Summary: The financial crisis of 2007/8 raises questions about the objectives of financial regulation and how such objectives may be met. Traditionally, capital requirements have been the cornerstone of the regulation of banks. However, the run on Bear Stearns in March 2008 led to its demise in spite of Bear Stearns meeting the letter of its regulatory capital requirements. Risk-based capital requirements that underpin the Basel approach fail to make the distinction between the inherent riskiness of the asset and its systemic importance. Liquidity requirements that constrain the composition of assets may be a necessary complement to traditional capital requirements. We also examine the rationale for imposing a maximum leverage ratio that has gained favor in the US and more recently in Switzerland. The rationale for a leverage ratio derives not from the traditional view that capital is a buffer against loss on the assets, but rather from the stability of liabilities in an interrelated financial system.
1. Introduction

The run on Bear Stearns in March 2008 and the subsequent opening of the Federal Reserve’s discount window to primary dealers has ignited a debate on the role of major broker-dealers in the smooth functioning of the financial system. The Bear Stearns episode has also opened up a debate on the rationale for financial regulation, and what tools may be used to achieve the objectives of financial regulation. The most pressing policy question has been whether broker-dealers (including the major investment banks) should fall under banking regulation overseen by the Federal Reserve, and if so how they should be regulated.

Traditionally, capital requirements have been the cornerstone of the regulation of banks. The rationale for capital requirements lies in maintaining the solvency of the regulated institution. By ensuring solvency, the interests of creditors – especially the retail depositors – can be protected. A creditor who has the ability to monitor the firm can protect his interests through the enforcement of covenants and other checks on the actions of the firm’s managers. However, in the case of a traditional deposit-funded bank, the creditors are the small retail depositors. Small depositors face a coordination problem in achieving the monitoring and other checks that large creditors are able to put in place. The purpose of bank regulation is seen as the protection of the interests of depositors by putting into place through regulation the restrictions on the manager’s actions that would arise in normal creditor-debtor relationships.

The traditional rationale for capital regulation leads naturally to the conclusion that the key determinant of the size of the regulatory capital buffer should be the riskiness of the assets. If the purpose of regulation is ensuring solvency, then the riskiness of assets determines the size of the regulatory capital buffer required. This is because the degree to which solvency can be ensured depends on the likelihood that the realized value of assets falls below the notional value of the creditors’ claim. The original Basel capital accord of 1988 introduced coarse risk buckets into which assets could be classified, but the Basel II rules have taken the idea much further, by refining the gradations of the riskiness of the assets, and fine-tuning the regulatory capital to the risks of the assets held by the bank.

However, the turmoil in the financial system witnessed in the 2007/8 financial crisis poses a challenge to the traditional view of regulation. The Basel II regulations that are in the process of being adopted throughout most of the advanced economies have largely been a by-stander in the unfolding credit crisis that began with the subprime crisis in the United States.

In particular, recent events raise a fundamental challenge to the traditional approach to financial regulation that rest on identifying solvency with capital. In particular, solvency regulation has come up short in ensuring stability of the financial system. The issue is highlighted in a recent open letter written by Christopher Cox, the chairman of the

---

1 Dewatripont and Tirole (1993) discuss the underlying contact theory principles behind the prudential regulation of banks.
Securities and Exchange Commission explaining the background and circumstances of the run on Bear Stearns in March 2008.  

[T]he fate of Bear Stearns was the result of a lack of confidence, not a lack of capital. When the tumult began last week, and at all times until its agreement to be acquired by JP Morgan Chase during the weekend, the firm had a capital cushion well above what is required to meet supervisory standards calculated using the Basel II standard.

Specifically, even at the time of its sale on Sunday, Bear Stearns' capital, and its broker-dealers' capital, exceeded supervisory standards. Counterparty withdrawals and credit denials, resulting in a loss of liquidity - not inadequate capital - caused Bear's demise.

Thus, in spite of Bear Stearns meeting the letter of its regulatory capital requirements, it got into trouble because its lenders stopped lending. Put differently, the problem was on the liabilities side of the balance sheet, rather than on the asset side.

One possible counterargument by Basel traditionalists might be to question the sharp distinction between solvency and liquidity in the SEC Chairman’s letter. They may argue that the run was triggered by concerns over asset quality, and a rapid sale of assets to meet the run would have revealed that Bear Stearns was insolvent. The coordination failure scenario painted in Bryant (1980) and Diamond and Dybvig (1983) raise the possibility of a sound bank succumbing to a self-fulfilling run, but in practice, runs happen to weak banks. Strong banks do not typically suffer runs, even though the run outcome is always a logical possibility.

Runs are typically associated with weak fundamentals. We will deal with this point in some detail in what follows. However, for policy purposes we need to distinguish equilibrium outcomes from efficient outcomes. Even if it is difficult to distinguish solvent banks from illiquid banks in practice, the distinction is nevertheless useful from a policy perspective since we can discuss the desirability of alternative policy measures to nudge the outcome in one direction or another.

Thus, the SEC Chairman’s point still needs to be addressed. The run on Bear Stearns happened in spite of Bear Stearns meeting the letter of its capital requirement. We need to understand why, and whether better rules can be put in place.

In this paper, we will argue that if the purpose of financial regulation is to ensure the stability of the financial system as a whole, then the traditional approach to financial regulation built around risk-based capital requirements is inadequate. In a system context, actions taken by financial institutions have spillover effects that affect the interests of other financial institutions. System stability then takes on the attributes of a public good. As with any other public good, market failure entails inefficient provision of the public

---

good, where actions that are individually rational for each market participant may lead to an inefficient outcome overall. The objective of financial regulation in a system context is to levy the appropriate Pigovian taxes that mitigate the externalities as far as possible, and thereby take the financial system as a whole closer to an efficient outcome.

In particular, there is a difference between the riskiness of the asset and the systemic importance of the asset. Sometimes, as we will see shortly, even a very safe asset under the Basel approach may have important systemic impact that arises from the way that financial intermediaries’ claims are interwoven, and how they react to unfolding events.

The system approach to financial regulation suggests that the current capital regime will need to be overhauled to reflect two further components in the regulatory toolkit.

- First, there is a case for liquidity regulation, which places limits on the composition of the balance sheet, not merely the relative size of equity to total assets.
- Second, even assets that have traditionally been viewed as very safe may justifiably face a regulatory capital charge. Indeed, a simple leverage constraint that does not take account of the riskiness of assets may be a better way to ensure system stability as compared to the traditional risk-based capital charge.

In what follows, we will expand on these points. Both elements have much in common with several recent proposals for the reform of financial regulation, such as Kashyap, Rajan and Stein (2008). Although there are significant overlaps in the motivation, there are also some differences in the rationale and focus. Our discussion here is motivated by the debates on the regulation of the broker dealer sector, and hence focuses on liquidity crises of the kind that undermined Bear Stearns, rather than the general shortage of capital during a downturn. However, it is clear that the two issues should be considered together in an overall agenda for regulatory reform.

We begin with some general remarks on the importance of a system perspective in financial regulation.

2. System Perspective

Financial stability is best viewed from a system perspective, rather than from the point of view of each individual financial institution. Andrew Crockett (2000) has argued for a distinction between the micro-prudential dimension of financial stability from the macro-prudential one. The micro-prudential dimension has to do with the soundness of individual institutions. The macro-prudential dimension has to do with the stability of the system as a whole. In his opening speech at the recent Jackson Hole conference, Fed chairman Ben Bernanke (2008) has argued for the superiority of the macro-prudential perspective.

It is a truism that ensuring the soundness of each individual institution ensures the soundness of the system as a whole. Crockett (2000) makes the point that this statement
is unhelpful, since it does not address how the soundness of all individual institutions are achieved. Actions that ensure the soundness of one institution may not be consistent with ensuring the soundness of another. Unless there are good reasons to believe that policies that ensure the soundness of a particular individual institution will invariably promote overall stability of the system, the prescription is a vacuous one.

Take a simple example, illustrated by figure 1 below in the spirit of Allen and Gale (2000). Bank 1 has borrowed from Bank 2. Bank 2 has other assets, as well as its loans to Bank 1. Suppose that Bank 2 suffers credit losses on these other loans, but that the creditworthiness of Bank 1 remains unchanged. The loss suffered by Bank 2 depletes its equity capital. In the face of such a shock, a prudent course of action by Bank 2 is to reduce its overall exposure, so that its asset book is trimmed to a size that can be carried comfortably with the smaller equity capital.

Figure 1

![Diagram of Bank 1 and Bank 2]

The micro-prudential imperative is for Bank 2 to reduce its overall lending, including its lending to Bank 1. By reducing its lending, Bank 2 achieves its micro-prudential objective of reducing its risk exposure. However, from Bank 1’s perspective, the reduction of lending by Bank 2 is a withdrawal of funding. Unless Bank 1 can find alternative sources of funding, it will have to reduce its own asset holdings, either by curtailing its lending, or by selling marketable assets.

In the case where we have the combination of (i) Bank 1 not having alternative sources of funding, (ii) the reduction in Bank 2’s lending being severe, and (iii) Bank 1’s assets being so illiquid that they can only be sold at fire sale prices, then the withdrawal of lending by Bank 2 will feel like a run from the point of view of Bank 1. In other words, a prudent shedding of exposures from the point of view of Bank 2 is a run from the point of view of Bank 1. Arguably, this type of run is what happened to the UK bank Northern Rock, which failed in 2007. The same perspective is useful in thinking about the run on Bear Stearns. In the letter written by SEC chairman Christopher Cox cited above, there is the following paragraph.

It is worth noting, however, that net capital rules are designed to preserve investors' funds and securities in times of market stress, and they served that purpose in this case. This investor protection objective was amply
satisfied by the current net capital regime, which - together with the protection provided by the Securities Investor Protection Corporation (SIPC) and the requirement that SEC-regulated broker-dealers segregate customer funds and fully-paid securities from those of the firm - worked in this case to fully protect Bear's customers.

Indeed, the run on Bear Stearns helped to protect investors’ funds. But this way of securing investors funds had the undesirable effect of undermining Bear Stearns itself. From a system perspective, the run is an undesirable outcome, even if it conforms to the micro-prudential objectives of the creditors to Bear Stearns.

The lesson is that the micro-prudential imperative of ensuring solvency at the level of the individual institution may not ensure the objective of achieving system stability. The truism that ensuring the solvency of each individual institution achieves overall system solvency is unhelpful as a policy prescription, since there may be a conflict between enhancing the solvency of one institution and maintaining the stability of the system as a whole. Therefore, as a practical matter, it is important for policy to be formulated from a system vantage point from the outset.

Thus, our starting point in the discussion is the following proposition.

**Proposition 1.** It is possible that actions that enhance the soundness of an individual financial institution undermine the stability of the system as a whole.

The system perspective has the virtue of opening up for scrutiny the motivation of the creditors to a bank suffering a run, as well as the fundamentals of the bank itself. Consider the situation depicted in figure 2 below, where bank 0 has \(N\) creditors. Among them are other banks, hedge fund clients who hold deposits at bank 0, or money market mutual funds. For convenience, we have labeled all creditors as “banks”.

![Figure 2](image)
A run outcome is associated with the self-confirming belief held by a creditor bank that other creditors will take the prudent course of action and cut funding to the stricken bank. Belief that others will cut funding justifies cutting funding oneself. However, in practice, runs are associated with weak fundamentals and more jittery creditors. If the fundamentals were stronger or the creditors were more relaxed (or both), the run outcome might be averted.

There is more at stake here than just the methodological point about equilibrium selection. The hope is that if policy makers could engineer the initial conditions through appropriate regulation so that fundamentals of the bank were stronger and the creditors were less jittery, then they could induce the stable non-run outcome, instead of the run outcome.

A useful framework for thinking about this problem is provided by an example given by Lawrence Summers (2000) in his Ely Lecture. In his lecture, Summers proposes the following thought experiment.

Imagine that everyone who has invested $10 with me can expect to earn $1, assuming that I stay solvent. Suppose that if I go bankrupt, investors who remain lose their whole $10 investment, but that an investor who withdraws today neither gains nor loses. What would you do?

…

Suppose, first, that my foreign reserves, ability to mobilize resources, and economic strength are so limited that if any investor withdraws I will go bankrupt. It would be a Nash equilibrium (indeed, a Pareto-dominant one) for everyone to remain, but (I expect) not an attainable one. Someone would reason that someone else would decide to be cautious and withdraw, or at least that someone would reason that someone would reason that someone would withdraw, and so forth.

…

Now suppose that my fundamental situation were such that everyone would be paid off as long as no more than one-third of the investors chose to withdraw. What would you do then? Again, there are multiple equilibria: everyone should stay if everyone else does, and everyone should pull out if everyone else does, but the more favorable equilibria seems much more robust.

I think that this thought experiment captures something real. On the one hand, bank runs or their international analogues do happen. On the other hand, they are not driven by sunspots: their likelihood is driven and determined by the extent of fundamental weaknesses. ³

There are two dimensions to the Summers thought experiment – the strength of fundamentals and how jittery the creditors are.

³ Summers (2000 p. 7)
The first has to do with the threshold for the proportion of creditors who need to coordinate in order to attain the good outcome. The weaker are the fundamentals, the more fragile is the borrower’s balance sheet to withdrawals. Summer appeals to our strong intuition that if the threshold value for coordination is close to 1, coordination is very difficult to achieve. However, if the threshold is much less, then coordination is easier.

The second dimension is the potential cost of mis-coordination. In the Summers example, the potential cost of failing to coordinate is losing one’s stake of $10, compared to the reward to successful coordination of $11. The higher is the cost of mis-coordination, the more jittery are the creditors. Again, our intuition would be that when the costs of miscooordination are high, coordination is more difficult to achieve.

It is possible to solve the Summers game using global game methods and verify that the two dimensions of the problem determine the unique equilibrium outcome. The unique outcome depends on the fundamentals along the two dimensions – the strength of fundamentals and the costs of miscoordination – and can be depicted diagrammatically in terms of the unit square below.

![Figure 3](image)

The horizontal axis indicates the coordination threshold $k$ of the proportion of creditors who need to remain invested in order to achieve the good outcome. The vertical axis measures the cost of miscoordination $c$ expressed as a proportion of the payoff to successful coordination.

The global game analysis confirms the strong intuition articulated by Summers that in the bottom left hand corner of the unit square (when the threshold for coordination is low and the cost of miscoordination is low), we achieve successful coordination. In the opposite

---

4 See Morris and Shin (2002).
corner, when both the threshold for coordination and the cost of miscoordination are high, the good outcome cannot be achieved. The exact dividing line between the good and bad regions depends on other parameters of the global game, but the benchmark dividing line is the straight line that cuts the unit square through the diagonal, as depicted in Figure 3 above.

One value-added of the global game analysis is that we can confront our intuitions with the formal analysis. Indeed, for the parameter values given by Summers in his thought experiment, the corresponding point in the unit square above is (2/3, 10/11), since at least two thirds of investors need to keep their money in, and the cost of miscoordiation is $10 dollars, compared to the good payoff of $11. This point lies in the failure region. The global game analysis suggests that Summers may have been too sanguine about the possibility of forestalling the run. But leaving aside specific cases, the more general lesson is that coordination failures can be remedied by changes in the environment that makes banks more robust to withdrawals, and which lowers the opportunity cost of miscoordination. This was a line of argument we pursued in our earlier currency attacks paper (Morris and Shin (1998)).

In the banking context, if the debtor bank held more cash in place of illiquid assets, it could meet the withdrawals more easily, thereby lowering the threshold for coordination among the creditor banks. The cost of miscoordination $c$ for the creditor banks could also be reduced if they held more cash, since they would be less vulnerable to a run themselves. A more liquid creditor bank would be less jittery. It is possible to formalize within a global game the idea that greater cash holdings by creditors will reduce the cost $c$ of miscoordination and we pursue such a formal analysis elsewhere.

Thus, to anticipate one of our conclusions later, liquidity requirements on banks may reduce the potential for runs through two channels. First, it makes debtor banks more robust to withdrawals. Second, it makes creditor banks less trigger-happy. The second channel (through the motivation of creditor banks) is an insight that can only be gained in a system context.

On the same theme, any institutional feature that makes creditors more constrained to curtail lending in reaction to events will undermine system stability. A prominent example in the context of Bear Stearns is the role of the tri-party repo agreement that Bear Stearns entered into with money market mutual funds. In these agreements, Bear Stearns pledged illiquid securities as collateral in return for money market mutual funds providing them short-term funding. The transaction was overseen by a central counterparty who held the collateral and administered the payments.

By their charter, most money market funds are prevented from holding the types of the illiquid securities pledged by Bear Stearns. If Bear Stearns had become illiquid, the collateral assets would have reverted to the creditors – i.e. the money market mutual funds, who would have been forced to sell the collateral assets possibly at a large loss, making them “break the buck” – i.e. suffer a loss that pushed the value of their assets
below par value. Such losses would have opened up the prospect of a run by retail investors on the money market mutual fund sector itself.

Some commentators have described the money market mutual funds as being extremely risk-averse, but a more accurate description of the motivation is in terms of the cost of miscoordination. In effect, for the money market mutual funds, the cost of miscoordination was extremely high. Hence, they heeded the imperative to be prudent, and withdrew their funding to Bear Stearns. Other creditors, such as Bear’s hedge fund clients would thus have been less willing to leave their money in.

The involvement of money market mutual funds in the triparty repo is an instance where institutional constraints made the run outcome more likely. It suggests that reforms of institutional arrangements could change the underlying payoffs in the coordination game in the direction of making the system less fragile. We summarize this lesson as follows.

**Proposition 2.** A creditor run is more likely if the coordination threshold is high or when the cost of miscoordination is high. Policies that lower the coordination threshold and the cost of miscoordination are likely to promote system stability.

The system perspective also raises an important distinction between the fundamental riskiness of an asset from its systemic importance. Even if an asset is very safe from the point of view of its credit risk profile, it may have a large impact on the stability of the system as a whole.

Consider the example illustrated in figure 4 below. In this example, Bank 1 holds mortgage backed securities (MBSs) as its assets, and funds the holding of the securities by pledging them as collateral in overnight repurchase agreements (overnight repos). Specifically, Bank 1 sells the securities to Bank 2 with the understanding that it will buy them back the next day at a pre-arranged price. Then, at the end of each day, the transaction is repeated. The repo enters as a liability on Bank 1’s balance sheet, and enters as a reverse repo on the asset side of Bank 2’s balance sheet.

**Figure 4**
Bank 2, for its part, funds its lending to Bank 1 by pledging the same securities to another bank (Bank 3) further down the chain. Thus, Bank 2 has the following short-term, safe balance sheet. On the asset side, it has overnight reverse repos to Bank 1. On the liabilities side, it has overnight repos to Bank 3.

From Bank 2’s perspective, its assets are extremely safe, for two reasons. First, the assets are short-term, and so the range of possible realizations of the asset value is small. Second, the loan is fully collateralized, so that even if Bank 1 is not able to repay, Bank 2’s claim is protected. Provided that the “haircut” on the repo is sufficiently large (i.e. the collateral value exceeds the loan value by enough), then Bank 2’s asset value is protected.

Furthermore, the two sides of Bank 2’s balance sheet have a maturity profile that matches perfectly. Both are overnight transactions. Hence, Bank 2 can react to any change in the environment by flexibly reducing the size of its balance sheet. By reducing the amount of the reverse repo to Bank 1 that it is willing to roll over, it can reduce its asset exposure. Similarly, Bank 2 is in a good position to meet any run on its liabilities. If Bank 3 refuses for some reason to roll over the overnight repos, Bank 2 can immediately respond by refusing to roll over its reverse repos to Bank 1. In this sense, Bank 2 is a very safe bank. From the Basel capital perspective, the required capital on Bank 2 would be extremely low. Therefore, the leverage that Bank 2 can attain would be very high. The Basel perspective justifies the high leverage by appealing to the safe nature of Bank 2’s assets when viewed in isolation.

However, when viewed from a system perspective, Bank 2’s assets are systemically important, since Bank 2’s assets are the mirror image of Bank 1’s liabilities. If Bank 1’s assets are illiquid, so that they cannot realize much value in a fire sale, then the impact on Bank 1’s solvency of a run on its liabilities would be severe. If Bank 2 refuses to roll over the overnight reverse repos to Bank 1, this will force Bank 1 to sell its MBS assets, unless it can find alternative sources of funding (say, from the central bank). Thus, from a system perspective, Bank 2’s assets have high systemic impact, even though they are very safe from the point of view of credit risk. This leads to our third proposition:

**Proposition 3.** There is a distinction between risky assets and systemically important assets. Safe assets can be systemically important.

The distinction between risky assets and systemically important assets take on added significance in a market-based financial system built around the practice of secured lending through repurchase agreements. In a repurchase agreement, the borrower sells a security today for a price below the current market price on the understanding that it will buy it back in the future at a pre-agreed price. The difference between the current market price of the security and the price at which it is sold is called the “haircut” in the repo, and fluctuates together with general funding conditions in the market.
The systemic impact of collateralized lending is especially strong when the haircut on the repo contract fluctuates in response to market conditions. The reason is that the haircut determines the maximum permissible leverage achieved by the parties involved in the transaction. In figure 4, suppose that the haircut on Bank 1’s repos is 2%, so that Bank 1 can borrow 98 dollars for 100 dollars worth of securities pledged. Then, to hold 100 dollars worth of securities, Bank 1 must come up with 2 dollars of equity. Thus, if the repo haircut is 2%, the maximum permissible leverage (ratio of assets to equity) is 50.

Suppose that Bank 1 leverages up the maximum permitted level. Such an action would be consistent with the objective of maximizing the return on equity, since leverage magnifies return on equity. Bank 1 thus has a highly leveraged balance sheet with leverage of 50. If at this time, a shock to the financial system raises the market haircut, then Bank 1 faces a predicament. Suppose that the haircut rises to 4%. Then, the permitted leverage halves to 25, from 50. Bank 1 then faces a hard choice. Either it must raise new equity so that its equity doubles from its previous level, or it must sell half its assets, or some combination of both.

Times of financial stress are associated with sharply higher haircuts. The increase in haircuts entail very substantial reductions in leverage, necessitating asset disposals or raising of new equity. Raising new equity or cutting assets entail painful adjustments for Bank 1. Raising new equity is notoriously difficult in distressed market conditions. But selling assets in a depressed market is not much better. The evidence points to banks adjusting leverage by adjusting the size of their balance sheets, leaving equity intact (Adrian and Shin (2007, 2008)). Thus, fluctuations in leverage are associated with pronounced fluctuations in the willingness to lend.

To the extent that the financial system as a whole holds long-term, illiquid assets financed by short-term liabilities, any tensions resulting from a sharp increase in repo haircuts will show up somewhere in the system. Even if some institutions can adjust down their balance sheets flexibly in response to the greater stress, there will be some pinch points in the system that will be exposed by the distressed conditions. The example in Figure 4 is an illustration.

The fluctuations in leverage in the context of widespread secured lending exposes the myth of the “lump of liquidity” in the financial system. It is tempting to be misled by our use of language into thinking that “liquidity” refers to a stock of available funding in the financial system which will be redistributed to those who need it most. So, for instance, in the examples given in Figures 1 or 2, the idea would be that when the funding tries up from one creditor, the borrower can tap alternative sources.

However, the fluctuations in funding conditions associated with fluctuations in repo haircuts exposes the idea of the “lump of liquidity” as a myth. When liquidity dries up, it disappears altogether rather than being re-allocated elsewhere. When haircuts rise, all balance sheets shrink in unison. Thus, there is a generalized decline in the willingness to lend. Liquidity should be understood in terms of the growth of balance sheets (i.e. as a flow), rather than as a stock (Adrian and Shin (2007), Fisher (2008)).
Indeed, the very term “secured lending” suggests that the assets are safe in terms of credit risk, since the loan is secured on collateral. However, funding conditions overall will vary substantially as haircuts fluctuate. The example in Figure 4 illustrates the fact that the fluctuations in Bank 2’s assets have systemic impact, even though it is safe from credit risk. In this way, it is possible that there is a divergence between the riskiness of the asset from its systemic impact. The Basel perspective of focusing only on the credit risk of the asset obscures this important distinction.

The distinction between risky assets and systemically important assets turns out to be crucial for broker dealers, since many of the items on the balance sheet of an investment bank are precisely those that are (i) collateralized and (ii) short term. Thus, as a prelude to our main discussion, we first examine the balance sheet characteristics of broker dealers.

3. Broker Dealer Balance Sheets

Broker dealers, the financial sector that includes the major investment banks, have balance sheets that are in sharp contrast to the archetypal deposit-funded bank. Below, we summarize the balance sheet of Lehman Brothers, as at the end of the 2007 financial year (November 30th, 2007).

The two largest classes of assets are (i) long positions in trading assets and other financial inventories and (ii) collateralized lending. The collateralized lending reflects Lehman’s role as prime broker to hedge funds and other borrowers, and includes reverse repos and other types of collateralized lending. Much of the collateralized lending will be short term, and therefore very safe from a credit risk perspective. They are precisely the types of assets that could be systemically important even though their credit risk may be small.

The other feature of the asset side of the balance sheet is how small the holding of cash is. The cash holding is $7.29 billion out of a total balance sheet size of $691 billion. However, the cash holding would be an underestimate of the cash that could be raised at short notice if the securities holdings include liquid securities that can be sold to raise cash.

The liabilities of Lehman Brothers reflect the short-term nature of much of its liabilities. The largest component is collateralized borrowing, including repos. Short positions (“financial instruments and other inventory positions sold but not yet purchased”) is the next largest component. Long-term debt is only 18% of total liabilities.

One notable item is the “payables” category, which is 12% of total balance sheet size. Payables include the cash deposits of Lehman’s customers, especially its hedge fund clientele. It is for this reason that “payables” are much larger than “receivables” on the asset side of the balance sheet (only 6%). Hedge fund customers’ deposits are subject to withdrawal on demand, and hence may be an important source of funding instability. We
will return to this issue below, where we discuss Bear Stearns’s balance sheet, and more prominent reliance on payables to customers.

Finally, we note that the equity of Lehman is only 3% of total assets ($22.5 billion), implying a leverage of 30.7. This is a much higher number than for commercial banks, which typically maintains a leverage ratio of 10 to 12. The higher leverage of investment banks reflects both the relatively low credit risk of the assets held, as well as the short-term nature of much of the claims and obligations. Indeed, the balance sheet as a whole
consists precisely the type of assets and liabilities that have low credit risk, but have high systemic impact.

The balance sheet for Bear Stearns has many of the same characteristics noted for Lehman Brothers, although there are also some notable differences.

On the asset side, long positions in securities and collateralized lending form the bulk of the assets. The long positions also include assets of special purpose entities that have been consolidated in accordance with accounting rules. As a counterpart to the asset holding in the special purpose entity, the liabilities of the SPE are also consolidated. Since SPEs fund themselves mainly with short-term borrowing (such as commercial paper), the liability item “short term debt” includes the liabilities of such entities.

One notable feature of Bear Stearns is the large proportion of funding that comes from payables – fully 22% of total balance sheet size. As with Lehman Brothers, the bulk of the payables are deposits of hedge fund customers, and reflect the large prime brokerage business at Bear Stearns. Because hedge fund customer deposits are payable on demand, they are vulnerable to a classic run that reflects coordination failure among the hedge fund customers. Such a coordination failure may reinforce whatever increase in repo haircuts that already prevail in the markets.

During the run on Bear Stearns in March 2008, the defection of its hedge fund clients was one of the contributory factors in the funding shortage that eventually led to Bear Stearns approaching Federal Reserve support. The Wall Street Journal’s special feature on Bear

---

5 The rules stipulate that when the sponsor is the main beneficiary of the special purpose entity and exercises substantial control of the vehicle, then the special purpose entity should be consolidated as being part of the sponsoring bank.
Stearns\(^6\) cites several hedge funds and other customers pulling out their funds from Bear Stearns at the height of the crisis in March.

The following chart plots the cash holdings of Bear Stearns in the days leading up to its demise. The data are taken from the SEC chairman’s letter, cited earlier.

---

In the three days from March 10th to 13th, the cash holdings of Bear Stearns dropped sharply from $18.1 billion on the 10th to only $2 billion on the 13th. This chart is a dramatic illustration of the run that led to the failure of Bear Stearns. The speed with which the cash is exhausted shows the role played by the instability of liabilities in leading to the failure of an institution. Thus, contrary to the traditional focus on credit risk on the asset side of the balance sheet, it is the run on the liabilities side of the balance sheet that is illustrated most amply by this episode. To the extent that broker dealer balance sheets consist primarily of such liquid and short-term claims and obligations, the run on Bear Stearns holds many useful lessons.

4. Implications for Financial Regulation: Liquidity Regulation

We now turn to the policy implications of our analysis so far. Our discussion is organized around the examples and propositions presented in section 2. Take the case discussed in Figure 2.

Consider Bank 0 as a bank such as Northern Rock, the UK bank that failed in 2007, or Bear Stearns that finances long-lived, illiquid assets by relying on short-term wholesale funding in the capital market. The exact identity of the lenders will differ from case to case, but the essence of the problem is that bank 0 has a maturity mismatch on its balance sheet, and relies on the capital market to provide short-term wholesale funding.

In this context, liquidity requirements on all banks (both debtor and creditors) may reduce the potential for runs through two channels. First, it makes debtor banks more robust to withdrawals. Second, it makes creditor banks less trigger-happy. Figure 6 illustrates.

Consider the Summers game described in section 2. Point A on the parameter space is associated with a run outcome. Point A depicts a fragile arrangement where the coordination threshold \( k \) is high, and the cost of miscoordination \( c \) is also high.
A liquidity requirement on the debtor bank (Bank 0) lowers the critical threshold $k$, by making the debtor more robust to withdrawals, since moderately sized withdrawals can be met by the liquid asset holdings of the debtor bank. In addition, the liquidity requirement on the creditor banks (Banks 1 to $N$) lowers the cost of miscoordination $c$ by making the creditor banks themselves less vulnerable to a deterioration of funding conditions.

The liquidity requirements may not have to be very onerous in order to be effective. Just as the fragility of the arrangement sets in motion a vicious circle of reasoning that leads to the run, increased robustness sets in motion a virtuous circle of reasoning that leads to a stable outcome. A more robust debtor bank instills confidence in the creditor banks, who in any case are more relaxed about the actions of other creditor banks in the face of potential deterioration of funding conditions.

The particular context will determine how onerous the liquidity requirements must be in order to achieve the stable outcome. A more systematic investigation – both theoretical and using numerical simulations – would be worth pursuing. However, the important principle is that liquidity requirements work by harnessing precisely those externalities that cause a run in the first place. The simple setting of the Summers game makes the comparison rather simple, but the same type of analysis can be brought to bear in a market context, as shown by Morris and Shin (2004). Here, the model is set in a more complex environment with more parameters to take account of – such as the elasticity of the residual demand curve that absorbs concerted selling. However, the underlying principles are identical as in the Summers game, and a unique outcome can be associated with each parameter configuration.

In addition, the underlying principle of distinguishing the credit risk of assets from their systemic impact seems important in any exercise of this sort, since the holding of cash buffers will affect the actions of interrelated players in subtle ways. As can be seen in
Figure 6, the (unique) equilibrium outcome can shift abruptly to small shifts in the underlying parameters of the problem that vary the susceptibility of the system to runs.

Recognizing the mutually reinforcing nature of actions holds out some hope that the liquidity requirements that would preclude financial system runs may not be too onerous. Even a modest liquidity requirement, if it is widely adopted in the system, may be sufficient to ensure stability. More systematic investigations will reveal precisely how onerous the liquidity requirements need to be in order to achieve robustness of the financial system to large shocks. In such an exercise, there will be inevitable trade-offs between the size of the shocks contemplated and the liquidity requirements needed to meet those shocks. Numerical simulations will reveal the terms of such a trade-off.

The cash holding of US broker dealers has been relatively stable over the last 25 years of so, fluctuating between 2 and 4 percent in recent years. The chart above shows the ratio of cash assets to total assets of the US broker dealer sector, as given by the Flow of Funds accounts. The sharp peak in 1987 and 1988 is associated with stock market crash of 1987. There are also increases in cash holdings in 2000 to 2002 (the bursting of the dotcom bubble) and in the last couple of quarters associated with the current credit crisis.

Interestingly, the relatively stable path for cash holdings for broker dealers is not present for US commercial banks. The chart below is from the H8 database of US commercial banks, as posted on the Federal Reserve’s website. The chart plots the ratio of cash assets\(^7\) to total financial assets of US commercial banks. The data are monthly.

\(^7\) Cash assets include vault cash, cash items in process of collection and balances at the Federal Reserve.
It is striking how the cash holding of commercial banks has declined steadily in recent decades. Congdon (2007) has noted that the decline in cash holdings for UK banks has, if anything, been even more dramatic. In the 1960s, it was typical for banks in the UK to hold as much as 30% of its assets in liquid form. However, Congdon reports that the ratio of liquid assets to total assets has fallen to 0.5% in recent years. The Bank of England’s Financial Stability Report of April 2008 charts the liquidity ratios according to several definitions. The report confirms the general trend that the liquid asset holdings of UK banks have undergone very large falls in recent decades (Bank of England (2008)).

Although liquidity requirements may mitigate the potential for fragility to runs, the institutional constraints imposed on particular types of market players should also be taken into account. The triparty repo agreement involving money market mutual funds injects elements of greater fragility by involving players who are constrained to cut back on lending when financial conditions deteriorate.

We have already mentioned that money market mutual funds can be seen as creditors whose cost of miscoordination, the parameter $c$, can be seen as being extremely large. Since the cost of miscoordination for these entities arises from the nature of the business and the charter that constrains their actions, liquidity requirements are unlikely to have much effect. Therefore, there would be a strong case for regulating their role in the triparty repo market.

Liquidity requirements would be complementary to the other reforms of capital regulations that mitigate the cyclical nature of risk-taking by financial intermediaries and shortage of capital during a downturn. Although we have focused our discussion on runs on the liability side, it would be important to place such liquidity crises into the overall context of the credit cycle.
Kashyap, Rajan and Stein (2008) have recently proposed a regulatory scheme that incorporates an element of funded capital insurance for banks, where banks pay into a fund that holds safe securities in a “lock box” that is opened in the event of tripping of aggregate triggers for financial distress. The time taken for verifying the aggregate triggers would make the capital insurance scheme better suited for addressing the shortage of capital in the down-phase of the financial cycle, and could be seen as having a longer-term focus than the very short-term acute liquidity crises envisaged in a liquidity crisis. However, liquidity requirements and capital should be considered together in any reform of the regulatory framework.

5. Implications for Financial Regulation: Leverage Constraints

We now turn our attention to another possible tool – a constraint on overall leverage of banks and brokers. We organize our discussion around the case illustrated in figure 4, reproduced below.

The scenario we consider is a generalized increase in haircuts in the capital markets. Specifically, Bank 1 experiences funding problems that results from the increase in the haircut in its repo transactions with Bank 2. The increase in repo haircuts does damage because repo haircuts were previously very low, and had encouraged all the banks in the system to lever up.

Any policy discussion should be based on a clear set of objectives that the policies are intended to achieve. We will take as a working assumption that the purpose of financial regulation is to reduce the amplitude of financial booms and busts, particularly the externalities that are generated in the boom and bust dynamics.

One policy measure that has attracted considerable attention already is that of imposing a maximum leverage constraint on banks, where the assets that figure in the leverage
calculation are the *unweighted* assets. The US has been at the forefront of such an initiative.  

Although such constraints have been criticized as being too blunt, the system perspective provides a rationale. The leverage constraint has the potential to prevent the build-up in leverage that leaves the system vulnerable to a sudden reversal. The idea is that the maximum leverage constraint is a binding constraint “on the way up” when funding conditions are ample, and banks can lever up easily. The build up of excessive leverage makes the system vulnerable to an increase in haircuts.

Note that the increase in haircuts will do most harm when starting from very low levels. A percentage point increase from 1% to 2% will mean leverage has to fall from 100 to 50. But a percentage point increase from 20% to 21% will have only a marginal effect on the initial leverage of 5. In this sense, the “chasing of yield” at the peak of the financial cycle is especially precarious, since the unwinding of leverage will be that much more potent.

By preventing the build-up of leverage during good times, the leverage constraint could act as a dampener in the financial system. As with any constraint that binds on the banks, the banks themselves will complain of being prevented from pursuing greater individual profits. However, this is as it should be, since any Pigovian tax is just that – a tax.

The leverage constraint works both at the level of the debtor, as well as the creditor. In the flow chart above, from the point of view of Bank 1 (the debtor), the leverage constraint will prevent Bank 1 building up excessive leverage, and thereby making Bank 1 less susceptible to a tick-up in the repo haircut. From the point of view of Bank 2 (the creditor), the leverage constraint binds “on the way up”, so that when eventually the tide turns, there is slack in the balance sheet capacity of Bank 2. Hence, its lending to Bank 1 will suffer a smaller shock to any rise in repo haircuts. Thus, for both the lender and the borrower the leverage constraint binds during boom times so that the imperative to reduce leverage is less strong in the bust. Indeed, the bust may be averted altogether, as the initial boom is dampened.

The fluctuations in leverage implied by the haircut in secured lending transactions suggest that banks and brokers expand their balance sheets to the maximum extent allowed by prevailing market conditions, only to cut back their balance sheets when funding conditions deteriorate. The imperative to maximize return on equity could be one reason for such behavior. The externalities are manifested “on the way down”, but the potential for externalities were created “on the way up”. The rationale for a leverage constraint is that it binds during the expansion stage of the cycle, inviting the banks either to raise new equity or to slow balance sheet growth.

It is also important to bear in mind that the sharp increase in repo haircuts during a crisis episode is endogenous. The severity of the crisis depends both on the extent of the

---

8 See the speech by FDIC chairman Sheila Bair (2006).
preceding boom and the actions of market participants.\textsuperscript{9} When leverage unwinds, the force of the unwinding will be stronger when the boom has extended for longer, and excesses have been allowed to continue for longer. One of the desired effects of the leverage constraint is to dampen the fluctuations in repo haircuts themselves.

In effect, a leverage constraint can be considered as a capital requirement that is not risk-sensitive. Safe assets attract a regulatory capital requirement that is equal to a risky asset. It is important to emphasize the difference in rationale between the leverage constraint considered here and the traditional risk-based capital requirement. As discussed already, the riskiness of reverse repos is small from a credit risk standpoint. Under a Basel style capital requirement, the required capital would be small. However, the leverage constraint will have the effect of mitigating the externalities generated by the fluctuations in funding conditions in a market based financial system built around secured lending. Instead of risks on the asset side of the balance sheet, the focus is on the liabilities side of balance sheets, and the potential spillover effects that result when financial institutions withdraw funding from each other. Thus, it is raw assets, rather than risk-weighted assets that matter.

The US authorities have continued to impose a leverage constraint on its regulated banks, at variance with the minimum capital requirements as laid down in Pillar I of the Basel II capital requirements. Recently, however, the authorities in Switzerland have announced their intention to introduce a US style leverage constraint. This announcement has generated a fierce controversy.

The most commonly encountered criticism against a raw leverage ratio constraint is that it does not take account of the riskiness of the assets. Basel II rules specify a very finely graduated capital requirement that depends on minute shifts in measured risks of the asset portfolio. A simple leverage ratio is seen as throwing away all the finely calibrated calculations of asset risk. The Financial Times has quoted the chief risk officer of Credit Suisse speaking in reaction to an announcement by Philipp Hildebrand, the vice governor of the Swiss National Bank, by saying “we manage banks according to Basel II, not Hildebrand I”.\textsuperscript{10} However, when viewed through the lens of systemic stability, the leverage ratio constraint has desirable properties that cannot be replicated by risk-based capital ratios alone.

A leverage ratio constraint seems particularly appropriate for Switzerland. The two large banks – UBS and Credit Suisse are both highly leveraged, even by the standards of the US investment banks. We saw earlier that the US investments banks were leveraged around 30 times. For UBS, its total assets at the end of 2007 were 2.27 trillion Swiss Francs. Its equity was only 42.5 billion Swiss Francs, implying a leverage ratio of 53. This is a very high number, even in comparison to the US investment banks. As noted already, commercial banks in the US have a leverage ratio of 10 to 12. Although most of the assets on UBS’s balance sheet are “safe” assets that attract a low capital requirement,

\textsuperscript{9} Brunnermeier (2008) and Brunnermeier and Pedersen (2005) describe the mechanisms at play when funding and market liquidity combine to amplify the financial distress.

\textsuperscript{10} Financial Times, Lex Column, July 1\textsuperscript{st}, 2008.
we have seen that the credit risk weights do not always reflect the strength of the externalities in a financial system.

There are, however, two conceptual issues that need to be tackled in the implementation of a leverage ratio constraint having to do with measurement of the two quantities involved in the definition of leverage. Leverage is the ratio of total assets to equity.

On the numerator, we face the issue of what to count as the total assets of the bank. In jurisdictions that apply the International Financial Reporting Standards (IFRS) of the IASB (International Accounting Standards Board), the assets held in securitization vehicles are counted as being part of the consolidated balance sheet. Inclusion of securitization vehicles raises the raw balance sheet size for European banks. For instance, Figure 7 taken from Shin (2008) shows the liabilities side of Northern Rock’s balance sheet in the ten years between its demutualization in 1997 to its failure in 2007. Much of the rapid increase in its balance sheet is accounted for by securitized assets (indicated by the red wedge).11 The rapid growth of Northern Rock’s assets therefore reflects the active securitization it had engaged in since demutualization.

Figure 7

Composition of Northern Rock's Liabilities
(June 1998 - June 2007)

Another accounting issue on the measurement of assets is how to deal with the fair value of derivatives contracts. Under IFRS, both the marked-to-market gains and marked-to-market losses are included on the consolidated balance sheet, making the balance sheet

---

11 See Yorulmazer (2008) for an empirical analysis of the UK banking sector at the time of the Northern Rock crisis.
appear much larger than it would otherwise be. The very high leverage of UBS is therefore partly an accounting phenomenon.\footnote{12 See “Banks According to GAAP” Financial Times, Lex column July 29th 2008.}

In addition, there is the issue of how to measure the denominator in the calculation of leverage – i.e. how to measure equity. Figure 8 plots the leverage series from June 1998 to December 2007 according to three different measures of equity, drawn from Shin (2008). In the figure, “shareholder equity” is common equity plus preferred shares. “Total equity” refers to shareholder equity plus subordinated debt.

Under the Basel approach to capital requirement, subordinated debt and preferred shares count as bank capital, since they are buffers against loss. However, repo haircuts provide another interpretation of equity – that of the stake that the controlling equity holder must have in order to borrow credibly from creditors who worry about moral hazard. Adrian and Shin (2008) and Krishnamurthy (2008) provide theoretical rationales for such an approach to leverage, drawing on the work of Holmstrom and Tirole (1997).

Under this alternative viewpoint, subordinated debt holders and preferred shareholders are just another class of creditor to the bank. They do not have control of the bank’s operations as the common equity holders do. For the purpose of calculating the permissible leverage in a moral hazard context, where the equity holders must have sufficient equity at stake so as to prevent them from engaging in moral hazard, it is the common equity that matters, rather than the equity enhanced by subordinated debt or preferred shares. This is because the common equity holders will take account of the

Figure 8

Northern Rock’s Leverage

June 1998 - December 2007

[Graph showing Northern Rock’s leverage from June 1998 to December 2007 with three different measures of equity: leverage on total equity, leverage on shareholder equity, and leverage on common equity.]
possible losses that result from their portfolio decisions, rather than the interests of the subordinated debt holders or preferred shareholders.

When leverage is interpreted strictly as the ratio of total assets to common equity, then we can see from figure 8 that Northern Rock’s leverage continued to climb throughout its history as a public company, rising from 22.8 in June 1998, just after its floatation, to 58.2 in June 2007, on the eve of its liquidity crisis. This could be considered a very large number, even by the standards of US investment banks who hold very liquid and short-term assets. Of course, Northern Rock’s leverage jumped even higher in December 2007 after its run, following the depletion of its common equity from losses suffered in the second half of 2007. The leverage on common equity at the end of 2007 was 86.3.

6. Implications for Financial Regulation: System-Weighted Capital Requirements

If we take seriously the idea that the exercise in computing leverage ratios as an exercise in computing Pigovian taxes to limit externalities, then a natural follow-up question is how we may assess the impact of the negative externalities imposed by one bank on the financial system as a whole. In the textbook example of the smoky factory that is located next to the laundry, the externality is the pollution emitted by the factory that soils the washing laid out to dry by the laundry. In this case, calculation of marginal costs will enter into the appropriate Pigovian tax on the factory.

For the financial system, the negative externalities are those that one institution imposes on another through fluctuations in funding conditions. In the example illustrated above, Bank 2 imposes a negative externality on Bank 1 if it decides to curb the scale of the collateralized lending. However, the impact of a similar decision by Bank 3 can be even greater, since a reduction in lending by Bank 3 will cause a reduction in lending by Bank 2, which in turn will induce a reduction in lending by Bank 1. If Bank 1’s assets are illiquid, the withdrawal of funding could cause even larger damage.

More generally, we can assign a “systemic impact factor” to each bank in the financial system that corresponds to the degree of spillover effect on other banks. The exact calculation of the systemic impact factor will depend on the network structure, and the nature of the assets held by each bank. However, the principle should be that any bank that lends a lot to banks that have high systemic impact factors should, itself, have a high systemic impact factor.

Such a principle could be implemented through the type of fixed point calculation that is used in journal citations ratings, or for calculating impact weights that Google uses for ranking websites. The impact weights for journals are such that a high impact journal receives many citations from other high-impact journals. Similarly, the rankings of websites have the feature that a high ranking website has links that point to it from other high-ranking websites. Indeed, our use of the term “impact factor” is intended to draw the analogy with calculations that are used for journal citations and Google webpage rankings.
In practice, however, complex calculations on impact factors will be practically difficult to implement. The lack of detailed balance sheet information that gives cross exposures will be an insurmountable hurdle. Even so, the principle of giving high systemic weight to institutions that have the potential to affect *others’* actions seems to be a sound one.

In any case, the impact factors associated with similar categories of financial institutions may naturally be clustered. Thus, broker dealers may have a high impact factor as compared to a small, locally based savings institution that deal primarily with retail customers and household borrowers. Thus, the fragmented way in which financial intermediaries in the US have been regulated may turn out to have a deeper, unintended rationale in terms of the economic principles. In any case, however the systemic impact factors are calculated, the principles of Pigovian taxation – that negative externalities should be taxed appropriately – can be a guide for our thinking.

An alternative approach in the practical implementation of system-weighted capital requirements could be tied to summary statistical measures of spillovers, if reliable and summary measures could be obtained. A promising line of research is given in Adrian and Brummermeier (2008) who consider the concept of “CoVar”, defined as the value at risk of an institution’s portfolio of assets conditional on some aggregate measure of distress. Although conceptual issues to do with aggregation and endogeneity of the portfolios in response to incentives to game the regulatory system still need to be worked out, the approach gives a promising line of research.

7. Concluding Remarks

The traditional approach to financial regulation that rests on risk-based capital requirements is not well suited to address the issue of stability of the financial system as a whole. The truism that taking care of the solvency of each individual institution ensures the stability of the system is not useful because it does not address the spillover effects.

Most importantly, the system perspective delivers a distinction between risky assets and systemically important assets. Even safe assets can be systemically important. The recognition of this distinction gives some rationale for two policy ideas that have attracted much attention – the idea of imposing a raw leverage constraint, and the idea of having a liquidity requirement that limits the composition of the asset portfolio, not merely its size. More systematic study will reveal how onerous the Pigovian taxes will have to be. The severity of the current financial crisis suggests that the optimal Pigovian taxes will not be zero.
References


http://www.princeton.edu/~hsshin/working.htm


http://www.bankofengland.co.uk/publications/fsr/2008/fsrfull0804.pdf

http://www.federalreserve.gov/newsevents/speech/bernanke20080822a.htm


http://www.bis.org/review/rr000921b.pdf


