Prone to Fail: The Pre-Crisis Financial System

Darrell Duffie
Graduate School of Business, Stanford University
July 19, 2018

In the years leading up to the financial crisis that began in 2007, the core of the financial system was vulnerable to major shocks emanating from any of a variety of sources. While this particular crisis was triggered by over-levered home owners and a severe downturn in U.S. housing markets (Mian and Sufi, 2015), a reasonably well supervised financial system would have been much more resilient to this and other types of severe shocks.

Instead, the core of the financial system was a key channel of propagation and magnification of losses suffered in the housing market (Aikman, Bridges, Kashyap, and Siegert, 2018). Critical financial intermediaries failed, or were bailed out, or dramatically reduced their provision of liquidity and credit to the economy. In the deepest stage of the crisis, as shown by Bernanke (2018), the failure of Lehman Brothers was accompanied by large, sudden, and widespread increases in the cost of credit to the economy and significant adverse impacts on real aggregate variables.

*Dean Witter Distinguished Professor of Finance, Graduate School of Business, Stanford University, and a Research Associate of the National Bureau of Economic Research. This paper is in preparation for the Journal of Economic Perspectives and for presentation at “The Financial Crisis at 10,” a symposium of the 2018 Summer Institute of the National Bureau of Economic Research. Because I was on the board of directors of Moody’s Corporation from October 2008 to April 2018, I avoid a discussion of credit rating agencies. I am grateful for research assistance from Marco Lorenzon, Yang Song, and David Yang, and for conversations with and comments from Thomas Eisenbach, Gary Gorton, Joe Grundfest, Anil Kashyap, Michael Ohlrogge, Hyun Shin, Andrei Shleifer, Jeremy Stein, Larry Summers, and Paul Tucker.
The core financial system ceased to perform its intended functions for the real economy at a reasonable level of effectiveness. The impact of the housing-market shock on the rest of the economy was correspondingly much larger than necessary.

In the decade before the crisis, the available record suggests an assumption by U.S. regulators that market discipline would support adequate levels of capital and liquidity at the major banks and investment banks. But perceptions by creditors that these firms were “too big to fail” implied a likely failure of market discipline. At the same time, regulatory supervision of these firms was insufficient at both micro-prudential and macro-prudential levels. In particular, the SEC had little focus or capabilities in the prudential supervision of investment banks, money market mutual funds, the commercial paper market, and financial market infrastructure. These were the greatest points of vulnerability in the core of the financial system.

Market discipline did not work. For example, the one-year credit spreads of large banks shown in Figure 1 show that, in the years leading up to the crisis, the largest banks were offered debt financing on amazingly generous terms, presumably because their creditors did not believe that they would be likely to take losses if any of the largest banks were to approach insolvency. Creditors apparently assumed that the biggest banks were too important to be allowed by the government to fail.

In a recent University of Chicago poll, U.S. and European economists were asked to gauge the relative importance of twelve factors contributing to
the financial crisis. The factor receiving the highest average importance rating\textsuperscript{1} in both the European and the American polls was “flawed financial sector regulation and supervision.”

Figure 1. The average one-year credit spread of large banks borrowing US dollars, as measured by the difference between the one-year U.S. Dollar London Interbank Offered Rate (LIBOR) and the one-year overnight index swap (OIS) rate based on the Fed Funds rate. Data source: Bloomberg.

Rich Spillenkothen (2010), director of banking supervision and regulation at the Federal Reserve Board from 1991 to 2006, wrote that “prior to the crisis, career supervisors in the regions and at agency headquarters -- primarily at the Federal Reserve, Office of the Comptroller of the Currency (OCC), and SEC -- failed to adequately identify and prevent the build-up of

\textsuperscript{1}See IGM Forum (2017). I was one of those polled. The other listed factors, in order of assessed average importance among all economists, beginning with the second-most important, were: underestimated risks (financial engineering), mortgages (fraud and bad incentives), funding runs (ST liabilities), rating agency failures, housing price beliefs, household debt levels, too-big-to-fail beliefs, government subsidies (mortgages, home owning), savings and investment imbalances, loose monetary policy, and fair-value accounting.
extreme leverage and risk in the financial system, particularly in large financial institutions.” Oversight by the SEC of the capital adequacy of the largest investment banks was particularly lax.\(^2\) AIG was not effectively supervised by the Office of Thrift Supervision.\(^3\) The Office of Federal Housing Enterprise Oversight placed few limits on the risks taken by the two giant housing finance intermediaries, Fannie Mae and Freddie Mac.\(^4\) Relative to other regulators, the Fed had significantly greater supervisory resources and focus on financial stability, yet failed to uncover solvency and liquidity threats that, with the benefit of hindsight, now seem clear.

The greatest danger to the functionality of the core of the financial system was posed by five systemically large dealers: Bear Stearns, Lehman Brothers, Merrill Lynch, Goldman Sachs, and Morgan Stanley. These firms, called “investment banks,” were exceptionally highly levered and dependent on flight-prone sources of short-term liquidity, including funding from under-regulated money market mutual funds, funding based on fragile collateral, and funding provided indirectly by their prime-brokerage clients.\(^5\)

---


\(^3\) See Polakoff (2009) and Finn (2010).

\(^4\) See Acharya, Richardson, Van Nieuwerburgh, and White (2011) and Stanton (2009).

\(^5\) In his speech “More Lessons from the Crisis,” the President of the Federal Reserve Bank of New York, William Dudley (2009), wrote that “A key vulnerability turned out to be the misplaced assumption that securities dealers and others would be able to obtain very large amounts of short-term funding even in times of stress. Indeed, one particularly destabilizing factor in this collapse was the speed with which liquidity buffers at the large independent security dealers were exhausted. To take just one illustrative example, Bear Stearns saw a complete loss of its short-term secured funding virtually overnight. As a consequence, the firm’s liquidity pool dropped by 83 percent in a two-day span. ….. The second factor contributing to the liquidity crisis was the dependence of dealers on short-term funding to finance illiquid assets. This short-term funding came mainly from two sources, the tri-party repo system and customer balances in prime brokerage accounts. By relying on these sources of funding, dealers were much more vulnerable to runs than was generally appreciated.” For details, including the extent of liquidity provided by prime-brokerage clients, see Duffie (2010).
Admati and Hellwig (2013) emphasize the socially excessive and weakly supervised leverage of the largest financial institutions. The debt of these firms was excessive because it was subsidized by the government through the presumption by creditors that these firms were too big to fail.

In the remainder, I will review the key sources of fragility in the core financial system. Section 1 focuses on the weakly supervised balance sheets of the largest banks and investment banks. Sections 2 and 3 review the run-prone designs and weak regulation of the markets for securities financing and over-the-counter derivatives, respectively. This is not to downplay other sources of systemic risk within the financial system. In particular, weaknesses that allowed the collapses of AIG, Fannie Mae, and Freddie Mac were disastrous. But these three firms were less critical to the day-to-day functionality of the financial system, especially with respect to the continued operation of backbone payments and settlements systems and the provision of liquidity to financial markets. In the final section, I examine in more depth the interplay of too-big-to-fail and the failure of market discipline.

1. **Regulators failed to safeguard financial stability**

   In hindsight, essentially all relevant authorities agree that the largest U.S. financial intermediaries, especially five large investment banks, were permitted by regulators in the years leading up to the crisis to have insufficient capital and liquidity, relative to the risks they took. Authoritative
voices supporting this view after the crisis included successive chairs and other governors of the Federal Reserve Board, Presidents of the Federal Reserve Banks of Boston and New York, SEC Chair Mary Schapiro, the Inspector General of the SEC, the Financial Crisis Inquiry Commission, the Lehman Examiner, the U.S. General Accountability Office, supervisory experts for the Board of Governors of the Fed and the Federal Reserve Bank of New York, and country-report examiners at the International Monetary Fund.

Yet, in the pre-crisis years, there was no apparent urgency to act. I am unable to offer a simple and convincing explanation for this failure. Calomiris and Haber (2015) ascribe the relatively high frequency of U.S. financial crises to broad themes of political economy, including the historical U.S. emphasis on a decentralized banking system. In their words, “A country does not choose its banking system: rather, it gets a banking system that is consistent with the institutions that govern its distribution of political power.”

---


7 Rich Spillenkothen, director of banking supervision and regulation at the Federal Reserve Board from 1991 to 2006, described that “Numerous factors have been cited or suggested to explain the shortcomings of supervision leading up to the crisis. These include: the absence of appropriately robust rules and standards; lack of attention to macro-prudential factors affecting the financial system as a whole; insufficient input from specialists, such as economists and capital and financial markets experts; and a failure of financial regulation to adapt to dramatic changes over time in the structure and activities of the financial system. All of these factors played some role, but they do not fully explain the shortcomings of supervision.” Spillenkothen (2010) also wrote: “Well into the crisis, when the severity and depth of some large banks’ problems were well-known, it appears (based upon FDIC’s published aggregate problem bank assets) that none of the very largest commercial banks, including those that received exceptional government assistance during the crisis, had their bank supervisory (CAMELS) ratings downgraded to problem bank status – a surprising situation that can only be explained by concern over the impact this could have had on financial markets.”

8 Calomiris and Haber later add that financial crises “occur when banking systems are made vulnerable by construction, as the result of political choices.”
For the specific case of the SEC’s weak oversight of the capital and liquidity of the largest investment banks, I am drawn to consider whether the failure to prudentially supervise this risk lies with the original and persistent mission of the SEC to protect the customers of financial firms, to the point of crowding out a focus on financial stability.\(^9\) Perino (2010) describes the political impetus for the creation of the SEC, to protect investors from abuses by financial intermediaries that were brought to light by the depression-era Pecora Commission. As a simple clue to the continuing emphasis by the SEC on investor protection over financial stability, its Inspector General (IG) filed a voluminous (457-page) report on the SEC’s failure to uncover the Madoff Ponzi scheme, but a mere 27-page report on the SEC’s failure to adequately supervise the largest investment banks.\(^10\)

Supervision of the capital and liquidity of the investment banks was done by the SEC’s Division of Trading and Markets. By comparison with notable post-crisis criticism by the Fed of its own supervisory work before the crisis, the reactions of the Division of Trading and Markets to the SEC’s IG report, to the report of the General Accountability Office (2009) on the financial crisis,\(^11\) to criticisms by the Financial Crisis Inquiry Commission,\(^12\) and in other public defenses of its pre-crisis supervision,\(^13\) seem narrow and

\(^9\) Kohn (2014) writes that “the Securities and Exchange Commission is tasked with to protecting investors and maintaining fair, orderly, and efficient securities markets, with a focus on getting adequate information to all investors at the same time. Achieving its objectives may well be necessary for financial stability, but they are not sufficient for the system, even in the areas under its jurisdiction.”


\(^12\) See Sirri (2010).

\(^13\) See Sirri (2009).
grudging. Since the crisis, the Fed has added substantial resources and focus to its supervision of the largest financial institutions.\(^\text{14}\)

An alternative hypothesis for the ineffectiveness of pre-crisis supervision is that it was simply too difficult, within reason, for regulators to detect the excessive buildup of risk and flight-prone short-run debt and derivatives in the core of the pre-crisis financial system, especially given significant financial innovation and complexity,\(^\text{15}\) and given the tendency of some regulated firms to hide their true financial conditions, as exemplified by Lehman’s infamous Repo 105 practice.\(^\text{16}\)

As yet another plausible explanation for the failure of regulators to control the buildup of systemic risk, Gennaioli and Shleifer (2018) propose that investors and policymakers assigned irrationally low probabilities to disaster outcomes, especially with respect to the performance of the housing market. They write: “The Lehman bankruptcy and the fire sales it ignited showed investors and policymakers that the financial system was more vulnerable, fragile, and interconnected than they previously thought. Their lack of appreciation of extreme downside risks was mistaken.”\(^\text{17}\)


\(^{15}\) See Spillenkothen (2010). Eisenbach, Lucca, and Townsend (2012), however, point to the “existence of economies of scale in bank supervision that are sufficiently strong to outweigh the effect of enhanced supervision for larger banks. This result also suggests that, in terms of realized hour allocations, banks in our sample do not appear to have grown to be ’too large to be supervised.’ ”


\(^{17}\) Consistent with the perspective of Gennaioli and Shleifer (2018), an internal review of pre-crisis supervision conducted at the Federal Reserve Bank of New York by Beim and McCurdy (2009), found that “Banks were not pushed too far out into the tail of the risk distribution or asked to review their plans for..."
and Shleifer “put inaccurate beliefs at the center of the analysis of financial fragility.” They note that the second-most important crisis factor according to a poll of leading economists conducted by the IGM Forum (2017), after “flawed financial sector regulation and supervision,” is “underestimated risks.”

Regardless of the relative weights placed on these various explanations for pre-crisis supervisory failures, I will argue here that (i) regulators placed undue reliance on market discipline and (ii) a requirement for reasonable financial stability is that all key financial regulators clearly accept a financial-stability mandate. The second point has been forcefully made by former Fed Vice Chair Donald Kohn (2014), whose primary recommendation is to assign every regulatory agency participating in the U.S. Financial Systemic Oversight Council “a financial stability objective – in carrying out its primary missions it should give weight to any risks posed by the institutions and markets it oversees to the overall stability of the US financial system.” Indeed, an internal review of failures of the Fed’s pre-crisis supervision conducted for the Federal Reserve Bank of New York by Beim and McCurdy (2009) resolved that “From now on systemic risk must be the most important single issue in bank supervision.”

Figure 2 shows the asset-weighted average leverage (the ratio of total accounting assets to accounting equity) of the holding companies of the largest four bank holding companies (J.P. Morgan Chase,\(^{18}\) Bank of America, dealing with an industry-wide liquidity or credit risk event, or to demonstrate their ability to handle a significant loss of confidence in the industry or loss of funding industry-wide.”

\(^{18}\) J.P. Morgan Chase merged during the sample period with Bank One and Chase Manhattan. For these calculations, it was treated on a consolidated basis throughout, pro forma, as though these mergers had occurred at the beginning of the sample period.
Citigroup, and Wells Fargo) and likewise of the five large investment banks. The pre-crisis leverage of the investment banks, Bear Stearns, Lehman, Merrill Lynch, Goldman Sachs, and Morgan Stanley, is especially high. Within each quarter, the leverage of these firms was likely much higher than shown in the figure because they were monitored for compliance only at the end of each quarter (Financial Crisis Inquiry Commission, 2011).

**Figure 2.** Average leverage (weighting by assets) of the holding companies of the investment banks (Goldman Sachs, Morgan Stanley, Lehman, Bear Stearns, Merrill Lynch) and the largest bank holding companies (J.P. Morgan Chase, Bank of America, Citigroup, and Wells Fargo). See Footnote 2 for the treatment of the mergers that created J.P. Morgan Chase. Data source: SEC 10K filings.

In 2002, the European Union introduced rules that would require financial intermediaries operating in the EU to have a consolidated supervisor. For business reasons, all five of these investment banks therefore needed to become supervised at the holding-company level. In 2004 and
2005, they elected to be supervised for this purpose by the SEC under its new Consolidated Supervised Entity (CSE) program.\footnote{In his written testimony before the Financial Crisis Inquiry Commission, the head of the Trading and Markets Division, Eric Sirri, wrote “The CSE program relied on the SEC’s authority under the Securities Exchange Act of 1934 to determine net capital rules for regulated broker-dealer subsidiaries of investment banks. In essence, the entire CSE program was constructed around an alternative net capital regime for the broker-dealer subsidiary, which carried as a condition the affiliated holding company’s consent to group-wide supervision by the Commission.” See Sirri (2010). A weakness of the CSE program was its non-statutory nature (Bhatia, 2011).}

In 2008, as the brewing financial crisis came to a full boil, Bear Stearns and Merrill Lynch were forced into mergers with J.P. Morgan and Bank of America, respectively. Lehman Brothers failed. To support their survival, Goldman Sachs and Morgan Stanley became licensed as bank holding companies, giving them direct access to the banking system’s “safety net.” As a result, the SEC shut down its CSE program.

The introduction of the SEC’s CSE program was probably not directly responsible for a significant increase in leverage among the investment banks. Revealing the prior lack of oversight of these firms, the SEC’s Associate Director of Trading and Markets, Michael Macchiaroli emphasized that “the Commission did not relax and requirements at the holding company level because previously there had been no requirements.” The Director of the Division of Trading Markets, Erik Sirri, also explained that the CSE program was not responsible for a major weakening of capital requirements for the investment banks.\footnote{A former director of Trading and Markets, Lee Pickard suggested in a 2008 that a 2004 change in the SEC’s minimum net capital rule, \textsection{15c-3}, was responsible for a significant increase in leverage of the investment banks (Securities and Exchange Commission, 2004a). This assertion is contradicted by Sirri (2009), Lo (2012), and McLean (2012).} Indeed, Figure 2 shows that the leverage of the investment banks was about as high a decade before the crisis as it was on the opening of the crisis.
The extreme leverage of the five investment banks, the existential crises faced by all of them in 2008, and the big post-crisis drop in leverage of the two survivors, Goldman Sachs and Morgan Stanley, all support a view that the SEC had never supervised the investment banks (or their subsidiaries) adequately from the viewpoint of solvency, even relative to the largest banks. The Inspector General of the SEC found\(^{21}\) that the SEC’s Division of Trading and Markets “became aware of numerous potential red flags prior to Bear Stearns’ collapse, regarding its concentration of mortgage securities, high leverage, shortcomings of risk management in mortgage-backed securities and lack of compliance with the spirit of certain Basel II standards, but did not take actions to limit these risk factors.”

As a further illustration of the limited focus of the SEC on the solvency of the investment banks, Figure 3 shows the SEC’s net capital requirement for each of the five large investment banks in 2005, and their actual net capital levels. The figure makes it obvious that the net capital rule (Katz, 2004) did not constrain the investment banks. The findings of Ohlrogge and Giesecke (2018) imply that during 2001-2007 the SEC’s net capital requirements\(^{22}\) represented an average of under 13% of the actual net capital reported by the five investment banks and the broker-dealer subsidiary of Citigroup. Although the investment banks and their subsidiaries had supplementary forms of capital requirements, none of these


\(^{22}\) Ohlrogge and Giesecke (2016) note the de-facto emphasis of regulators on the early-warning trigger, which is 2.5 times the actual net capital requirement, and find that earlying warning triggers represented 28.9% of the average in their sample of reported net capital.
were effective in controlling solvency risk or emphasized in SEC supervision.  

From a financial-stability perspective, a key concern is that the SEC’s supervision of risk taking by the investment banks focused mainly on the protection of the customers of the investment banks from losses, rather than on the solvency of their balance sheets and the attendant systemic risks. 

For example, a member of the IMF’s country examination staff for the U.S., Bhatia (2011), wrote that the SEC’s mission “stresses ex post enforcement over ex ante prudential guidance.” As another illustration, by my count, only

---

23 The required net capital is 2% of “aggregate debt items” (ADI), which is essentially a measure of customer-related claims on the broker dealer subsidiary of the I-bank. As shown in Figure 3, the net capital requirement is enhanced by an early warning trigger, which is 5% of ADI. The reporting firm is required to notify the SEC whenever the firm’s net capital has breached this 5% ADI early-warning level. Supplementary forms of capital requirement are discussed by Ohlrogge and Giesecke (2018). In written testimony to the Financial Crisis Inquiry Commission, the Inspector General of the SEC, David Kotz, stated that “the [CSE] program did not require CSE firms to have a leverage ratio limit. Further, despite TM [The Division of Trading and Markets] being aware that Bear Stearns’ leverage was high and some authoritative sources describing a linkage between leverage and liquidity risk, TM made no efforts to require Bear Stearns to reduce its leverage. … Bear Stearns was not compliant with the spirit of certain Basel II standards and we did not find sufficient evidence that TM required Bear Stearns to comply with these standards. … Without an appropriate delegation of authority, TM authorized the CSE firms’ internal audit staff to perform critical audit work involving risk management systems, instead of this work being performed by the firms’ external auditors, as the rule that created the CSE program required.” See Kotz (2010).

24 Giesecke and Ohlrogge (2016) write that “a key feature of net capital for broker-dealers is its focus on liquidity, rather than solvency as is the case for bank capital. Calculations of net capital for broker-dealers start with a computation of net worth as defined under generally accepted accounting principles (which thus roughly covers assets minus liabilities, but does not deduct equity). Afterwards, a broker-dealer makes certain adjustments to net worth by adding qualifying subordinated loans, deducting illiquid assets, and then finally applying specified haircuts to the remaining liquid assets in consideration of the market risk they bear. As a result, as the SEC put it, ‘net capital essentially means . . . net liquid assets.’ ” A further indication of the attitude of SEC at the time is contained in testimony to the Financial Crisis Inquiring Commission by the head of Trading and Markets of the SEC, Eric Sirri, who wrote “Although the investment bank holding companies elected to be supervised by the Commission under the CSE program, thereby complying with applicable capital or capital reporting standards at the holding company and regulated entity level, a number of these firms ultimately were overwhelmed by unprecedented demands for liquidity in a crisis of confidence. Despite these extraordinary occurrences, it is important to note that the cash and securities of customers of the broker-dealer were never imperiled, and remained protected by the Commission’s financial responsibility requirements.” See Sirri (2010).
one of a list of 545 pre-crisis\textsuperscript{25} SEC regulatory actions reported in Gadinis (2012) was related to the adequacy of capital or liquidity.

Figure 3. Net capital (blue) and required net capital (red), in 2005, for each of the five largest investment banks: Morgan Stanley (MS), Lehman (LEH), Bear Stearns (BSC), Merrill Lynch (MER), and Goldman Sachs (GS). Data source: SEC 10K filings. An “early warning requirement” is also triggered when net capital falls below the level shown in pink.

According to the Financial Crisis Inquiry Commission (2011), “Michael Halloran, a senior adviser to SEC Chairman Christopher Cox, told the FCIC the SEC had ample information and authority to require Bear Stearns to decrease leverage and sell mortgage-backed securities, as other financial institutions were doing. Halloran said that as early as the first

\textsuperscript{25} Gadinis’ dataset includes all SEC enforcement actions against broker-dealers, for any violation of the securities laws, that was finalized in 1998, 2005, 2006, and the first four months of 2007.
quarter of 2007, he had asked Erik Sirri, in charge of the SEC’s Consolidated Supervised Entities program, about Bear Stearns (and Lehman Brothers), ‘Why can’t we make them reduce risk?’ According to Halloran, Sirri said the SEC’s job was not to tell the banks how to run their companies but to protect their customers’ assets.”

In post-crisis congressional testimony, SEC Chair Mary Schapiro remarked: “It is also clear that the SEC did not do enough as consolidated supervisor to identify certain risks and require additional capital and liquidity commensurate with the risks. As stated previously, the program was in my view insufficiently resourced, staffed, and managed from its inception.”

Indeed, the SEC devoted exceptionally few resources to the supervision of the five large investment banks. In September, 2008, the SEC’s CSE program had a total of only 21 employees supervising these five huge firms, or about four staff members per firm. By comparison, a very rough estimate based on data from staff reports of the Federal Reserve Bank of New York is that the Fed devoted about 19 supervisory staff, on

---

26 See Schapiro (2010).
27 In her testimony before the House Financial Services Committee concerning the Lehman Brothers Examiners Report, Schapiro (2010) stated that “At the time the program was terminated in September 2008, it had approximately 21 staff, including 10 monitoring staff.” According to the Financial Crisis Inquiry Commission (2011), “only 10 ‘monitors’ were responsible for the five investment banks; 3 monitors were assigned to each firm, with some overlap.”

28 Table 1 of Eisenbach, Haughwort, Hirtle, Kovner, Lucca, and Plosser (2017) shows that in 2014 the Fed had 22 supervisory staff for each of its “complex financial institutions,” which at the time were The Bank of New York Mellon Corporation, Citigroup Inc., The Goldman Sachs Group, Inc., JP Morgan Chase & Co., Morgan Stanley, and the U.S. operations of Barclays PLC, Credit Suisse Group AG, Deutsche Bank AG, and UBS AG, as well as the nonbank firms American International Group, Inc., General Electric Capital Corporation, and MetLife, Inc. From the data underlying Figure 1 of Eisenbach, Lucca, and Townsend (2016), I arrive at a rough estimate of 19 staff per firm in 2008 by multiplying the 2014 number, 22, by the ratio of the total number of full-time equivalent supervisory staff at the Fed in 2008 (which was 583) to the corresponding number in 2014 (which was 671).
average, to each of the systemically important financial firms that it
oversaw.\textsuperscript{29}

The Financial Crisis Inquiry Commission (2011) noted that, “In
January 2008, Fed staff had prepared an internal study to find out why none
of the investment banks had chosen the Fed as its consolidated supervisor.
The staff interviewed five firms that already were supervised by the Fed and
four that had chosen the SEC. According to the report, the biggest reason
firms opted not to be supervised by the Fed was the ‘comprehensiveness’ of
the Fed’s supervisory approach, ‘particularly when compared to alternatives
such as Office of Thrift Supervision (OTS) or Securities & Exchange
Commission (SEC) holding company supervision.’ ”

\textbf{2. Securities financing markets: core meltdown risks}

Figure 4 illustrates that, relative to other major economies and in an
absolute sense, credit provision in the United States is significantly more
dependent on capital markets than on conventional bank lending. The
intermediation of U.S. capital markets relies heavily on the largest dealers
and on the markets for financing their large securities inventories, especially
the market for repurchase agreements, or “repos.”

\textsuperscript{29} The Office of the Comptroller of the Currency (OCC) also devotes substantial supervisory resources to
Figure 4. Fraction of credit via capital markets is defined as 100% less the ratio of total credit provided by banks to total credit. Data source: BIS long series on total credit.\(^\text{30}\)

As famously remarked by Diamond (2013), “private financial crises are everywhere and always due to problems of short-term debt.” This particular crisis manifested itself in new forms of short-term debt runs in which repos played a major role. Figure 5 shows a significant increase between 2001 and 2008 in the reliance by dealers on one-day repo financing, both in absolute terms and also relative to longer-term repos. This is consistent with the central hypothesis of Gorton, Metrick, and Xie (2014), that as financial fragility increased over time, wholesale creditors became more and more anxious to have a quick option to cut their exposures.

\(^{30}\) The underlying data can be found in the BIS Statistics Warehouse, at https://stats.bis.org/df=BIS:WEBSTATS_TOTAL_CREDIT_DATAFLOW(2.0):dq=CN+GB+JP+US+XM.P.A+B.M+N.XDC.A%3FstartPeriod=1985-01-01&endPeriod=2017-12-31:pv=1,3~0.0,0~both
Before the crisis, each of the major dealers obtained hundreds of billions of dollars in overnight credit in the repo market. On each repo, a dealer transfers collateralizing securities to its creditor and receives cash. When the repo matures, typically the next morning, the dealer is responsible for returning the cash with interest, and is given back its securities collateral.

Figure 5. Total repo outstanding of U.S. primary dealers, quarterly rolling averages. “Overnight and continuing” repos are those whose original maturity is one day or which are renewed on a daily basis. Term repos are those with an original maturity of more than one day. Data source: Federal Reserve Bank of New York. \(^{31}\)

![Figure 5. Total repo outstanding of U.S. primary dealers, quarterly rolling averages.](image)

Figure 6 is a schematic diagram of the core financing market for U.S. securities, “ground zero” for the new run dynamics. As shown, money-market mutual funds were a mainstay of this funding. Not shown in this diagram, but also important to dealers as a source of secured funding, were

\(^{31}\) The underlying data can be found at https://www.newyorkfed.org/markets/gsds/search.html#
securities-lending firms.\textsuperscript{32} Money funds, sec-lending firms, and other cash investors in repos often held the collateral securities provided to them by dealers in accounts at two “tri-party” agent banks, J.P. Morgan Chase and Bank of New York Mellon. Likewise, the repo investors transferred their cash to the dealers’ deposit accounts at the same two tri-party banks.

\textbf{Figure 6.} A schematic diagram of securities financing at the core of the pre-crisis U.S. financial system.

In the pre-crisis period, each morning, when the dealers’ repos matured and the dealers repaid the cash investors, the dealers needed intra-day financing for their securities inventories until new repos could be arranged and settled near the end of the same day. This intra-day credit was provided by the tri-party agent banks. Even term repos that had not matured on a given day were temporarily cashed out in the morning and financed during the day by the tri-party banks. This practice offered operational simplicity. In this manner, up to $2.8 trillion in intra-day financing was provided to the

\textsuperscript{32} See Copeland, Martin, and Walker (2014a). Securities lending and repo are close substitutes.
dealers every day by the two tri-party agent banks (Copeland, Martin, and Walker, 2014b).

Systemic risk is dramatically magnified when key infrastructure providers such as these two tri-party banks are also large sources of credit to their users. This “wrong-way” systemic risk was further heightened by the practice of settling the cash side of tri-party repos with unsecured commercial bank deposits in the same two tri-party agent banks. These tri-party repo practices exposed the core of the securities funding market to extreme threats in crisis scenarios, and are contrary to well recognized international (CPSS-IOSCO) standards for financial market infrastructure. Since the crisis, a senior industry task force forced the provision of intra-day credit by the tri-party clearing banks to be almost entirely eliminated (Federal Reserve Bank of New York, 2010).

In the event that a dealer’s solvency or liquidity were to come under suspicion, money market funds and other cash investors could decide not to renew the daily financing of the dealer’s securities. Copeland, Martin, and

---

33The settlement of FMI transactions in commercial bank deposits is naturally contrary to principles set down by Committee on Payment and Settlement Systems, Technical Committee of the International Organization of Securities Commissions (CPSS-IOSCO) (2012), whose Principle 9 for financial market infrastructure (FMI) states: “An FMI should conduct its money settlements in central bank money where practical and available. If central bank money is not used, an FMI should minimize and strictly control the credit and liquidity risk arising from the use of commercial bank money.” CPSS-IOSCO (2012) continues by stating, “One way an FMI could minimize these risks is to limit its activities and operations to clearing and settlement and closely related processes.” For more details, see Duffie (2013). Department of the Treasury (2008) writes that “In the United States major payment and settlement systems are generally not subject to any uniform, specifically designed, and overarching regulatory system. Moreover, there is no defined category within financial regulation focused on payment and settlement systems. As a result, regulation of major payment and settlement systems is idiosyncratic, reflecting choices made by payment and settlement systems based on options available at some previous time.” The practice of settling tri-party repos in unsecured commercial bank deposits persists to this day.
Walker (2014a) and Krishnamurthy, Nagel, and Orlov (2014) document that this actually happened to Lehman.

Even if money-fund managers were willing to finance the dealers on a given day, the money fund’s own institutional cash investors could run at the first sign of trouble. Moreover, a key SEC regulation governing the composition of money fund assets, Rule 2a7, precludes investment by money funds in the bonds and other assets that they were assigned as repo collateral. Thus, when a dealer fails, its money-fund counterparties could be forced to firesale the collateral.

As explained by Duffie (2014), if a major dealer were unable to roll over its secured funding during a pre-crisis business day, a tri-party bank’s balance sheet would suddenly become imbalanced by the risk of revaluation of hundreds of billions of dollars worth of securities provided by that dealer as intra-day collateral. This raised several contagion channels, outlined as follows.

First, the tri-party agent banks would have had an incentive, or have been forced, to firesale the securities collateral, causing a sudden drop in the prices of the weaker collateral, of which there was a large amount during the pre-crisis period, including equities and a significant amount of asset-backed securities (Begalle, Martin, McAndrews, and McLaughlin, 2015). The spillover price impact of a firesale into security markets and thus onto other investors could have been severe.
Second, under the stress of an intra-day failure by a client dealer, a tri-party agent bank could easily have been prevented from offering tri-party clearing services or intra-day financing to other major dealers. Both operationally and in terms of access to intra-day credit, access to tri-party services is existential for the major dealers. With no obvious alternative source of financing, a dealer could have been forced to join the firesale.

Third, the entire system depended on the willingness of money fund managers and their own sophisticated institutional investors to remain exposed to dealers. Institutional investors in “prime” money market funds, those permitted to hold non-government securities, are particularly flight prone. In actuality, on September 16, 2008, the Reserve Primary Fund disclosed significant losses on investments in commercial paper issued by Lehman Brothers. The Fund’s net asset value dropped to 97 cents per share, “breaking the buck.”34 Within a few days, according to analysis by Schmidt, Timmermann, and Wermers (2016), over $300 billion of investments in prime money market funds had been redeemed, mainly by “fast” institutional investors.35 These redemptions occurred even at money funds with little or no exposure to Lehman Brothers.

This run on prime money market funds grew in the ensuing days. Absent a halt to this massive flight of one of the main sources of short-term

34 Under post-crisis pressure from the newly created Financial Stability Oversight Council, the SEC changed its rules governing money market mutual funds, allowing only those funds investing exclusively in U.S.-government-quality assets to apply “constant net asset value” (CNAV) accounting, which amounts to a fixed price of a dollar a share until rounding forces a fund’s net asset value per share below one dollar, thus “breaking the buck.” SEC rules were changed to prevent prime money market funds from using CNAV accounting, and forced these funds to have the ability to apply redemption gates and fees. As a result, over $700 billion in prime fund investments shifted to government-only money market funds.
35 See, also, Kacperczyk and Schnabl (2010).
credit to the securities dealers, some or all of these dealers might have been unable to continue financing a substantial fraction of their securities inventories.

Securities dealers, including the huge dealer subsidiaries of bank holding companies such as Citibank, Bank of America, and J.P Morgan, have no direct access to financing from the Fed because they are not banks. The Fed’s discount window can provide financing only to regulated banks, and only for “Fed-eligible” collateral, which does not include a significant portion of the assets that were financed in the repo market before the crisis.

The largest dealers and banks also obtained substantial amounts of short-term funding from the commercial paper market, either directly or indirectly through off-balance-sheet structured investment vehicles (SIVs). As documented by Gorton and Metrick (2010, 2012), Gorton, Metrick, and Xie (2014), and Schroth, Suarez, and Taylor (2014), the asset-backed commercial paper market was particularly prone to runs. The run on prime funds, on other (non-tri-party) sources of repo financing, and on the asset-backed commercial paper market could have caused a complete meltdown of the securities financing market.

Sections 23A and 23B of the Federal Reserve Act effectively prevent the securities dealer subsidiary of a bank holding company from taking indirect advantage of Fed liquidity that is obtained through the bank subsidiary of the same holding company.

Baily, Litan, and Johnson (2008) describe the liquidity risk associated with SIV-sourced credit as follows. “Until the credit crunch hit in August 2007, this business model worked smoothly: a SIV could typically rollover its short term liabilities automatically. Liquidity risk was not perceived as a problem, as SIVs could consistently obtain cheap and reliable funding, even as they turned to shorter term borrowing (see Figure 6). Technically, the SIVs were separate from the banks, constituting as a “clean break” from a bank’s balance sheet as defined by the Basel II Accord (an international agreement on bank supervision and capital reserve levels), and hence did not add to the banks’ capital or reserve requirements. Once the SIVs ran into financial trouble, however, the banks took them back onto their balance sheets for reputational reasons, to avoid alienating investors and perhaps to avoid law suits.”
Only aggressive action by the Fed and the U.S. Treasury averted an enormous collapse of core financial markets and even deeper panic. The Fed invoked its emergency lending authority to provide liberal lender-of-last-resort funding to dealers through the Primary Dealer Credit Facility, the Term Auction Facility, and a host of other new emergency lending facilities. On September 19, 2008, the U.S. Treasury Department offered a guarantee to any money market mutual fund.

Shockingly, without this aggressive fiscal and central-bank lender-of-last-resort support to securities financing markets, the impact of the financial crisis on the real economy would have been far deeper than it actually was.

3. The opaque and unstable pre-crisis swap market

The enormous pre-crisis over-the-counter (OTC) derivatives market contributed significantly to the fragility of the financial system, particularly through its lack of transparency and low extent of collateralization.39

38 The new Fed facilities included, on September 18, 2008, the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility, on October 7, 2008, the Commercial Paper Funding Facility, and on October 21, the Money Market Investor Funding Facility. For more details, see Kacperczyk and Schnabl (2010, 2013). In Chapter 2 of Federal Deposit Insurance Corporation (2017), Lee Davison of the FDIC describes the substantial emergency support by FDIC to banks under the Temporary Liquidity Guarantee Program (TLGP). The TLGP included a substantial volume of guarantees of commercial paper issued by bank holding companies, thus probably adding significantly to emergency government support for the securities financing market. The U.S. government and Federal Reserve offered crisis support through a vast array of other programs that were not directly related to maintaining liquidity in securities financing markets.

39 Cunliffe (2018) remarks that “The financial crisis exposed complex and opaque webs of bilateral derivatives contracts both between financial firms and with real economy end users. These were often poorly collateralised or not collateralised at all.”
Across the entire OTC derivatives market, there were essentially no regulations governing minimum margin, central clearing, and trade reporting. The actual amount of margin provided in practice was low (Financial Stability Board, 2017). Counterparty exposures and the degree to which they were protected by collateral were generally not observable by anyone (including regulators) other than the two counterparties to each individual position. This lack of transparency contributed to run risk.

In the pre-crisis OTC derivatives market, runs could occur in the form of novations (transfers of existing derivatives positions from one counterparty to another) and through the option to terminate derivatives contracts whenever a counterparty experiences an insolvency, a failure to pay, or a change of control. These run options played important roles in the failures of Bear Stearns and Lehman Brothers (Duffie, 2010).

During the financial crisis, as sub-prime-related asset prices fell sharply and concern about counterparty creditworthiness grew, margin calls on derivatives acted as another stress amplifier.

For example, in addition to its direct losses on sub-prime mortgage securities, AIG had sudden heavy cash demands associated with margin calls on the credit-default-swap (CDS) sub-prime mortgage protection that it had provided to a number of major dealers (McDonald and Paulson, 2015). The dependence of these dealers on AIG’s performance on these CDS was an
important factor in the decision by the Fed and then the Treasury to rescue AIG.\textsuperscript{40}

Huge opaque uncleared derivatives exposures added to the atmosphere of extreme concern when the largest dealers began to fail. These fears were ultimately well founded. For example, Cunliffe (2018) notes that “Following its collapse, Lehman’s uncleared derivatives counterparties filed claims totalling $51 billion in relation to its derivatives business. In the event, it was four years before the first payments were made to these uncleared derivatives creditors, and claims against Lehman’s are still ongoing.” At its failure, Lehman had a relatively small book of swap positions in comparison with the largest of the other dealers.

Adding to the systemic risk, participants in the OTC derivatives market generally lacked the ability to bypass dealer balance sheets by trading on exchanges and by clearing their positions at central counterparties. This infrastructure was essentially unavailable outside of the dealer community. Central counterparties were only lightly used by dealers.

Figure 7 shows the huge pre-crisis buildup in the aggregate gross market value\textsuperscript{41} of outstanding over-the-counter (OTC) derivatives, peaking

\textsuperscript{40}McDonald and Paulson (2015) write: “if these six banks [Goldman Sachs, Société Générale, Merrill Lynch, UBS, DZ Bank, and Rabobank] had chosen to respond by selling assets to get back to their pre-AIG default debt to equity ratios, they would have needed to sell $312 billion in assets. Second, the cancellation of the credit default swaps would leave many of the counterparties with unhedged exposure to real estate risk. Retaining this risk could reduce the capacity for other risk-taking. Third, even if one concludes that counterparties could have absorbed losses due to an AIG failure, other market participants would not have known at the time who was exposed and in what amount. For this reason, the failure of any large financial firm can be stressful for the financial system—a conclusion that is not particular to credit default swaps or AIG.”
in 2008 at roughly $35 trillion dollars. As reflected in the figure, post-crisis regulatory collateral and capital requirements subsequently encouraged a major decline in gross outstanding market values through the increased use of central clearing and efficient new methods for conserving space on dealer balance sheets.\footnote{The BIS definition of aggregate gross market value is “Sum of the absolute values of all outstanding derivatives contracts with either positive or negative replacement values evaluated at market prices prevailing on the reporting date. Thus, the gross positive market value of a dealer’s outstanding contracts is the sum of the replacement values of all contracts that are in a current gain position to the reporter at current market prices (and therefore, if they were settled immediately, would represent claims on counterparties). The gross negative market value is the sum of the values of all contracts that have a negative value on the reporting date (ie those that are in a current loss position and therefore, if they were settled immediately, would represent liabilities of the dealer to its counterparties). The term ‘gross’ indicates that contracts with positive and negative replacement values with the same counterparty are not netted.” [https://www.bis.org/statistics/glossary.htm?&selection=312&scope=Statistics&c=a&base=term]\textsuperscript{41}}

There was ample opportunity before the crisis for regulators to control the buildup of systemic risk in the OTC derivatives market. When, in 1998, the Commodity Futures Trading Commission (CFTC) made a move to regulate this market,\footnote{See Commodity Futures Trading Commission (1998).} other regulators pushed back. Treasury Secretary Robert Rubin, Fed Chair Alan Greenspan, and SEC Chairman Arthur Levitt immediately urged Congress to block the proposed regulation,\footnote{See Rubin, Greenspan, and Levitt (1998) and President’s Working Group (1999). Rubin, Greenspan, and Levitt proposed alternative legislation called “Broker-Dealer Lite,” under which the SEC, and not the CFTC, would regulate the OTC derivatives market. Levitt (1998) wrote: “If adopted, the proposed [Broker-Dealer Lite] rules would provide U.S. securities firms with greater flexibility in structuring their OTC derivatives activities by allowing them to conduct transactions involving both securities and non-securities derivative products through one entity. It should be emphasized here that flexibility is the goal.” McCaffrey (2016) writes: “Many observers view the deregulation of OTC derivatives in 2000, through the Commodity Futures Modernization Act, as a serious mistake contributing to the financial crisis. However, no widespread support for external regulation of OTC derivatives existed until after the financial crisis began in 2007. Rather, most analysts accepted on substantive and/or political grounds that the system of private regulation of the OTC derivatives, with informal government oversight, would continue...”\textsuperscript{44}} stating “We have grave concerns about this action and its possible consequences. … We

\footnote{For example, Duffie (2017) explains how compression trading eliminated over $1 quadrillion notional of redundant OTC derivatives.\textsuperscript{42}}
are very concerned about reports that the CFTC’s action may increase the legal uncertainty concerning certain types of OTC derivatives.”

Figure 7. Aggregate gross market values of over-the-counter derivatives, globally. Data source: Bank for International Settlements.

This contretemps between the CFTC and other regulators was more than a typical jurisdictional turf battle. Those blocking the CFTC’s regulatory impulse were concerned that new regulations would reduce the legal certainty of over-the-counter derivatives contracts, or would merely encourage a migration of derivatives trading to London, a competing financial center where the regulation of OTC markets was also extremely light. With significant impetus from Congress, these concerns lead to the passage of de-regulatory legislation, the Commodity Futures Modernization Act of 2000. This was a key step in the striking failure to regulate the enormous build-up of risk in the OTC derivatives market at any time before
the crisis.\textsuperscript{45} From this point, the size of this market grew exponentially, and with almost no oversight by regulators. One of the “major regulatory and supervisory policy mistakes” identified by Spillenkothen (2010) was the “unwillingness to directly regulate the over-the-counter (OTC) derivatives market, relying instead on counterparty and market discipline and on supervisors’ assessments of regulated entities’ risk management practices.”

We now know that it is possible to add substantial prudential regulation to the OTC derivatives market without stamping out market activity because this has actually been done in the post-crisis period. Roughly three quarters of swaps are now centrally cleared, all inter-dealer swaps have minimum margin requirements, and all swap transactions must be reported publicly, with details provided to regulatory data repositories that allow the supervision of exposures to individual market participants. Under the Basel-III regulatory capital accord, the largest dealers are now subject to markedly higher capital requirements on their OTC derivatives exposures. Yet, despite these stringent new regulations, Figure 8 shows significant post-crisis growth in trading activity in the OTC market for interest-rate derivatives, by far the largest segment of the OTC market.

There remain, however, important concerns over the ability to safely resolve the failure of central counterparties (CCPs), which have become enormous concentrations of risk under post-crisis regulations.\textsuperscript{46}

\textsuperscript{45} See Greenberger (2010).
4. **Too-big-to-fail eviscerates market discipline**

Evidence from the crisis of 2007-2009 soundly rejects prior assumptions by regulators concerning the power of market discipline\(^{47}\) and self-preservation to maintain financial stability.

Those old assumptions are encapsulated in a speech in 2000 to the National Bureau of Economic Research by Fed Governor Laurence Meyer, who stated that “As large banking institutions become increasingly complex -- and fund themselves more from non-insured sources -- market discipline

\(^{47}\) In its consultative paper on capital adequacy, the Basel Committee on Banking Supervision (1999) wrote that market discipline “imposes strong incentives on banks to conduct their business in a safe, sound and efficient manner.”
and its prerequisite, public disclosure, must play a greater role. Indeed, increased transparency and market discipline can also help substantially to address concerns about increased systemic risk associated with ever-larger institutions and to avoid the potentially greater moral hazard associated with more-intrusive supervision and regulation.”

In 1997, Fed Chair Alan Greenspan claimed that48 “As we move into a new century, the market-stabilizing private regulatory forces should gradually displace many cumbersome, increasingly ineffective government structures. This is a likely outcome since governments, by their nature, cannot adjust sufficiently quickly to a changing environment, which too often veers in unforeseen directions.”

In a post-crisis hearing, Henry Waxman, Chairman of the House Committee on Oversight and Government Reform, asked49 Greenspan, “Well, where did you make a mistake then?” Greenspan replied, “I made a mistake in presuming that the self-interest of organizations, specifically banks and others, were such that they were best capable of protecting their own shareholders and their equity in the firms.”

With successive changes in the leadership of the Fed, the tone of reliance on market discipline adjusted. In 2007, Fed Chair Bernanke50

48 See Greenspan (1997).
49 See House of Representatives, Committee on Oversight and Government Reform (2008) at page 33. In his prepared remarks, at page 17, Greenspan similarly commented, “Those of us who have looked to the self-interest of lending institutions to protect shareholders’ equity, myself included, are in a state of shocked disbelief. Such counterparty surveillance is a central pillar of our financial markets state of balance.”
50 See Bernanke (2007).
remarked that “The lesson of history appears to be that neither market
discipline nor regulatory oversight alone is completely adequate for keeping
the banking system safe and sound. However, regulators have increasingly
come to appreciate the value of a hybrid system that supplements direct
regulation with a substantial amount of market discipline. Fortunately,
regulators have a variety of ways to restore and strengthen market discipline
for banks, notwithstanding the existence of the federal safety net.” In her
post-crisis revision of that view, Fed Chair Yellen (2015) stated that “The
checks and balances that were widely expected to prevent excessive risk-
taking by large financial firms -- regulatory oversight and market discipline -
- did not do so.”

In his memo on lessons about prudential supervision learned from the
crisis, Rich Spillenkothen, director of banking supervision and regulation at
the Federal Reserve Board from 1991 to 2006, remarked51 that the Fed’s
“strong institutional bias in favor of counterparty and market discipline,
repeated expressions of skepticism regarding the efficacy of regulation, and
an acute sensitivity to regulatory burden did have an effect.”

Reliance on market discipline implies an assumption that excessive
risk taking by a financial intermediary is governed by the intermediary’s cost

supervisory weaknesses conducted by Beim and McCurdy (2009) offers similar and more pointed
criticisms. They describe two “basic assumptions [that] are wrong: 1. ‘Banks can be relied upon to provide
rigorous risk control.’ In reality banks’ internal risk management and control functions were often
ineffective in the run-up to the crisis and were usually trumped by the pressure to do profitable business. 2.
‘Markets will always self-correct.’ A deference to the self-correcting property of markets inhibited
supervisors from imposing prescriptive views on banks.” They wrote that “Interviewees noted the common
expectation that market forces would efficiently price risks and prompt banks to control exposures in a
more effective way than regulators.”
of debt financing, based on creditors’ perceived risk of losses at insolvency. However, before the crisis, there was nothing close to a realistic plan for how to safely resolve the insolvency of systemically important financial firms. Any such failure was likely assumed, correctly as we now know, to cause a major deepening of any crisis or to constitute a crisis in and of itself. This created a presumption among creditors that the largest banks were “too big to fail.”

Thus, despite their thin pre-crisis solvency buffers, the big banks and investment banks were offered amazingly low costs of credit, as already indicated in Figure 1 at the one-year maturity point, and as shown again in Figure 9, which illustrates the average five-year credit-default-swap rates of the largest US and European banks. (A CDS rate is reasonably well approximated, both in theory and empirically, to the market yield compensation paid to creditors for bearing default risk.) Because of this, the pre-crisis cost to big-bank shareholders of expanding their balance sheets with debt financing was much lower than the associated social costs stemming from systemic failure risk. Their trading desks jumped at almost any opportunity to borrow that allowed them to grab a few basis points of profit, because their funding costs were so small.

As an example of this, Andersen, Duffie and Song (2018) model how pre-crisis banks could exploit their exceptionally low credit spreads to capture shareholder profits from even small violations of covered interest parity (CIP). In the post-crisis era, however, much larger CIP violations remain unexploited because of substantially higher big-bank debt funding spreads.
Figure 9. Three-month rolling averages of the average of the five-year credit default swap rates of the holding companies of five large U.S. banks (JPMorgan Chase, Citi Group, Bank of America Merrill Lynch, Morgan Stanley, and Goldman Sachs) and of five large European banks (Deutsche Bank, BNP Paribas, Société Générale, Barclays, and RBS). Figure source: Duffie (2018).

Figure 10 shows a tripling of the total assets of the five largest investment banks and the four largest banks during the decade leading up to the crisis. The incentive to borrow caused by being too big to fail and the lack of methods for safely resolving an insolvency of any of these firms, combined with the forbearance of regulators, created an increasingly toxic brew of systemic risk.

Genaioli and Shleifer (2018) argue against conventional moral-hazard explanations of the excessive pre-crisis leverage of the big banks. I agree. A more plausible cause of this leverage is the borrowing incentives created by
distorted financing costs. In blowing up their balance sheets with debt, these firms did not even need to think about the moral hazard of government bailouts – they merely needed to observe the exceptionally low costs of debt financing offered by creditors who were apparently convinced that these firms would not be allowed to fail. When Lehman ultimately did fail, the surprise of creditors exacerbated the ensuing panic (Bernanke, 2018; Gennaioli and Shleifer, 2018).

Figure 10. Total assets, by year, of Goldman Sachs, Morgan Stanley, Lehman, Bear Stearns, Merrill Lynch, J.P. Morgan Chase, Bank of America, Citigroup, and Wells Fargo. See Footnote 2 for the treatment of the mergers that created J.P. Morgan Chase. Data source: SEC 10K filings.

That the largest financial intermediaries were too big to fail was predicated on the absence of an insolvency resolution methodology that could be applied without cratering the economy. A particular source of intractability to the safe failure resolution of large banks and investment
banks was their huge books of OTC derivatives and repos. These “qualified financial contracts” are legally exempt from the bankruptcy code. Because of this exemption, counterparties to failing firms are not subject to automatic stays or voidable preferences. They can therefore quickly terminate their contracts and keep their collateral. The impact of the resulting fire sale, especially on bond and derivatives markets, would have been far bigger for most of these systemically important firms than even that caused by the failure of Lehman Brothers, which was relatively small by comparison.

In order for market discipline to limit failure risk, creditors would have needed to believe that they could be forced to experience a significant loss at insolvency. Post-crisis legislation, Title II of the U.S. Dodd-Frank Act and the European Union’s Bank Recovery and Resolution Directive, is designed to force wholesale “loss-absorbing” creditors to give up their debt claims the next time a large bank fails. In effect, these creditor claims are to be cancelled and replaced with equity claims. The threat of invoking this resolution scheme, called “bail-in,” is made more credible because the enabling legislation includes a temporary stay on the termination of qualified financial contracts.

52 A revision of the bankruptcy code known as Chapter 14, proposed by Jackson (2016), would eliminate these exemptions for systemically important financial firms. See, also, United States Department of the Treasury (2018). For the largest U.S. bank holding companies, Title II of the Dodd-Frank Act (Library of Congress, 2010) allows the failure resolution administrator to impose a stay on swaps, repos, and other qualified financial contracts.

53 For details, see Duffie and Skeel (2012).

54 United States Department of the Treasury (2018) gives an update on progress with failure resolution methodology and the proposal by Jackson (2016) to amend the U.S. bankruptcy code with a new Chapter 14, which is designed to better address the failure of systemically important financial institutions. Like Title II of the Dodd Frank Act, Chapter 14 would impose a temporary stay on qualified financial contracts such as repos and over-the-counter derivatives.
Whether or not bail-in actually works reasonably well in practice, what matters for big-bank borrowing costs is that creditors believe that it would be tried. It appears that they do now believe this. As shown in Figures 1 and 9, the cost of wholesale unsecured credit for the largest banks has increased dramatically, despite the significant improvements in capital and liquidity achieved under post-crisis regulations. Rosengren (2013), Carney (2014), and Tucker (2014) estimate a full order of magnitude increase in the capital buffers of the largest banks. In support of this conclusion, Figure 11 shows a dramatic increase in the asset-weighted average solvency ratio of the largest financial firms over their pre-crisis levels. This “solvency ratio” is defined as the ratio of tangible common equity to an estimate of the standard deviation of the annual change in the market value of the firm’s assets.55

Correspondingly, Berndt, Duffie, and Zhu (2018) estimate a substantial post-crisis reduction in the average big-bank likelihood of nearing insolvency. But, conditional on that event, they estimate a large post-crisis increase in the probability that the government would force wholesale creditors to take a significant loss. The latter effect dominates, by far, leading to a material post-crisis increase in credit spreads.

This explanation for high post-crisis big-bank credit spreads is at odds with that put forward by Sarin and Summers (2016), who instead argue for a continuing failure of these firms to improve their solvency. Sarin and Summers suggest that their high post-crisis credit spreads reflect the reduced

55 The estimated standard deviation of annual changes in firm asset value is obtained by inverting the Merton-Black-Scholes model to obtain the implied volatility of assets, based on the modeling in Berndt, Duffie and Zhu (2018).
franchise values of their business operating models, rather than a reduced reliance by creditors on too-big-to-fail.

**Figure 11.** The average, weighted by accounting assets, of the solvency ratios of the financial holding companies JPMorgan Chase, Citi Group, Bank of America Merrill Lynch, Morgan Stanley, and Goldman Sachs. The solvency ratio is defined as the ratio of tangible common equity to an estimate of the standard deviation of the change in market value of the firm’s assets. Source: Calculations by Berndt, Duffie, and Zhu (2018).

A belief by creditors that the largest banks are no longer too big to fail leads to a better alignment of the risk-taking incentives of these banks with social incentives to control systemic risk. The greater is the credit spread of a financial intermediary, the greater is the impact of debt overhang in reducing the incentives of its shareholders to expand the intermediary’s balance sheet with debt financing. Indeed, since the crisis, significant increases in unsecured dealer credit spreads have forced the trading desks of
the largest dealers to restrict access to their balance sheets, and to charge
their trading clients for newly designated “funding value adjustments”
(FVAs). Andersen, Duffie, and Song (2018) explain these FVAs as debt-
overhang costs to bank shareholders for enlarging their balance sheets. That
is, because of new failure resolution rules, market discipline has to some
extent finally begun to work.

Although the incentives of big-bank shareholders to expand their
balance sheets to provide immediacy to the market are now more aligned
with social incentives, day-to-day market liquidity has in some cases
suffered, a different form of social cost. There remains an under-exploited
opportunity to bypass big-bank balance sheets with greater use of
centralized “all-to-all” trade venues, such as swap and bond exchanges.

References

Acharya, Viral, Matthew Richardson, Stijn Van Nieuwerburgh, and Lawrence J.
White. 2011. Guaranteed to Fail, Fannie Mae, Freddie Mac, and the Debacle of

Acharya, Viral, and Philipp Schnabl. 2009. “How Banks Played the Leverage
‘Game’? In Restoring Financial Stability: How to Repair a Failed System, edited by V.
Acharya and M. Richardson, Chapter 2, pp. 83-100, New York: Wiley.

Admati, Anat, and Martin Hellwig. 2013. The Bankers’ New Clothes: What’s Wrong

macroprudential regulation have prevented the last crisis?” Working paper, Bank of
England. June. [URL to be supplied]

Andersen, Leif, Duffie, Darrell, and Yang Song. 2018. “Funding Value Adjustments,”
NBER Working Paper Number w23680, Graduate School of Business, Stanford


Further examples are provided by Duffie (2018).


Government Accountability Office. 1998. “Risk Based Capital: Regulatory and Industry Approaches to Capital and Risk,” Report to the Chairman, Committee on Banking, Housing, and Urban Affairs, U. S. Senate, and the Chairman, Committee on


Islamorada, Florida, Federal Reserve Board of Governors, January. 


