2 Nature of Systemic Risk

Before considering the details of future regulation, it is desirable to have a good understanding about the causes of liquidity and solvency problems. If a financial institution is insolvent, it should be closed down. However, if the financial problems stem primarily from temporary liquidity problems, then intervention might be justified to save the financial institution. In this Chapter we outline how liquidity problems can lead to solvency problems and how relatively small shocks can cause liquidity suddenly to dry up , carrying the potential for a full-blown financial crisis. We first look at the problems from an individual financial institution's perspective, and then highlight the importance of looking at it from a systemic context. As we outline several amplification mechanisms, it will become apparent that the current philosophy of banking regulation – that you can make the system safe by making individual institutions safe – is an unsatisfactory basis for insuring the stability of the system as a whole.

2.1. Solvency, Liquidity and Maturity Mismatch

A financial institution is insolvent when its 'going concern' value does not exceed the expected value of its liabilities. In normal times, when financial markets are strong, it is fairly easy to identify insolvent financial firms. However, at times of crisis, it is difficult since solvency becomes so co-mingled with liquidity issues. Prices of assets become disconnected from estimates of expected cash flows and, instead, reflect the prices that could be obtained if the assets had to be sold tomorrow to the few investors prepared to buy such assets at such time (the liquidity price).[12]

The mechanisms that explain why liquidity can suddenly evaporate operate through the interaction of funding illiquidity due to maturity mismatches and market illiquidity.

As long as a financial institution's assets pay off whenever its debt is due, it cannot suffer from funding liquidity problems even if it is highly levered. However, financial institutions typically have an asset-liability maturity mismatch and hence are exposed to funding liquidity risk. A funding shortage arises when it is prohibitively expensive both to (i) borrow more funds (low funding liquidity) and (ii) sell off its assets (low market liquidity). In short, problems only arise if both funding liquidity dries up (high margins/haircuts, restrained lending) and market liquidity evaporates (fire sale discounts).

More specifically, *funding liquidity* describes the ease with which investors and arbitrageurs can obtain funding from financiers. Funding liquidity is high-and markets are said to be 'awash with liquidity'-when it is easy to raise money. Typically, when a leveraged trader, such as a bank, dealer, or hedge fund, purchases an asset, he uses the purchased asset as collateral and borrows (short-term) against it. However, he cannot borrow the entire price. The difference between the security's price and its value as collateral-the margin or haircut – must be financed by the trader's own equity capital. Margin lending is short-term since margins and haircuts can be adapted to market conditions on a daily basis.

Financial institutions that rely substantially on short-term (commercial) paper or repo contracts have to roll over their debt. An inability to roll over this debt-if, for example, the market for commercial paper dries up-is equivalent to margins/haircuts increasing to 100 percent, because the firm becomes unable to use its assets as a basis for raising funds. Similarly, withdrawals of demand deposits or capital redemptions from an investment fund have the same effect as an increase in margins. Funding liquidity risk is due to maturity mismatches and can thus take three forms: 1) margin/haircut funding risk, or the risk that margins and haircuts will change; 2) rollover risk, or the risk that it will be more costly or impossible to roll over short-term borrowing; and 3) redemption risk, or the risk that demand depositors of banks or even equity holders withdraw funds. All three incarnations of funding liquidity risk are only detrimental when assets must be sold only at fire-sale prices-that is, when market liquidity is low.

Market liquidity is low when it is difficult to raise money by selling the asset at reasonable prices. In other words, market liquidity is low when selling the asset depresses the sale price. When market liquidity is low, it is very costly to shrink a firm's balance sheet.

These two liquidity concepts do not exist in a vacuum; they are influenced by the financial soundness of other financial institutions.

Traditionally, capital requirements have been the cornerstone of financial regulation – especially so for banks. The current thinking behind the use of capital requirements is that maintaining a capital buffer allows an institution to absorb losses on its assets and remain solvent, thereby protecting its creditors – notably retail depositors. Moreover, that thinking relies on the reasoning that the solvency of each individual institution ensures the soundness of the financial system as a whole. This thinking leads naturally to the conclusion that the key determinant of the size of the regulatory capital buffer should be some measure of risks associated with the assets of that institution. 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 each bank. Protagonists of Basel II argue that its essential difference with Basel I is that it is far more 'risk-sensitive'.

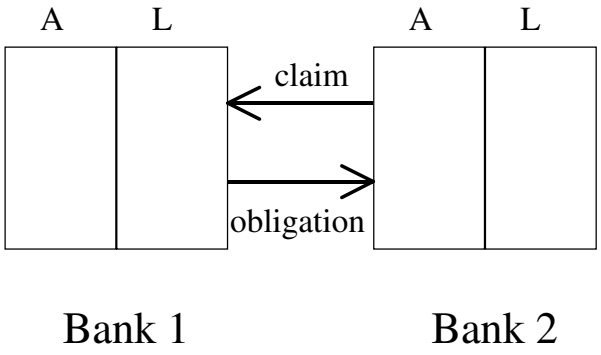
While this seems reasonable from an individual bank's perspective, it is clear that the level of market and funding liquidity is not exogenously given but determined in the economy as a whole and hence, important adverse feedback effects might arise. This requires a more systemic view of liquidity crises.

2.2 Funding Liquidity and the Domino Model

It is a truism that ensuring the soundness of each individual institution ensures the soundness of the system as a whole. However, for this proposition to be a good prescriptive guide for actions, we need to have confidence that actions that enhance the soundness of a particular institution will invariably promote overall stability. However, the proposition is vulnerable to the fallacy of composition.[13] It is possible, indeed often likely, that attempts by individual institutions to remain solvent can push the system into collapse.

Take a simple example, illustrated by figure 2. 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 2



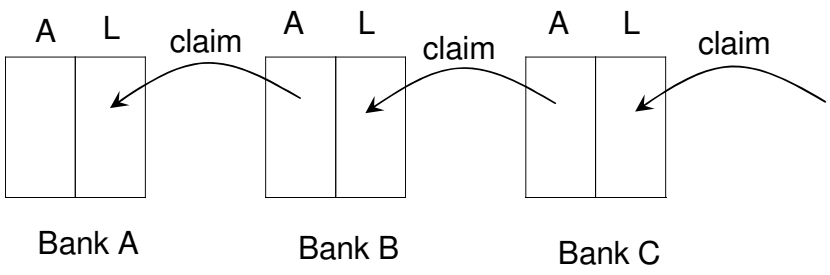
One way to ensure the solvency of Bank 2 is for it to reduce its overall lending, including its lending to Bank 1. By reducing its lending, Bank 2 reduces its risk exposure. However, from Bank 1's perspective, the reduction of lending by Bank 2 is a withdrawal of funding. Unless Bank 1 can find alternative 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, as well as the US securities houses Bear Stearns and Lehman Brothers, both of which suffered crippling runs in 2008.

The importance of the liabilities side perspective puts into question the traditional view of how systemic risk propagates throughout the financial system. A naive version of such a view could be depicted in Figure 3.

Here, bank A has borrowed from bank B, and bank B has borrowed from bank

Figure 3



C, etc. Then, if A takes a hit and defaults, then bank B will suffer a loss. If the loss is large enough to wipe out B's capital, then B defaults. Bank C then takes a hit. In turn, if the loss is big enough, bank C defaults, etc. We could dub this the 'domino' model of financial contagion.

The domino model of contagion has been examined in numerous simulation studies conducted at central banks, but the universal conclusion has been that the impact of the domino model of contagion is very small. It is only with implausibly large shocks that the simulations generate any meaningful contagion. The reason is that the domino model paints a picture of passive financial institutions who stand by and do nothing as the sequence of defaults unfolds. In practice, however, they will take actions in reaction to unfolding events, and in anticipation of impending defaults.

2.3 Loss Spiral – Asset Price Effect

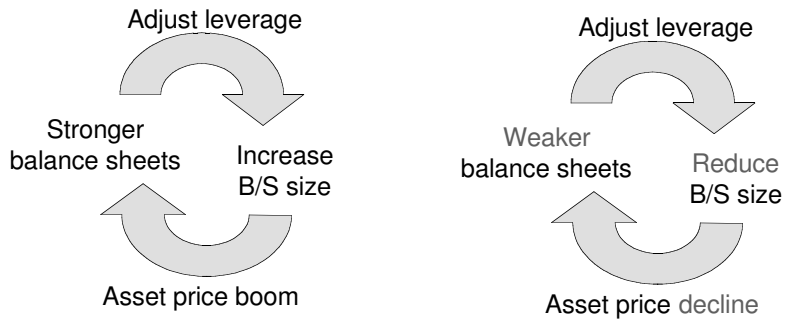
Thus, the domino model does not take sufficient account of how prices and measured risks change, and how such changes impact on the behaviour of market participants. In the simplest scenario of the domino model, asset prices are fixed at their book values, and balance sheets take a hit only with default. Such a view is obsolete in the market-based financial system where balance sheets are marked to market and where financial institutions react to changes in measured risks.

Indeed, defaults need not even be *necessary* to generate contagion. Price changes themselves may be enough. When financial institutions mark their balance sheets to market, changes in prices lead to losses that may be sufficient to transmit the shocks to other institutions even when they do not hold claims against each other. Losses worsen funding liquidity for many financial institutions, forcing them to shed even more assets which further depresses prices and increases losses, and so on. The loss spiral leads to sharp asset price movements especially at times of financial crisis.

If greater demand for the asset puts upward pressure on its price, then there is the potential for a feedback effect in which stronger balance sheets feed greater demand for the asset, which in turn raises the asset's price and lead to stronger balance sheets. Having come full circle, the feedback process goes through another turn. The circular figure on the left in Figure 4 illustrates the feedback during a boom. Note the critical role played by procyclical leverage.

The mechanism works in reverse in downturns. Consider a fall in the price of an asset held widely by hedge funds and banks. Then, the net worth of such an

Figure 4



institution falls faster than the rate at which the asset falls in value, eroding its equity cushion. One way that the bank can restore its equity cushion is to sell some of its assets, and use the proceeds to pay down its debt. The circular chart on the right in Figure 4 illustrates the feedback during a bust. Note the importance of marking to market. By synchronizing the actions of market participants, the feedback effects are amplified.

Take the episode of the distress suffered by European life insurance companies in the summer of 2002. By the nature of insurers' balance sheets, they did not borrow from each other as banks do. However, when stock prices plumbed new lows in the summer of 2002, the European life insurers found that their regulatory constraints were beginning to bind. In the U.K., for instance, the usual 'resilience test' applied to life insurance companies in which the firm has to demonstrate solvency in the face of a further 25% stock market decline was beginning to bind. German and Swiss insurers were even more constrained. The remedy for these insurers was to sell stocks, so as to reduce their exposures to them. However, large scale sales merely served to depress prices further, making the constraints bind harder. This generated a further round of selling, and so on. The regulators in the affected countries suspended the solvency tests for several weeks until the crisis abated. For instance, the U.K. Financial Services Authority diluted the resilience test so as to preempt the destabilizing forced sales of stocks by the major market players.[14]

The domino model of contagion is flawed, and is not useful for understanding financial contagion in a modern, market-based financial system. Instead, the key to understanding the events of the global liquidity and credit crunch in 2007-08 is to follow the reactions of the financial institutions themselves to price changes, and to shifts in the measured risks.

The transition to a market-based financial system is most advanced in the United States, but its influence has been very profound for the global financial system as a whole. Even for traditional deposit-taking banks, their marginal source of funding has been the capital markets, for example through repurchase agreements or commercial paper. This is because the traditional source of funding such as retail deposits are usually insufficiently flexible to fund expansions of lending. Moreover, the spreading of funding to include capital markets was often seen by banks, regulators and shareholders as increasing the liquidity and hence the solvency of a financial institution To this extent, the traditional distinction between banking and capital markets has become very difficult to draw. Indeed, the leit-

motif for the crisis of 2007-8 has been the amplification of the banking crisis through capital market conditions, spurred on by the pervasive use of mark-to-market accounting and market-sensitive risk management systems.

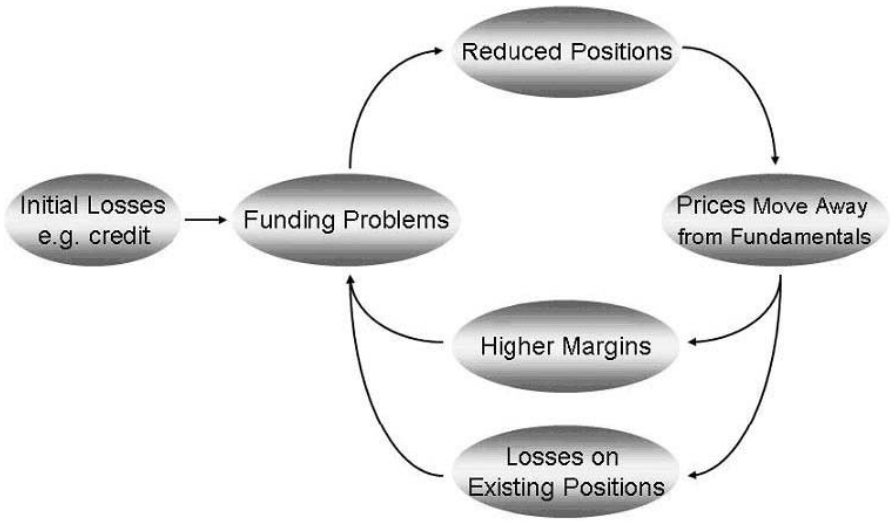
When financial institutions are integrated into the capital markets, market conditions dictate overall funding conditions. The balance sheet dynamics of financial intermediaries that mark their balance sheets to market and use market sensitive risk measures have some distinctive features.

2.4 Margin/Haircut Spiral

The loss spiral is not purely due to asset price effects, since a leveraged institution that suffers mark-to-market losses of $x has to reduce its position by $x times its leverage ratio.

The *margin/haircut spiral* reinforces the loss spiral since it forces the financial institution to reduce its leverage ratio on top of it. Margins and haircuts implicitly determine the maximum leverage a financial institution can adopt. Margins/haircuts spike in times of large price drops and thereby lead to a general tightening of lending. Brunnermeier and Pedersen (2009) – see Figure 5 -- show that a vicious cycle emerges, where higher margins and haircuts force de-leveraging and more sales, which increase margins further and force more sales, leading to the possibility of multiple equilibria.[15] As asset prices drop, risk measures (like Value-at-Risk) increase, which not only lead to higher margins and external funding costs, but also reduce risk-appetite within banks. Risk managers step on the brakes and force traders within a bank to de-lever their positions. Leverage is procyclical. When many market participants de-lever in stressed environments, liquidity disappears down a black hole.[16]

Figure 5

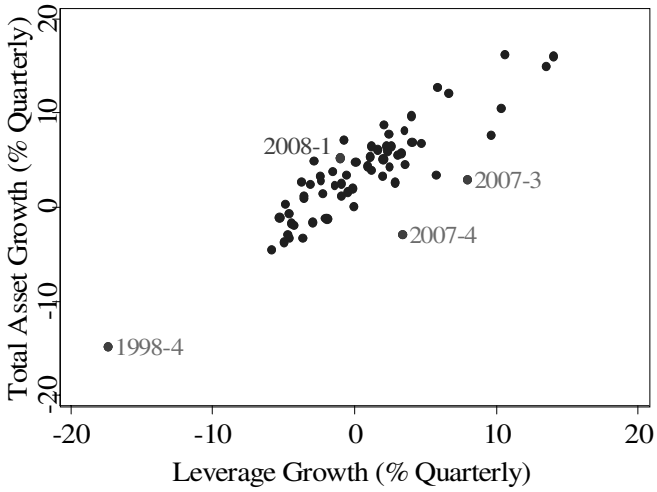


During downturns both spirals force leveraged investors to unwind their positions causing a) more losses and b) higher margins/haircuts and tighter lending standards, which in turn exacerbate the funding problems, and so on. Both spirals lead to procyclicality.

Figure 6 below shows empirical evidence for the margin spiral for the then US investment banks.[17] It shows the scatter chart of the weighted average of the quarterly change in assets against the quarterly change in leverage of the (then) five stand-alone US investment banks – Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch and Morgan Stanley.

Leverage is high when balance sheets are large, while leverage is low when balance sheets are small. This is exactly the opposite of the traditional finding for households, whose leverage is high when balance sheets are *small*. For instance, if a household owns a house that is financed by a mortgage, leverage falls when the house price increases, since the equity of the household is increasing at a much faster rate than assets[18]. For investment banks, however, the relationship is

Figure 6 Leverage Growth and Asset Growth of US Investment Banks



Source: SEC; Adrian and Shin (2007)

reversed. It is as if the householder responded to an increase in house prices by increasing the mortgage loan to value ratio so that leverage increases in spite of the increased value of his house.

A procyclical leverage ratio offers a window on the notion of financial system liquidity. When leverage is procyclical, the demand and supply response to asset price changes can amplify shocks. To see this, consider an increase in the price of assets held widely by leveraged market players and intermediaries. The increase in the price of assets strengthens the players' balance sheets, since the net worth of levered players increases as a proportion of their total assets.

When balance sheets become stronger, leverage falls. To the extent that the intermediary wants to avoid holding too much equity (for instance, because return on equity becomes too low), it will attempt to restore leverage. One way it can do so is by borrowing more, and using the proceeds to buy more of the assets

it already holds.

There is a more subtle feature of Figure 6 which tells us much about the financing decisions of financial intermediaries. Recall that the horizontal axis measures the (quarterly) change in leverage, as measured by the change in log assets minus the change in log equity. The vertical axis measures the change in log assets. Hence, the 45-degree line indicates the set of points where equity is unchanged. Above the 45-degree line equity is increasing, while below the 45-degree line, equity is decreasing. Any straight line with slope equal to 1 indicates constant growth of equity, with the intercept giving the growth rate of equity.

A feature to note from Figure 6 is that the slope of the scatter chart is close to 1, implying that equity is increasing at a constant rate on average. Thus, equity seems to play the role of the forcing variable, and all the adjustment in leverage takes place through expansions and contractions of the balance sheet rather than through the raising or paying out of equity. Said differently, it shows how the margin spiral and loss spiral reinforce each other.

A closer look at repo haircuts, which determine the implicit maximum leverage that is permitted in collateralized borrowing transactions such as repurchase agreements (repos), is instructive since repos are the primary source of funding for market-based banking institutions. 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 funding conditions in the market.

The fluctuations in the haircut largely determine the degree of funding available to a leveraged institution. The reason is that the haircut determines the maximum permissible leverage achieved by the borrower. If the haircut is 2%, the borrower can borrow 98 dollars for 100 dollars worth of securities pledged. Then, to hold 100 dollars worth of securities, the borrower 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 the borrower 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. The borrower 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 the borrower faces a predicament. Suppose that the haircut rises to 4%. Then, the permitted leverage halves to 25, from 50. The borrower 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.

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.

Times of financial stress are associated with sharply higher haircuts, necessitating substantial reductions in leverage through asset disposals or raising of new

Figure 7 Haircuts on Repo Agreements (percent)

Securities	April-07	August-08
U.S. treasuries	0.25	3
Investment-grade bonds	0–3	8–12
High-yield bonds	10–15	25–40
Equities	15	20
Senior leveraged loans	10–12	15–20
Mezzanine leveraged loans	18–25	35+
Prime MBS	2–4	10–20
ABS	3–5	50–60

Source: IMF Global Financial Stability Report, October 2008

equity. The table below is taken from the October 2008 issue of the Global Financial Stability Report of the International Monetary Fund (IMF (2008)), and shows the haircuts in secured lending transactions at two dates – in April 2007 before the financial crisis and in August 2008 in the midst of the crisis. Haircuts are substantially higher during the crises than before.

Raising new equity or cutting assets entail adjustments for the borrower. Raising new equity is notoriously difficult in distressed market conditions. But selling assets in a depressed market is not much better. The evidence from the scatter chart in Figure 5 above is that borrowers tend to adjust leverage primarily through adjustments in the size of the balance sheet, leaving equity unchanged, rather than through changes in equity directly.

2.5 Procyclicality and Margin Spirals

These liquidity spirals are the underlying cause of procyclicality. As asset prices drop, losses mount and margins/haircuts increase.

So far we have not explained why a drop in asset prices leads to higher margins, haircuts and a more cautious attitude towards lending. Should not a lower price reduce the probability of a further decline in the near future? Is not a price reduction that results from a lack of liquidity likely to be temporary, so that investors with the necessary expertise face a great buying opportunity? Hence, one might think that lenders would be willing to lend more freely by lowering margins after prices have dropped. There are at least three reasons why one observes exactly the opposite in the data:

a) Backward-looking risk measures

b) Time-varying volatility

c) Adverse selection.

Margins, haircuts and a bank's internal risk tolerance are typically obtained from risk-measures like Value-at-Risk (VaR). While the definitions of these measures have their own shortcomings, the bigger problem is how they are estimated. Typically these risk measures are estimated naively using past data. Hence, a sharp temporary price drop leads to a sharp increase in the estimates of these risk measures. This hikes margins/haircuts, constrains investors, and may force them to sell

off their assets. Paradoxically, the forced fire-sale might, justify the sharp increase in the risk-measure ex-post. For example, as in a boom phase volatility and default estimates are low, margins will be low which allows higher leverage and supports the expansionary phase. When the first adverse shocks hit, the volatility estimates shoot up leading to a deleveraging process described by the margin spiral. In short, if the objective function of individual institutions is to maintain return on equity, or value at risk, leverage will be procyclical. Ideally, one should take such endogenous effects due to risk mismeasurement into account.

Second, the volatility of a price process could be time-varying. A sharp price decline may signal that we are about to enter more volatile times. Consequently, margins and haircuts should be larger and lending should be reduced after such a price decline. An extreme example was the situation in August 2007, when the asset-backed commercial paper market dried up completely. Prior to the crisis, asset-backed commercial paper was almost risk-free because of overcollateralization – i.e. first losses would be assumed by lower tranches. However, in August 2007, the overcollateralization cushion evaporated, making such assets much more risky. Consequently, investors were unwilling to let structured investment vehicles roll-over their debt.

The third reason why margins increase when prices drop is that asymmetric-information frictions emerge. As losses mount, debt becomes more risky and hence more 'information sensitive.' Also, financiers become more careful about whether to accept a pool of assets as collateral since they fear receiving a particularly bad selection of assets. They might, for example, be worried that structured investment vehicles sold the good, 'sellable' assets and left as collateral only the bad, less valuable, 'lemons.'

2.6 Externalities – Rationale for Regulation

The presence of liquidity spirals per se does not justify government interventions. One must argue from a social welfare perspective that financial institutions overexpose themselves to the risk of getting caught in a liquidity spiral by holding highly levered positions with excessive maturity-mismatches. We argue that this is indeed the case due to the following two risk-spillover externalities that we alluded to in Chapter 1:

a) Fire-sale externalities
b) Interconnectedness externalities

The fire-sale externality arises since each individual financial institution does not take into account the price impact its own fire-sales will have on asset prices in a possible future liquidity crunch. Hence, fire-sales by some institutions spillover, and adversely affect the balance sheet of others, causing a negative externality. This externality is pointed out in Stiglitz (1982) and Geanakoplos and Polemarchakis (1986) and subsequently appeared in numerous academic papers.[19] It is arguably the main rationale for bank regulation.[20]

In general, a financial institution is also not concerned how many others it will drag down, should it fail. Especially the failure of big and interconnected institu-

tions would bring down these negative risk-spillover effects on others. An opaque market structure, as for example in over-the-counter markets (OTC markets), exacerbates these effects.[21]

What makes matters even worse is that the potential prospect of a government bailout gives institutions the incentive to become 'too big to fail' and 'too interconnected to fail.' The larger an institution, or the more interconnected it is, the higher the probability that a financial institution will be bailed out in times of crisis. In short, the current system implicitly subsidizes institutions that cause negative externalities on others. Hence, we will argue in the subsequent Chapter that the regulatory framework has to focus on risk spillovers, i.e., externalities.

In general it might be desirable for the monetary authority to step in after a 'once in a blue moon' liquidity shock[22], since it is socially not optimal for each bank to be required to provision against those shocks. However, since financial institutions expect this, they will alter their behavior – which provides another rationale for financial regulations.

2.7 Aggregate Liquidity Expansions and Contractions

We conclude this section by recalling that institutions that hold assets with high market liquidity (or short-term assets) can adjust their balance sheet size flexibly by reducing lending and not rolling over debt. However, when the financial system as a whole holds long-term, illiquid assets financed by short-term liabilities, any tensions resulting from a sharp, synchronized contraction of balance sheets will show up somewhere in the system. Even if some institutions can adjust down their balance sheets flexibly, there will be pinch points in the system that will be exposed by such de-leveraging.

Fluctuations in leverage in the context of widespread secured lending exposes the myth of '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 could be redistributed to those who need it most. 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. When a bank such as Northern Rock finds itself at the receiving end of a run by its creditors, it cannot simply turn to another creditor to take up the slack, for all other creditors are simultaneously curtailing their lending. In this sense, liquidity should be understood in terms of the growth of balance sheets (i.e. as a flow), rather than as a stock.