the total market capitalization of the index in May 2011.
Data source: Bloomberg Financial L.P.

the dispersion of the implied vols among one another each day. Volatility dispersion rose sharply during the subprime crisis, reaching a peak in March 2009, as the equity index reached its crisis low.

14.3.3 Correlations

Correlation between asset returns can change abruptly during crises. Typical patterns of what "moves together" break down, potentially wreaking havoc with the prices of correlation-sensitive securities and with hedging strategies. We look next at the behavior of realized and implied correlations during crises and extreme events.

Historical Correlations During crises and episodes of stress, realized correlation, that is, correlations of historical returns, can change rapidly. This can have a large impact on the settings and performance of hedges based on correlations, and on portfolio allocations based on closely related statistics such as beta. Historical correlations and betas, like volatilities, are generally sensitive to such measurement choices as the number of days of historical returns included. This sensitivity of results can increase greatly during periods of stress.

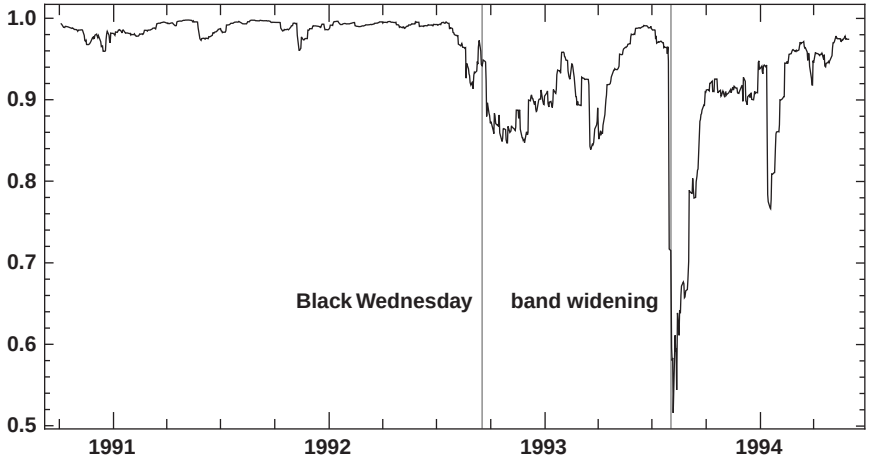

FIGURE 14.19 Proxy Hedging and the ERM Crisis 1992–1993
Daily correlation between logarithmic changes in the USD-DEM and USD-FRF exchange rates, computed using EWMA model with decay factor 0.94, October 5, 1990, to May 31, 1995.
Data source: Bloomberg Financial L.P.

A dramatic example of the potential for changes in correlation to affect hedging occurred during the European Monetary System's Exchange Rate Mechanism (ERM) crisis of 1992 and 1993. As noted above, "weak" currencies typically have higher interest rates than strong ones. The French franc was generally perceived as weak relative to the deutsche mark, and therefore had higher international money market rates. In consequence, the practice became widespread for European importers and exporters to hedge franc-denominated payables against the U.S. dollar with marks rather than francs. Because the U.S. dollar is the predominant vehicle currency in international trade, a substantial portion of French export trade made use of "proxy hedging." The correlation between the two currencies' exchange rates against the dollar were highly correlated under the ERM fixed exchange-rate system, as seen in Figure 14.19, so proxy hedging appeared to be a safe practice until the 1992 crisis. Since it effectively was a short mark/long franc position, proxy hedgers suffered losses in 1992–1993 when the franc/mark exchange rate first became more volatile and eventually, in August 1993, depreciated sharply against the mark.

Another example occurred during the second phase of the 1997–1998 Asian crisis, following Russia's default on its foreign debt in August 1998 and the failure of the hedge fund Long Term Capital Management (LTCM) the next month. Among the many calamities that befell LTCM were losses

on trades involving recently issued U.S. Treasury bonds. The most recently issued bond or note is called *on-the-run*, while those issued earlier are called *off-the-run*. Typically, on-the-run securities trade slightly tighter than off-the-run, since their market liquidity is somewhat better. LTCM had exploited this phenomenon and its apparent stability to earn the small spread between on- and off-the-run bonds, taking long positions in off-the-run and short positions in on-the-run notes and bonds. Because the spread is so narrow, large positions are needed to earn a significant returns in dollars; to earn $1,000,000 per year with a spread of 6 basis points, for example, requires long and short positions of $1,000,000,000. High leverage is needed to earn a significant rate of return.

High leverage, however, also significantly increases risk. The on- versus off-the-run spread can widen sharply during a crisis. In a flight to quality, the difference in the market liquidity of on- and off-the-run bonds increases, and the value of liquidity to investors rises, causing investors to shun off-the-run bonds unless there is a larger discount relative to on-the-run. When this occurs, the correlation between the yields drops from nearly 1, and the risk of the long-short "arbitrage" as well as the value of the short position increase dramatically, as occured in September 1998 (see Figure 14.20). For LTCM, the losses were magnified by its high leverage in this trade.

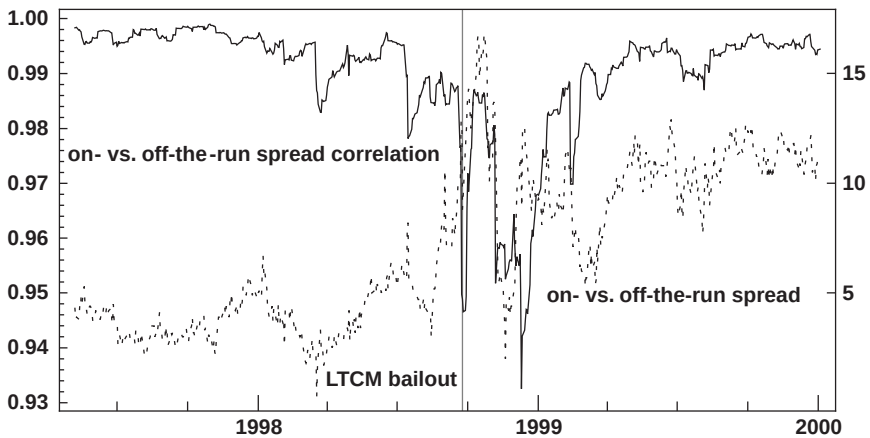

FIGURE 14.20 On- vs. Off-the-Run Rate Correlation
Correlation coefficient of daily changes in (solid line, left axis) and spread between (in basis points, dotted line, right axis) the yields to maturity of the on-the-run and first off-the-run 30-year Treasury bond. Correlation computed using EWMA model with decay factor 0.94, May 7, 1995, to December 31, 1999.
Data source: Bloomberg Financial L.P.

Betas are a function of both the volatilities and correlations of individual equities and equity indexes, and therefore can also fluctuate widely during crises. Asset managers using equity betas as a measure of risk or as a tool to determine the amount to invest in a stock can find this challenging.

An example is provided by the behavior of SLM Corp. ("Sallie Mae," ticker symbol SLM) stock price over the period 2007 to 2009. SLM is a student loan originator that began as a government-sponsored enterprise, but was entirely privatized in the mid-1990s. Its business model is heavily dependent on securitization markets: It originates loans, but then sells them into ABS rather than financing the loans permanently on its own balance sheet. In the spring of 2007, before the full breakout of the subprime crisis, a purchase of SLM by private equity firm J.C. Flowers was announced, leading to an immediate 20 percent increase in SLM's stock price. When the crisis deepened, Flowers withdrew from the leveraged buyout, and SLM's stock price, like those of other financial firms, dropped sharply. Its decline from its high point in the summer of 2007 to its low in early 2009 was about 95 percent.

Until the buyout deal was announced, SLM had a beta to the S&P 500 of close to 1, as seen in Figure 14.21. When the deal was announced, its beta was initially driven higher, in excess of 3, by higher volatility. It then traded

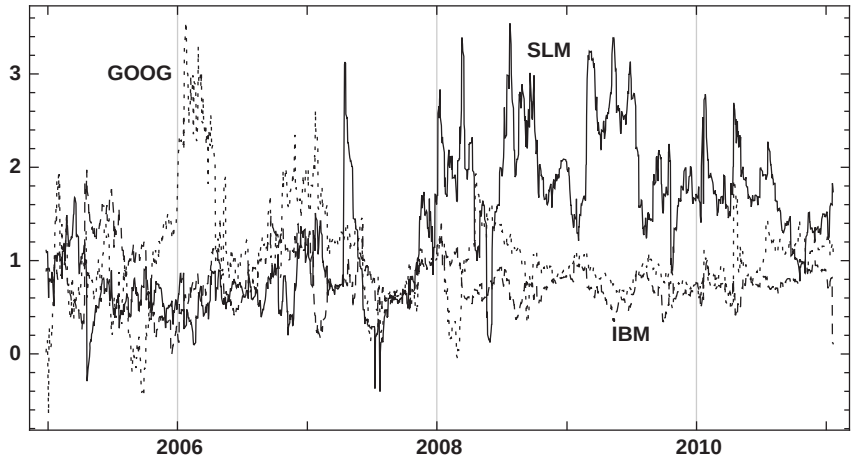

FIGURE 14.21 Changing Equity Betas during the Subprime Crisis
Rolling betas of IBM Corp (IBM), Google (GOOG) and SLM Corp. to the S&P 500 equity index, Dec. 27, 2004 to Jan. 21, 2011. The betas on each day are computed using the EWMA covariance matrix with a decay factor of 0.94.
Data source: Bloomberg Financial L.P.

like a typical "deal stock," more responsive to idiosyncratic news about the firm than to economy-wide news that moved the index, with a beta close to zero. As the buyout deal began to unravel, SLM began to trade more on general news about the financial system and the state of the economy, and its beta gravitated back toward unity. Finally, with the onset of the subprime crisis, SLM's beta again rose past 3. As an intermediary particularly dependent on the financing of its student loans through securitization, it was subject to high systematic risk, but also to increased idiosyncratic risk presented by the fragility of its funding.

By comparison, IBM Corp., an established blue-chip company, has had a beta within a fairly narrow range close to one throughout the pre-crisis and crisis periods. A different contrast is provided by Google Inc. (GOOG), a new technology company, and generally a high-beta stock. It has a more variable beta that tends to rise in excess of one during rallies and declines when the broad market is declining. The betas of both stocks were much lower and less variable than of SLM at most times during the crisis.

An old saw in trading and risk measurement states that, during crises, "all correlations go to one." Caution is warranted in interpreting this. While historical correlations do rise in stress periods, it may not be because of a regime switch to a higher-correlation joint return distribution. Boyer, Gibson, and Loretan (1997) and Loretan and English (2000) point out that sample correlation from an unchanged distribution can be much higher when estimated in periods of higher subsample volatility.

The Impact of High Correlation in Crises Not only sample asset return correlations, but also correlations implied by options and other derivative securities can change dramatically during crises. We have discussed implied correlations in credit markets in Chapters 9 and 11 and equity implied correlation in Chapter 10. Both these measures of systematic risk rise sharply during crises.

Equity base correlation is the most common measure of implied credit correlation, and is a measure of the extent to which the market believes defaults will coincide. It is derived from the prices of the equity (0–3%) standard tranche of credit index CDS, and the prices of CDS on individual companies' senior unsecured debt. Suppose equity tranche prices are relatively high and single-firm CDS spreads are wide. In other words, the market is willing to pay a relatively high amount for a lottery ticket–like security, the equity tranche, that survives only in the low-likelihood event that almost no firms default, and is at the same time willing to pay a wide spread to protect against default by any individual firm. The market is then implicitly betting that not only does each firm have a high default probability viewed

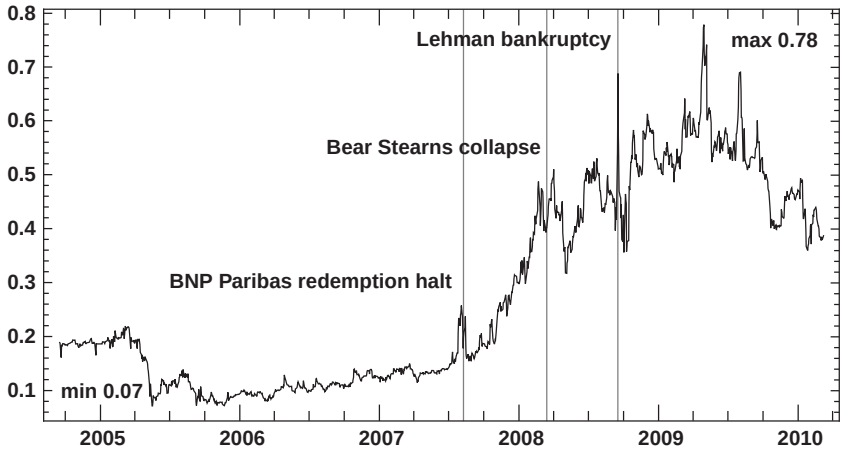

FIGURE 14.22 Implied Credit Correlation 2004–2010
Base correlation of equity (0–3%) tranche of the five year CDX.NA.IG.
Source: JPMorgan Chase.

in isolation, but that there is a high likelihood that, if even a few firms default, many will default. Conversely, the market implicitly believes there is at least a small chance that widespread defaults will be avoided, and the equity tranche return will thus be high. This is essentially a view that systematic credit and systemic risks are high.

Suppose, in contrast, that equity tranche prices are a bit lower, but spreads are relatively wide. This would indicate that the market believes it likely that, even if there are surprisingly few defaults, there will be at least enough to wipe out the equity tranche. In other words, it views default as less likely to cluster and systematic risk as low. Wide spreads of individual firm CDS then reflect firm-specific more than economy-wide drivers of default.

Figure 14.22 displays equity base correlation of the benchmark 125-name investment-grade CDS index over the past six years. Prior to the crisis, implied correlation was relatively low, always below 0.25 and generally in the low teens. We described the 2005 episode of extremely low base correlation in Chapter 11. As the crisis took hold, implied correlation rose steadily. Interestingly, it spiked just prior to the Paribas redemption halt, as markets worried about "hung loans" on large intermediaries' balance sheets. It reached a high of 0.78 just after the nadir of most asset prices in the early spring of 2009.

High correlation in crises drastically reduces diversification benefits and increases return volatility in investment portfolios. It is not surprising that a diversified long equity portfolio suffers losses during a crisis, since the high

correlation will make it trade like an equivalent amount of the declining broad equity market. However, diversified high-grade credit portfolios such as traditional bond portfolios can also suffer large mark-to-market losses, but low or no credit writedowns. Even though the portfolio is diversified from a credit point of view, the *spread correlation* may approach unity. Some institutional investors do not have to report mark-to-market losses, and do not rely on short-term borrowing to finance portfolios, and will be relatively unaffected by such an event, as long as the credit of the portfolio remains largely unimpaired. Other investors may have to report losses.

Long-short equity portfolios can have high idiosyncratic risk, even if their net exposure to the market is small. Although the systematic risk of the portfolio will remain low, the increase in idiosyncratic risk can lead to much higher volatility than expected during a crisis.

Implied credit correlations provided some surprises during the subprime crisis. In Chapter 11, we saw that ignoring potential changes in implied correlation led to losses for certain CDX trades. Placing too much reliance on implied correlation has also led to problems, in this case for investors selling equity standard tranches to protect portfolios against widespread defaults. The hedging effect was muted by the failure of equity tranches to fall as much as expected when credit spreads widened.

14.4 CAUSES OF FINANCIAL CRISES

Financial crises are a phenomenon at least as old as commercial banking, and the analysis of financial crises is nearly as old as crises themselves. It is impossible to adequately summarize the vast literature on crises, known in an earlier day as business cycle theory. Most contemporary debates on the origins of financial crises have echoes and roots in similar debates dating back, in some cases, centuries. In this section, we look at some macroeconomic policies and phenomena that have been identified as typical causes of crises.

The macroeconomic causes—or at least antecedents—of crises can be categorized under four main headings. We use the term "macroeconomic" to distinguish these factors from financial factors and credit conditions, but they are, of course, closely interrelated:

- Excessive government debt
- Imbalances in international payments
- Excessively loose or stringent money and credit emission
- Large shocks to the real economy

The macroeconomic causes of crises are among the most controversial areas in economics and economic history. The debate, for example, on the causes of the Depression crisis that began in 1929 remains unresolved. The theses advanced include monetary policy errors, international imbalances related to the U.S. net creditor position, and a widespread, sharp fall in commodity prices during the late 1920s. Few students of crises take a monocausal view, but most emphasize certain factors. We look at a few examples of the macroeconomic background to crises, and briefly describe some approaches to modelling financial crises.

The study of crises is fraught with difficulties of measurement and definition. Most of these difficulties are aspects of the fact that crises are unusual, comparatively rare, large-scale events that happen in history, rather than in a laboratory, and are therefore hard to capture statistically. They are unique, like unhappy families. There is neither a standard definition of crises, nor a standard measure of their duration or intensity. And while there are a number of models of crises, there is no generally agreed view on their causes. The study of crises remains an area in which researchers spend a relatively high amount of effort on the orderly presentation of facts rather than unitary theories.

14.4.1 Debt, International Payments, and Crises

Many, if not most, financial crises have an important international aspect. These problems include default or difficulty servicing private or public debts to foreigners, currency crises, and international transmission of banking crises.

Debt and International Crises Public debt and monetary policies are related, and have an ancient history, long preceding the advent of central banks and bond markets. The recurrence of financial crises has long been attributed to excessive borrowing by governments, and their mismanagement of money and credit. In centuries preceding the establishment of central banks, sovereigns controlled the money supply directly through the monopoly of coinage rights. Sovereigns had two paths to spending in excess of tax revenues and income from their private domains. They could subject the money they issued to *debasement*, that is, issue coins with less precious metal than their nominal value called for, or borrow from banks, which began to emerge in their modern form during the High Middle Ages. Fiscally motivated extreme debasement occurred in ancient times, for example, under the late Roman emperor Diocletian, and until the advent of paper and book-entry monetary systems. Crises occasioned by debasement took a simple form; by impeding monetary exchange and throwing society back

on barter, debasements drastically reduced the gains from trade and specialization. They reduced standards of living until indirect exchange could be rebuilt.

True financial crises have occurred since the origins of modern banking. Bank runs and failures were frequent in the High Middle Ages. Episodes of widespread bank failures, particularly those affecting a wide geographical area, were often associated with defaults on sovereign borrowing. A general banking crisis took place in Florence and Venice in the 1340s. Edward III of England and the Florentine city-state had borrowed on a large scale for war finance and were unable to repay in full. The crisis began in Venice, where a number of banks were weakened by large fluctuations in the relative values of gold and silver coinage, and then spread to Florence, where the much larger Bardi and Peruzzi banking houses were forced into bankruptcy.

In the modern era, a public-debt background to many crises can be identified. Foreign-currency indebtedness, whether public or private, is particularly prominent. The Barings crisis of 1890, the Mexican peso ("Tequila") crisis of 1994–1995, and the Asian crises of 1997–1999 are examples.

We can draw a few observations from a vast historical and analytical literature on international capital markets, debt, and financial crises:

- Countries incur large public-sector debts appear to be more prone to financial crises.
- International financial crises are frequently associated with large imbalances in international payments. Large public-sector debts often coincide with large international net debtor positions. Apart from government debts, countries that have run protracted current account deficits have a greater propensity to experience crises. The exception has been the United States, which, as issuer of the world's primary reserve currency, has had greater leeway, termed in the 1960s the "exorbitant privilege," to run balance-of-payments deficits.
- Just as with private debtors, when public-sector imbalances have been financed at short-term, a crisis is more likely to occur and to coincide with banking-sector stress. Countries as well as financial intermediaries can face credit difficulties arising from short-term financing. A typical trigger of foreign-exchange crises is sudden difficulty in rolling over short-term foreign-currency funding of capital-account deficits. This phenomenon was at work, for example, in the Mexican peso crisis and in the Russian crisis of 1998. In both cases, the governments involved faced a rollover risk event in financing short-term dollar-denominated government bonds.
- International transmission or contagion is an important aspect of international crises. For example, sovereign defaults by a number of countries

tend to coincide. Large public-sector debt of some highly-indebted governments appears to make international financial crises that affect less indebted countries more likely. The East Asian crisis of 1997 began with the Thai baht devaluation of July 1997, but spread very rapidly to other countries in the region. The next year, Russia defaulted on its debts and devalued. In the very last phase of the extended crisis, in early 1999, some Latin American countries' debt and currencies came under pressure: Brazil abandoned its peg of the real to the dollar on February 1.

An important mechanism of international transmission is *wrong-way exposure*, in which a local enterprise borrows in a foreign currency that can be expected to appreciate when the local economy is weak. The local borrower is then obligated to pay more local currency to meet debt obligations just as its local currency revenues are declining. Foreign banks may have significant losses on such exposures and become an additional contagion vector.

- Domestic banking crises or stresses, such as runs and widespread bank failures, are closely associated with international crises. The "twin crises" literature focuses on interactions between currency crises and banking system weakness. Bank holdings of sovereign debt contributed to financial crises in the Middle Ages, as we have seen. Awareness of the large holdings of European sovereign debt by both local and foreign banks was among the factors behind the strong increase in liquidity preference and risk aversion in response to the European debt crisis in 2010.
- A number of crises with international ramifications are associated with *financial liberalization*. A number of countries with strict capital controls began in the 1980s to permit free flows of capital. This was often associated with the privatization of government-owned enterprises, including banks. Another wave of financial liberalization occurred when formerly socialist countries introduced private ownership and market economies in the late 1980s and 1990s. As experience with liberalization accumulated, policy makers began somewhat belatedly to recognize that the sequence of liberalization of a financially repressed economy was crucial to the success of the project.

Speculative Attacks on Fixed Exchange-Rate Systems Currency crises have played a disproportionate role in postwar crises, though only a minor role in the subprime crisis. Currencies are among the few asset prices that have routinely been subjected to price controls by developed countries. Currencies that are perceived as overvalued are prone to *speculative attack*. These occur when market participants take large long positions in currencies

that are expected to appreciate and short positions in currencies that are expected to depreciate. Speculative attacks on currencies are an important example of the peso problem, the phenomenon of large, sudden changes in the "regime" governing asset prices, discussed in Chapter 10.

Currency crises are also prone to contagion. There is a clear channel of transmission among a set of countries in a region with currencies pegged either to one another, such as Europe in the 1980s and 1990s, or to a common anchor currency, such as the East Asian pegs to the dollar prior to 1997. If one or a few countries devalue or break the peg, the countries that do not devalue will immediately be viewed as having overvalued currencies and themselves become vulnerable to speculative attack.

The economics of currency speculation involve interest as well as exchange rates. Suppose an investor takes a short position in a foreign currency that has what he deems to be an unsustainable fixed exchange rate against the domestic currency $\bar{S}$, measured in domestic currency units per foreign currency unit. The investor believes that the exchange rate in τ years will be $\mathbf{E}[S_{t+1}]$, with $\bar{S} > \mathbf{E}[S_{t+1}]$ and $\mathbf{E}[S_{t+1}]$ denoting a subjective expected value. That is, investors expect the foreign currency to become cheaper in domestic currency units. The investor can short the foreign currency by borrowing it for τ years in the international money markets at a rate r^*, exchanging it for the domestic currency, and investing the domestic currency proceeds for τ years at a rate r. If forward foreign exchange markets exist, this position can also be established via a short forward position in the foreign currency.

The expected capital gains from this position, per foreign currency unit shorted, are $\bar{S} - \mathbf{E}[S_{t+\tau}]$, the domestic money market proceeds are $\tau r \bar{S}$, and the borrowing cost of one foreign currency unit, measured in domestic currency, is $\tau r^* \bar{S}$. The net interest earned on the position is $(r - r^*)\tau \bar{S}$. This is typically negative, as the interest rate of a weak currency is generally higher than that of a strong one. As we will see later on, this differential is a useful way to identify currencies with the potential for sudden depreciation.

The expected profit on the speculative position, in domestic currency, is

$$\bar{S} - \mathbf{E}[S_{t+\tau}] + (r - r^*)\tau \bar{S}$$

The rate of return depends on the capital required to borrow foreign currency. For forward foreign exchange transactions, this can be quite low, providing high leverage and potential returns. The risk of the trade depends on the entire distribution of $S_{t+\tau}$. If the distribution is skewed to higher rather than lower values of $S_{t+\tau}$, that is, if there is a high probability that, conditional on the foreign currency not being devalued, it will appreciate, then a risk-averse investor may choose to stay out of the trade, even though it has positive expected value. Conversely, if the distribution is skewed to

lower rather than higher values of $S_{t+\tau}$, a less risk-averse investor may put on the trade, even though it has negative expected value.

The main risk in the trade is that the foreign currency appreciates rather than depreciates. Because of this risk, such a foreign exchange strategy trade is called an "open position." However, this outcome has a low probability in many episodes of exchange rate tension, in which even quite aggressive moves by central banks to defend a parity are patently ineffective. The expected loss in the event the speculative attack doesn't result in devaluation is small; the conditional expected value $\mathbf{E}\left[S_{t+\tau}|S_{t+\tau} > \bar{S}\right]$ will only be slightly higher than $\bar{S}$, and the probability of a more-than-trivial appreciation is likely to be low.

In the typical case of a speculative attack, $\bar{S} - \mathbf{E}[S_{t+\tau}]$ is a large positive and $(r - r^*)\tau\bar{S}$ a small negative quantity, especially if the time frame τ within which the devaluation is expected to occur is short. Fixed exchange rates that are widely perceived to be unsustainable are therefore said to offer "one-way bets" to investors shorting the weak currency.

Example 14.1 (Currency Speculation) Suppose the exchange rate is fixed at 1.25 domestic currency units per foreign currency unit, but a devaluation to an exchange rate of 1.00 is expected within a year. If this view proves correct, the capital gains on each unit of foreign currency equal 0.25 domestic currency units. Suppose the domestic and foreign one-year money market rates are 3 and 6 percent. The net cost of maintaining the short position is then $(0.03 - 0.06)1.25 = -0.0375$ domestic currency units per year and the total profit is $0.25 - 0.0375 = 0.2125$. If the exchange rate strengthens to 1.30 instead, the loss on the trade is 0.0875 per foreign currency unit. If the trade horizon is one month, with the same expected conditional future exchange rates of 1.00 and 1.30, the expected gain conditional on devaluation is 0.24685 and the loss conditional on no devaluation is 0.053125, an even more lopsided balance.

Countries maintain pegged exchange rates to foster exchange rate stability and stability of their domestic price levels. But currencies with fixed exchange rates become vulnerable to speculative attack if they are perceived to be conducting expansive monetary and fiscal policies that will make the pegged exchange rate unsustainable. In models of speculative attack in the genre of Krugman (1979), speculative attacks occur when governments peg the exchange rate and in addition run budget deficits financed by either domestic money creation or out of foreign exchange reserves. No finite amount of reserves can stave off the attack indefinitely; at some point, the prospect of reserve exhaustion draws close enough that remaining reserves are depleted instantly.

Exchange rates, via short-term capital flows, affect the banking and financial system in an essential way. Currency crises can be transmitted through international trade and capital flows and interest-rate relationships among currency zone members, and therefore have a particular tendency to spread beyond the initial target of a speculative attack and impact other markets. Countries with a fixed-exchange rate, but somewhat higher interest rates, may experience large short-term capital inflows, which can abruptly reverse in a speculative attack.

We next describe two major episodes of speculative attacks on pegged currencies that created the framework for the contemporary managed floating regime among major currencies. The end of the Bretton Woods system, the framework governing the international financial system in the aftermath of World War II (see Chapter 1), was marked by a speculative attack on the U.S. dollar. The dollar was pegged; that is, it had a fixed price against other currencies, maintained by the central banks involved through currency purchases and sales. The dollar price of gold was also fixed, maintained by the readiness of the U.S. Treasury to sell gold for U.S. dollars to overseas monetary authorities at a fixed price per ounce of $35. But there had long been doubts about the sustainability of these fixed prices, and skepticism grew rampant in the 1960s as U.S. trade and current account surpluses declined and U.S. inflation increased. The United States sold its gold reserves at an accelerating pace at the fixed $35 price, and short-term international claims against the U.S. dollar increased. Foreign central banks could soak these up for a time with the assistance of an occasional modest devaluation of the dollar against their currencies. But this "adjustable peg" system further eroded the credibility of the fixed exchange rates and motivated investors to short the dollar. The system collapsed in two stages: In August 1971, the United States ended gold convertibility, and in March 1973, following an attempt to stabilize rates with another large dollar devaluation, fixed exchange rates were abandoned (see Figure 14.23).

As described in Chapter 1, West European countries, for which international trade is a much higher fraction of output than for the United States, were quite uncomfortable with the end of the Bretton Woods system and of fixed exchange rates. Their major experiment in attaining intra-European exchange rate stability, the *European Monetary System* (EMS), lasted from 1979 until the introduction of a common currency in 1999. The EMS employed fluctuation limits around the pegged exchange rates, called "central parities." If an exchange rate reached one of the fluctuation limits, both central banks were obliged to intervene to support the weakening currency. This system experienced a series of large scale-speculative attacks in 1992 and 1993.

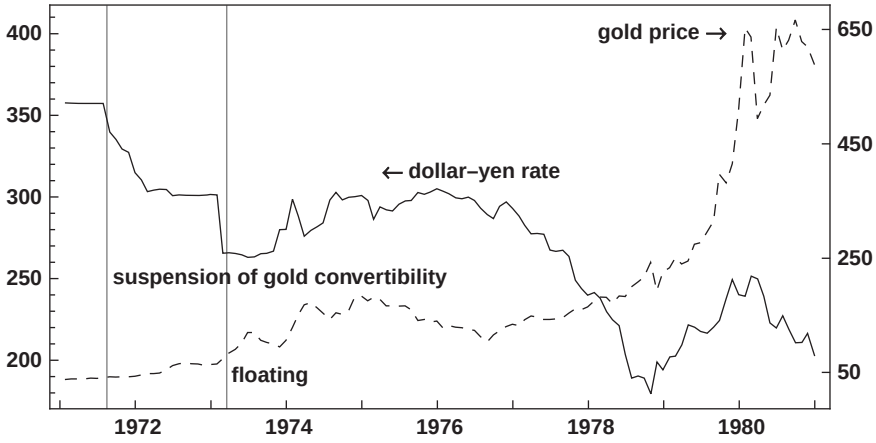

FIGURE 14.23 Gold and the U.S. Dollar at the End of Bretton Woods
The dollar-yen exchange rate (left *y*-axis) and the U.S. dollar price of gold (right *y*-axis), month-end, 1971 to 1980.
Source: Bloomberg Financial L.P.

The system had periodically undertaken changes in the central parities, as divergent macroeconomic performance created selling pressure on currencies of countries with higher inflation rates than the "core" countries, West Germany and the Netherlands. German reunification in 1989 committed Germany to large public expenditures and a potential increase in the domestic price level stemming from its commitment to convert East German nominal wages to deutsche mark at a ratio of 1:1. The Bundesbank, Germany's central bank, in an effort to offset the inflationary consequences, raised interest rates sharply, and the central parities began to appear unsustainable.

Speculative attacks began on EMS currencies other than the mark and the guilder, as well as on Scandinavian currencies such as the Swedish krona which, while not part of the EMS, had pegged rates. In an ultimately unsuccessful effort to stem the pressure, the Riksbank raised its overnight lending rate to 500 percent before devaluing. We referred earlier to Sweden's guarantee of bank liabilities to stem the attendant banking crisis. To avert a devaluation of sterling, the Bank of England intervened massively in foreign exchange markets, and eventually raised rates twice in one day, before withdrawing sterling from the EMS on the same day, September 16, 1992, or "Black Wednesday." The subsequent exchange rate volatility was astonishing. By the close of the following day, the pound had depreciated over

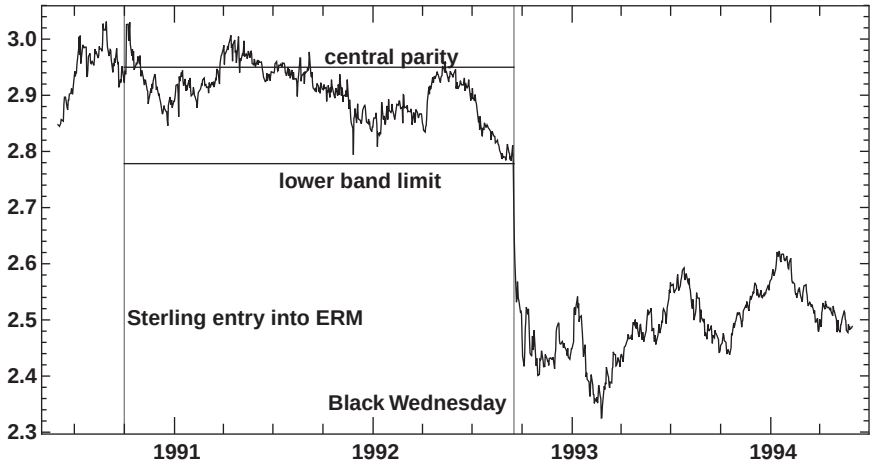

FIGURE 14.24 Sterling in the European Monetary System
The sterling-mark exchange rate, deutsche mark per pound sterling, daily close, June 1, 1990, to May 31, 1995.
Source: Exchange rate data from Bloomberg Financial L.P.

5 percent below its lower fluctuation limit; by the end of February 1993, it had lost almost 20 percent of its value against the mark (see Figure 14.24).

14.4.2 Interest Rates and Credit Expansion

Apart from debt and international imbalances, credit and monetary policies are the most frequently-cited causes or background factors of financial and the economic expansions that precede them. Sweden's Riksbank, established in 1668, and the Bank of England, founded in 1694, are among the world's first central banks, though city-states such as Amsterdam and Hamburg had established similar institutions earlier in the seventeenth century. Central banks gradually assumed responsibility for management of internal and external monetary policies.

At the risk of drastic oversimplification, two frameworks for thinking about these issues had been formulated by the turn of the twentieth century. The first, associated with Irving Fisher focused on the demand and supply of money, and is ancestrally related to *monetarism*. In this view, the demand for money is generally stable, and booms and busts are caused primarily by fluctuations in the money supply. A recognized exception is that, during a crisis, the demand for money may suddenly rise sharply.

A second framework, formulated by Knut Wicksell, was based on a distinction between the *natural* and the *money rate of interest*. The natural rate of interest, expressed in real rather than nominal terms, is the rate of interest that equilibrates the supply of savings and demand for investment capital. At the natural rate, the capital market would clear in the present, and the markets for the future goods produced would clear in the future, with no disappointments and no surprises. The money rate of interest is that actually set by the banking system. Unsustainable booms are set off when the money rate is set below the natural rate, and deflationary spirals when the money rate is set to high.

Wicksell's paradigm remains as a ghostly presence, identifiable behind the scenes in the *Taylor rule*, but is not an explicit element of most models of fluctuations and crises. The Taylor rule states a central bank operating procedure: They should—or actually do, in another interpretation of the rule and of the post-1970s consensus approach to monetary policy—set short-term nominal interest rates equal to a constant related to the long-term equilibrium real interest rate, the real rate consistent with sustainable non-inflationary long-term growth, plus the current inflation rate. The real rate is commonly estimated at 2 to 3 percent. But the central bank should temporarily raise or lower rates if inflation or output growth deviate from their target rates. The parameters of the Taylor rule are set so that, when inflation is not at its target level, policy makers follow a *feedback rule*, "leaning against the wind" by raising the real interest rate. Implicit application of this approach, and its impact on expectations, was considered key to success in reducing inflation at a surprisingly low output cost from 1979 on.

The Taylor rule is challenging to estimate for a number of reasons: its dual interpretation as both descriptive and prescriptive, the variety of ways in which the gap between actual and potential output or employment, and current or forecast inflation can be measured, and because the Federal Reserve is seen as carrying out *interest-rate smoothing*, changing target interest rates incrementally and over longer periods of time.

During the "Great Moderation," discussed in Chapter 1, both short- and long-term interest rates were quite low. Two stories predominate about the role of low interest rates in the macroeconomic background to the subprime crisis: in one, interest rates at a four-decade low were at least a contributing factor to the crisis; in the *global savings glut* story, they were the result of an unusual combination of international asset market imbalances and historical circumstances.

The low interest rate story begins by noting that that, according to many estimates, the short-term interest rates set by policy fell below Taylor rule levels during the decade preceding the subprime crisis. Central banks generally operate via short-term interest rates, as long as they are above zero,

but long-term interest rates also fell to unusually low levels. In fact, long-term rates kept falling even after short-term rates began rising from mid-2004. This unusual phenomenon was dubbed the "yield curve conundrum."

Low interest rates didn't, however, lead to rising inflation, the crucial signal that rates were "too low." As we saw in Figure 1.11, U.S. core CPI dropped below 5 percent in 1991 and below 3 percent in 1996, and hasn't returned to those levels since. Economists were divided even at the time about the reasons for the apparently benign economic climate. Some considered it a success of economic policy, while others remained uncomfortable with very low interest rates, in spite of low inflation. Yet others felt the reports of moderation of volatility were greatly exaggerated, observing, for example, that a reduction in the volatility of a single output sector, inventory investment in durable goods, appeared to be responsible for the entire reduction in GDP growth volatility seen in Figure 1.15.

The undercurrent of discomfort with low interest rates focused on credit expansion and rising asset prices. The low volatility of macroeconomic variables we noted in Chapter 1 also extended to asset returns; spreads and yield levels fell as well. We have seen that low volatility can increase fragility by increasing asset values and risk limits on positions, and, in this view, may have played a role in fostering instability. Concerns had been raised before the crises that, while volatility might have declined, financial instability had become more pervasive. These concerns were supported by the rising frequency of larger and smaller financial crises from about 1970 onwards.

In the half-decade prior to the subprime crisis, Federal Reserve policy makers articulated a "risk management approach" to monetary policy, sometimes explicitly couched in terms of a policy maker loss function. It meant that the Federal Reserve would act more aggressively to avert low-probability events with severe consequences. In the low-interest rate view, the risk management approach, combined with low consumer price inflation, contributed to an asymmetry of policy response, in which monetary policy was not tightened in periods of expanding credit and rising asset prices, but loosened in periods of declining credit aggregates and valuations. The perception of asymmetric macroeconomic policy contributed, in this interpretation, to the compression of risk premiums.

Several related channels have been noted by which low interest rates might increase the fragility of the financial system:

First, low rates increase the incentive to enter into leveraged trades. We described how this incentive works in Chapter 12. Low interest rates encourage trades in expectation of rising asset prices, such as housing purchases with low down payments. Low rates, by facilitating maturity and liquidity transformation, also encourage carry trades and leveraged positions in modestly higher-yielding securities, such as senior structured credit products funded at money market rates or hedged with lower-yielding securities.

Interest rate smoothing is also thought to have played a role: Any increases in the level of short-term rates would proceed gradually, limiting the downside and liquidity risks of investments that depended on leverage via short-term funding.

Second, by reducing all-in bond yields, low rates increase the incentives of some market participants to seek out riskier investments to attain return targets. At the same time, it drives risk premiums narrower, so that tail risks, such as the systematic risks of senior securitization bonds, are priced low. This effect is important because it encompasses important market participants, such as insurance companies funding fixed-rate annuities and defined-benefit pension funds. This "search for yield," or "yield panic" reflects agency cost and externality problems, as the institutional investor increases the probability of being able to meet return requirements, but also increases systematic risk and system fragility. Low rates also strengthen market participants' financial positions by increasing the value of their assets and lead them to expand leverage and balance sheets. Finally, and particularly in its interaction with interest-rate smoothing, the perception that central bank policy seeks to avoiding bear markets can diminish the perception of risk by market participants.

These effects of low interest rates have been called the *risk-taking channel*. Diamond and Rajan (2011) present a formal model in which banks are incentivized to fund themselves to some extent with short-term debt because it adds to discipline over borrowers and because the banks anticipate low interest rates in crises. Tighter monetary policy reduces the banks' short-term funding incentive and thus the probability of crises. The increase in risk-taking induced by low rates can manifest itself in increasing balance-sheet leverage, but also in off-balance-sheet transactions that are more difficult to observe.

U.S. broker-dealers provide an example of how low interest rates work through the risk-taking channel. As we saw in Chapter 12, certain forms of securities lending, for example, particularly where higher-quality securities are exchanged for lower-quality ones ("borrowed versus pledged"), are not necessarily reflected in financial reports, or are at least not easy to discern, though it can be indirectly observed in data such as repo volumes (see Figure 12.3).

Low interest rates may interact with other factors, too. For example, in 2004, the SEC changed regulatory capital requirements for major-broker dealers. This permitted the broker-dealers to take on greater leverage, much of it in short-term forms such as commercial paper, repo and securities lending. The assets being financed often lacked market liquidity. The presumption for broker-dealers, as opposed to commercial banks, is that they can be funded short-term, since the asset side of their balance sheets consists largely of inventories of liquid securities. This presumption was no longer

accurate by 2007, as broker-dealers became deeply involved in the process of creating securitized credit products. Their balance sheets were employed in part as warehouses for the illiquid intermediate products of the securitization process, including loans held for future securitizations, and bonds created in recent securitizations, but not yet sold to the broader markets, or held as proprietary investments.[3]

The global savings glut view of low interest rates focuses more on the causes than the potential consequences. In this view, the increase in capital flows and the large volume of investment inflows to developed countries, especially the United States, were important factors in the decline in long-term interest rates in developed countries. These flows have their origin in two underlying phenomena: the low U.S. savings rate and persistent U.S. current account deficit, on the one hand, and the successful export-based development strategy pursued by a number of developing countries, especially in Asia, on the other. Each of these phenomena would be unsustainable for long on its own. Together, they have persisted for several decades and intensified during the decade preceding the subprime crisis.

The current-account deficit could be explained either by U.S. investment outlays in excess of U.S. saving, which would pull capital in and raise interest rates, or eagerness by foreigners to buy U.S. assets, which would push capital into the U.S. and lower interest rates. Proponents of the latter, savings glut, view, point to falling U.S. interest rates as evidence that the "capital push" is predominant.

The U.S. current account deficit is made up primarily by developing countries, particularly in East Asia. This is surprising in itself, since one would expect industrial countries to invest in less-developed countries rather than the reverse. One explanation is that these surpluses are part of an overall development strategy. Though counterintuitive, China is said to run a current account surplus in order to foster internal investment. The key to the apparent contradiction is the role of collateral. By investing surpluses and holding reserves in the form of U.S. assets, China conveys collateral to the U.S. The liability that China collateralizes in this fashion is foreign direct investment in China. This direct investment is valuable because it brings technological progress with it, and is worth obtaining even at the cost of locking up a considerable portion of its net assets as collateral. If it were to expropriate the direct investments made by the industrial countries, it would risk having its dollar-denominated assets frozen. In other words, China runs a large current account surplus in order to facilitate the small

[3]See U.S. Securities and Exchange Commission (2004) for the text and background discussion of the change to Rule 15c3-1.

direct-investment deficit embedded in the overall capital flows. The overall development strategy is to build a technologically advanced capital stock that will enable a large rural labor force to eventually be employed at a much higher marginal productivity of labor.

An alternative explanation of the counterintuitive, "uphill" flow of capital from less- to more-developed regions points out that, although emerging markets have grown rapidly, they have not developed financially nearly as rapidly. Investors in these countries need the safe assets issued in advanced countries, depressing their rates of return. This line of reasoning dovetails with the story, laid out in Chapter 12 and earlier in this chapter, pointing out the effect of the demand for collateral in fostering the growth in investment-grade securitized credit products. In all of its variants, the global savings glut view sees limits on the ability of U.S. monetary policy to influence longer-term interest rates and the risk-taking channel.

14.4.3 Procyclicality: Financial Causes of Crises

Earlier in this chapter, we examined the self-reinforcing mechanisms that amplify volatility and the credit and liquidity crunch during financial distress. These mechanisms can be viewed as manifestations of the financial system's tendency to procyclical behavior; booms and busts, once underway, tend to amplify until an unsustainable extreme is reached, at which point the reversal, a recovery or a crash, commences.

The risk-taking channel of monetary policy which we have just described is another source of procyclicality. Attention to possible procyclical tendencies in the financial system is an important part of research on the risk-taking channel of monetary policy: The subprime crisis has many observers asking themselves what they had missed. As noted, there is also evidence that the frequency and severity of crises has been increasing in recent decades.

The wide range of views on procyclicality can be categorized along two dimensions

The strength of procyclicality: Some identify tendencies that can be readily counteracted by appropriate policy, while others see large-amplitude, low-frequency fluctuations in activity and financial aggregates as nearly uncontainable in the modern financial system.

The sources of procyclicality: Some see fluctuations as triggered exogenously, that is, by random shocks or policy errors, while others view procyclicality as endogenous, that is, inherent in the organization and functioning of the financial system.

These views on financial crises can be divided along the lines of perennial issue in debates on crises, and economics generally: How reliable is a market economy's tendency toward equilibrium with full employment of economic resources?

A parallel debate, also dating back centuries, concerns the question of whether the financial system automatically controls the volume of money and credit. In the early nineteenth century, two schools of thought on this subject emerged in the United Kingdom. In the *Currency School* view, the banking system would be prone to excessive issuance of credit without externally-imposed limitations. The *Banking School* taught that credit regulated itself automatically: If banks collectively issued too much credit to business, financed by note issuance, the holders would begin redeeming the notes for gold, nipping excessive credit expansion in the bud. As far as British monetary policy was concern, the debate was settled by Peel's Act of 1844, a milestone in the development of modern central banking, which put note issue into the exclusive hands of the Bank of England and restricted its volume. But the debate around the propensity of the banking system to permit credit aggregates to grow excessively has recurred.

We have seen in this chapter how fluctuations in firms' net worth and in the value of collateral they can offer can amplify crises. Observers have been aware of this mechanism since at least the Great Depression, when it found expression in Fisher's (1933) schema of "debt-deflations." Because it puts firms' net worth in the forefront as a driver of boom and bust, and is attentive to the interactions of net worth and asset prices, the debt-deflation approach has resonated with many post-subprime crisis observers. As we saw earlier in this chapter, leverage in new forms was extremely high for many financial intermediaries prior to the subprime crisis, making net worth a key driver of the crisis. In Fisher's schema, an initial state of excessive leverage leads to liquidations, losses, and a drain of liquidity that reduces net worth, leading to fresh rounds of liquidation. These effects are brought about by the reduction in prices, initially those of assets, but eventually also of commodities and final goods. Apart from lower prices, the debt-deflation process brings about a reduction in output, a reduction in nominal interest rates, and a rise in real interest rates.

Another way in which net worth constraints increase instability is to impair the markets' ability to carry out arbitrage and near-arbitrage. Effecting even a riskless arbitrage generally requires some minimum of capital and time. Investors may withdraw capital from vehicles such as hedge funds that specialize in carrying out arbitrage operations. Investors' as well as lenders' urge to withdraw capital may be spurred by a general desire to regain liquidity rather than avoid capital losses. Chapter 12 illustrated the potential effect of this mechanism of procyclicality with the sharp fluctuations in

convertible bond values (see Figure 12.2), Chapter 13 with the extreme bond-CDS basis, and this chapter with negative swap spreads in late 2008.

Withdrawal of capital may also arise as a consequence of asymmetric information. Fund investors, like lenders to banks, have less information than the fund's (or bank's) managers about the investments made with their capital. They may therefore rationally rely on fund returns in order to determine whether the fund managers have skill in carrying out arbitrage. When returns are low or negative, which with some probability is due to lack of skill rather than market conditions, investors withdraw capital more eagerly. The reduction in capital, from both the desire for liquidity and the negative return signal, can amplify asset sales.

A similar effect may occur in asset markets. In Chapter 12, we noted that some investors, called "noise traders," may make investment decisions based at least in part on the prices of the assets themselves, rather than on an analysis of their fundamental values relative to the price. A decline in price may indicate that "someone else" knows something negative about fundamental value, and induce them to sell. Earlier, we saw that dynamic hedging by investors can induce self-perpetuating declines in asset prices. One mechanism by which this can happen is for hedging sales in response to a small initial decline in price, for example by option sellers, to amplify it by incorrectly signaling a change in fundamentals. The price decline is amplified further by market participants selling purely in response to earlier price declines.

These mechanisms are to some extent symmetrical, so price increases can also become self-sustaining. But capital and funding constraints will introduce asymmetries between boom and bust, and make asset price declines more powerful and difficult to arrest.

The view that crises are inevitable is associated with the view that there are large low-frequency fluctuations in risk appetite. Keynes referred to risk appetite as "animal spirits," relating its fluctuations to those of economic activity,[4] and the label has stuck. Characteristically, as in his descriptions of thrift, Keynes found a distanced and ironic phrase for a socially useful but unaesthetic quality. The contrasting expression "irrational exuberance" came into usage in late 1996 to characterize the stock markets of the time.

The view that crises are inevitable starts from the observation that risk aversion is subdued—animal spirits are high—during quiet times. This leads to expansion of business activity and bubbles in asset prices. At some point it

[4]See Chapter 12 of Keynes (1936). Keynes' alternative formulations are "spontaneous urge to action rather than inaction," and "the delicate balance of spontaneous optimism."

becomes clear, often after some relatively minor triggering event, that not all these business plans can actually be realized. A panic ensues, risk appetites shrink, asset prices collapse, and business activity remains low until the cycle gradually restarts. No economic forces exist that would bring the system to a steadier equilibrium path.

A formulation that combines the two elements of net worth and risk aversion was put forward by Hyman Minsky in the "financial instability hypothesis." The hypothesis is set out as a three-stage theory of boom and bust, each corresponding to a type of finance and a type of borrower (or "unit") that predominates during that phase:

1. *Hedge finance* can be repaid out of the cash flows of the borrower, and does not need to rolled over or refinanced.
2. *Speculative finance* units can pay interest on debt, but cannot repay principal, out of cash flows, and will have to refinance or sell the investment at a higher price to remain solvent.
3. *Ponzi finance* refers to debt on which even interest cannot be paid out of cash flows. Further borrowing is needed to service existing debt. It will be necessary in addition to sell the asset at a profit to repay the loan.

Only a financial system in which hedge financing predominates can be stable in the Minsky paradigm. Enough elements of speculative or Ponzi finance can—and in a free-market system are likely to—set in motion an endogenous cycle, which will either be contained via public-sector intervention or culminate in a bust. The Minsky paradigm is uncannily similar to the credit creation that preceded the subprime crisis, and has therefore garnered renewed attention.

In macroeconomics, the impact of capital and net worth constraints on economic activity, particularly during crises, is called the *credit channel* of transmission of monetary shocks and policy. It is similar to the risk-taking channel in its focus on the financial condition of market participants. In contrast to the risk-taking channel, which focuses on the encouragement low interest rates give to risk-taking, the credit channel focuses on the impact of monetary policy on firms' balance sheet health.

14.4.4 Models of Bubbles and Crashes

Financial crises are related to, but distinct from *bubbles*. A bubble occurs when asset prices become detached from fundamentals and trade on the basis of expected future capital gains, as in Minsky's Ponzi form of finance. Bubbles have been defined more clearly in models than in history. Even in hindsight, it is hard to be sure that any particular episode of rapid asset

price increases constitutes a bubble. Two examples are the *tulip mania* of seventeenth-century Europe and the stock market boom of the 1920s. In both these cases, a rapid and sharp runup in prices was followed by an even more abrupt decline. Yet historians have been able to marshal considerable evidence that the asset price had arguably been rational *ex ante* in view of the fundamentals, even if the view of the bulls was not borne out in retrospect.

It is all the more difficult to identify bubbles, in real time, as they occur, and one of the challenges facing public policy is to determine the best means of addressing them. Should policy attempt to dampen swings in asset prices, or focus on the effects of such swings on economic activity? If the focus is to be on asset prices, by what criteria can we identify "bubbles"? We return to these issues in the next chapter.

Models in which the current asset price depends on expectations of future fundamentals, for example dividend-growth models of stock prices, and which also assume rational expectations, which we define in a moment, generally have an infinite number of solutions. The asset price, in these solutions, is a function of current expectations of all the future fundamentals, and an arbitrary constant. If the constant is not zero, the price may veer off ever further from the fundamentals-based value, and still be consistent with rational expectations and with the model.

We can illustrate this using the basic security pricing model of Chapter 2. We look at asset prices in a discrete-time framework, but with infinitely many periods. The price S_t at time t of any asset is equal to the sum of two components: the expected value of its price next period, and its expected cash flow d_t, set at time t but paid at time $t + 1$, which we call a "dividend." The asset return is

$$\frac{S_{t+1} - S_t + d_t}{S_t} \quad t = 1, 2, \ldots$$

Now, let's assume *rational expectations*, as is common in economics. It sets asset prices equal to their actuarial or "physical" expected values, and abstracts from risk appetites and risk premiums. In a rational expectations model, the asset price must be set so that its expected return is equal to the discretely compounded risk free return r_t:

$$r_t = \mathbf{E}\left[\frac{S_{t+1} - S_t + d_{t+1}}{S_t}\right] \quad t = 1, 2, \ldots$$

The only element that is not known at time t is the future price S_{t+1}, so we can pull everything else out of the expectation notation:

$$r_t S_t = \mathbf{E}\left[S_{t+1}\right] - S_t + d_t \quad t = 1, 2, \ldots$$

to get

$$S_t = \frac{\mathbf{E}[S_{t+1}] + d_t}{1 + r_t} \quad t = 1, 2, \ldots \tag{14.1}$$

Now, we can use Equation (14.1) to solve for the expected value of the asset price, from the point of view of time t, in a more distant future than $t + 1$. For convenience, let the risk-free interest rate and the dividend be constant, putting the focus entirely on asset-price risk. The next-period asset price is

$$S_{t+1} = \frac{\mathbf{E}[S_{t+2}] + d}{1 + r} \quad t = 1, 2, \ldots$$

so

$$S_t = \frac{1}{1 + r} \left\{ \mathbf{E} \left[\frac{\mathbf{E}[S_{t+2}] + d}{1 + r} \right] + d \right\} \quad t = 1, 2, \ldots$$

Notice that we have the expected value of an expected value in this expression. A standard trick in solving rational expectations models is to note that, at time t, we know no more about *what we will know in the future* than we know right now about the future.[5] So from the time-t point of view,

$$\mathbf{E}[\mathbf{E}[S_{t+2}]] = \mathbf{E}[S_{t+2}]$$

Make this substitution, and take all the nonstochastic terms out, to obtain

$$S_t = \frac{1}{(1 + r)^2} \mathbf{E}[S_{t+2}] + \left[\frac{1}{1 + r} + \frac{1}{(1 + r)^2} \right] d \quad t = 1, 2, \ldots$$

[5]Technically, this step relies on a property of conditional expectations known as the *law of iterated expectations* or *tower rule*. In essence, we have been working here with conditional expectations of the future asset price. The law of iterated expectations states that if the conditioning set—the set of possible outcomes in the "outer" expectation—is the same as the conditioning set of the "inner" expectation, we can collapse it into one conditional expectation.

We can now successively substitute out the next period's asset price. After doing so τ times, we get

$$\begin{aligned} S_t &= \frac{1}{(1+r)^\tau}\mathrm{E}\left[S_{t+\tau}\right] + d\sum_{\theta=1}^{\tau}\frac{1}{(1+r)^\theta} \\ &= \frac{1}{(1+r)^\tau}\mathrm{E}\left[S_{t+\tau}\right] + \frac{1}{r}\left[1-\frac{1}{(1+r)^\tau}\right]d \quad t=1,2,\ldots \end{aligned}$$

Now, if the asset matures at time T, then we know its terminal value, such as the par amount of a bond, for certain. If we set $S_T = 0$ for simplicity, then S_t is just the present value of the dividend flows:

$$S_t = \frac{1}{r}\left[1-\frac{1}{(1+r)^\tau}\right]d \approx \frac{d}{r}$$

The approximation holds if the time to maturity τ is long.

But suppose the asset is infinitely lived, such as an equity or a commodity. The time-t asset price is now seen to have two components:

$$S_t = \lim_{\tau\to\infty}\frac{1}{(1+r)^\tau}\mathrm{E}\left[S_{t+\tau}\right] + \frac{d}{r} \tag{14.2}$$

The first is the present value of the future asset price, and the second, called the *fundamental value* of the asset S^*, is the present value of the future cash flows:

$$S^* = \frac{d}{r}$$

That first component

$$b_t \equiv \lim_{\tau\to\infty}\frac{1}{(1+r)^\tau}\mathrm{E}\left[S_{t+\tau}\right]$$

is not "nailed down." As long as b_t, the "bubble" component, grows (in expected value) at the risk-free rate, that is,

$$\frac{\mathrm{E}\left[b_{t+1}\right]}{b_t} = 1+r$$

it is an equilibrium price, since it is consistent with rational expectations, the absence of arbitrage and with our model of the asset price. This phenomenon is called *nonuniqueness* or *multiple equilibria* and is ubiquitous in economic and financial models in which market participants' expectations are equal to physical expectations of future economic variables or asset prices. Situations in which there are several possible outcomes, depending on what market participants expect to happen, or what impact they expect their actions to have on the expectations and actions of others, are common in such models.

One way to get a unique solution is to simply rule out bubbles via the *transversality condition* $b_t = 0$. Modelers dislike doing this because it's arbitrary and doesn't address the fact that the models admit multiple solutions. This is particularly the case since at least one specification of a bubble is eerily similar to our intuitive notion of bubbles and crashes in financial markets. Suppose

$$b_{t+1} = \left\{ \begin{array}{c} \frac{1+r}{1-\pi}\, b_t \\ 0 \end{array} \right\} \text{ with probability } \left\{ \begin{array}{c} 1-\pi \\ \pi \end{array} \right\}$$

with $0 < \pi < 1$. Since $\frac{1+r}{1-\pi} > 0$, the bubble component b_t will grow over time, until a crash occurs and it goes instantly to zero. The return on the asset will exceed the risk-free rate by enough to compensate for the possibility of a crash. A crash is not inevitable, but the longer the time horizon, the likelier it is to occur. The lower the probability of a crash, the longer the bubble can last. Once a crash occurs, though, the asset price will remain at its fundamental value forever.

In this specification, which equilibrium is realized is random. Models in which equilibrium is random, but unrelated to economic fundamentals are said to display *self-fulling crises* or *sunspots*. Recall that this is precisely what characterizes endogenous liquidity risk, described earlier in this chapter and in Chapter 12.

There is much controversy over whether sunspots and multiple equilibria are real-world phenomena. But the idea is useful in understanding some of the phenomena we have discussed, such as panics, runs, and the role of animal spirits. In the models of bank runs we've described, for example, a run occurs because depositors are afraid others will run. An equilibrium state in which nobody runs can occur as easily as a run; a state in which asset owners "all hold hands" and refrain from selling can occur as easily as a fire sale. A number of models of balance-of-payments crises and currency crashes, such as that of Krugman (1979), cited above, also display multiple equilibria.

In another class of models of bubbles and crashes, credit is an integral part of the process. In an approach set out in Kiyotaki and Moore (1997, 2002), cycles are generated by the net-worth effect. In periods of rising asset prices, firms are able to borrow more and expand; periods of falling asset price persist analogously. The approach of Kiyotaki and Moore (1997) rests on the risk-shifting impact of credit contracts, in which the borrower keeps most of the upside, but has a limited potential loss. In this type of model, intermediation and the agency problems it can generate play an essential role.

Gennotte and Leland (1990) provide a model in which asymmetric information interacts with market liquidity to amplify price swings. In their model, some investors rely on asset price information to draw inferences about value. This behavior is analogous to that of noise traders, or of the providers of capital in Shleifer and Vishny (1997), who draw inferences about the investment skill of their agent-managers from returns. When uninformed investors see a decline in asset price, they interpret it as a decline in fundamental value, and join the selloff, amplifying it. Gennotte and Leland apply the model to the October 1987 market decline.

14.5 ANTICIPATING FINANCIAL CRISES

Market participants as well as policy makers have a keen and growing interest in identifying the potential for crises and extreme events. This potential can be signaled by macroeconomic indicators, at relatively low frequencies, and by the behavior of asset prices.

14.5.1 Identifying Financial Fragility

Ideally, but perhaps unrealistically in practice, one would like to predict financial crises, or at least identify data that help discern whether financial fragility is growing. For example, identifying sharp or sustained increases in asset prices that are not matched by improvements in fundamentals, so that prospective returns or risk premiums are shrinking, might help identify bubbles. We discuss the difficulties of doing so here and in the next chapter.

Another set of indicators germane to financial fragility are measures of leverage. Figure 14.25 shows one simple measure of leverage for the United States, the ratio of credit market borrowing to GDP, showing the total as well as the contributions of the government, private nonfinancial, and financial sectors to the total. The data are drawn from the U.S. Flow of Funds Accounts, and do not adequately represent leverage generated via derivatives, off-balance-sheet vehicles, and collateral markets along the lines

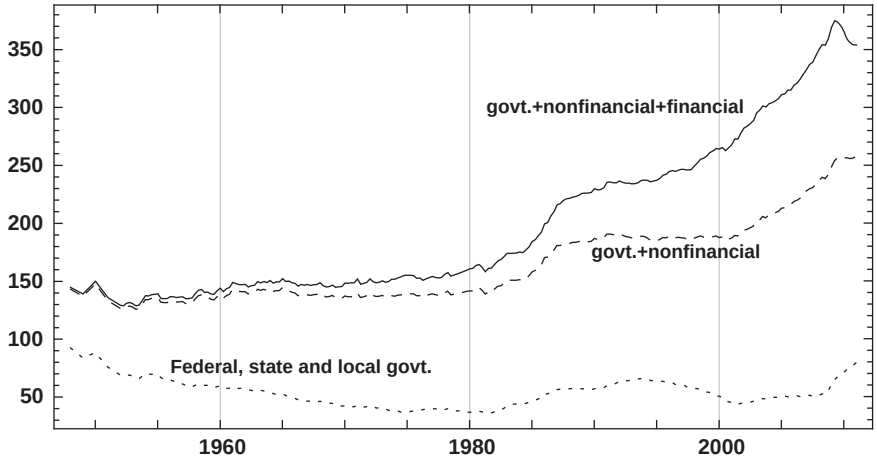

FIGURE 14.25 U.S. Leverage 1947–2010
Ratio of credit market debt outstanding to gross domestic product in current dollars, percent, quarterly. "Nonfinancial" is the sum of private nonfinancial and rest of world. "Financial" includes all financial sectors.
Source: Federal Reserve Board, Flow of Funds Accounts of the United States (Z.1), Table L.1, lines 1–10; U.S. Bureau of Economic Analysis, National Income and Product Accounts, Table 1.1.5, line 1.

laid out in Chapter 12, but are at least indicative of general trends over the past few decades.

Leverage increased steadily from the early 1980s until the subprime crisis. Most of the increase in leverage has been in the private sector, especially households and the financial sector. Household leverage increased particularly rapidly during two phases, the mid-1980s and the decade preceding the subprime crisis. Financial-sector leverage has been increasing throughout, due primarily to increasing leverage by the nonbank sector. As noted, banks' rising leverage took forms not reflected clearly in this data set. The trends reversed with the onset of the panic phase of the crisis. Financial and private nonfinancial leverage dropped sharply. Government debt rose, but not enough to offset private-sector deleveraging.

There is no widely accepted standard for what constitutes "high" leverage relative to GDP. One might argue that leverage, as an aspect of financial development, should increase to some extent with economic growth. Nonetheless, the unabated and rapid rise of the leverage-to-GDP ratio over the past three decades provided an indicator of excessive credit creation and fragility.

14.5.2 Macroeconomic Predictors of Financial Crises

The difficulty of generalizing about the nature and causes of crises extends to the search for indicators and warning signals. Macroeconomic data, apart from asset prices, because they have low frequency, pose special problems in this respect. Much work has been done by international organizations such as the International Monetary Fund (IMF) to identify macroeconomic and other low-frequency indicators of financial and balance-of-payments stress, such as data on the condition of the local financial sector. These are especially important for developing countries, which have less well-developed financial markets and a narrower range of observable asset prices. The evidence, however, is weak that such indicators do a good job at predicting crises.

14.5.3 Asset-Price Predictors of Financial Crises

Warning signals of crises may be obtained from asset prices. We've seen many examples throughout this book of asset price dislocations caused by extreme events and crises. Some of these asset price phenomena may appear early enough to serve as useful warning signals. Certain asset prices, particularly the interest-rate term structure and credit spreads, have been used with considerable accuracy to forecast fluctuations in economic activity. For example, the spread between yields on 10-year Treasuries and 3-month T-bills has been a reliable precursor of recessions.

Crises generally coincide with economic downturns, but occur much less frequently, so it is more challenging to identify leading indicators and verify their information content statistically. We can single out three categories of asset prices that appear to move early in response to stresses:

1. "Liquidity spreads," a diffuse group including on-the-run and off-the-run government bond spreads and money-market risk spreads
2. Foreign exchange forward markets
3. Option implied volatilities

Most liquidity spreads, such as those related to the TED spread, the money market spread between a risk-free and credit-risky money market instrument, contain both liquidity and credit risk components. Examples are displayed in Figures 12.4 (repo spreads), 13.4 (the CDS basis), 14.9 (the TED spread), 14.10 (Libor-OIS), 14.11 (Libor-Fed funds), and 14.15 (the swap spread). Liquidity spreads tend to widen out before other credit spreads, such as those on corporate bonds, as some bankers and dealers first become more reticent about granting short-term credit to others. As we

have seen, the spread between Libor and OIS began to widen dramatically in August 2007, as concern over banks' liquidity and capital positions began to affect money markets. The behavior of swap spreads, however, was sharply at odds with the past, tightening rather than widening. Credit spreads did not widen out dramatically until 2008.

An interesting example of money market spreads possibly having provided a warning signal of the 1929 stock market crash is presented by Rappoport and White (1993). They obtained time-series data on different types of money market rates: ordinary short-term loans by banks to nonfinancial customers, commercial paper, and loans to brokers, who in turn used the funding to make secured margin loans to buy stock. In February 1928, the Federal Reserve began to tighten monetary policy to address growing concern about the rising stock market. All money market rates rose, but the spread between margin loans and short-term commercial loans and its volatility rose particularly sharply (see, especially Rappoport and White, 1993, Fig. 2 on p. 555).

Another type of signal is provided by extreme moves in a low-volatility environment, an intriguing example of which occurred in February 2007. Implied volatility, and credit and liquidity spreads were still just barely higher than the all-time lows they had recently touched, but a sense of foreboding about the market had begun to take hold. A number of mortgage lenders had defaulted (though others were still being acquired by large banks and broker-dealers). HSBC had in mid-February made the first announcement by a major banking house of losses due to subprime lending. Toward the end of February, there was a brief spasm of volatility in several markets.[6] It is not clear what was the immediate trigger: in the prior day or two, the Chinese stock market had plunged 9 percent in response to a tightening of policy, Freddie Mac had announced that it would no longer purchase subprime mortgages, and yet another subprime mortgage originator, Fremont General, had signalled distress by delaying its earnings filing. By later standards, the February moves in prices and spreads were quite small, but they represented very large moves relative to the low volatility at the time.

To see how extreme the end-February 2007 episode was, and to put it in context, look again at Figure 10.2. It displays conditional exceedances by the S&P index of a 99.8 percent confidence interval, using a prior-day EWMA volatility estimate. Suppose we set an even higher confidence level, say, 99.99999 percent, so that there is a one-in-10-million chance of an excession on any given day under the normality assumption. This

[6]See Vikas Bajaj "Freddie Mac tightens standards," *New York Times*, February 28, 2007.

corresponds to a ±5.32 vol move. There have been only 19 such events since 1928. The following table displays the dates, daily returns, annualized conditional (EWMA) volatility computed via EWMA, and the number of prior-day vols by which the S&P moved, in percent. The last column is computed as the ratio of that day's return to the prior-day estimate of volatility at a daily rate.

Date	Return	Annual vol	Return/daily vol
28Oct1929	−13.9	40.7	−5.4
15Mar1933	15.4	44.7	5.5
23Jan1939	−6.6	19.5	−5.4
05Sep1939	11.2	26.1	6.8
13May1940	−5.7	15.2	−5.9
03Sep1946	−10.4	16.3	−10.2
04Feb1948	−4.3	10.7	−6.4
03Nov1948	−4.7	12.5	−6.0
26Jun1950	−5.5	10.6	−8.3
06Jul1955	3.5	9.8	5.7
26Sep1955	−6.8	9.5	−11.5
28May1962	−6.9	17.6	−6.2
17Aug1982	4.6	13.7	5.4
11Sep1986	−4.9	14.6	−5.3
19Oct1987	−22.9	30.1	−12.1
13Oct1989	−6.3	9.2	−10.9
15Nov1991	−3.7	9.3	−6.4
27Oct1997	−7.1	16.2	−7.0
27Feb2007	−3.5	6.6	−8.5

Viewed this way, February 27, 2007, stands out as exceptional. Only four outliers in over 80 years of data are larger on this normalized basis. One is the October 1987 market crash. Another is the day President Eisenhower suffered a heart attack. Yet the February 2007 event preceeded the onset of the "hot" phase of the crisis by over five months. Other indicators of market fear, such as equity implied correlation, an indicator of concerns about systematic risk (see Chapter 10), also jumped sharply on this date.

There is strong association of these very large outliers with low volatility. The correlation of the volatility with the absolute value of the data in the last column, the ratio of return to volatility, is −18 percent, and if we exclude the great outlier of October 13, 1987, the correlation is −40 percent. This is another bit of evidence that risk managers and policy makers should be wariest when volatility is low.

The Forward and Term Premiums Perceived risks of large asset price moves may be reflected in futures and forward prices of those assets. Similarly, anticipations of large changes in interest rates will be reflected in the behavior of the yield curve, since expectations of short-term rates are important in the setting of longer-term rates.

An example of information appearing in forward rates has occurred regularly in currency markets. Selling pressure on currencies with pegged exchange rates is often reflected in the forward premium, which shows a more pronounced expectation of depreciation, particularly if foreign-exchange intervention keeps spot rates from responding fully to the pressure in the short term. An implication of this phenomenon is exploited in Svensson's (1991) "simplest test" of target zone credibility. If the monetary authorities set fluctuation limits for the currency, the currency's forward rate trading outside those limits is evidence that the limits are not credible.

Another example is the behavior of yield curves during "inflation scares," episodes in which market participants doubt the credibility of the central bank's commitment to keep inflation low (see Goodfriend, 1993). The yield curve tends to steepen sharply in such episodes; the increase in expected future inflation increases the forward interest rate for longer terms to maturity.

Option Implied Volatility-Based Crisis Predictors Option prices are also potentially useful indicators of financial stress. Like the prices of futures and forwards—or of any cash asset, for that matter—option values are based on expectations for the future. Options lend themselves particularly well to this purpose because their prices are sensitive to the size and likelihood of future asset returns, and to how risk-averse and ill-positioned to absorb them at least some market participants are. An extreme return scenario with a low probability may have little impact on the price of a cash security or a futures, since these are expected values across all outcomes. But it may have a large impact on the values of options that pay off if that scenario is realized. Indicators of the likelihood of an extreme event can be derived from the option-based risk-neutral probability distributions we studied in Chapter 10, employing options across the exercise price axis. We will focus here on indicators based solely on time series of at-the-money implied volatility.

In the standard Black-Scholes theory, implied volatility is equal to the known, constant diffusion volatility parameter. If the theory were true, but the diffusion parameter happened to be unknown, it would be easy to estimate and would provide accurate volatility forecasts. Implied volatility is a Black-Scholes model concept, so if Black-Scholes does not hold, an alternative model is needed to determine the expected relationship between implied volatility and future volatility. In most empirical work on this relationship,

an alternative model or an ad hoc specification establishes a null hypothesis of unbiasedness, that is, that today's implied volatility is the expected value of future realized volatility.

Such tests of "second-moment efficiency" are the analogue of tests of market efficiency that compare mean realized asset returns to the forecast implied by the current forward or futures price. Futures and forwards are often found to be biased predictors of future prices. For example, as we have seen in Chapter 10 and earlier in this chapter, emerging-market foreign-exchange forward premiums often persistently "forecast" large exchange rate changes that simply don't take place. Rather, the level of the forward rate moves up and down with the spot rate, following a path that is hard to distinguish from a random walk.

The evidence, in contrast, is that options do have some short-term predictive value, at least compared to alternative predictors such as recent historical volatility. In other words, "second-moment efficiency" appears to be rejected less robustly than "first-moment efficiency." However, implied volatilities also appear to be biased upward, that is, they tend to forecast higher volatility than is actually realized, giving rise to the notion of a "volatility risk premium." Implied volatility also seems to have some value for predicting large moves. Implied volatility is typically mean-reverting, so that neither periods of low volatility nor sharp spikes in volatility are expected to persist for very long.

Figure 14.26 illustrates one episode in which options appear to have provided early warning of an extreme event. It compares the warning signals provided by foreign exchange forwards and options prior to the Thai baht float of July 1997. This episode marked the beginning of the Asian financial crisis of 1997. A speculative attack against the baht was initiated on May 14, 1997. Implied volatility began to rise just before the attack, and rose sharply as soon as it began. The forward rate began to rise only some time after the beginning of the attack. It was caught in the tug of war between traders shorting the baht and the Bank of Thailand's interventions in both the spot and forward currency markets. Official intervention inhibited the rise of the forward rate but did not affect the options market and the implied volatility signal.

Once the baht was permitted to float, both historical and implied volatility rose sharply. This behavior was not unusual: In a crisis, implied volatility often leads historical volatility. Realized volatility then spikes much higher than implied, but also "decays" rapidly. Implied volatility continues to rise, and decays slowly.

Asset price–based warning signals can be very useful, but a number of caveats are necessary. One major drawback is the number of false positives or Type II errors. This problem is not unique to derivatives prices, as

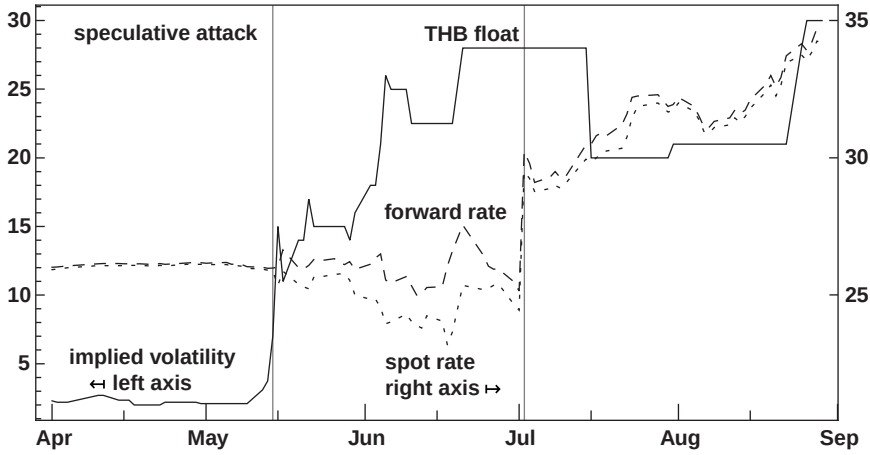

FIGURE 14.26 Behavior of Implied and Historical Volatility in Crises
The plot describes the behavior of the Thai baht (USD-THB) and implied volatility on USD-THB during the Asian crisis, daily, April through August 1997. Spot rate in THB per USD. Implied volatility of one month at-the-money forward options. Historical volatility computed using the EWMA/RiskMetrics approach with decay factor 0.94 and 90 business days of data.
Data source: Bloomberg Financial L.P.

attested by Paul Samuelson's timeless remark that "the stock market has predicted nine out of the last five recessions" (quoted in Siegel, 1991, p. 27). Liquidity spreads widen, and implied volatilities rise, much more often than crises actually occur. Growing awareness by market participants of the relationship between asset-price indicators and the risk of extreme events may change their relationship, an example of the operation of Goodhart's Law (see Chapter 15). Every historical situation is unique, rather than following an unvarying pattern. Warning signals therefore have value only when used in conjunction with other data and with nonquantitative analyses. It should, finally, be borne in mind that interpretations of asset prices as signals of market perceptions share an important limitation: They are risk-neutral indicators, and can be distorted by changing risk premiums.

This has motivated the development of risk aversion indicators based on asset prices. These include the VIX, often called a "fear gauge," based on at-the-money volatilities, and the recently-introduced CBOE Skew Index, which is based on the prices of at- as well as out-of-the-money options of S&P 500 constituents. Chapter 10 described equity implied correlation. Indexes based on other asset prices and on historical volatility, as well as on implied volatility, have been developed in the trading research departments

of several large banks and dealers, and a number are presented throughout this book.

The search for indicators of crises and extreme events is hardly new, but has become a lively research area in the wake of the subprime crisis. We'll return to this challenging subject in the next chapter, as part of our discussion of public policy to prevent crises.

FURTHER READING

Kindleberger (1978, several subsequent editions) combines a great deal of historical data with an analytical classification of crisis events. Mackay (1962), first published over 150 years ago, is an early work recognizing financial crises as a distinct phenomenon. Grant (1994) is a narrative of selected American banking and capital markets panics. Brown (1987) is a data-rich study of international financial panics, including the interwar era. Kindleberger (1986) and Temin (1989) are a good starting point for studying the macroeconomic background of the Great Depression. Reinhart and Rogoff (2009) is a general history of financial crises, but with a particular focus on crises with sovereign debt and international payments roots. Kaufman (2000*b*) is an analytical survey of systemic risk events.

Although far from ended, there are already a great many valuable studies of the subprime crisis in print and on the Web. Their focus ranges widely, from narratives, to attempts to identify the causes of the crisis, to essays on the implications for financial practitioners. For a detailed chronicle of the crisis, the International Monetary Fund's semiannual *Global Financial Stability Reports* and the Bank for International Settlements' *Annual Reports* and *Quarterly Reviews* issued from mid-2007 onward are invaluable. The U.S. Congress established a number of bodies to study the subprime crisis and report on the public-sector response to it. These bodies maintain web sites with many useful reports and summaries of testimony expressing a wide range of views. The Congressional Oversight Panel home page is http://cop.senate.gov/ and the Office of the Special Inspector General for the Troubled Asset Relief Program home page is http://www.sigtarp.gov/. The Financial Crisis Inquiry Commission has published its final report as U.S. Financial Crisis Inquiry Commission (2011), and maintains its trove of background documentation at http://fcic.law.stanford.edu/. The New York Fed maintains timelines of the crisis at http://www.newyorkfed.org/research/global_economy/policyresponses.html.

Apart from those we cite in reference to specific issues, there are a number of general essays on the crisis worth reading. Crouhy, Jarrow, and Turnbull (2008) and Greenlaw, Hatzius, Kashyap, and Shin (2008) were

early assessments. Golub and Crum (2010) discusses risk management implications, with an emphasis on liquidity risk. Swagel (2009) is a narrative by a former senior Treasury official.

Many of the references for Chapter 12 are pertinent to this chapter as well. Pozsar, Adrian, Ashcraft, and Boesky (2010), Adrian and Shin (2009*b*), Tucker (2010), and U.S. Financial Crisis Inquiry Commission (2010) discuss the relationship between the nonbank sector and the subprime crisis. Singh and Aitken (2009*a*, 2009*b*) describe and quantify the decline in credit extension collateralized by securities on the part of banks and broker-dealers during the subprime crisis. Acharya, Cooley, Richardson, and Walter (2009) ties the shadow banking story together with the observation in Chapter 9 that structured products are exposed to hard-to-discern systematic risks. These papers complement the perspective brought by Gorton (2009) and Gorton and Metrick (2010). In addition to the references at the end of Chapter 12 on SIVs and ABCP conduits, see Chapter II of International Monetary Fund (2008a) on the role of these off-balance sheet vehicles and of securitization in the subprime crisis. See also Hördahl and King (2008) on the behavior of repo markets during the subprime crisis.

Bernanke and Lown (1991) and Wojnilower (1980) and (1985) review credit crunch episodes in the postwar U.S. economy. Schularick and Taylor (2009) and Bordo and Haubrich (2010) provide long-term evidence that the severest economic contractions coincide with credit expansions and subsequent credit crises. Ivashina and Scharfstein (2010) details the overall decline in bank lending and "run" on bank lines following the Lehman bankruptcy.

Calomiris and Gorton (1991) uses historical data to better understand the triggers and background of bank runs and panics. Duffie (2010) draws together several new forms a run can take, including the risk of being shut out of the triparty repo system, in describing the collapses of Bear and Lehman. Shin (2009*a*) describes the run on Northern Rock. See also Pedersen (2009). Laeven and Valencia (2008) analyze a database chronicling in detail systemically important banking crises globally from 1970 through 2008.

Longstaff (2004) attempts to estimate the crisis liquidity premium in U.S. Treasury prices. Fleming and Garbade (2002, 2004, 2005) provide institutional and analytical background on settlement fails in repo markets. Liu, McConnell and Saretto (2010) discuss an episode that has been widely interpreted as a failure of market functioning. Kiff, Elliott, Kazarian, Scarlata and Spackman (2009) and Stulz (2010) contribute to the debate over the role of OTC credit derivatives in the crisis.

Brunnermeier (2009), Brunnermeier and Pedersen (2009), and Shleifer and Vishny (2011) discuss self-reinforcing mechanisms in crises. Kaufman (2000*a*) discusses the concept of contagion. Aharony and Swary (1983)

is an empirical study of U.S. bank failures and finds no evidence of pure contagion in the data. Krishnamurthy (2010) chronicles the effect of self-propogating mechanisms based on risk and liquidity during the subprime crisis. Daníelsson, Shin, and Zigrand (2004) and Chapter II of International Monetary Fund (2007) discuss risk triggers. See Goodhart (2004), Laux and Leuz (2010), Brunnermeier, Crockett, Goodhart, Persaud and Shin (2009), and Chapter III of International Monetary Fund (2008*b*) on accounting triggers.

Bennett, Keane, and Mosser (1999) describe the phenomenon of mortgage coupon concentration and its potential impact on interest-rate dynamics. Perli and Sack (2003) and Krause (2003) discuss mortgage hedging with Treasuries and its impact on the 2003 Treasury market volatility episode. Kambhu (1998) and Kambhu and Mosser (2001) estimate the effects of interest-rate option hedging on interest-rate volatility. Malz (1995) describes an episode in which exotic option dynamics had a pronounced impact on the spot foreign exchange market.

See Jones (1953) on the Diocletian inflation, and Mueller (1997) and Hunt (1994) on the fourteenth-century financial crisis. Schnabel and Shin (2004) describe an early example of the fire-sale mechanism in crises. Bresciani-Turroni (1937) describes a paradigmatic inflation crisis and the "flight to real values." Solomon (1982) chronicles the decline of the Bretton Woods system. Kaminsky and Schmukler (1999) analyzes the behavior of market prices during the Asian crisis of 1997–1999. Schwert (1990*a*) notes the increase in equity market volatility during recessions and following extreme negative returns. Longin and Solnik (1995) study correlation behavior in equity markets. Engle (2009) discusses the reasons different asset-return correlations vary in different ways over time.

Kaminsky and Reinhart (1999) and Kaminsky, Reinhart, and Vegh (2003) analyze the causes and interaction of currency and other international financial crises. Garber and Svensson (1995) is a survey of models of speculative attacks on pegged exchange rates. *Brookings Papers on Economic Activity* has published a number of papers offering rich institutional and historical detail on the financial crises of the past three decades within a theoretical framework. For example, Kharas, Pinto, and Ulatov (2001) studies the Russian crisis of 1998; Radelet and Sachs (1998) the East Asian crisis of 1997; and Dornbusch, Goldfajn, Valdés, Edwards, and Bruno (1995) a number of currency crises of the 1980s and 1990s.

A number of other articles and essays cover recent crisis episodes. Desai (2000) summarizes the background of the Russian crisis and provides a critical analysis of the Russian government's response. Committee on the Global Financial System (1999*a*) is a detailed summary of the 1998 Asian crisis. The LTCM crisis is described in President's Working Group on

Financial Markets (1999). See also the references provided in Chapter 1 for some of the specific episodes listed in Table 1.1.

Calomiris (2004) is a survey on financial fragility and the background of crises. Bernanke (1983) presents evidence for the key role, alongside monetary and other nonfinancial macroeconomic factors, played by loss of access to credit in deepening the Great Depression. As runs and closures caused the banking industry to shrink, nonfinancial firms were thrown back on internally-generated sources of capital. Bernanke and Gertler (1995) discusses the credit channel of monetary policy transmission in modern economies.

The role of collateral markets in the contemporary financial system and in the subprime crisis is explored in Gorton (2009). Data on the contraction of liquidity brought about by the collapse of collateral markets are provided in Gorton (2009) and Gorton and Metrick (2010). Gorton and Metrick (2010) also provides information on available repo data. See also Adrian and Shin (2009*b*).

Still-germane nineteenth-century discussions of interest rates and financial crises are summarized in O'Brien (2007), including the Banking and Currency School controversy. Haberler (1958) covers the early twentieth-century debates that culminated in the Keynesian revolution. Wicksell's approach was originally presented in Wicksell (1907) and Wicksell (1967). These works make fascinating reading in light of the liquidity and credit expansion leading up to the subprime crisis, and also serve as a useful reminder of how old issues in economics resurface again and again in new forms.

On postwar macroeconomic policy generally, consult the references provided at the end of Chapter 1. The original presentation of the Taylor rule is Taylor (1993). Taylor (2007) and Kahn (2010) present the case that interest rates were set below Taylor-rule levels prior to the subprime crisis, and relate leverage and the housing bubble to low rates. Greenspan (2004) and Mishkin (2008) describe the risk management approach in monetary policy. McConnell and Perez-Quiros (2000) present a sectoral analysis of the decline in output volatility in the 1980s. Dooley and Garber (2005), Caballero (2006) and Obstfeld and Rogoff (2010) present three views of the global savings glut hypothesis; the first as a stable, the latter two as an unstable equilibrium. Caballero, Farhi, and Gourinchas (2008), Greenspan (2010), and Bernanke (2011) argue that global payments imbalances and the savings glut were an important cause of the crisis.

Borio, Kennedy, and Prowse (1994) is an important early study of the impact of monetary policy on asset prices and an early document in the literature on the risk-taking channel. Borio and Zhu (2008) and Adrian and Shin (2011) are comprehensive treatments of the risk-taking channel, and Dell'Ariccia, Laeven and Marquez (2010) present a theoretical model of the

interaction of leverage, the level of rates, and monitoring in its workings. Gambacorta (2009) is a survey with ample references on this literature. Examinations of the potential for and sources of greater procyclicality in the innovated financial system include Allen and Saunders (2004), Borio and White (2003), and White (2006, 2007, 2010). The relationship between leverage and interest rates is discussed in Adrian and Shin (2009*a*). Search or reaching for yield is discussed in Rajan (2005) and Hellwig (2009).

Bordo, Eichengreen, Klingebiel, and Martinez-Peria (2001) present evidence that international crises have been more frequent in recent decades. Minsky (1992) is a succinct presentation of Minsky's views. Other presentations include, in the context of the postwar American financial system, Minsky (1986), and, in the context of the boom of the 1920s and subsequent Depression, Minsky (1984).

The basic model of bubbles presented in this chapter is adapted from the Blanchard (1979) and Blanchard and Watson (1982) "generic" version. Camerer (1989) and Brunnermeier (2008) provide surveys of bubble models. Garber (1989, 1990) offer a skeptical analysis, revisiting some classic bubbles and offering evidence that they were considerably less detached from reality than is commonly assumed. In addition to the bubble models based on credit and liquidity discussed in the text, see Allen and Gale (2000) and Geanakoplos (2009).

Friedman and Kuttner (1993) and Estrella and Mishkin (1998) are examples of relatively successful asset-price predictors of business cycle turning points. Kaminsky, Lizondo, and Reinhart (1998) is a survey of macroeconomic warning signals of currency crises. Berg and Patillo (1999) is a statistical study of models intended to help predict crises. Bustelo (2000) takes a skeptical view of the potential of macroeconomic indicators to predict emerging markets crises.

Söderlind and Svensson (1997) provide a survey of market-based warning signals. Coudert and Gex (2008) and Kumar and Persaud (2002) study indicators of risk aversion and their predictive power for financial crises. Chen, Hong, and Stein (2001) and Finger (2008) study crash warning signals in U.S. equity markets. Malz (2001*a*) provides statistical evidence for the predictive power of implied volatility for extreme asset returns.

Some references to studies of the ability of implied volatility to forecast future realized volatility are provided at the end of Chapter 10. Fleming (1998) summarizes the modeling issues involved. Canina and Figlewski (1993) are more skeptical than most of the literature on second-moment efficiency. Studies of how the risk of financial crises and extreme events might be anticipated by risk-neutral distributions based on option prices include Bates (1991, 1996, 2000), Melick and Thomas (1997), and Malz (1996).