



# The Fundamental Principles of Financial Regulation

Markus Brunnermeier, Princeton University
Andrew Crocket, JPMorgan Chase
Charles Goodhart, London School of Economics
Martin Hellwig, Max Planck Institute
Avinash D. Persaud, Chairman. Intelligence Capital Limited
Hyun Shin, Princeton University

Preliminary Conference Draft
Please do not quote without permission
January 6, 2009
Geneva Report on the World Economy 11

# Contents

# Foreword

The authors of this Geneva Report are predominantly macro-economists. In our view macro-economic analysis and insight has, in the past, been insufficiently applied to the design of financial regulation. The purpose of this paper is to help rectify that lacuna.

The crisis which began in the US sub-prime mortgage market in early 2007 and then spread broadly and deeply was not the first banking crisis. It was closer to the 100th. We can draw a few important implications from this observation. If an event with widespread and severe economic and social consequences keeps on repeating itself, the onus is surely on the authorities to change something. Chiding bankers is satisfying; but insufficient. When a regulatory mechanism has failed to mitigate boom/bust cycles, simply reinforcing its basic structure is not likely to be a successful strategy. Moreover, a type of crisis that repeats itself cannot easily be put down to new, complex, instruments. In this report, we set our sights on moderating the recurring cycle of financial crises, cycles that in our view are not wedded to particular instruments, institutions, individuals or information.

The prevention of crises in the banking system is more important than in the case of other industries. As outlined in Chapter 1, the externalities from an individual bank failure both to other banks and thence to the wider economy are just so much greater. One of the key purposes of bank regulation is to internalize the social costs of potential bank failures via capital adequacy requirements. The regulation of banks must do more than instil best practice amongst bankers, or converge regulatory capital to the capital a prudential bank would otherwise hold. The current approach to systemic regulation implicitly assumes that we can make the system as a whole safe by simply trying to make sure that individual banks are safe. This sounds like a truism, but in practice it represents a fallacy of composition. In trying to make themselves safer, banks, and other highly leveraged financial intermediaries, can behave in a way that collectively undermines the system. Selling an asset when the price of risk increases, is a prudent response from the perspective of an individual bank. But if many banks act in this way, the asset price will collapse, forcing institutions to take yet further steps to rectify the situation. It is, in part, the responses of the banks themselves to such pressures that leads to generalised declines in asset prices, and enhanced correlations and volatility in asset markets. Such endogeneity of risk, described further in Chapter 2, is greater the more there is a common driver of behaviour.

Financial crashes do not occur randomly, but generally follow booms. Through a number of avenues, some regulatory, some not, though often in the name of risk-sensitivity, sophistication and modernity, the role of current market prices on behaviour has intensified. These avenues include mark-to-market valuation of assets; regulatory approved market-based measures of risk, such as credit default swap spreads in internal credit models or price volatility in market risk models; and the increasing use of credit ratings, which tend to be correlated, directionally at least, with market prices.

In the up-phase of the economic cycle, price-based measures of asset values rise, price-based measures of risk fall and competition to grow bank profits increases. Market discipline encourages financial institutions to respond to these three related developments by some combination of (i) expanding their balance sheets to take advantage of the fixed costs of banking franchises and regulation (ii) trying to lower the cost of funding by using short-term funding from the money markets and (iii) increasing leverage. Those that do not do so are seen as underutilizing their equity and are punished by the stock markets. When the boom ends, and asset prices fall and short-term funding to institutions with impaired and uncertain assets or high leverage dries up, leading to forced sales of assets which drives up their measured risk, the boom turns to bust.

In Chapter 3, we distinguish between micro and macro-prudential regulation. Micro prudential regulation concerns itself with factors that affect the stability of individual institutions. Macro-prudential regulation concerns itself with factors that affect the stability of the financial system as a whole. As we will attempt to show, the nature of the regulation applied to an individual financial institution depends crucially on how "systemic" its activities are. This in turn is related, inter alia, to its size, degree of leverage and interconnectedness with the rest of the system.

A critical component of macro-prudential regulation must be to act as a countervailing force to the natural decline in measured risks in a boom and the subsequent rise in measured risks in the subsequent collapse. This countervailing force has to be as much rule based as possible. Supervisors have plenty of discretion, but their ability to utilize it is limited by the general short-sighted desire to prolong a boom and by bankers pleading for equality of treatment. In a boom, lending, leverage and reliance on short-term liquidity become mutually reinforcing and excessive. To counter this we propose, in Chapter 4, counter-cyclical capital charges. Regulators should increase the existing capital adequacy requirements (based on an assessment of inherent risks) by two multiples. The first is related to above average growth of credit expansion and leverage. Regulators should agree on the degree of bank asset growth and leverage that is consistent with the long-run target for nominal GDP, so that the multiple on capital charges rises the more credit expansion exceeds this target. The purpose of this capital charge is not to eliminate the economic cycle – something which would be unrealistically ambitious - but to ensure that in a boom, when risk measures are suggesting banks can safely leverage or lend more, banks are putting aside an increasing amount of capital which can then be released when the boom ends and asset prices fall back.

The second multiple on capital charges should be related to the mis-match in the maturity of assets and liabilities, as discussed in Chapter 5. One of the significant lessons of the Crash of 2007/8 is that the risk of an asset is largely determined by the maturity of its funding. Our proposed adjustment to mark-to-market accounting should provide a further incentive to reduce maturity mismatch. Northern Rock and other casualties of the crash might well have survived with the same assets, if the average maturity of their funding had been longer. When regulators make little distinction how assets are funded, there is a tendency for financial institutions to rely on cheaper, short-term funding, which increases systemic fragility. If short-term funding of long-term assets carries a capital cost – because it weighs on systemic

stability – it will moderate banks' reliance on systemically adverse short-term funding and encourage them to seek longer-term funding.

A combination of these charges should push banks to develop incentive packages that are more encouraging of longer-term behaviour, as we outline in Chapter 6. A little more is required on this front, though we do not share the zeal of some for governments to be involved in the micro-decisions of private firms.

There is a tendency, commonly observed amongst politicians, to review the structure of the regulatory system before considering the potential instruments to achieve better regulatory control. Our position, Chapter 7, is the reverse. The structure of regulation should reflect the purposes and powers of the regulatory authorities. Macro-prudential, and micro-prudential, instruments are both needed, but differ in focus and in their needed professionalism. Hence, they should be carried out separately, respectively by Central Banks and by Financial Services Authorities. Again, financial and asset-price cycles differ from country to country. So contra-cyclical policy needs to be assumed more by the host country, thereby shifting some of the emphasis in regulation from the home to the host country.

Besides our key recommendations on macro-prudential measures and mark-to-funding, we make proposals on a whole series of minor issues, such as the role of stress tests, the adoption of maximum loan-to-value ratios in mortgage markets, etc., etc. These are gathered up and reported in our conclusion, Chapter 8. Really busy readers could skip straight there. Overall our intention is to develop a program of practical initiatives that could better attack the key features of externalities and systemic failure in financial markets.

# Chapter 1: Analytical Background

There is a vast body of financial regulation in existence. This is normally extended incrementally, frequently to close a loophole which some earlier fraud or financial disaster has exposed. Even such measures as may have seemed to involve a discrete jump in the regulatory process, such as the passage of the Basel I Accord in 1988, turn out, after closer inspection, to have been largely an attempt to agree on, and to harmonise, pre-existing 'best practices' in the key nation states, without much overt attempt to rationalise them against fundamental principles, or underlying theory. Exceptions occur only after major crises, as in the USA after 1929-33, (with Glass-Steagall and deposit insurance), and, perhaps, now.

There are good reasons for such an incremental approach under normal circumstances. Like the common law, it builds on the accrued wisdom of generations. It is practicable, do-able and (generally) common-sensical. Yet it is possible for such an incremental, and generally reactive, process to migrate over time in wrong, or just inferior, directions. When a major crisis erupts, such as that which has roiled financial systems in the world since August 2007, there is both a case and an opportunity for revisiting the underlying principles of (financial) regulation to examine whether the existing system is appropriately designed. There is a general willingness now to question existing regulatory practices and to consider, without prejudice, a wide range of alternative proposals. Nothing at this juncture is too hallowed by tradition and usage to escape questioning and to be off-limits to reform.[1] In particular, the regulatory system stands accused of having failed to mitigate the recent cycle in leverage, credit expansion and housing prices. Nothing was done to tighten regulations (e.g. on capital, liquidity or remuneration) in the upswing, nor, until recently, to relax the pro-cyclical implications of the accounting/regulatory framework in the downturn. Regulation, in effect, provided little or no check, nor barrier, to the decisions taken by banks, and other financial operators, in their pursuit of (short-term) profit maximisation. It was not adapted to changes in the underlying

[1] The temper of the times is illustrated by this passage from a paper by Davis, Polk and Wardwell, (the US law firm), to provide guidance on 'The Emergency Economic Stabilization Act of 2008', (October 4, 2008). They write, p. 43, that,

> "The Act is only the first step in the return to health for the US financial system. The idea of restructuring the archaic US financial regulatory system has been in the academic air for sometime and was recently also taken up by Treasury in its Blueprint, (The Department of the Treasury Blueprint for a Modernized Financial Regulatory Structure issued on March 31, 2008.) It should be clear to all by now that the fragmented nature of the current US regulatory system was a co-conspirator in the creation of the mess. None of this invalidates the critiques of those who have pointed out problems with the Sarbanes-Oxley Act and its impact on the competitiveness of the US capital markets. The problems are, in fact, larger than the false dichotomy between regulation and deregulation. They are worse than that. The problem is one of ineffective regulation leading to the wrong types of incentives within a fragmented regulatory structure that was unable to cope with new products and new circumstances in a changed world. What is needed is a complete reordering of the system, including both deregulation and re-regulation, depending upon which is more effective for the stability of the financial system, the competitiveness of the US capital markets and the economic health of the country. Naturally, there will be many interests to balance and the ability of our political leaders to make those changes cannot be assumed."

vulnerabilities in the system as a whole, and allowed financial engineering to avoid its impact, e.g. SIVs and other methods of deconsolidating risks.

What is needed is, first, a restatement of the basic objectives of financial regulation and, then, an assessment of whether the current regulatory framework is well structured to attain such objectives, and, if not, to explore what can be done to restructure such regulation so that it does.

So let us start by asking what should be the purposes of regulation. Traditional economic theory suggests that there are three main purposes.

1. to constrain the use of monopoly power and the prevention of serious distortions to competition and the maintenance of market integrity;
2. to protect the essential needs of ordinary people in cases where information is hard or costly to obtain, and mistakes could devastate welfare; and
3. where there are sufficient externalities that the social, and overall, costs of market failure exceed both the private costs of failure and the extra costs of regulation.

(1) above has been a main rationale for the regulation of private utilities, but, until recently,[2] has only entered the financial scene in a few rare cases, e.g. where the network economies of having a single market procedure, e.g. a clearing house, are so great that those who control access to the network could potentially extort huge rents from those trying to join.

The effect of the recent crisis, as it was also in Japan, has, however, been to reduce competition in the banking industry. In order to prevent weaker banks from failing, they have been folded into stronger, and generally larger, banks, thereby creating a small number of national 'champions' in each country.[3] Japanese City banks have been reduced to three. JPMorgan Chase, Bank of America, Wells Fargo/Wachovia and Citibank now bestride the US scene, (though all remain nominally subject to the 10% deposit cap). Concerns about reductions in competition are brushed aside, as with the Lloyds/HBOS merger, in the rush to shore up a fragile system. The result is an oligopolistic system, dominated by 'champions' who are far too large to fail, in some smaller countries (e.g. Iceland) perhaps too large to save, and are in a position to wield great influence and power.

---

[2] The 10% limit for deposit concentration in the USA could soon become relevant.

[3] The Daily Telegraph, October 1, 2008. Our thanks are due to Russell Taylor, writer of the Alex Cartoon, for permission to reprint this.

How serious are the dangers posed by such greater concentration? Some banking markets have become more contestable, notably via IT techniques, such as bidding for time deposits; others are less so, e.g. loans to SMEs. Our proposals may very slightly mitigate the trend towards concentration by proposing tougher regulation for large, systemic banks. Beyond that, however, we advocate reinforced scrutiny by the competition authorities of potential anti-competitive practice by the larger banks.

(2) above has come to mean that bank deposits have become implicitly, or explicitly, fully (100%) insured and guaranteed, at least up to some upper limit. This has now gone further in the current crisis. By the same token there are controls on the behaviour of insurance companies and pension funds. Mutual funds, unit trusts, money market funds, etc., are not guaranteed except in extreme cases, such as recently observed in the USA, but are required to behave in certain required ways. The debate in these cases is not whether they should be regulated, but how this might best be done. The point is that the political process works to protect the interest of small (and sometimes large) clients of financial institutions, regardless of the formal legal position. This is not going to change.

Such 100% deposit insurance, up to a now more elevated ceiling, creates moral hazard, both amongst depositors, and also amongst banks, so long as the premium paid by each bank is not accurately and immediately adjusted in alignment with such a bank's riskiness, (and this is difficult to apply, though the Canadian deposit insurance corporation, CDIC, has made a successful attempt along such lines). . However awful, and risky, such a bank may be, or may become, it can always raise extra funds, once 100% insured, by raising interest rates slightly, (until and unless the supervisors close it down). In order to prevent the worst excesses of moral hazard, and to protect the tax payer, there is a need for a <u>minimum</u> level of capital, which, if breached, acts as a trigger for prompt corrective action (PCA). Such a minimum level of capital does <u>not</u> provide any protection for the shareholders and bank officers, rather the reverse as it is an intentional hostage, giving them necessary 'skin in the game'. Nor does it provide any resilience to the banking system, (only the buffer above such a minimum provides that), except in so far as PCA allows for an orderly run-down, rather than distressed sales and a news-worthy bankruptcy, of a bank in severe difficulties.

As has been clearly seen in 2007/8, (3) is by far the most important reason why banks, and certain other key financial intermediaries and markets, need regulation. But why does the failure of banks, and of some other financial institutions, involve systemic externalities that are not present when an ordinary manufacturing or service-sector firm goes bust. The basic answer comes from the fact that the failure of a banking-type institution, say Lehman Bros, Northern Rock or Glitnir, weakens the other banks and financial markets with which they were involved, whereas the failure of, say, a car company or a laundry tends to strengthen the remaining companies in the same sector, by removing a competitor. And lying behind this is the even more important consideration that the continued health of the financial system, and even more so of the banking sector within it, is key to the satisfactory functioning of the wider economy, to a qualitatively different extent from most other sectors.

There are, at least, five reasons for such negative externalities. The first is pure informational contagion, particularly in the context of intermediaries with a maturity mismatch between liabilities and assets (see Chapter 5). If bank A fails, this throws more doubt on the continued solvency of bank B, when B is perceived as being of the same type as A. When such doubt arises, depositors and lenders to B lose confidence, withdraw their funds, causing a sudden liquidity problem for B; this moves relative interest rates, and access to funds, against B, making its future solvency even more threatened. Thus the failure of Lehman Bros rapidly led to the end of the US Securities House model, (with Merrill Lynch being forced into a merger with Bank of America, and Goldman Sachs and Morgan Stanley becoming banks). If Northern Rock had been allowed to fail, there might have been runs on Bradford and Bingley and on Alliance and Leicester the next day, and on HBOS on the following day.[4] The demise of Glitnir in Iceland was rapidly translated into the collapse of Landsbanki and Kaupthing. These last two examples, however, indicate that while the size of the bank in difficulties plays a role in the spread of resulting contagion, it is not the only factor. If the failing bank is (perceived as) similar to other banks, and the cause of its failure may apply to them also, then like Northern Rock and Glitnir it will be contagious. If, however, it is perceived as being a unique outlier (e.g. BCCI), or if the cause of loss is particular to that bank, and not applicable to its close competitors, (notably when arising from fraud, e.g. Barings and Nick Leeson, Soc Gen and Jerome Kerviel), then there is much less risk of direct contagion.

But will not the same argument apply to other non-bank companies? If car producer A fails, it is likely to be due to a generalised fall in demand for cars. Will not this lead lenders to car producer B, whether on commercial paper, trade credit and bonds to refuse to renew or roll over, even despite the greater demand for B's products? The failure of the car company, however, does not have such an important signalling effect for its competitors. If the demand for cars falls, this is evident to lenders well before the first company fails. The failure of one car company means that the remaining companies will do better, not worse.[5] Moreover, most corporates and many households have a contingent line of credit with their bank, to tide them over such difficulties, until they can restore their own position by cutting cost, or increasing profits on sales, after their competitor's failure. So non-banks rely on banks in a crisis, while banks in turn have to rely on the Central Bank.

This leads on to the reason for the second externality arising from bank failures, which is a loss of access to future funding for the failed bank's customers. Of course, a client of failed bank A can try to transfer her custom to surviving bank B, but bank B will have less direct information on this client, and is likely, especially in the likely conditions of fear and panic surrounding major instances of bank failure, only to provide replacement credit facilities on much tougher terms. A bank failure causes an externality in the guise of the loss of specific information links between the failed bank and its customers. While this is also the case in other industries which have long-term relationships with their customers, it is especially pronounced for banks.

---

[4] While the cynical will note that such runs were only delayed by a few quarters, the breathing space gained might have allowed the crisis to be resolved with much less loss.

[5] GM, however, has argued that if there should be common suppliers, then its failure could adversely rebound on Ford and Chrysler by causing the bankruptcy of such suppliers.

The third externality is that banks, and financial intermediaries, trade much more amongst themselves than do other corporates. Hotels and steel mills do not have significant inter-hotel or inter-furnace markets. Such interactions between banks and other financial intermediaries relate not only to the straight-forward interbank market, but also to an increasing range of other derivative markets, involving guarantees (mono-line insurers), credit default swaps, as well as prime brokerage services, etc., etc. It was because of the ramifications of such connections that both Bear Stearns and AIG were provided with public sector support, and, in part, because of such ramifications that the failure of Lehman Bros proved so devastating.[6]

In the longer run when the dust has settled, a failing bank (like Continental Illinois) can often pay back a large percentage of its inter-bank borrowing, and the need to rearrange derivative contracts, to which the failed bank was a party, can ultimately come out close to a zero-sum game. But in the immediate aftermath of the failure of an inter-connected bank, there is much uncertainty about how much creditors of that bank will get back, and by what date. This will lead analysts to try to make instant assessments of who potentially stands most at risk, and this will then feed directly back to our first externality, informational contagion. Thus the fact that Continental Illinois did ultimately pay back to its correspondent banks over 90 cents in the dollar is <u>not</u> proof that its (abrupt and disorderly) failure might not have triggered runs on at least some of its correspondent banks.

So far we have been concerned only with the failures of banks, and some other systemic institutions. But that failure is generally triggered by a decline in the value of the assets held by the bank, and by a run on the bank, itself usually primarily caused by a perceived decline in the bank's asset values. Liquidity problems usually generate underlying solvency worries, (though the illiquid bank will attempt to deny this, as in Northern Rock, the Icelandic banks, etc.). In order to deal with such liquidity problems prior to failure, and in the course of liquidation after failure, the bank in difficulties will often be forced to sell assets (fire sales). But such sales will drive down the current market price of the same assets held on other banks' books, when these are valued on a mark-to-market basis. And, of course, the same is true the other way around; solvency is not exogenous to liquidity. When there is a generalised liquidity problem attempts to deal with it will lead to declines in asset values, creating a solvency problem, even where none existed before. In short, there is an internal amplifying process (liquiditiy spirals) whereby a falling asset market leads banks, investment houses, etc., to make more sales (deleveraging), which further drives down asset prices and financial intermediaries' assessed profit and loss and balance sheet net worth.

We believe that it is this internal, self-amplifying dynamic that has lain at the root of both the recent, and virtually all prior, financial crises. The argumentation and analysis for this claim is set out at greater length in Chapter 2. Thus we believe that

[6] An additional related problem is the mutual impact of declines in the credit standing of counterparties. If bank A has an OTC claim on bank B, whose credit is downgraded, the net worth of bank A will decline, and vice versa. If both should occur simultaneously, then in pure mark-to-market accounting, there should be no change in net worth, as the value of each bank's liabilities has declined by the same amount as its assets, but it is artificial for a bank to take credit for the view that it is less likely to repay its liabilities.

financial crises are predominantly caused by market dynamics, not just by external shocks, though such shocks, e.g. the downturn in the US housing market in 2006, the quadrupling of oil prices in 1973/74, the Stock Market collapse in 1929, may well have been the trigger.

One immediate implication of this is that the standard format of banking stress tests is fundamentally insufficient. These stress tests review the effect on each bank's profits and capital of some (historically-based) exogenous shock. But, if financial crises are primarily caused by endogenous risk, whereby the banks' reactions to such a trigger sets off an amplifying spiral, via declines in asset prices and reductions in credit expansion, such stress tests, focussing on exogenous risk, will miss out on the (more important) second, and higher, round effects. Attempts to adjust stress tests for endogenous risk have not yet borne much fruit. It may be that the best way to assess the implications of endogenous risk is via new endogenous "Co-risk-measures" that measure the increase in overall risk after conditioning on the fact that one bank is in trouble (possibly for endogenous reasons). Another way to go is develop a model to explore the likely actions, reactions and interactions within the banking system, but that remains largely an exercise for the future, (though the papers of Goodhart, Sunirand and Tsomocos represent a start).

The fifth, and final, form of externality is akin to the fourth. Instead of, or as well as, selling financial assets to regain liquidity, and to improve capital ratios, a bank, or financial intermediary, may seek to restrict new credit extension, e.g. by rationing via higher margins/haircuts or by raising interest rates, or other costs, to borrowers. Such deleveraging, via credit restriction, will have the general effect of lowering output and prices, whether of goods, or services, or assets in the economy. This will raise the probability of default for all other borrowers. Thus there is yet a further self-amplifying spiral whereby credit restriction weakens the economy, which leads to more default and asset price declines, which causes yet more credit restrictions.

The implication of all this is that the appropriate regulatory concern, caused by externalities, lies with the impact of the difficulties of the individual financial intermediary, whether by failure or large-scale forced deleveraging, on the wider system. And, of course, market failures (in the guise of resource misallocations) also occur during the boom phase, with excessive credit expansion and investment in the 'bubble' assets. That is, in principle, separate from the risk management practices of the individual bank. The individual small depositor is protected by deposit insurance (where the insurance fund in turn requires a minimum capital ratio and prompt corrective action as safeguard).[7] For the rest, the riskiness of an individual bank, or any other financial intermediary, should properly remain the province of the bank's managers, owners and debt holders, subject to market discipline, except in so far as that institution's demise should impinge on the wider system, via spill-overs and externalities.

[7] The historical record suggests that bank failures can occur quite suddenly without an observable prior steady decline in capital ratios. So the combination of PCA and a minimum capital ratio may not be enough to protect the deposit insurance fund, and hence the taxpayer. This is one of the arguments, discussed later in Chapter 3, for extending micro-prudential supervision beyond the ranks of large and system institutions.

So, the first claim of this paper is that regulation has been excessively focussed on seeking to improve the behaviour and risk management practices of individual banks, too micro-prudential, for which we would assert that it has slight justification in the theory of regulation. By the same token it has been far too little focussed on wider systemic issues, insufficiently macro-prudential, where it does have a locus. By consciously seeking to make prudential capital move more closely in accord with banks' own choice of economic capital, regulation did too little to restrain bank expansion in the upswing, nor has it been able to provide any support against the current implosion of the system as a whole.

Let us take two key examples of the difference between the macro and the micro perspective. First let us consider liquidity. In one of his earlier papers, Hellwig considers a banking system consisting of n banks, where n is quite large. A demand deposit is placed with the first bank, which lends it on in the interbank market for one week to bank 2, which lends it on for two weeks to bank 3, and so on, until finally bank n gets an interbank deposit for n-1 weeks and lends it on to an end-user for n weeks. No bank has a serious maturity mis-match, but the system as a whole does. It could unravel quite quickly. What occurred in 2007/8 was just such an unravelling of wholesale financial markets. This experience shows that neither the system as a whole, nor individual banks, such as Northern Rock, can put their faith for maintaining liquidity in continued, unquestioned, access to wholesale markets on reasonable terms.

From an individual bank/micro perspective, it was reasonable and efficient for each bank to assume that, in normal times, they would have access to the wholesale money markets. Once some banks made this assumption, banks that did not do so were put at a competitive disadvantage. This was one of the forces behind the de-mutualisation of building societies and their evolution into banks, so that they could tap wholesale money markets. At a micro-level this was not viewed as increasing risk, but reducing it by providing alternative and more flexible sources of funding. But the exploitation of market access by almost all banks in normal times, increased the likelihood of disaster in abnormal times.

This then raises the key question of which parties should stand behind the system to provide access to liquidity in the case of a failure of markets to function adequately. There are four present candidates. The first candidate is the banks themselves. Market failure most often occurs (ignoring physical problems, such as computer failure) because of credit counterparty risk. Government debt, when denominated in that Government's own currency, has no such risk. Banks with ample quantities of government paper amongst their assets can withstand temporary liquidity problems. But this requires that banks carry large quantities of government paper – providing less room for private lending – or that the credit and liquidity problems are contained in scope and time.[8] Yet credit and liquidity problems have a way of running along far reaching fault lines. If we are to rely solely on the banks, and do not wish to have an overly repressed credit system, we would have to induce banks to behave in a

[8] Minimum liquidity, or cash ratios, are, of course, poor ideas since the assets satisfying that minimum cannot be used. Holding assets as a proportion of wholesale funding is better, but what is actually needed is a counter-cyclical measure. Goodhart (2008) suggests one such possibility.

much more conservative, risk averse way – probably more so than would be consistent with an innovative, dynamic economic system.

Second, there could be private insurance. For example, the credit counter-party risk of an asset, bought on the basis of whole-sale funding, could be insured by another financial intermediary, e.g. a mono-line insurer, via the CDS market, etc. The problem with this is that the overall risk is not eliminated but simply transferred and often concentrated in (insurance) intermediaries whose own position would be threatened by a major shock. The problems of AIG, mono-lines and Fannie and Freddie are cases in point. Moreover, as touched on above, risks that appear uncorrelated in normal times become highly correlated in stress situations. Indeed rather than relying more on private insurance, the lesson of the recent crash is that bodies and markets that purport to provide credit insurance need to be brought more directly into the macro-prudential net, and have their ability to take on and concentrate such risk more closely controlled.

Third, the Central Bank could become the market maker of last resort, to use Willem Buiter's apt phrase. When markets dry up, the Central Bank, in some extended Special Liquidity Scheme, takes the assets off the hands of the banks. If banks are leery of lending to each other, the Central Bank interposes itself as the central clearing house, taking in deposits from surplus banks and lending to deficit banks. There is not that much difference between being a 'lender of last resort' in a primarily bank-based system and being a "market-maker of last resort' in a predominantly capital market system. Both are subject to the same kind of 'runs', that call for official intervention. Of course, the Central Bank might suffer some loss, if conditions become really dire, but with the Government and Treasury behind it, it can always be recapitalised.[9] A greater concern than loss (we believe) is moral hazard. Should the Central Bank step in as market maker of last resort at the first whiff of difficulty, would that not lead the banks, and other financial intermediaries, to take on much more risk in normal times in the belief that they could unload it on the Central Bank in bad times?

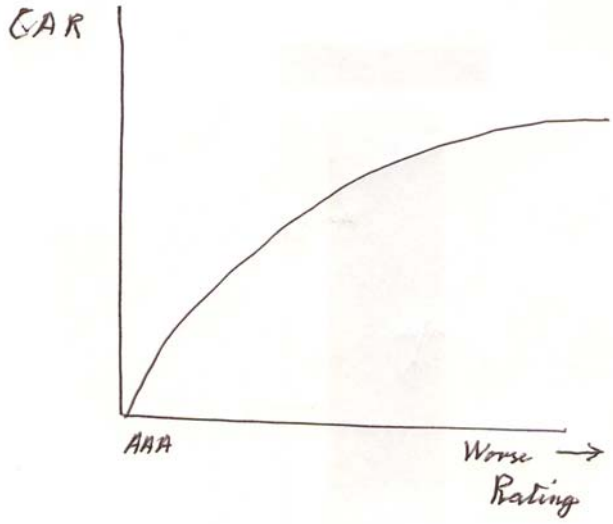
The fourth of our candidates is the Government (Treasury) which could provide public sector insurance against credit counter-party risk. This has now been done on a wide scale, and was the essence of the Kotlikoff/Mehrling/Milne (2008) proposal that the government guarantees the highest grade mortgage-based securities against such risk, thereby transforming them effectively into public sector debt. The questions then obviously arise on what terms and on what occasions the public sector should provide such insurance, and the prior issue of moral hazard recurs.

So, there are four potential sources of protecting the financial system against the failure of wholesale financial markets and, hence, of illiquidity. These are: (1) the banks, and the other financial institutions, themselves; (2) private insurance ; (3) the Central Bank; (4) public insurance. The question to be determined is what weight should be placed on each.

[9] The ECB would be recapitalised by its constituent NCBs.

In addition to the case of liquidity, a second example of the difference between macro and micro prudential behaviour relates to capital adequacy. Indeed one reason why regulators paid little attention to the liquidity problems discussed above was from a belief that if a bank had adequate capital, then it could always raise extra funding on wholesale markets. The micro-prudential approach suggested a risk-weighted capital adequacy requirement, as has indeed been introduced under Basel I and II. Surely a bank holding AAA assets is safer than a bank holding BBB assets, and therefore needs to hold significantly less capital, as prudential backing. Obviously in one sense, but not in another. Regulatory capital is meant to be held against unexpected loss, and not against expected loss, which should be met by a higher interest rate spread. The rating (should) measure the expected probability of default, whereas what matters is the likelihood of migration (downwards) of the rating, and the loss of value should that occur. Assume that both banks have the same risk-weighted tier one ratios, with say a similar buffer of 2% above the 4% requirement (i.e. 6%),[10] and that the risk of downwards migration (of say two notches) is the same for both assets, (AAA and BBB). Then which bank has more systemic risk? The answer generally is the AAA bank. This is for three reasons. First, AAA assets (many of which are mortgage-based structured products) are truly systemic, in the sense that they only lose value in a system-wide crisis, whereas BBB assets generally incorporate considerable idiosyncratic (i.e. diversifiable) risk. Second, the mark-to-market decline in value from the (assumed equal) migration may well be greater. Third, the relationship between rating and CAR is curved, see Figure 1, so that an equivalent horizontal migration leads to a greater proportionate requirement for extra capital at the 'best' end. So, for a given equal migration and equal capital buffer, the AAA bank will find itself in greater difficulties than the BBB bank.

Figure 1




[10] Having a minimum capital ratio, as a protection for banks, is just as silly as a minimum cash ratio, since it becomes unusable. We do advocate having such a minimum, but as a protection for the deposit insurer, and a trigger for prompt corrective action, and not in any way related to the need for resilience in the banking sector.

But perhaps this is no more than to record that risk-weightings are, inevitably flawed, and fail to reflect risk properly. More important is the point that micro-prudential measures, such as Basel II, and macro-prudential measures, such as we will advocate in Chapters 3 to 5, have differing purposes. The objective of a micro-prudential measure is to keep the individual institution behaving prudently, while that of the macro-prudential measure is to safeguard the system as a whole.

These two roles are often quite dissimilar. The micro-prudential concern is about individual risk; the macro-prudential with common, herd behaviour, and with shifts in generalised attitudes to risk. Individual institutional risk can often be seen to be low, or falling, as in 2004-6, when common macro-prudential risk is rising (and vice versa, as in 1992/93). Similarly, micro-prudential risk is concerned about risk concentration within individual institutions; macro-prudential-prudential risk relates more to similar portfolio holdings amongst institutions in the system. Indeed if all the individual institutions should be concentrated each in dissimilar portfolios (diversified into similar portfolios), the micro risk would be high, but the macro risk low (vice versa with low micro, but high macro risk). All this is set out in greater detail in Chapters 2 and 3.

Whereas we do claim that not enough attention has been given to macro-prudential risk alleviation, this is not to suggest that the present micro-prudential measures are unnecessary or wrongly designed (though they can be improved), but just insufficient on their own. Indeed where an institution, or market, is sufficiently large or strategic, so that its failure by itself would cause externalities, then it does need individually-targeted micro-prudential controls; in the case of banks this would be the Basel II risk-weighted CARs. Our point is rather that the micro-prudential regulations are not sufficient by themselves. They need to be supplemented by macro-prudential controls. We propose alternative measures whereby the Basel II risk-weighted CAR is interacted with macro-prudential measures to achieve a counter-cyclical overall effect and to penalise systemically dangerous funding mismatches. This issue, and the design of such counter-cyclical measures, is taken up in Chapter 4.

The current Basel II requirements for capital adequacy are pro-cyclical; as ratings migrate downwards in a bust, CARs rise, at a time when profits fall, write-offs increase and capital markets are unwelcoming to additional issues of equity, as evidenced recently (though the value of new capital raised exceeded our initial fears). What had not been sufficiently appreciated beforehand was the extent of interaction between the pro-cyclicality of the CARs and of the emerging mark-to-market, fair value, accounting system, IFRS, especially IFS 39, and FASB, especially FAS 157. That interaction is now well understood and under the US Emergency Economic Act of October 2008, a commission in the USA will study the wider financial implications of using mark-to-market for the financial system, whether its use should be amended, and, if so, how.

We are not in a position to second-guess the outcome of that study, and accounting practices are not our central focus, though we do see problems in moving

to any alternative procedure.[11] The point that we do want to emphasize is that the less that can be done to lessen the pro-cyclicality of 'mark-to-market', the more urgent it becomes to put greater weight on switching the effects of macro-prudential regulation from being pro-cyclical to becoming counter-cyclical. We also propose a mark-to-funding framework that (i) reduces procyclicality and (ii) provides incentives to reduce maturity-mismatch. Nevertheless we shall make some further brief comments on this topic in Chapter 5. we shall also deal there, and at rather greater length, with two other current issues, bankers' remuneration and limits on loan to value (and/or loan to income) ratios, plus a few words on other related topics.

A chief criticism of the current system of CARs, Basel I shifting over to Basel II, has been that it appeared to do too little to limit bankers' credit expansion in the boom, nor to help offset the wave of panic, failures and deleveraging in the subsequent crisis. CARs never seemed to bite, and financiers seemed to be able to do as they pleased, aided by much regulatory arbitrage via the shadow banking system and derivative markets, e.g. hedging counter-party risk via CDS.

An effective counter-cyclical macro-prudential policy will be an unpopular policy, since its purpose is to constrain the regulated from doing what they want to do when they want, by legal prohibition or by making it much more expensive. There is a natural incentive to avoid the regulation via a shift of business into the unregulated sector. We describe this as the 'boundary problem', which is described in more detail in Appendix A, largely a reprise of Goodhart (2008). There are two aspects of the boundary problem; the shift of activity to unregulated players; and the use of financial engineering to enable given capital to support more credit. Both are important. In the last boom the use of off-balance-sheet entities was arguably as, or more, serious as the shift to unregulated institutions.

The main point is that the 'boundary problem' is so pervasive that either financial regulation has to be fairly light-touch, so as to avoid massive avoidance via disintermediation, or to be restrictive and prescriptive in the sense of preventing disintermediation via legal prohibition. Our preference is for light-touch regulation, (with one exception on housing loan-to-value ratios, to be discussed later in Chapter 6). In general, restrictive control of financial intermediation stifles innovation and, especially if government starts to intervene with direct controls over bank lending, interferes with the appropriate allocation of capital.

This poses quite a problem. How do you make regulation counter-cyclical, effective (and hence unpopular), and yet at the same time relatively light touch? This is not easy; indeed if the solution was easy, it would have been discovered and applied long ago. We believe that our proposals, taken as a whole, would help to resolve this dilemma.

A second criticism of the Basel approach to CARs is that they did not do such detailed thinking about incentives and sanctions. Instead, they simply suggested preferred forms of (bank) behaviour. Thus they came out with proposed capital ratios, which then became translated into rigid minima. But such minima became in

[11] Under what circumstances would mark-to-market be suspended and by whose say-so? What alternative would be applied? How would that square with the ideal of transparency?

practice unavailable at times of need. For example banks currently cannot allow their published tier 1 capital to fall below 4%, despite one of the greatest unexpected shocks of all time. In fact, no significant bank would now dare to allow its ratio drop much below about 7-8%, because the market punishes banks even more effectively than supervisors. Far from adding to the resilience of the banking system, such required minima just represent a burden, and may even indeed exacerbate risk-taking by making it harder for bankers to obtain their target return on assets (ROA). Instead what is essential is to devise a calibrated ladder of increasing penalty, as the CAR falls below the 'well capitalised' level; again this largely involves following the lead of the authors of the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991.

# Chapter 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

[12] Today the divergence is as much as 50% of the asset. Assets where delinquency rates are less than 20%, are trading with an 80% discount to par.

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.

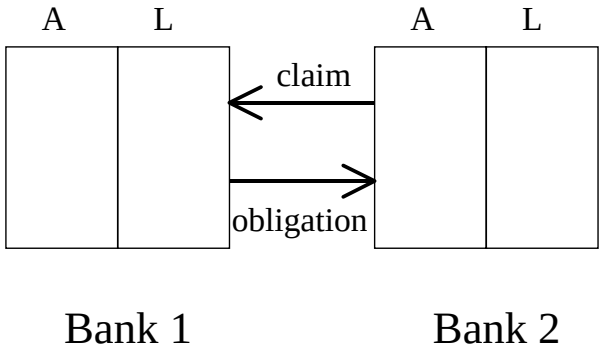
### 2.1.1. Systemic View Point – Funding Liquidity and the Domino Model

It may seem reasonable to believe that ensuring the soundness of each individual institution ensures the soundness of the system as a whole. However, this suffers from 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.

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.

**Figure 2**

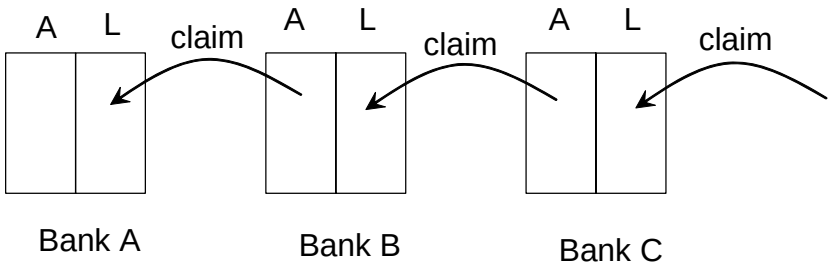



In the case where we have the combination of (i) Bank 1 not having alternative sources of funding, (ii) the reduction in Bank 2's lending being severe, and (iii) Bank 1's assets being so illiquid that they can only be sold at fire sale prices, then the withdrawal of lending by Bank 2 will feel like a 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.

[13] A fallacy of composition arises when one infers that something is true for the whole from the fact that it is true for each of the individual components of the whole.

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 the following chart.

**Figure 3**




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

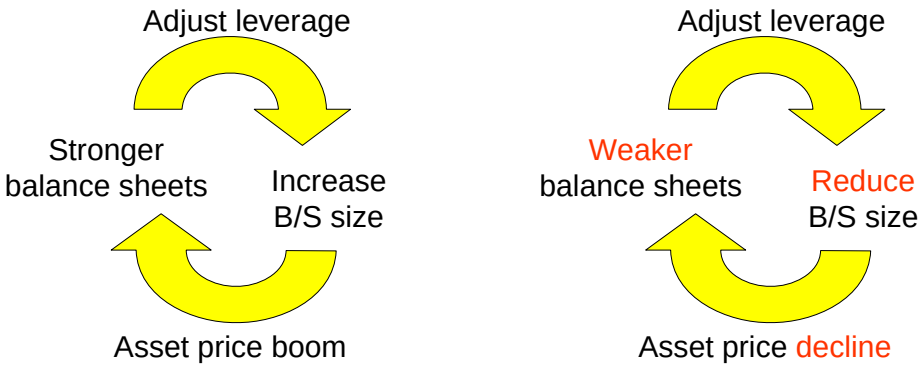
### 2.2. Loss Spiral – Asset Price Effect

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 illustrates the feedback during a boom. Note the critical role played by procyclical leverage.

Figure 4




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 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 above on the right 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

[14] FSA Guidance Note 4 (2002), "Resilience test for insurers". See also FSA Press Release, June 28th 2002, no FSA/PN/071/2002, "FSA introduces new element to life insurers' resilience tests".

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 leitmotif 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.

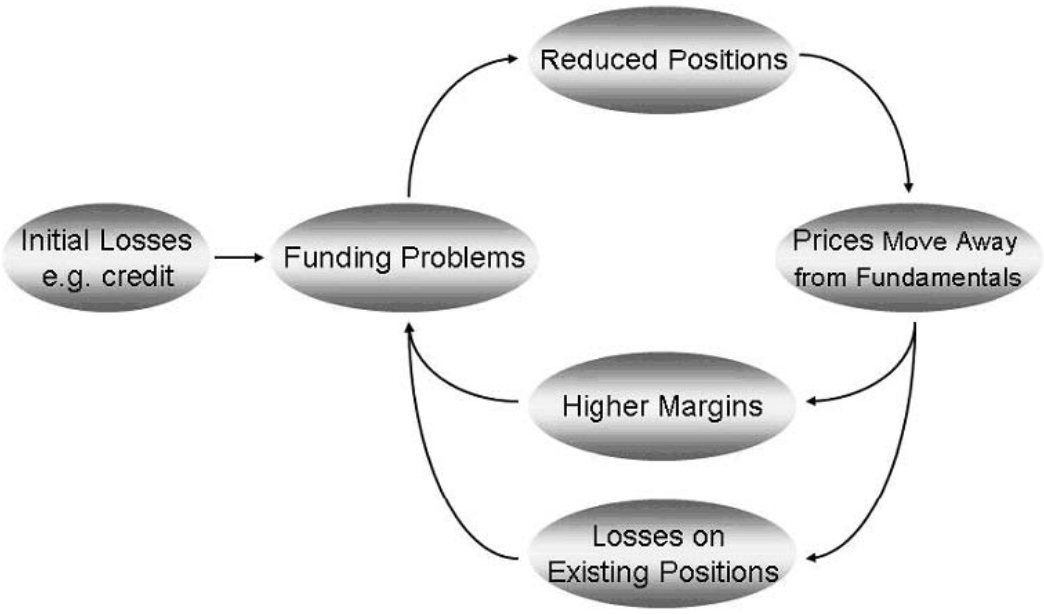
### 2.3. Margin/Haircut Spiral

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]

[15] Market Liquidity and Funding Liquidity, M. Brunnermeier and L. Pedersen (2009), Review of Financial Studies.
[16] Liquidity Black Holes, A. Persaud, Wider Discussion Paper No 2002/31, http://62.237.131.23/publications/dps/dps2002/dps2002-31.pdf.

**Figure 5**



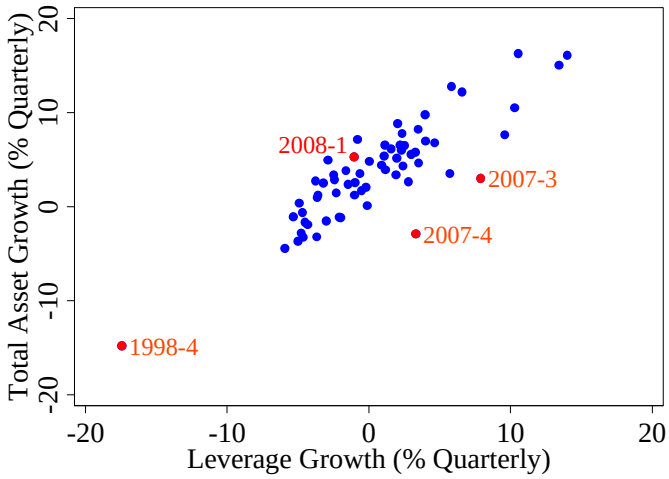

During downturns both spirals force leveraged investors to unwind their positions causing a) more losses and b) higher margins/haircuts and tighter lending standards, which in turn exacerbate the funding problems, and so on. Both spirals lead to procyclicality.

Figure 6 below is taken from Adrian and Shin (2007) providing nice empirical evidence for the margin spiral for the then US investment banks. 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[17]. For investment banks, however, the relationship is 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.

[17] This traditional response, may have been eroded by the trend in the most recent boom for home owners to try to benefit from lower interest rates and higher home values by frequent resetting of mortgages.

**Figure 6.**




**Leverage Growth and Asset Growth of US Investment Banks**
**(Source SEC; Adrian and Shin (2007))**

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. Indeed, as we see below, the evidence points to broker-dealers adjusting leverage by adjusting the size of their balance sheets, leaving equity intact.

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 equity. The table below is taken from the April 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.

**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, April 2008**

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.

## 2.4. 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."

### 2.5. 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.[18] It is arguably the main rationale for bank regulation. [19]

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 institutions 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.[20]

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[21], 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.

## 2.6. 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

[18] Geanakoplos and Polemarchakis (1986), "Existence, Regularity, and Constrained Suboptimality of Competitive Allocation When the Asset Market is Incomplete" in Heller, Starr, Starrett (Eds.) Uncertainty, Information and Communication, Essays in Honor of Kenneth J. Arrow, Vol. 3, Cambridge University Press, Cambridge, UK.

[19] While most current risk measures like Value-at-Risk (VaR) focus on the risk of an individual financial institution, Adrian and Brunnermeier (2008) develop a new risk measure, "CoVaR," that explicitly takes the risk spillovers into account.

[20] On network effects, see "Deciphering the Liquidity and Credit Crunch" Brunnermeier (2009), Journal of Economic Perspective 27(1).

[21] Today, this notion has been popularised by Nasim Taleb as a "black swan."

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.

# Chapter 3: Who Should be Regulated (by Whom)?

In Chapter 2 we provided a theoretical foundations for the regulation of financial firms by outlining an underlying mechanism which leads endogenously to financial instability. The first question one has to address is: Who should be regulated? Since any effective regulation forces firms to deviate from their preferred option, they always have an incentive to move their business outside the boundary of regulation. It is then no surprise that the adverse mechanism described in Chapter 2 reappears in the unregulated sector, which calls then for government support when a crisis hits. Commercial banks setting up associated conduits, SIVs and hedge funds in the last credit bubble is a vivid reminder of this "boundary problem", which is discussed in further detail in Appendix A.

In this section we propose some guiding principles on the scope of regulation, before emphasizing the importance of counter-cyclical financial regulation (Chapter 4) and liquidity and maturity mismatches (Chapter 5).

## 3.1. Classification of Financial Institutions based on Objective Risk-Spillover Measures

First, the classification of financial institutions should be based on objective risk- measures that capture the risk-spillovers from one institution to the next. This is especially important for macro-prudential regulation. Any financial institution that is subject to systemic risk, not only banks but also other interconnected financial players, like mono-line insurers, insurance companies (like AIG which turn out to be large-scale sellers of credit default-swaps), should be covered by regulation. The fault line of regulation should be primarily determined by the institution's actions and asset-liability structure, while its legal identity as bank, insurance company, SIV etc. should only play a secondary role.

Among others, CoVaR is one such spillover risk measure. It quantifies how financial difficulties of one institution can increase the tail risk of others. [22] Unlike the typical Value-at-Risk (VaR) measure, which captures the risk of a single institution, CoVaR captures the links across several institutions. More specifically, bank X's CoVaR is the conditional VaR of bank X's counterparty or the whole financial sector after conditioning that bank X is in difficulty.[23] . In particular, if bank X causes some

[22] Regressing an index of financial institutions or bank X's counterparties on bank X with quantile regressions is one tractable way to estimate the (non-timevarying) CoVaR. For more details see Adrian and Brunnermeier's working paper titled "CoVaR", http://www.princeton.edu/~markus/research/papers/CoVaR.
The work by Segoviano, of the IMF, and Goodhart examining the effect on the Probability of Default of other banks in the system, conditioned on the failure of any specific bank, is another exercise in this same genre.

[23] The same "conditioning-method" can be used for any risk-measure and is not restricted to VaR, which is not the ideal risk measure. The "co-expected shortfall" is an alternative spillover risk measure.

risk spillovers on its counterparties, then the CoVaR exceeds the unconditional VaR. Such a risk spillover measure should be supplemented with stress tests scenarios.

The risk-spillover of a financial player can be high if it (i) *causes* financial difficulties at other institutions or it is (ii) *simply correlated* with financial difficulties amongst other financial institutions. A good risk-spillover measure should encompass both channels, but this distinctions helps us to group financial institutions as:

1. **"individually systemic"**:

These institutions *cause* risk spillovers and include institutions that are so large, so massively interconnected, and so iconic as `national champions' that no government would ever allow them to fail. They require macro-prudential regulation and also micro-prudential regulation (e.g. Solvency II) due to their size.

2. **"systemic as part of a herd"**.(e.g., highly levered hedge funds)

These may be sufficiently small, and insignificant, for their individual condition not to be of great concern to the authorities, particularly when this is driven primarily by idiosyncratic factors, but when they move together as part of a larger group, their correlated fluctuations may well be systemic. Hence, they require some macro-prudential regulation but very limited micro-prudential regulation which might, in the case of hedge funds, be executed via their prime-brokers.[24]

3. **Non-systemic large** and not highly levered (e.g., Insurance Companies and Pension funds)

These institutions need full micro-prudential regulation, but no additional macro-prudential regulation.

4. **Tinies,** especially if they are unlevered, should have minimal conduct of business regulations.

The risk-spillover measure should determine whether a firm needs macro-prudential regulation (group 1 and 2) or not (group 3 and 4) and influence the extent of the capital and liquidity charges. For example, an insurance company like AIG that sells credit default swaps (CDS) on a large scale belongs to group 1 instead of 3.

Each year, and on the occasion of each major market event, the relevant regulators and supervisors should, as a matter of course, draw up a list of which financial institutions they consider to be 'systemic', and share such information with other supervisors.

At one stage we proposed that the regulators/supervisors should publish that list, in order to enhance the discipline that would impose on regulators/supervisors to do this exercise carefully. But we have been persuaded that the disadvantages of that course outweigh the advantages. Not only is the concept of 'systemic' fuzzy and state-contingent, but also it could lead to moral hazard and to artificial behaviour among the regulated in their attempt to achieve their desired status; this may represent

[24] Madoff's fund did not use a separate prime broker, since he had his own broker/dealer business. This was akin to allowing front and back offices to merge, and should never have been allowed.

a case for 'constructive ambiguity'.[25]

### 3.2. Rules for Individually Systemic Institutions

Let us revert to institutions that are individually systemic. There should be serious concern about allowing any of them individually to get into trouble. That means that micro-prudential regulation and supervision remain relevant. The established form of such micro-prudential regulation for banks remains Basel II. We advocate that this continues to be applied, as originally proposed, to large systemic, international banks, and that the appropriate micro-prudential controls continue also to be applied to large unlevered institutions.

This is not to say that the methods for assessing risk-weighted assets (RWAs) do not need reform; they do. As has been exemplified throughout the current financial crisis, the risks to a bank's liquidity and capital arise in large part from that bank's off-balance-sheet exposures and contingent liabilities, e.g., in back-up lines of credit to connected conduits, SPVs, SIVs, etc., and in derivative markets. The Basel Committee has done good work, which needs to be extended further now, in assessing how such contingent commitments should be incorporated into measures of RWAs.

Moreover, credit ratings, whether by CRAs or by the banks themselves, are a measure of expected loss, not of unexpected loss. For the latter, what one needs is some estimate of the likelihood of downwards rating migration, and of the resulting scale of loss in that event: a difficult exercise, but not impossible.

More attention should also be given to diversification, but care should be given to distinguishing between idiosyncratic and systemic diversification.[26] A bank which concentrates, for example, totally on loans to a particular category of borrowers in a particular region has no idiosyncratic diversification; i.e., it is subject to certain obvious risks, (i.e. if demand for such borrowers' products declines). However, being unlike most other banks, it is less likely to cause contagion, or spill-over risks, should it fail. Such a bank is, therefore, less dangerous to the system as a whole, and needs no particular extra macro-prudential regulation. It can make its own choices, without external pressure. Idiosyncratic concentration does not cause systemic spill-over risk.

Rules should also be designed in such a way that banks have no incentive to move assets into off-balance sheets vehicles and conduits. As outlined in Brunnermeier (2009) under the current system banks had an incentive to park and diversify assets in off-balance sheet vehicles in order to maintain a lower capital charge.[27]

### 3.3. Rules for Institutions that are" Systemic in a Herd"

[25] One possible way to increase the discipline on the regulators/supervisors would be for them to disclose and discuss their list in a closed session before a Select Committee of their Legislature once every year.

[26] On this topic, recent papers by Acharya and Yorulmazer, and by W. Wagner are germane.

[27] See Brunnermeier (2009), "Deciphering the Liquidity and Credit Crunch 2007-08", *Journal of Economic Perspectives*, 27(1).

We now consider the larger number of smaller institutions, mostly small banks, foreign branches and subsidiaries, most hedge funds, private equity, etc., who are not individually systemic, but may become so when they move together as a group (or herd). Consider a system of *n* different banks, where each has a portfolio consisting of an identical fraction (*1/n*) of the aggregate portfolio, (a representative bank system in effect). If one bank fails, the likelihood is that all will. A risk-spillover measure would calculate the aggregate bank portfolio in the country, and then compute how much the individual banks' portfolios correlate with the aggregate. The higher such correlation, the less diversified the system. As usual with risk assessment, co-variance is more important than variance. Since we focus on (left) tail-events, co-risk measures are superior to simple covariance measures.

The problem with Basel II was not so much that it was an incorrect metric of micro-prudential risk (though, of course, it was to some extent), but that it took insufficient account of macro-prudential risk, as set out and argued in Chapter 2. We argue that better measures of macro-prudential risk are to be found in leverage ratios, maturity mismatches and estimates of bank credit expansion and asset price expansions.

This group is currently divided in its view how to regulate this group of financial institutions for micro-prudential purposes. One view is that this wider range of (individually non-systemic) intermediaries should still be subject to individual micro-prudential regulation. Controlling their individual assumption of risk remains, on this view, both a desirable and appropriate function of regulation/ supervision. If such micro-prudential supervision is in force, then the macro-prudential factors can be interacted with them along exactly the same lines, as we propose in the following Chapters for systemic institutions.

The other view is that, if an institution is not itself 'systemic', there are no theoretical grounds for external interference in its own chosen risk profile; so, on this view, there would be no need for <u>any</u> micro-prudential regulation. Instead, all that would be required would be some simplified macro-prudential requirements, relating core capital to leverage and, perhaps, the intermediary's rate of asset expansion (growth).

Perhaps each country's regulators/supervisors could choose between these alternatives, which could also differ between types of intermediaries, thus banks could be treated differently from hedge funds.

### 3.4. International Considerations for International Entities

Typically one of the greatest concerns amongst both banks and politicians relates not so much to competitive inequalities amongst and between domestic banks, but those between domestic owned banks and subsidiaries and branches of foreign banks. Our principle is that the country or entity that bears the burden in case of a bailout of a financial institution should also be in charge of regulating that financial institution. Hence, we would suggest that any branch (of foreign-owned) banks designated as 'systemic' by a host country should automatically be required to change

its status to being a separately capitalised subsidiary. Then exactly the same capital and liquidity adequacy requirement calculations would apply to foreign-owned systemic subsidiaries[28] as to domestic banks .

As a consequence, if a bank should choose to open branches/subsidiaries in a foreign country that are large enough to be defined by that host country as 'systemic', then it will have to hold a separate pot of capital in that country, according to the host country's calculations. That will reduce the synergies of cross-border banking. It also raises the question of whether (and how far) such separate pots should/could go to satisfy the CAR of the consolidated bank. Our view is that regulators/supervisors are primarily concerned about conditions in their own countries (as has certainly appeared to be the case in practice in the current crisis). If so, the home regulator would relate the CAR to the RWA/Leverage ratio in the home country plus the 'non-systemic' branches/subsidiaries in other countries. We realise that this would be unpicking a part of the prior principle of consolidation on the home regulator.

We would encourage a more European approach within the Euro area. If burden sharing could be agreed upon within Euroland, regulation could be transferred to a European institution.

[28] If a subsidiary, or a branch, of a foreign-owned bank sited in some country, (especially in an off-shore centre), was primarily engaged in foreign-currency, entrepot, business, (e.g. intermediating in euro-markets), we would assume that the host country would not categorize that subsidiary/branch as being 'systemic' in its own market. Perhaps foreign-owned banks might want to divide their operations in a host country into two parts, a non-systemic branch running f.c. entrepot operations, and a, possibly systemic, subsidiary dealing mainly in host country business.

# Chapter 4: Counter-cyclical Regulation

In Chapters 1 and 2, we noted the pro-cyclicality that follows from banks chasing returns on equity, maintaining value at risk, and using mark-to-market valuation and risk approaches. We also highlighted the substantial risks for the financial and economic system when the cycle turns down. In this Chapter and the next, we describe how counter-cyclical regulation may be put in place. There are two main strands of financial regulation, one on capital and one on liquidity. We discuss capital regulation in this section and new approaches to the regulation of liquidity in Chapter 5. The main principles that we invoke are:

1. The main objective of counter-cyclical regulation should be to reduce the systemic risk that fluctuations in the conditions of an institution, or market, would have on the rest of the system. Systemic institutions (markets) should be regulated in direct proportion to their systemic risk. To achieve this, CARs are needed that are based on better risk spillover measures that take leverage, maturity mismatch and financing into account.
2. The measures have to be counter-cyclical, i.e. tough during a credit boom and more relaxed during a crisis. We propose a laddered response to ensure a prompt resolution of emerging problems before they can spill over to the wider financial system.
3. To ensure strict adherence and implementation of such rules, it is important to put an incentive structure for regulators in place and guarantee their independence from political and lobbying pressure.

## 4.1. Focus on Systemic Risk Spillovers

Capital charges should focus on the risk spillovers an institution causes, or is correlated with, rather than simply the institution's individual risk. As noted earlier, current required minimum capital adequacy ratios provide very little resilience and support to the system. Even so, we still advocate a low, fixed minimum ratio for capital, not as a protection for the regulated banks, but as a protection for the deposit insurance fund and the taxpayer, and as a trigger for prompt corrective action. This could be set as a low percentage of total balance-sheet assets and a low liquidity risk measure (as outlined in Chapter 5). Under FDICIA (1991) this percentage was set as 2%, though the arguments for choosing this rather than another number are not clear (at least to us). Overall, however, the focus should be on risk spillovers that potentially undermine the financial system as a whole.

## 4.2. When to Look Out for Systemic Risk?

Most financial crises are preceded by asset price bubbles. Bubbles often emerge after financial liberalizations or innovations and can persist since even rational sophisticated investors find it more profitable to ride a bubble rather than to go against it. This is in sharp contrast to efficient market hypothesis, but supported by

empirical findings.[29] Herding behaviour among financial institutions which are evaluated against the same benchmark are further contributing factors.

Counter-cyclical regulation should be most constraining during the height of a bubble. In the past, regulation followed a "benign neglect" policy with respect to bubbles. One justification for this approach was the argument that bubbles are difficult to identify with certainty. We find this reasoning unconvincing, since such an argument could be brought forward for almost any important policy decision. We favour a "lean against the wind" risk-management approach. We argue that such a leaning should be primarily done by counter-cyclical regulatory measures, such as we propose here, not solely via interest rates.[30] Financial authorities should be alerted when clear indications of a bubble emerge, even if the bubble cannot be identified for certain.

The regulation should be particularly effective for bubbles whose bursting might adversely affect the financial intermediation sector. While the bursting of the technology bubble in the early 2000s caused a lot of localised disruption, it bears no comparison to the turmoil which the bursting of the credit and housing bubble has caused. The big difference between them was that the technology bubble did not severely damage the lending sector. Said differently, it is important to determine whether a current funding and credit expansion is sustainable or subject to sudden reversals, with detrimental consequences for the economy. While monetary policy can have some role in "leaning against the wind" approach, bank regulation is central in controlling excesses in lending practices.

### 4.3. How to modify CAR?

The problem with Basel II was not so much that it was an incorrect metric of micro-prudential risk (though, of course, it was to some extent), but that it took insufficient account of macro-prudential and systemic risk, as set out and argued in Chapter 2. Better measures of macro-prudential risk are to be found. We argue that leverage ratios, maturity mismatch and estimates of bank credit expansion should be taken into account. Since both RWAs (Basel II) and measures of macro-prudential risk have strengths and weaknesses, our proposal is to <u>combine</u> the two approaches. How would we do this? We propose multiplying the basic CAR, estimated under Basel II, by a factor, or factors, relating to macro-prudential/systemic risk.

We have to determine two elements:

(i) the factor and how to compute it
(ii) which capital ratio (tier 1 or tier 2) to multiply.

With respect to computing the macro-prudential *factor*, quantitative impact studies should determine the weights on leverage, maturity mismatch, credit and asset price expansion; and also the time periods over which such expansion should be estimated. Highly levered and fast growing 'systemic' institutions would be subject to

[29] Abreu and Brunnermeier (2003) "Bubbles and Crashes" in Econometrica provide theoretical reasoning why rational traders prefer to ride the bubble rather attack against it. Brunnermeier and Nagel's (2004) Journal of Finance article "Hedge Funds and the Technology Bubble" provides empirical support for this finding.

[30] We continue to support the adoption and maintenance of inflation targets, and interest rate policy is primarily predicated to the achievement of such targets.

higher capital requirements than the rest. The idea is that when there is increasing systemic risk, with increasing leverage, maturity mismatch, credit expansion and asset price increases, the multiplication factor would be greater than unity, and less than unity during periods of deleveraging.

One specific measure that achieves this is CoVaR, outlined in Chapter 3.[31] However, many risk measures suffer from the short-coming that they may be pro-cyclical if naively applied, since after a boom phase estimated volatility and correlations are low, and a subsequent price drop would lead to a drastic jump in such an estimated risk measure. This would trigger a procyclical margin spiral as described in Chapter 2. Hence, we propose to estimate risk measures such as coVaR using primarily past "crisis data" and underweighting recent and current data. This problem is not so acute when using simple leverage ratios and measures of credit expansion or asset price fluctuations.

With respect to the relevant *capital ratio*, one problem is that there are several Basel II ratios, notably the core Tier 1 ratio, the 4% Tier 1 ratio and the 8% Tier 1 plus Tier 2 ratio. In the 1980s, largely in the context of the Basle Committee on Banking Supervision (BCBS) deliberations on the Basel I Capital Accord, there were extensive and bruising discussions on the definition of capital. Few since then, have had any enthusiasm or stomach for re-opening this issue.[32] Nevertheless if our intention is to interact a Basel II CAR with macro-prudential, counter-cyclical factors, we have to choose (or to propose) what should be the relevant ratio for doing so. It is our view that, as the recent crisis has progressed, the market has come to place most weight on the core Tier 1 ratio, and, for reasons already set out in Chapter 2, we tend to concur with that. So we would propose interacting macro-prudential factors with the core Tier 1 ratio for each bank (and/or systemic bank subsidiary). But the choice of micro-prudential ratio for inter-active purposes is not a matter on which our group has strong views. In the light of the use by governments of preference shares for the recapitalisation of banks, tier 1 might be preferred to core tier 1. The critical point is that such a ratio can be adopted as the basis for interaction.

### 4.4. More on Bank Capital: Two Notions

In the discussion of which Basel II ratio to interact with our macro-prudential, counter-cyclical, controls, it may help to draw a conceptual distinction between two notions of bank capital. There is, first, the notion of bank capital (implicit in the Basel approach) as a buffer against loss that protects depositors. Under this first notion of bank capital, hybrid claims such as preferred equity or subordinated debt are counted as bank capital, since both are claims that are junior to depositors. Indeed, under the Basel capital accord, subordinated debt counts as Tier 2 capital.

However, there is a second, contrasting notion of bank capital as the claim held by the owners of the bank who have control over the bank's operations.

[31] For more details see the working paper "CoVaR" by Adrian and Brunnermeier.

[32] There are enthusiasts for requiring 'systemic' banks to issue more marketable subordinated debt, and to relate regulatory responses to the market price of such debt. It has not, however, been demonstrated why this approach would not be just as pro-cyclical as the present Basel II RWA system.

Arguably, hybrid claims such as preferred shares or subordinated debt do not qualify as bank capital under this second notion of bank capital, as they can be seen as junior forms of debt. When the bank has too little capital in this second sense, the owners' incentives reflect their highly leveraged balance sheet. When faced with a dwindling stake in a leveraged entity, the owners with control have little to lose, and everything to gain by engaging in risk-shifting bets on the bank. The increased repo haircut imposed by the capital market during distress episodes could be seen as the increased margin demanded by creditors in the capital market to changed circumstances.

The key point is that the repo haircut and the implied maximum leverage is a constraint imposed by the capital market, and reflects the terms on which creditors are willing to lend to those with control over the leveraged entity. One plausible channel through which the constraint operates is the wish by creditors to avoid being embroiled in a lengthy and costly bankruptcy settlement after the borrower has defaulted. When a bank breaches the maximum leverage ratio permitted by the market, the bank must take remedial action to reduce its leverage, or face a run by its creditors.

Northern Rock's demise illustrates these issues starkly. Northern Rock was a highly leveraged institution when considering the leverage on common equity. Its high leverage made it especially vulnerable to a deterioration in overall funding conditions for the financial system as a whole. See Figure 8.

**Figure 8**

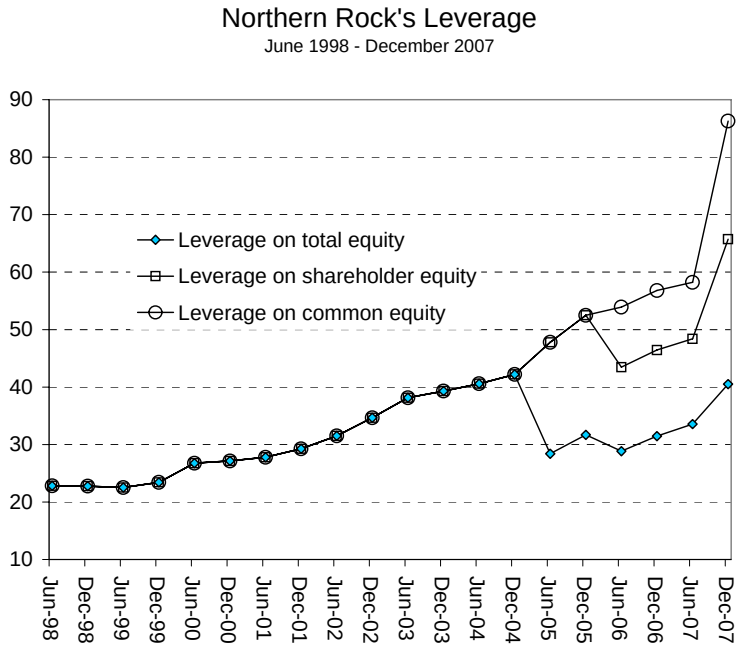



Figure 8 plots the leverage of Northern Rock from June 1998 to December 2007, using three different measures of equity. Common equity is the most basic form of equity – it is the stake held by the owners of the bank with voting power and hence who have the right to exercise control over the bank. "Shareholder equity" in

Figure 8 is defined as common equity plus preferred shares. Finally, "total equity" in Figure 8 is shareholder equity plus subordinated debt, a class of debt that is senior to the common and preferred equity, but which is junior to other types of debt taken on by the bank, including deposits.

Figure 8 shows very explicitly how important it is to calculate the leverage ratio with common equity in the denominator. Leverage based on total equity creates the false impression that Northern Rock might have no difficulty rolling over its short-term repo-funding.

### 4.5. Ladder of Responses

With regards to the second main principle, that regulation should be most constraining during credit booms, we propose a propose a prompt and laddered response when a financial institution violates our much more stringent capital and liquidity requirements starting at an earlier stage. We follow the successful implementation of laddered approach introduced by the FDICIA. For example, when the capital and liquidity requirements are, say 1% below target value, supervision could be enhanced. When the target is missed by 2%, the institution could in addition be forbidden to pay out dividends or make other forms of equity payouts. A missed target by 3% might disallow any bonus payments to the CEO and other board members. A miss by 4 % could require recapitalization or closure within two months.

### 4.6. Clear Incentives for Regulators: Rules versus Discretion

The third main principle that we propose is that objective criteria and pre-specified rules should be put forward to guarantee that financial regulation is strictly enforced. To ensure that the enforcement of these rules are credible, regulators must face the right incentive structure and enjoy a degree of independence that allows them to impose potentially unpopular steps. When everyone is calling for more regulation, e.g., as now, just after a crisis, it is not needed at all, since bank managers are timid and risk averse. When regulation is needed, no one wants it, because asset prices are rising, there is a boom, everyone is optimistic, and regulation just gets in the way (see Appendix A). Almost every regulator/supervisor will seek maximum discretion. Because of the above considerations, regulation should be based on pre-set rules; otherwise, few regulator/supervisors will actually dare to face the odium of tightening in boom conditions. There is actually little that we are proposing here that in principle and in theory could not have been accomplished under the discretionary Pillar 2 of Basel II. In practice it was not used that way at all, and probably never would be.

### 4.7. Cross-country Considerations

Countercyclical measures should be applied on a country-by-country basis, since cycles are not identical and matching across all parts of the world. Thus Germany and Italy did not share in the housing cycle that affected USA, UK, Spain, etc. Credit expansion took place at a very different pace in various countries. So we

cannot, and should not, as yet talk about a world cycle. However, for certain regions, notably the euro-zone, countries could coordinate countercyclical measures on a wider regional basis. Thus, even though the Basel II basic CARs are level across countries, the actual interacted counter-cyclical CARs would become higher in countries with an asset boom and rapid credit expansion, than in countries not exhibiting such conditions.

### 4.8. Contrast to Spanish Dynamic Provision Mechanism

It may be worthwhile comparing and contrasting our approach with that of the Spanish pre- (or dynamic) provisioning mechanism.[33] They are closely similar in intent and construction. Both use macro-prudential measures to interact with prudential requirements. We put more weight on leverage ratios and maturity mismatch;[34] the Spanish put more weight on credit growth (compared to longer-term average values); both would apply numerical coefficients, which have been pre-set, to adjust prudential requirements to the relevant individual banks, including subsidiaries of foreign banks in Spain. The Spanish scheme relates solely to credit growth and provisions in Spain. Just as in the case of our proposals, if the Spanish scheme were applied more widely to other countries, its application would differ from country to country depending on the state of the cycle in each country.

A complication of the Spanish pre-provisioning scheme is that it appears to run foul of the accountants and of IFRS (and the tax authorities?). A disadvantage of our proposal is that, by interacting (national) macro-prudential factors with the Basel II CAR, it ends the single level-playing-field for cross-border capital requirements and, with that, the simplicity of head-quarter management and of overall consolidation/control by the home regulator/supervisor. Under our proposal here much more control would flow back to each (systemic) subsidiary and to its host regulator.

While we applaud the principles of Spanish pre-provisioning, we do not think that its quantitative effect has been to moderate the credit cycle by as much as our mechanism could. Nevertheless should the proposal outlined here be regarded as too radical, the universal adoption of the Spanish pre-provisioning scheme (and the adjustment of IFRS to allow that to occur) would represent "counter-cyclical-lite."

---

[33] On the Spanish procedures, see G. Jimenez and J. Saurina, (2006), 'Credit Cycles, Credit Risk and Prudential Regulation', International Journal of Central Banking, 3 (2), (June), 65-98, and Fernández de Lis, S., Martínez Pagés, J. and J. Saurina, (2000), 'Credit Growth, Problem Loans and Credit Risk Provisioning in Spain', Working Paper No. 0018, Banco de Espana.

[34] The danger of leverage is greater, the worse the mismatch between funding and asset maturities. The weight to be placed on leverage ratios in macro-prudential CARs should vary inversely with the degree to which our proposals on Liquidity (Chapter 5) are adopted.

# Chapter 5: Regulation of Liquidity and Maturity Mismatches

We have argued in previous Chapters that the current philosophy of banking regulation - that you can make the system safe by making individual institutions safe[35] - is an unsatisfactory basis for insuring systemic stability. One of the principal reasons for regulating banks, over and above the way we regulate other businesses, is that their inter-relationships are such that the rational response of one prudent institution to an unexpected loss – a reduction in lending and a sale of assets - may have systemic implications for other prudent institutions. These systemic implications are compounded by herding behaviour and the use of contemporaneous prices in measures of asset value and risk, as discussed in Chapter 3.[36]

In Chapter 2 we indicated that, while traditional views of systemic risk are based on contagious bank failure, a key avenue through which systemic risk flows today is via funding liquidity combined with adverse asset price movements due to low market liquidity.[37] For example, it is worth reiterating that in the case of Northern Rock, Lehman Brothers and Bear Stearns, the failure was precipitated by the inability of these firms (i) to roll over their liabilities (funding illiquidity) and (ii) to sell mortgage products at non-fire sale-prices (market illiquidity), rather than their finding that their borrowers did not pay up. The freezing of the interbank market, the asset backed commercial paper market, and other sources of funding was more systemic than specific to certain institutions.

This section seeks to develop a regulatory approach to liquidity that may make systemic liquidity events less frequent or severe.More specifically, we propose the following two measures:

1. Mark-to-funding accounting rule

Pools of assets for which long-term funding is secured can be put in a "hold-to-funding account" and do not have to be marked-to-market.

2. Explicit capital charge for liquidity risk

Financial institutions who hold assets with low market liquidity and long-maturity and fund them with short-maturity assets should incur a higher capital charge.

Both measures should provide an incentive for institutions to find more stable long-term sources of funding. Before explaining the specifics of the proposal, we outline why it is inadequate solely to consider the payoff quality of assets and essential to consider liquidity aspects.

---

[35] For a more detailed explanation of this assessment of the current approach to regulation see Redesigning Regulation of Pensions and Other Financial Products, John Nugée and A. D. Persaud, (2006), Oxford Review of Economic Policy, 22, (1), pp. 66-77.

[36] See, "Sending the herd off the cliff edge: the disturbing interaction between herding and market sensitive risk management systems", A. D. Persaud, Jacques de Larosiere Prize, Institute of International Finance, (2000).

[37] See, "Market Liquidity and Funding Liquidity", (2009), Review of Financial Studies.

### 5.1. Focussing solely on Assets' Expected Payoffs is Insufficient

We argue that focusing exclusively on the quality of a bank's asset portfolio is insufficient, even though the bank's asset quality and its ability to access funding liquidity are related. Historically, bank supervisors had hoped that this relationship would be strong enough to make capital adequacy requirements sufficient for dealing with the issue of liquidity.[38] If there had been little doubt about the quality of the assets, a liquidity crisis might not have emerged. However, it seems also likely that the crisis would have been more modest in the first instance and potentially more containable if the same assets had been funded with longer-term liabilities. Assets would not have been sold in distressed fashion in such an environment

### 5.2. Funding Liquidity and Maturity Mismatch

The financial system's reliance on short-term funding of long-term assets with potentially low market liquidity has been the main source of financial instability. One of the most critical lessons of this crisis is that, while we have previously focused on asset quality, systemic risk has as much to do with how assets are funded. If two institutions have the same asset, but one funds with long-term debt and the other by borrowing overnight from the money markets, it makes a substantial difference to the potential for systemic risk. Yet current regulatory rules make little distinction between how the same assets are funded. The absence of distinction gives banks an incentive to fund assets cheaply. This incentive is most pronounced when the yield curve is upward sloping in a boom. This explains why there was a collective reliance on short-term, wholesale market funds in the run-up to the 2007 crash.

Is maturity mismatch an inevitable feature of a private banking system?After all, is not banking all about borrowing short and lending long? Not necessarily. There are many caveats to that generalization, and it is also a matter of degree. Small business will complain that, given the onerous covenants on bank loans, they do not feel that they have borrowed long-term from banks. And it is often said that depositors are more likely to have a divorce than leave their bank. The reality is that retail demand deposits are not as instant as they appear. The ever-declining ratio of private sector domestic deposits to total liabilities, as banks increasingly relied on short-term wholesale deposits in the run up to the crash of 2007, was one measure of increasing funding liquidity risk amongst banks. Effectively, banks' maturity mismatch got worse through wholesale financing.

[38] In the first meeting of the Basle Committee of Bank Supervisors (BCBS) in February 1975, the Chairman George Blunden, said '- the Committee's main objective was to help ensure bank solvency and liquidity'. "From the outset, the BCBS appreciated that solvency and liquidity were inter-related; both were essential for the stability and survival of a banking system. An illiquid bank (system) would not remain solvent for long, nor an insolvent bank (system) remain liquid", C. A. E. Goodhart, in a forthcoming volume on the history of Basle 1.

### 5.3. Mark-to-Funding – A New Accounting Rule

We propose a new accounting rule in order (i) to give financial institutions an immediate incentive to reduce their maturity mismatch and (ii) to reduce the procyclicality which mark-to-market induces in asset booms and bust due to the loss spiral. We are as much concerned about the over-expansion in booms as by the crashes in busts; indeed the former often causes the latter. In particular, we are worried that current price declines force institutions with medium-term funding or liabilities to sell them in order to comply with prudential rules based on mark-to-market valuations, leading to further price declines that in turn force further sales. We are also concerned about the role mark-to-market accounting rules may play in a crisis, in keeping buyers away from assets they consider well valued over the medium-term, for fear of the impact of volatile mark-to-market valuations over the short-term. Mark-to-market volatility is highlighted as the main reason why credit investors are not buyers of instruments today. It seems to us that in these specific cases, applying mark-to-market accounting, quite apart from its systemic implications, is not in the prudential interests of the firm or the economy.

However, we are also not comfortable with the idea that mark-to-market accounting should simply be suspended in periods of market decline, or for assets to be simply shifted from the available-for-sale and trading books onto the hold-to-maturity (banking) book, where they can be valued differently. It seems to us that this would worsen one of the main problems of pro-cyclicality, which is that bankers and investors pay insufficient attention in the boom to the possibility that prices may fall back when the boom is over. Moreover, it does not strike us as credible for a bank to declare in the middle of a crisis that it is their intention to hold an asset to maturity and so it is no longer necessary to value it using current market prices, if the asset is funded using short-term money market borrowing that may well dry up before the asset matures. It is also not responsive to the increased demands of transparency that accompany the confusion of a crisis.

We believe that there is a middle road that is broadly similar to the mark-to-market approach, but better reflects the prudential interests of financial firms and is more honest than a suspension of market-to-market or a wholesale shift of assets to "hold to maturity". The approach is called mark-to-funding,[39] and the principle is that assets should be valued and managed in a crisis, not according to the intention of the holder or the short-term vagaries of the market, but according to the capacity of the holder. Capacity to hold on to assets is driven by the maturity of the funding of the asset. In other words, if a bank has funded its twenty-year assets with one-month borrowings, whatever their intention, they should value the asset with the expected price of the asset in one month time. In particular, it can ignore temporary price movements within a month. If a bank is funded over the short-term, this approach will provide no more relief and would be no less pro-cyclical than mark-to-market value accounting. However, this would be a fair reflection of the price risk the firm faces if funding is not rolled over.

[39] See "Reason with the messenger; don't shoot him: value accounting, risk management and financial system resilience." Avinash Persaud, Vox.EU, 12 October 2008, and "New twist on market to market stirs debate, Financial Times, November 30, 2008.

If regulators emphasise short-term valuations irrespective of the funding of assets, the financial system's natural risk absorbers and market liquidity providers will not be able to act in that capacity. We would be left with an over-reliance on short-term funding and risk traders whose risk management strategy is the systemically dangerous behaviour of trying to reduce risks when they rise by selling them on to someone else and the system's resilience would have been unnecessarily reduced.[40]

Let us conclude our discussion on mark-to-funding with two caveats. First, while mark-to-funding alleviates pro-cyclicality, we are also aware that some pro-cyclicality remains since funding tends to be more short-term during financial crises. Second, allowing "hold-to-funding account" grants financial institutions some discretion how to value their assets. This introduces additional asymmetric information and might hinder future funding. In short, mark-to-funding alleviates the loss spiral, it can worsen the margin spiral. To overcome the asymmetric information problem we subscribe to Allen and Carletti (2008)[41] proposal that each financial institution should publish two balance sheets: one based on mark-to-funding and another one based on mark-to-market evaluations.

### 5.4. Capital Charges against Illiquidity

Our second measure to overcome liquidity risk is to impose a capital charge on it. Conceptually, regulatory capital should be set aside against the riskiness of the combination of an asset and its funding, since the riskiness of an asset depends to a large extent on the way it is funded. The goal of our objective of this liquidity adjusted capital charges is to encourage banks to find long-term funding, and dissuade them from greater leverage.

Let us first consider the wider merits of this principle before returning to more practical issues of estimating regulatory capital, since there will be many ways of applying this principle and the precise methodology of application is less important than the principle. The principle is in fact quite close in spirit to the mismatch ladder that had been considered by the Basle Committee several decades ago. To adopt the Basle language, if capital is to be risk sensitive, it must be sensitive not just to the risk of assets, but to the risk of the combination of the asset and its funding, which includes the leverage and maturity mismatch. We argue that if two banks hold the same asset, the one funding the asset with term deposits would set aside a lower amount of capital than the one funding the assets with overnight borrowing from the money markets. If funding markets dried up for three months, the short-term funded bank would be in difficulty and would be forced to sell assets that would worsen the liquidity and solvency environment for its competitors. Said differently, the bank would cause a fire-sale externality due to low market liquidity.

In practical terms, adjustments to capital to reflect the maturity mis-match between assets and liabilities could be done as simple multiples to the current requirements for capital which are based on the credit quality of assets. If the

[40] Regulation, valuation and systemic liquidity, A. D. Persaud, Banque de France, Financial Stability Review, No. 12, 2008,

[41] Allen and Carletti (2008), "Mark-to-Market Accounting and Liquidity Pricing", Journal of Accounting and Economics.

boundary lines of regulation are recast as we suggest in Chapter 4, the multiples could have a minimum below 1.0, allowing this new capital requirement to be applied to institutions that have assets funded by long-term capital and enabling them to put aside less capital than those who fund assets by borrowing.

We propose that the maturity mismatch multiple is a function of the months of *effective mismatch* between the asset maturity and the funding maturity. Estimating the effective funding maturity is marginally easier than estimating the effective asset maturity. It relates to the term of the borrowing. An issue arises where funding is through deposits which tend to be effectively long-term funding, but depositors often have instant access. In this case banks could be given an opportunity to prove to their supervisor that the effective maturity of their deposits is longer than a day, perhaps using past deposit behaviour in stressed environments as evidence[42].

The effective maturity of an asset takes the asset's market liquidity into account. We take as effective maturity of an asset to be the lower of the maturity of a loan[43] and the length of time it would take to sell the asset in a stressed environment without taking a significant haircut. Assets that are accepted by the central bank as collateral for loans without a significant hair cut would have an asset maturity of less than one day. One month loans have an asset maturity of a month or less.

The problem is assessing the effective maturity for long-term loans that are not (normally) eligible as collateral at the central bank. In the circumstances were the maturity of loans are in excess of two years, supervisors can apply a default range of 12 to 24 months and require banks to put 50% of their assets in a 12 month bucket and 50% in a 24 month bucket to be reviewed periodically. If we had not previously proposed counter-cyclical multiples to the existing capital charge (see Chapter 4) we would also recommend regulators to raise the proportion of assets held in the 24 month bucket in a methodical manner as a boom progresses.

If an asset can be sold to the central bank in a day and is funded with overnight borrowing, there is no maturity mismatch. On the other hand, if a twenty year mortgage may be sold in approximately 24 months, and is funded with overnight money, there is a significant maturity mis-match.

In our judgement, a reasonable range for the multiple might be 0.5 to 2.0, with a maturity mismatch of somewhere between 6 months carrying a multiple of 1.0. The current crisis inevitably calibrates our sense of what is an adequate degree of safety. Ten years ago, ensuring that funding maturities would allow institutions to survive a few weeks in the face of a disruption to money markets was considered adequate. In the 2007-08 crisis, wholesale money markets have been more or less closed for many borrowers for over a year.

[42] It should be noted that where banks have attracted deposits through above normal interest rates, they appear to have captured the more footloose depositors who are quicker to depart at the first sight of trouble. Perhaps a measure of the longevity of the deposit is how poor the competitive interest rate is, a Divisia approach.

[43] For assets other than loans we measure measured as the average time it takes to be paid back the current value of the asset

Finally, each bank can calculate its effective maturity mismatch for the whole balance sheet or also for several subpools. That is, capital could be assessed against pools of assets that have been assigned to pools of funding. The assignment of pools of funding can be changed as long as the capital requirement is then re-assessed. It is better to allow banks to form pools explicitly on their balance-sheets rather than give them the incentives to do it via off-balance sheet vehicles.

As always in regulation, there are issues to be faced concerning home/host responsibilities. Traditionally regulation of liquidity has been the responsibility of the host country, at least in the case of subsidiaries. Subsidiaries can, and do, apply to the host country central bank for liquidity assistance. On the other hand, centralisation of (and economising on) liquid assets has been a key aspect of cross-border banking. That conflict of interest was, until recently, camouflaged by the willingness of regulators to dispense with controls over liquidity.

But now that conflict is resurfacing, for example in the FSA's requirement in the UK that all banks, including foreign-owned subsidiaries, hold specified ratios of British public sector debt. We would propose relating such liquidity ratios only to those foreign-owned subsidiaries designated as 'systemic' (see Chapter 3). Beyond that, we would suggest that, whereas the host country retains the prime responsibility for liquidity requirements, this can, and should, be modified by bilateral, or multi-lateral MOUs, perhaps in Europe through the good offices of CEBS.

# Chapter 6: Other Regulatory Issues

## 6.1. Introduction

Capital and liquidity requirements are the main staple of financial regulation. Nevertheless there are a number of other issues that need some discussion, though in some cases only briefly.

A major focus of public anger has been the huge remuneration of senior banking officials, who now appear to have reaped inappropriate pecuniary benefits for taking risks that others have ended up shouldering. Moreover, did the structure of such a reward system itself induce these same executives knowingly to take enhanced risk? We discuss this in sub-section 6.2..

The boom-bust cycle in the housing market has been closely associated with the credit cycle amongst the banks. In sub-section 6.3. we ask whether another instrument that the authorities might use to dampen down such cycles could be (counter-cyclical) variations in Loan-to-Value (LTV) or loan-to-income ratios.

We touch, briefly, on issues relating to Credit Rating Agencies, in sub-section 6.4., before concluding with short comments on Central Clearing Houses for derivatives, and the need to dampen end-year (and end-quarter) spikes in financial markets (sub-section 6.5. and 6.6.).

## 6.2. Remuneration

Compensation practices at financial firms have become a topic of particular scrutiny in recent years. The general public has become increasingly aware of the high levels of remuneration in the financial sector and indignant when large severance packages are awarded to executives who have presided over meltdowns in franchise values. Both regulators and practitioners have recognized the potential for perverse incentives to lead to unjustified risk taking and thus to contribute to systemic instability. Political leaders, most recently at the G20 summit, have called for a reexamination of the incentive structure of compensation in the financial sector.

One benefit of a crisis is that it prompts action to deal with vulnerabilities that have been allowed to build up over an extended period of benign conditions. The drawback, however, is that hasty responses can have unintended consequences, and can focus on issues that have the greatest public profile, as opposed to the most significant practical impact.

The high level of public interest in the subject, combined with the financial turmoil of the current crisis means that some additional regulation of financial sector pay is inevitable.[44]

[44] To quote Thomas Huertas, Director of Banking Sector Regulation at the UK FSA:

In what follows, we attempt to separate which pay practices are a legitimate subject for financial regulation and which are not. In line with the rest of this report, we focus primarily on issues that are directly associated with financial stability.

We do not believe that a public perception that financial sector salaries are "too high" is a sufficient reason for regulation. A free market system generates wages and prices determined through the interaction of supply and demand. Unless there are reasons to believe that these forces are artificially distorted, in which case, as we argue below, there is a case for public intervention, trying to control pay is usually ineffective and frequently counterproductive.[45] Markets are adept at finding ways around regulations that attempt to fix wages and prices at levels that do not correspond to market equilibrium.

This is not to say there is no public policy interest in dealing with pay disparities. Far from it. Societies have devised numerous techniques to address the issue of income inequality, including most importantly the progressive income tax. The absolute level of pay in the financial sector, insofar as it raises social issues related to income disparities, should be dealt with by these mechanisms. In this respect, finance is no different from any other economic sector

There are, however, two aspects of financial sector pay that are legitimate objects of regulatory intervention. One relates to conflicts of interest and stakeholder protection, and the other relates to systemic stability.

Throughout our report, however, we are in essence talking about social externalities which for a host of reasons are not sufficiently internalized by banks. Bankers' remuneration has incorporated insufficient internalizing of the social costs of excessive lending. But we aim to deal with this through our additional capital charges. The response of banks to less profits in the boom should be smaller bonuses, so there would be less need for regulators to meddle in the overall level of remuneration.

*Stakeholder protection* issues arise when asymmetric information leads to levels of remuneration that depart from those that would prevail in a genuinely free market for talent. This can, in principle, occur when managers of businesses award themselves contracts that are more generous than the owners of the business would choose to award if they were aware of all the facts. The most immediate sufferers from such practices are the owners (shareholders) of the business, but insofar as excessive pay erodes the profitability of a financial firm, broader financial stability issues can also arise.

---

> "…Supervisors will tackle remuneration policies. Firms themselves have admitted that remuneration policies may have been a contributory factor to the financial crisis. We concur, and have therefore written to the CEOs of major firms to assure that firms' remuneration policies are consistent with sound risk management. We will also work with other regulators in bodies such as the Financial Stability Forum to assure that this problem is tackled on a global basis".

[45] Regulators appear to agree with this assessment. To quote Thomas Huertas again: "Our concern is not with the level of pay. We have no objection to people earning high compensation, provided they earn it in a way that is consistent with sound risk management..."

The answer to this potential distortion, we believe, lies in strengthened corporate governance. Boards of directors should have compensation committees that are appropriately reflective of shareholder interests. This means, at a minimum, that such committees should be composed of independent directors, with the capacity to reach informed judgments on pay levels and structures. We believe that transparency is also a valuable safeguard. Membership in Compensation Committees should be disclosed, and salaries and benefit levels for all senior executives should be made public.

To address the suspicion that compensation consultants play a role in ratcheting compensation levels upward, there could also be a case for requiring advice of compensation consultants to be divulged more widely. We believe that the supervisory authorities responsible for conduct of business regulation should formulate guidelines consistent with these principles and hold regulated institutions to account if they do not follow them.

Of much more moment from our perspective, however, is the potential for remuneration practices to adversely affect *systemic stability*. This will occur if the structure of remuneration encourages decision makers to take risks whose social costs diverge from the costs facing the individual decision taker. It is not hard to think of ways in which this could happen. Where a financial sector decision taker receives a portion of the profits generated in any time period, but does not absorb a corresponding share in the losses generated in other time periods, he/she has an incentive to take additional risks. The executive will obtain potentially large returns in good times, while simply receiving nothing in bad times.

There is a widespread belief that that was a general feature of remuneration packages in the period leading up to the present crisis, though quantified evidence is difficult to come by. It is certainly true that a number of company CEOs received large pay packages during the expansion phase and were not required to give back their gains when their companies ran into difficulties. Indeed, in a number of cases, it was reported that they received additional large severance packages when they were eventually forced to leave their positions. What is not known, however, is what impact this had on incentives for risk-taking, or how much of the remuneration was in forms that imposed losses on the recipient when stock prices plummeted.

How can society protect itself against the systemic consequences of distorted incentives in compensation structures? It can be argued that the first line of defense is companies' own self-interest. It is hard to believe that financial firms knowingly encourage excessive risk-taking. They presumably try to set up incentive structures that limit perverse incentives. And competitive pressures should ensure that the firms that are most successful in doing this are the long-term survivors. To some extent this happens, but the mechanism is self evidently insufficient, for at least two reasons.

*First*, it is very difficult for the owners of firms (ie shareholders) to effectively monitor the incentives facing managers. They do not have the information or expertise. And in a protracted period of good times, it is far from clear how risky certain activities actually are. So the shareholders of a firm are insufficiently aware of the risks being run by senior management, and senior management in turn may be

insufficiently aware of the risks being run by the bankers and traders whose activities they oversee.

The *second* reason why it is not possible to rely simply on the self-interest of firms' owners is that the costs of an individual company's failure will understate the social costs, for reasons developed elsewhere in this report. The failure of one financial institution typically weakens the position of all others and brings the prospect of systemic meltdown closer.

For both of these reasons, we accept that there is an *a priori* case for supervisory intervention in the field of compensation practices. But what form should such intervention take? There is a clear risk of unintended consequences in subjecting compensation practices to regulation based on political pressure and public opinion. If the consequence is to constrain salaries below market clearing levels, then practices will spring up that evade the constraints. An example is the law introduced in the US in the 1990s that excluded salary payments in excess of $1 million per year from expenses deductible for tax purposes. The consequence was the development of incentive compensation techniques (such as payment in restricted stock and options) that rendered the intended salary restraint ineffective, and that had their own perversities.

So the question is what kinds of regulatory rules or guidelines would improve the incentive structure of executive pay in the financial sector. We consider here five possible techniques:

The *first* relates to the way in which bonuses are paid. It should be feasible to require delayed payment of bonuses, to mandate the greater use of payment in company stock, and to require the vesting of this stock for longer periods. We support these techniques, particularly for senior executives whose decisions affect the overall health of the company. But we caution that too much should not be expected of these reforms. Many, indeed most, financial companies, have followed such practices for quite a long time.

A *second* technique relates to the way in which bonuses are calculated. Employee incentive compensation should be tied strictly to *risk-adjusted* returns. Here too, however, it would be a mistake to expect too much. Most companies already attempt to measure risk-adjusted returns: the problem lies in the difficulties of calculating risk. Measured risk usually seems low at the height of a boom so that *actual* risk is underestimated. Moreover, even if traders are rewarded on a risk-adjusted basis, they may still have an incentive to take excessive risk. Systemic risk is something that materializes only infrequently. Traders and others will benefit from taking excessive risk in all the years when risk does not materialize, and will not suffer corresponding loss in years when major reverses occur.

A *third approach* attempts to equalize the balance of returns and losses through "claw-back" provisions in bonus awards. Bonus payments could be placed in an escrow account, and only released after a suitable period had passed, during which no losses had been recorded. If such losses did occur, the employees claims on the compensation balance held in escrow would be reduced accordingly. As economists, this appeals to us on incentive grounds, though we doubt its practicality

and even its legal enforceability. Since such provisions would obviously be unattractive to employees, the consequence would presumably be even higher bonuses to compensate for the attendant uncertainty.

A *fourth* approach is to reduce the incentive for risk taking by individual decision takers. For example, it has been suggested that bonuses should be based on firm-wide performance, rather than the performance of individuals or their business units. We have some doubts about the effectiveness and desirability of this. It would certainly reduce the incentive to take risks in individual activities. But it would largely undermine the purpose of incentive pay. Furthermore, it might appear inequitable in a multi-functional firm, if employees in a business unit that was successful had their pay reduced because of the less successful activities of another business unit. It may even lead to the emergence of smaller, more specialized firms that would do little to change systemically dangerous behaviour, but would make the financial system even harder to observe.

The *fifth* approach seeks to utilize the checks and balances of corporate governance to bring individual and corporate interests into better alignment. Internal compensation committees could be required to have risk control staff as ex officio members, and such staff could be given veto powers over compensation structures deemed to be too risky.

In what form should regulation over pay be introduced? We have elsewhere expressed a preference for "bright line" rules that limit supervisory discretion, on the grounds that it is hard for supervisors to impose discretionary rules in times of general euphoria. In the case of compensation, however, we see little alternative but to rely on a measure of supervisory discretion. We believe that supervisors should formulate a set of remuneration guidelines, based on the principles set out above, and assess the degree of compliance of each supervised institution. Using a relatively simple scale (eg. "fully compliant", "largely compliant" and "partially compliant", (which could be published) automatic adjustments would be made to capital ratios. Those institutions that were judged to have compensation practices that failed to restrain excessive risk taking would therefore pay a penalty in additional required capital.

### 6.3. Loan-to-Value Ratios in Mortgages

The epicentre of the financial crisis occurred in the housing market, especially, but not only, in the USA and UK. The boom/bust cycle was exacerbated by the conditions for mortgage lending becoming ever easier in the boom and tightening in the bust. This was particularly so for loan-to-value (LTV) ratios. Time and state-varying LTVs have been used successfully in countries such as Estonia and Hong Kong.

They can, however, be easily avoided in most circumstances by taking out second mortgages or routing the mortgage borrowing abroad. Such avoidance can be discouraged, as is done in several European countries, by limiting the right of a lender to repossess property to first liens recorded in a register kept within the country. So second secured mortgages can be prevented as well as first mortgages made abroad, but not recorded locally, since they would be, in effect, unsecured.

Resort to legal prohibition in this fashion is undesirable. Nevertheless the impetus to the housing cycle caused by the competitive ratching up of LTVs in the boom (often to 100+, e.g. Northern Rock), and their abrupt decline (often to around 75%) in the bust, has been large. The lesser evil may be for the Central Bank to set a maximum LTV, say 90%, and even to lower it should house price increases appear to be getting out of hand.

More generally, the US regulators are proposing the creation of a separate mortgage origination authority that would set standards for such origination. We believe this would be a valuable innovation, that could usefully be copied by other countries. Appropriate mortgage origination standards would curb abuses and limit the extent to which unqualified buyers were enabled to bid up housing prices on the basis of lax lending standards.

### 6.4. Credit Rating Agencies

Credit Rating Agencies have lost much of their reputation in this crisis, mainly as a result of giving high initial ratings to securitized mortgage-backed securities that after the event appear to have been wrong. They are widely accused of nefarious behaviour, notably via a conflict of interest, since they get paid by the issuers of securities (the sell side), who naturally want higher ratings. We tend to think that this particular criticism is exaggerated. Credit rating agencies have a franchise value that depends on objective opinions. This would be undermined if they were known to shade their assessment in order to gain business.

What is of greater concern is the conflict of interest that arises in the advisory business of CRAs. The advisory arms of CRAs help potential issuers structure offers in such a way as to gain a desired rating. Having advised an issuer on debt structure, it is hardly likely that the rating arm of the CRA would fail to grant the promised rating. We therefore favour the legal separation of ratings business from ratings advisory services. We also favour enhanced transparency about the way in which CRAs assess the creditworthiness of structured products. What we do not favour, however, is formal oversight of the ratings process. This would, we believe, tend to give too much of an official endorsement to ratings.

Rather than accusing CRAs of sharp practice, our view is rather that the CRAs failed to appreciate the likelihood that US housing prices might decline across the board, and the extent that probabilities of default could migrate upwards in that event. Most forecasters have had poor, chequered records over the last two years. CRAs are just another group of forecasters, and they have done just as badly.

Anyhow, whether the CRAs were knaves (conflict of interest) or fools (poor forecasters, like most other forecasters), the question is what to do about their role in the general conduct of regulation. Tightening up, yet further, on potential conflicts of interest,[46] and, yet more, transparency in both methods and results is all very well; the greater problem is that the ratings provided by (fallible) CRAs, using fallible models, have been placed at the centre of the regulatory process itself, for example the Basel II

[46] The worse conflict is in their consultancy and advisory activity.

Standardised Approach. And given the strategic behaviour that rating disclosures can generate, placing CRAs at the centre of regulation may have added to their fallibility.

In this latter regard there are two alternative generic approaches. The first is to remove CRAs and their ratings as far as possible from the structure of formal regulation altogether. Investment managers and bankers should take responsibility for their own decisions, and they (and their regulators) should no more be allowed to hide behind CRA forecasts than, for example, behind government forecasts of future growth.

The alternative, second, approach is to register and to regulate the CRAs, but to leave them with a central role in the regulatory process.

Our preference is for the first approach. Regulation will not make the CRAs forecast better; but will mean that the authorities will be conjoined in the resulting condemnation as and when the CRAs get it wrong in future, as they inevitably will. Moreover, governments are themselves large-scale issuers of debt. Might government regulation trespass on the (fragile) independence of CRAs when it comes to rating such debt?

### 6.5. Centralized Clearing House Arrangements vs. OTC Markets

Certain financial markets are systemic in the sense that their closure or malfunction would cause adverse externalities and contagion. As a general matter what is needed is a centralized clearing house (CCH) in any such systemic market, to lessen the risk that the failure, or anticipated failure, of a counter-party might cause widespread financial problems; the CCH should have the power to determine and to adjust the conditions for trading, e.g. margin requirements.[47] In particular, when a market grows to a size when it becomes systemic, such as the Credit Default Swap (CDS) market, the relevant authorities should have the powers to require an Over the Counter (OTC) bilateral and unregulated market to be reshaped into a centralized, regulated market. Reforming, and improving the regulation of, the market infrastructure of the financial system, centralized counter parties for systemic markets, improving the clearing, settlement and payment systems, remains a key element in the whole exercise of reconstructing the regulation of the financial system.

While there are substantial benefits in centralizing the clearing and settlement processes of systemically important markets, we are less convinced that all OTC contracts should be forced on to an exchange. There are legitimate institutional difficulties in a market organizing a centralized clearing facility, but there are many trading venues and exchanges and we have some respect for the revealed preference for some instruments to trade off an exchange and some on, especially where there are idiosyncratic contracts. We believe the mechanism of differential capital haircuts could be used to shift a large proportion of relevant transactions onto regulated exchanges, while allowing the flexibility of OTC contracts where these are economically justified.

[47] See Brunnermeier (2009), "Deciphering the Liquidity and Credit Crunch 2007-08" for a theoretical explanation how network effects in over the counter-market arrangements can lead to adverse amplifications.

### 6.6. Year-end Spikes

Finally, quite a lot of the pressure to maintain adequate liquidity concentrates on the presentation of an end-year balance sheet, with minor spikes at the half-year and end-quarter. Concerns about end-year positions typically cast their shadow forward into the late autumn for weeks, if not months ahead. Each institution wants to window-dress their published end-year figure to show a degree of liquidity that they do not feel the need for at other times.

That this should be allowed to disrupt, or threaten to disrupt, financial systems appears absurd. There are several potential remedies. The institutions could be required to report an average figure, (over the last quarter), taken at monthly or weekly observations. Alternatively, the authorities could just facilitate whatever window-dressing the institutions wanted by offering one-day repos on a massive scale. The resulting balance sheet would often hardly be ‘fair and true’, but the end-year spike is economically damaging, especially at a time of frayed nerves. Some solution needs to be found.

# Chapter 7: The Structure of Regulation

A sensible maxim in this field is to defer delegation of responsibility to an institution for achieving some objective until, and unless, one can also equip that institution with sufficient powers and instruments to achieve that end. There has been some tendency in the past to allocate responsibility for financial stability to Central Banks without due consideration of what instruments they might use to achieve that objective.

The one instrument that they could wield, the short term interest rate, has been predicated to the achievement of (goods and services) inflation. The recent period of financial turmoil has raised queries whether the primary target of price stability should be widened to include asset prices. While we do suggest that an appropriate measure of housing prices should be in the price index used for the inflation target, we do believe that, wherever possible, a separate objective should be achieved by a separate instrument (the Tinbergen principle of relating instruments to targets). Thus the objective of financial stability should be achieved by the development and application of instruments designed for that purpose.

It has been a primary purpose of this paper, in sections 3 to 5 to sketch out macro-prudential instruments that can be used in this way. We like to think that we have proposed a sufficient armoury. An important question is the balance between discretion and rules in their application. The more that the utilisation of such instruments is likely to provoke opposition from major interest groups at the time of their application, the more such application needs to be based on pre-set, pre-announced, (even statutory) rules. 'Taking away the punch-bowl, just when the party gets going' is no more popular with respect to asset price booms, than to the macro-economic conjuncture. For this latter reason we would advocate that much of the counter-cyclical armoury that we have suggested becomes couched in presumptive, rule-based terms.

Throughout we have emphasized the differences between macro-prudential and micro-prudential regulatory measures. Such differences extend naturally to the ethos, discipline and cultures of the institutions involved. The macro-prudential institution should be macro, aggregate, systemic and economic in outlook; the micro-prudential institution(s) should be more micro, individual, prudential, legal and accounting-based. This is closely in accord with the US Treasury, 'Blueprint for Modernized Financial Regulatory Structure', (March 2008). Naturally the macro institution will be the national Central Bank and the micro institution(s) will be one, or more, Financial Services Supervisory institutions.

It is, however, a mistake to channel all direct supervisory contact with the individual regulated financial intermediaries through the micro-prudential institution(s), as has been done in the UK. The macro-prudential body, in effect the Central Bank, needs to maintain direct links, including on-site supervision when required, with all those institutions designated as 'systemic' and also with those which the Central Bank suspects may be becoming systemic. We repeat, and support, the concluding principal 'observation' (observation 9) of the G-30 (2008) paper on 'The Structure of Financial Supervision' that "Irrespective of structural approach, central

banks everywhere express the critical importance of their having information about, and a direct relationship with, large systemically important financial institutions."

In a sense what we are recommending is a reversion to the prior twin-peaks approach, with one peak being the macro, systemic, economic Central Bank, and the other being the micro, individual, prudential (and conduct of business), legal and accounting FSA. When the UK went for a unified, single peak, approach, there was no discussion of its advantages and disadvantages vis-à-vis the twin peaks approach. There should have been. The decision then was wrong, despite its apparent economising on scarce supervisory resources and its limitation on vexatious supervisory visits to the regulated. Macro and micro-prudential regulators/supervisors have essentially different viewpoints, and both are valid.

Dealing with the question of the structure of financial regulation within the individual nation state, however, is much easier than trying to review and to reform the international structure. Here we start with two considerations, indeed facts, that limit the application of an international level-playing-field. First asset price cycles, and the pace of credit expansion, differ between countries. So, counter-cyclical measures have to be applied by host countries to the (systemically important) financial institutions in their own countries. The inevitable implication is that, even though the principles of application may/should be the same across countries, the effective capital ratios applied to banks will differ depending on where their assets and liabilities are situated.

In our view, the ability to apply counter-cyclical regulation (both to capital and liquidity) implies, as its corollary, a shift of the balance of powers towards the host country, away from the home country, and also some departure from the level-playing-field ideal. How serious a drawback this might be is a matter for discussion. The large cross-border financial intermediaries might be slightly inconvenienced[48] but does that matter that much if the purpose of the exercise is to tailor the regulatory countervailing pressures to the financial stability conditions within each country? This would, of course, have numerous structural implications, implying for example a less pressing need for 'colleges of regulators'.

The second consideration, fact, that we would note is that crisis management is often very expensive, and that the main source of such funding clearly has to be national Treasuries, and ultimately the taxpayer.[49] He who pays the piper, calls the tune. So long as the Minister of Finance, and national taxpayers, are ultimately at risk of needing to pay out money from the results of regulatory failure, they will want to design and control their own regulatory and supervisory procedures, and rightly so.

But what about Europe, where the objective is a single financial system? Crisis management can be hugely expensive, as has been seen. As soon as, but not

[48] Any potential additional burden on the regulated could, in principle, be offset by a redoubled effort to harmonize reporting requirements and definitions across countries.

[49] The only source of international funding for crisis management is the IMF. But their available resources are relatively small, dwarfed in size by the funds recently applied in developed countries to recapitalise their own banking systems. Calls to give the IMF greater responsibilities have not been matched by measures to give them greater resources or other instruments of control. Without the latter, they cannot realistically assume the former.

before, the federal centre gets the power to raise funds (ultimately via taxation) to undertake such crisis management, then supervision can be transferred to the federal centre, and the euro-zone run for regulatory and financial stability purposes as a single financial entity. Absent such a fiscal centralisation, which the authors of this paper would like to happen, the fiscal powers, crisis management, supervision and the protection of national financial systems remain at the nation state level, (as has been seen in practice in this crisis), and the design of regulation has to reflect this.[50]

Although crisis management has to be done at the nation state level, absent a shift of fiscal competence for this purpose to the supra-national region, there remain considerable useful opportunities for international cooperation in crisis prevention.

At present, international cooperation is carried on through a network of supervisory committees and through international organizations such as the BIS, the IMF and, particularly, the Financial Stability Forum (FSF). Sectoral supervisory committees have a history stretching back over thirty years. The Basel Committee on Banking Supervision was formed in 1974 under the aegis of the G10 central bank governors in the wake of the Herstatt crisis. Its original goal was to clarify the areas of responsibility of home and host supervisors where there was a failure of an internationally active bank. Subsequently, however, the Committee became the source of supervisory rule-making more generally. It is now best known for setting minimum capital standard for internationally active banks ("Basel I" and "Basel II"). Although negotiated only among G10 regulators and central banks, these soon became global standards.

International cooperation in the securities and insurance fields took longer to crystallize, for both political and historical reasons. There was no "crisis" in these sectors to force cooperation; there was no pre-existing body, such as the G10 central bank Governors' committee, to act as a convening authority; and the responsibility for insurance and securities regulation was generally more fragmented. Still, following the model of the banking regulators, cooperative committees of standard setters were set up by the International Organisation of Securities' Commissions (IOSCO) and the International Association of Insurance Supervisors (IAIS).

The need for a more formal international body to take overall responsibility for global financial stability, and to bring together the relevant national authorities, was recognized in the wake of the Asian crisis. The G7 ministers and governors established the FSF in 1999, comprising Central Bank Deputy Governors, Deputy Finance Ministers and heads of regulation from G7 countries, along with senior representatives of the main international institutions. Subsequently, participation was extended to a few countries outside the G7 with important financial markets. (The Chairman of the Forum is appointed in a personal capacity; one of the authors of this report served as the first Chairman).

[50] The recent CEPS Task Force Report (December 1, 2008), on 'Concrete Steps towards More Integrated Financial Oversight', proposes using the European Investment Bank (EIR) for this purpose (Section 3.3). While we welcome their appreciation of this issue, we fear that the potential scale of fiscal requirement, as evidenced in the current crisis, could be well beyond the EIB's financial capacities, even if it were to call upon all its additional capital resources. Moreover, this would lead the EIB towards becoming the main regulatory authority; would that be a welcome development?

Although the FSF has done much useful work, and established closer relationships among key regulatory authorities, it has suffered from several handicaps, which need to be addressed. We believe that, if these handicaps can be satisfactorily dealt with, the FSF could play a key role in a revised global regulatory environment.

A *first* problem with the FSF, as currently structured, is that it does not have formal representation from key emerging markets. This is partly because the original G7 wanted to retain their control over a process they believed affected primarily their own markets. It is also partly due to the desire to keep the number of participants manageable in order to promote confidential discussions within the Forum.. (Even with the current country membership, multiple participants from countries and IFIs means that there were some 35 people around the table.)

Initially, the FSF tried to balance the need for emerging market input and small-group discussions through regional meetings with key authorities in Asia, Latin America and Eastern Europe. This had some success but was eventually abandoned, partly because of pressure on key participants' time. We believe the time has now come to formally expand the membership of the FSF by adding participants from the main emerging markets. It seems that the political leadership of the G20, in its declaration following the November 2008 summit, accepts this logic. The difficult political task will be deciding exactly which countries should be invited to join.

A *second* institutional shortcoming of the current FSF lies in the multiple representation of the G7 countries (and to a lesser extent, of international organisations). The G7 countries have three representatives each, to accommodate the competing claims of finance ministries, central banks and regulatory agencies. Not only does this add to the numbers around the table and thus inhibit discussion, it dilutes responsibility. We believe the major economies should designate a single top-level participant from the agency with overarching responsibility for financial stability at the national level. This individual should preferably be the head of the agency and should become the sole representative of the country in the FSF. A similar limitation could be imposed on international organisations, such as the IMF and World Bank. If this were done, it should be possible to bring into the FSF all key emerging markets without expanding its size beyond what is consistent with frank and confidential round-table discussion.

A *third* problem for any international body is the tendency of national authorities and the general public to resist warnings of vulnerability during good times. The FSF issued several warnings in the period leading up to the 2007 outbreak of turbulence, but these were not heeded. (A similar experience befell warnings issued by the IMF and BIS.) It is to deal with this tendency that we have advocated "hard-wiring" counter-cyclical regulatory guidelines that we believe might be hard to implement through supervisory discretion.

Nobody can mandate that the advice of the Financial Stability Forum should be accepted by governments. But we believe techniques should be explored to underwrite the independence of judgments by the Forum and to ensure that its judgments are given adequate publicity.

A *fourth* problem that can be ameliorated but not completely solved is the fact that the legal basis of regulation is national, while financial institutions and markets global in scope. We do not believe that international legal authority, however desirable, is a practical political possibility in the foreseeable future. Nor do we think it is desirable to constrain the activities of financial institutions and markets within national boundaries. For this reason, there will inevitably be gaps and overlaps in regulation. Any practical approach to this problem must therefore focus on strengthening mechanisms of cooperation among independent national authorities.

Even more important than the internal governance mechanism of the FSF is what it should actually do. We see key functions in three areas: (i) crisis prevention, through ensuring stronger, more appropriate and more consistent prudential standards; (ii) crisis warning, through the monitoring and highlighting of systemic vulnerabilities, and (iii) crisis management, in cases where financial turmoil breaks out.

With regard to *crisis prevention*, there is no need to eliminate the current role of international standard setters in the areas of their individual responsibilities. In other words, the Basel Committee, perhaps with a revised membership, should continue to be the lead grouping for formulating cross-border banking standards; and similarly with IOSCO and the IAIS for the securities and insurance industries, and IASB for accounting standards. But the FSF should have the responsibility of opining on the consistency of proposed supervisory standards with overall systemic stability.

Elsewhere in this report, we have pointed out examples of where a focus on institutional (or microprudential) safety and soundness may not be sufficient to ensure system-wide (or macroprudential) stability. Indeed such a focus may even be counterproductive in this respect. We see one task of a revamped FSF as being to verify that rules proposed by standard setters are consistent with overall system-wide stability Subjecting the proposals of supervisory groupings to a "consistency check" by a body such as the FSF could provide a useful mechanism to make sure that regulation does not become inadvertently procyclical.

With regard to *crisis warning*, the role we envisage for the FSF is not greatly different from the one it plays at the moment. But we do see a need for modifications to improve the chances that warnings do not go unheeded. One would be to give co-responsibility for early warnings to the IMF, enhancing the authority of the resulting judgments, and bringing in the IMF's macroeconomic expertise. To safeguard against political pressure to pull punches, Early Warning Reports should be produced under the authority of the Chairman of the FSF and the Managing Director of the IMF, without editing by member country representatives. With regard to the FSF's recommendations, countries would be asked to "comply or explain" their responses.

*Crisis resolution* presents perhaps the greatest challenge. As already noted, crisis resolution usually involves fiscal resources and thus involves governments and parliaments. In the case of a financial crisis such as the one we are now experiencing, it is hard to see governments delegating this authority to an international organisation, or using domestic budgetary resources to help support the cross-border operations of foreign-based financial institutions. We recognise, therefore, that the best that can be

expected is to use cooperative mechanisms, and existing discussion fora to improve information exchange and to promote negotiations on optimal resolution strategies.

The significance of this activity should not be minimised, however. Where there are global ramifications from a particular course of action, countries represented in the FSF should commit to avoid measures harmful to the interests of others. When a country is contemplating a measure to provide systemic support to its financial system, (for example, a guarantee of bank deposits) it should accept an obligation to inform partner countries of its intentions in advance of their implementation, and to provide for an adequate interval to consider the international implications of such actions. None of this, of course, would guarantee a cooperative outcome. But it would at least ensure that the mechanisms existed to permit such a resolution, if the political will was there.

# Chapter 8 : Conclusions

The design of financial regulation is not straight-forward. When everyone is baying for more, tougher regulation, it is not needed, (because everyone is risk averse). When such regulation is badly needed, no one wants it, (since the good times are expected to roll on). This suggests that financial regulation should be focussed, primarily rule-based, (because discretion will be hard to use during periods of boom/euphoria), and time and state-varying (light during normal periods, increasing as systemic threats build up). The Spanish dynamic pre-provisioning scheme is about the only current instrument that meets these criteria.

Our analysis of the state of financial regulation leads on to quite a lengthy list of key points and recommendations which can be divided into four main headings, to wit General, Capital, Liquidity and Other.

**8.1. General Conclusions and Recommendations** (primarily from Chapters 1 and 2)

1. Regulation (external intervention) should always be capable of justification as a consequence of some specified market failure
2. The main cause of externalities arises because the social cost of systemic financial collapse exceeds the private cost to the individual financial institutions (and markets). A collapse of a financial institution causes risk spillovers. Effective regulation should provide incentives for financial institutions to internalize these externalities (risk spillovers).
3. The main cause of systemic collapse is endogenous risk, the likelihood of self-amplifying spirals like the loss and margin spiral.
4. Stress tests examine the responses of banks to exogenous risks. By construction they do not incorporate endogenous risk. Completely new techniques, perhaps based on models and endogenous risk-spillover measures, like CoVaR, need to be devised to explore the implications of endogenous risk for the system.
5. Requirements based on minimum capital ratios do not provide resilience, since they cannot be breached. They represent a tax, not a source of strength.
6. Requirements should, instead, be normally restated in terms of higher target levels of capital, with a specific, statutory (i.e. not discretionary) and forceful ladder of increasing sanctions. This ladder should have a minimum point at which either the deficiency is satisfactorily redressed, or the institution is shut down, i.e. prompt corrective action, p.c.a.
7. The response to the current crisis has led in several countries to a further concentration of the banking system and, perhaps, elsewhere amongst hedge funds. Greater intervention to encourage competition and to prevent oligopolistic behaviour may well be warranted.

**8.2. Capital Requirements** (primarily from Chapters 3 and 4)

1. Banks, and any other financial institution subject to deposit insurance, should be subject to some (low) minimum capital requirement. This is not to be seen as a protection for the regulated institution (rather the reverse; it constrains the banks), but as a protection for the deposit insurance fund, and a trigger for p.c.a.

2. All regulators/supervisors in each country should agree their own list of systemic institutions and markets, and be prepared to exchange lists with supervisors in other countries. Although such lists should not be made public, (o.a. moral hazard and the fuzzy definition of 'systemic'), there should be mechanisms for ensuring that regulators/supervisors take this exercise most seriously.

3. All such systemic institutions should be subject both to micro-prudential regulation, examining their individual risk characteristics (along the lines of Basel II) and to macro-prudential regulation, related to their contribution to systemic risk. We suggest that this latter be done by adjusting the micro-prudential ratio by a co-efficient relating to the macro-prudential assessed risk.

4. Macro-prudential regulation should be countercyclical and lean especially against bubbles whose bursting can impair the financial intermediation sector.

5. We argue that the best measures of an institution's contribution to macro-prudential risk are its leverage, maturity mismatch and rate of expansion. More precise endogenous risk-spillover measures that also take liquidity aspects into account should be developed. So we would interact each, systemic, institution's tier 1 Basel II ratio by multiplicand, which could be below, as well as above, unity, based on a mixture of leverage, maturity mismatch and growth.

6. Institutions which are not individually systemic, but which are (i) highly leveraged with short-term debt and (ii) hold assets with low market liquidity (at times of a crisis), can nevertheless have systemic effects via joint herd-type behaviour. So they should all, except for the tinies, both report, and have some constraints (in the form again of a ladder of sanctions), on their macro-prudential riskiness, i.e. their leverage, maturity mismatch and credit expansion, (which could perhaps vary between kinds of institutions, e.g. banks and hedge funds).

It would be for discussion, (and our group could not agree), whether such highly-leveraged, but individually non-systemic, institutions should also have any additional micro-prudential regulation.

7. Asset-price and credit cycles differ from country to country, and from region to region. Although the principles of counter-cyclical regulation should be universal, its application would lead to differing ratios in each area applying the regulations, normally in the host country.

8. Each host country (region) should have the right to designate a cross-border subsidiary, or branch, as 'systemic'. Systemic branches should be required to become subsidiaries. Foreign-owned subsidiaries should be subject to the same capital requirement calculations, and hold that in domestic assets, as its own domestic banks.

9. An alternative approach, which might be less radical, would be to generalise and to extend the present Spanish dynamic pre-provisioning scheme to all countries, though this also would need to be applied on a country-by-country basis. If this were to be done, IFRS would have to be revised to permit this.

10. The application of macro-prudential measures should be by the Central Bank; for this purpose they should be able to undertake (on-site) supervision of individual systemic institutions, separately from the micro-prudential supervisor(s). Efforts should be made to limit the administrative burden of multiple supervisors, and reporting requirements and definitions should be harmonised.

**8.3. Liquidity** (primarily from Chapter 5)

1. We propose a 'mark to funding' approach to provide incentives for more long-term funding. This approach is, in effect, closely akin to the maturity mis-match ladder previously considered by the Basel Committee on Banking Supervision (BCBS) and by some Central Banks.

2. Unlike most such prior exercises we would also provide incentives to hold liquidity by, once again, interacting the assessed liquidity with the capital adequacy ratio. Liquidity is measured by an effective maturity mismatch which takes the market liquidity of assets (at times of crisis) into account. The more liquidity fell below the well-targeted level, the higher the CAR would have to be, and vice versa. The relationship (trade-off) need not, however, be linear.

3. We doubt whether additional private insurance can then help much on occasions when market and funding liquidity vanishes; the examples of the mono-lines and of AIG confirm our doubts. The answer would seem to be some combination of public sector market-making, (as now by the Fed in the CP market), and public sector insurance, (guarantees of one kind or another).

4. We suggest that mark-to-funding might be a principle that could apply to the portfolios of financial institutions for accounting purposes, either as an alternative, or a supplement, to the present categories, i.e. hold to maturity, available for sale, trading book.

**8.4. Other Considerations** (primarily from Chapters 6 and 7)

1. We propose that supervisors should formulate a set of remuneration guidelines, and, (as in other examples below), adjust capital ratios according to the degree of compliance.

2. We advocate the Central Bank setting maximum Loan-to-Value (LTV) ratios for residential mortgages as an additional macro-prudential measure. This would involve, as a corollary, outlawing several obvious avoidance measures, e.g. second mortgages.

3. We argue, at several points, that credit ratings are systematically misused in the regulatory process. Whereas we are happy to see further tightening of 'conflict-of-interest' and transparency regulations, we would otherwise seek to exclude CROs from the regulatory network

altogether. We regard both the Basel II approach to the use of credit ratings and the European proposals for their enhanced regulation as misconceived.

4. We support the efforts of the CRMPG to move systemically important derivative markets onto centralised clearing houses.

5. We cannot understand how, and why, the end-year spike in financial markets has been allowed to remain. It is both absurd and damaging. We suggest two alternative methods of eradicating it.

6. Because cycles (in asset prices and credit) vary from country to country, as well as from time to time, we propose a shift of emphasis in regulatory powers towards the host country.

7. Because crisis management is often extremely expensive, it has to be done by the (host) Central Bank in conjunction with its own Ministry of Finance. As soon as, but not before, the Eurozone obtains fiscal powers to manage any such crises, then macro-prudential management can be shifted from the National Central Banks to some federal euro-zone body.

8. Whereas crisis management has to be done at a (national) level consonant with the availability of fiscal (taxpayer) funding, crisis prevention can, and should, be done internationally. We make several proposals to reform both the structure and remit of the Financial Stability Forum.

We have put forward nearly 30 key points and recommendations. If adopted, they would change the present system radically and for the better.

## Appendix : The Boundary Problem in Financial Regulation

There are a number of fundamental, generic issues relating, at all times and everywhere (almost) to financial regulation. In particular if regulation is effective, it will constrain the regulated from achieving their preferred, unrestricted, position, often[51] by lowering their profitability and their return on capital. So the returns achievable within the regulated sector are likely to fall relative to those available on substitutes outside. There will be a switch of business from the regulated to the non-regulated sector. In order to protect their own businesses, those in the regulated sector will seek to open up connected operations in the non-regulated sector, in order to catch the better opportunities there. The example of commercial banks setting up associated conduits, SIVs and hedge funds in the last credit bubble is a case in point.

But this condition is quite general. One of the more common proposals, at least in the past,[52] for dealing with the various problems of financial regulation has been to try to limit deposit insurance and the safety net to a set of 'narrow banks', which would be constrained to hold only liquid and 'safe' assets. The idea is that this would provide safe deposits for the orphans and widows. Moreover, these narrow banks would run a clearing-house and keep the payments' system in operation, whatever happened elsewhere. For all other financial institutions outside the narrow banking system, it would be a case of 'caveat emptor'. They should be allowed to fail, without official support or taxpayer recapitalisation.

In fact, in the UK something akin to a narrow banking system was put in place in the 19th century with the Post Office Savings Bank and the Trustee Savings Bank. But the idea that the official safety net should have been restricted to POSB and TSB was never seriously entertained. Nor could it have been. When a 'narrow bank' is constrained to holding liquid, safe assets, it is simultaneously prevented from earning higher returns, and thus from offering as high interest rates, or other valuable services, (such as overdrafts), to its depositors. Nor could the authorities in good conscience prevent the broader banks from setting up their own clearing house. Thus the banking system outside the narrow banks would grow much faster under normal circumstances; it would provide most of the credit to the private sector, and participate in the key clearing and settlement processes in the economy.[53]

This might be prevented by law, taking legal steps to prohibit broader banks from providing means of payment or establishing clearing and settlement systems of their own. There are, at least, four problems with such a move. First, it runs afoul of political economy considerations. As soon as a significant body of voters has an

---

[51] Though not so in every case. In some cases regulation may raise profitability for good reasons (increasing demand as a result of enhanced confidence in the quality of the product) or for bad (creating barriers to entry and restricting competition). I am grateful to Gavin Bingham and to Eva Hupkes for this and other helpful comments.

[52] Though it keeps on being revisited; Telser, L.G. (2008) is a recent example.

[53] This does not rule out any role for quasi-public utilities in the financial system. Public sector narrow banks, like POSB in the UK or Postfinance in Switzerland, can continue to provide useful services, especially if their services are priced appropriately. Also there can often be a role for a quasi-public sector utility in financial market infrastructures.

interest in the preservation of a class of financial intermediaries, they will demand, and receive, protection. Witness money market funds and "breaking the buck" in the USA. Second, it is intrinsically illiberal. Third, it is often possible to get around such legal constraints, e.g. by having the broad bank pass all payment orders through an associated narrow bank. Fourth, the reasons for the authorities' concern with financial intermediaries, for better or worse, go well beyond insuring the maintenance of the basic payment system and the protection of small depositors. Neither Bear Stearns nor Fannie Mae had small depositors, or played an integral role in the basic payment system. Nevertheless, as has already been discussed in Section 5, with particular respect to time-varying loan to value ratios, there may in some instances be an argument for using legal prohibitions to help police the boundary between regulated and unregulated functions.

When a financial crisis does occur, it, usually, first attacks the unprotected sector, as occurred with SIVs and conduits in 2007. But the existence of the differential between the protected and unprotected sector then has the capacity to make the crisis worse. When panic and extreme risk aversion take hold, the depositors in, and creditors to, the unprotected, or weaker, sector seek to withdraw their funds, and place these in the protected, or stronger, sector, thereby redoubling the pressures on the weak and unprotected sectors, who are then forced into fire sales of assets, etc. The combination of a boundary between the protected and the unprotected, with greater constraints on the business of the regulated sector, almost guarantees a cycle of flows into the unregulated part of the system during cyclical expansions with sudden and dislocating reversals during crises.

Exactly the same arguments can be deployed against the proposals that we have made, in Section 3, for time-varying capital requirements.[54] And much the same criticism can also be applied to other proposals, such as the reversion to the use of a leverage ratio for capital adequacy requirements, that might limit credit expansion and leverage in the boom.

What we observe in this latest financial cycle has been, first, a huge expansion of credit, a massive rise in leveraging during the upswing, followed by the crisis, curtailment of credit expansion and major deleveraging with severe, and continuing, effects on the real economy. For reasons which are by now widely understood, the present regulatory system (comprising Basel II and the move to mark-to-market accounting practices) not only did too little to restrain the upswing, but is also exacerbating the downturn. In other words it is highly procyclical.

So the obvious answer would seem to be to switch to a system which restrains credit expansion and excess leverage in the upturn, and relaxes such regulatory requirements when managers are themselves more risk averse and cautious in the

[54] And in op-ed articles in the Financial Times, 'A proposal how to avoid the next crash', January 30, 2008, and 'A party pooper's guide to financial stability', June 4, 2008, by Goodhart and Persaud. Thus in the former paper,

> "We propose that bank capital requirements should not only be contra-cyclical but also related to the rate of change of bank lending and asset prices in the relevant sectors. The capital adequacy requirement on mortgage lending could be linked to the rise in both mortgage lending and housing prices, and lending to construction and property companies to the rise in such lending and in commercial property prices."

downturn. What regulation needs to do is to counter the natural proclivities of managers, (by the appropriate adjustment of incentives, sanctions and trade-offs), not to try to mimic them, (as in the boast that Basel II sets regulatory capital closer to the economic capital desired by bank managers). This is the rationale for having some mechanism, whether time-varying risk-weighted CARs or a leverage ratio, or any other such, that restrains the regulated from such credit expansion in an upswing. It is, at least, arguable that raising capital ratios in asset booms should not really be seen as a 'burden', since it actually corresponds to an underlying increase in latent risk. A really far-sighted manager would do the same, but the pressures of herd mentality, competition for market share, etc., make it just too difficult for most managers to sit out the dance.

Suppose that such counter-cyclical adjustment can be done successfully.[55] Capital requirements are ratcheted up enough in good times to prevent the regulated expanding as much as they otherwise would. The result, as with narrow banks, would be to lessen the profitability and returns on the regulated, relatively to the unregulated at such times. There is sure to be, or to develop, a boundary problem. During good times funds will flow from the regulated to the unregulated, and the regulated will seek to find ways of transferring business to unregulated associates. During crises the flow will reverse, likely with serious adverse consequences. Our own proposals are just as subject to this generic boundary problem as any other. If financial regulation is effective, it will have to face the boundary problem.

Since the problem is caused by boundaries between the more and the less regulated, one extreme solution might be to regulate either no financial institution, or all of them alike. Both proposals have some adherents, with some advocating 'free banking'[56], constrained by market discipline alone, and others a completely controlled financial system, as for example practiced in most countries until the 1960s. Neither extreme would seem feasible; 'free banking' would, we believe, lead to results that would be politically and socially unacceptable, whereas totally controlled financial intermediation is inconsistent with a free market capitalist economy, especially so if exchange controls on capital flows would be needed to prevent disintermediation abroad. So if the extremes are eliminated, financial regulators and supervisors will always operate in an interior space, in which there is certain to be a 'boundary problem'.

The unregulated, however, often depend on services, e.g. payment and administrative, and on back-up lines of credit from the regulated. Indeed, the unregulated are frequently associates, or off-shoots, to the regulated. So cannot one

[55] The Central Bank that has tried hardest to introduce counter-cyclical measures is the Banco de España, notably with its time-varying requirements for loan loss provisioning, see de Lis, et al (2000). Unfortunately the application of the latest international accounting rules means that this measure may now have to be abandoned, or at least completely recast. Nevertheless, despite having a construction boom far larger proportionality than in the USA or the UK, the banking system in Spain has, so far, been relatively unaffected by the international financial turmoil.

[56] Even here a boundary problem of a kind remains. Cash is guaranteed against default; under 'free banking' deposits are not. So in good times people place their money in bank deposits, withdrawing back to cash in crises, thereby exacerbating the crisis. Without this boundary problem, 'free banking' just might have been viable.

maintain the boundary without excessive difficulty by some combination of prohibition on the regulated maintaining associated unregulated entities, and of limitations on the regulated's ability to provide credit (even on a contingent basis), (and services?) to the unregulated? We have already noted some of the main arguments against legal prohibition. Moreover, how could one prevent a foreign bank providing such services? Even if one could, and wanted to, draw a strict dividing line between, say, regulated banks and unregulated hedge funds, would it not be possible for hedge funds jointly to establish a separate central institution to provide them with quasi-banking services, including the provision of credit? If the unregulated become, as a result of regulation, more profitable than the regulated, over a long enough run of years for this to become publicly apparent, the unregulated will, one way or another, always be able to attract enough funding for extra expansion, however severely the dividing line between the regulated and the unregulated may be drawn.

Recognition that such a problem is generic may help to mitigate it. Many regulators/supervisors appeared to have been taken largely unawares by banks' reliance on associated SIVs, etc., in 2007. They should not have been. Any new regulation, such as Basel I, Basel II, or time-varying CARs such as here, will bring with it new boundary problems. Any supervisor must try to learn how the regulated are seeking to avoid the constraints placed upon them, (and if the regulated are not doing so, it may be an indication that the regulation is just ineffective!).

Although boundary problems are a generic consequence of effective financial regulation, it does not mean that all such regulation is a waste of time, nor that such problems cannot be mitigated by sensible design. We turn next to some proposals for setting the boundary in a manner that can help to lessen such problems.

So what should regulators/supervisors do in this respect? They should start by trying to list the key financial markets and systems in their own country. Having done so, they should review whether and which financial institutions are so important to the functioning of that market, or system, that their downfall, whether in the form of bankruptcy or major deleveraging, would seriously disrupt the operations of that market or system. Having done so, they could give the financial institutions involved a choice, either to reduce their exposure to this market (or system), or to be regulated. For example, any hedge fund with a total size beyond some limit, or involvement in any market beyond some scale, would be more closely supervised; otherwise not. Most would avoid any such supervision. But that would be all to the good. The aim would be to keep funds small and diversified enough so that they can be allowed to fail.[57] This, however, ignores the problem of herd behaviour among small, and medium, sized financial institutions. A way of dealing with this was discussed in Section 3.

[57] This solution is similar to the one proposed in Hüpkes (2004), 'Protect functions, not institutions'. In addition, that article suggests that some critical functions could be performed by quasi-public utilities (e.g. CLS bank). Note that the US does have a limit on the market share of banks (5% of the national market), plus additional limits for business in individual states. One consequence of a rule that requires an activity to be regulated once a certain size is reached is a clustering of activities just below the ceiling. This suggests that it serves as a boundary. Such rules exist for some trading activities.

In essence the financial supervisors have got to ask themselves, which financial institutions can be allowed to fail, and which cannot. Those that they claim cannot be allowed to fail, should be specifically regulated. The criteria for regulation should be made public. Any institution which is regulated as too important to fail, should be allowed to appeal against that ruling, and should also have the option to avoid regulation by downsizing. Besides occasions of institutional downfall, regulators need to be concerned with such market failures as may lead to resource misallocations, e.g. in the guise of asset bubbles and busts.

What difference would this approach make? Probably not much. The few remaining large US investment houses have already come under the Fed's umbrella, but it is equally important that the myriad of small broker/dealers do not get lumbered with unnecessary regulation. Similarly, supervisors need to assess when hedge funds, and other financial institutions, e.g. monoline insurers or private equity funds, become so large and prominent in certain key markets that their failure could completely disrupt the functioning of those markets.[58] A further problem arising from the activities of hedge funds, private equity, prop trading, etc., is the 'crowded trade'. Apart from more required transparency, it is hard to see how this can be avoided.

This could, perhaps, be as market-related regulation. Moreover as markets change and develop, so should the Boundary change. Also note that the Boundary does depend on the estimated effects of failure. Only those institutions so big and connected that their failure would dislocate the key financial markets would be regulated. There is no case for regulating all broker/dealers or hedge funds, only a handful of those that are so large that their failure would disrupt the financial system.

A major problem is that the more effective regulation becomes, the more unpopular it will be, since it will prevent the regulated from doing what they want to do.[59] The Boundary problem will worsen such unpopularity. It leads to the following claims; that such regulation is:-

a) Ineffective and unfair, resulting in disintermediation;
b) Inefficient and cost enhancing;
c) Complex and capable of being subverted.

Let us take each charge in turn. If there is a Boundary problem, (and regulation within the Boundary is effective), then, almost by definition, there can be no level-playing-field. The unregulated outside the Boundary have a stronger competitive position than those within. Not only could this be described as unfair, but business will be bound to flow from the regulated to the unregulated, in other words disintermediation will occur. Moreover, there will be less information about the unregulated, and their risk management may be even worse. So the resulting financial booms and crises could even be enhanced. What the regulators will have done is to

[58] Although credit ratings agencies have played an influential role, the failure of one of the big ones would be a nuisance, but would <u>not</u> completely disrupt financial markets. There is no case for their regulation under this criterion.

[59] For such reasons Basel II was rather popular with the large international banks. During booms when these banks wanted to expand, Basel II provided no constraint. During the latest financial crisis, when a combination of panic, market forces and self-preservation has been causing banks to cut back on lending and to delever anyhow they can then blame the regulators for their restrictive policies.

take the business away from the regulated (the good guys in white hats) and given it to the unregulated (the bad guys in black hats).

And all that is true up to a point. The point is that the aim of the exercise is to prevent the key financial institutions from overstretching themselves, and so failing, rather than preventing any financial institution from doing the business, if it thinks it profitable. During property boom and bubbles a local national regulator ought to be thankful if lending into such a boom does become diverted elsewhere.[60] An example was when Canary Wharf, the large London city office project, was financed by foreign, not British banks; the British bank regulators felt relief.

An example of the difficulty of the Boundary problem is the British fringe bank crisis in 1973/74. Because of prior constraints (only partially regulatory), much property finance was then done by the 'fringe banks' outside the framework of controls, and financed in the wholesale market. When the British property market collapsed in Autumn 1973, so did the fringe banks. The Bank organised a 'Lifeboat' to save the better elements of the Fringe; having discovered that it was felt in the event necessary to rescue these, the logical next step was to extend the Boundary to cover all banks, as done in the 1979 Banking Act.

So the first problem with effective regulation is that it will induce an unlevel-playing-field, which is unfair, and will cause disintermediation, which will negate some of the purpose of the exercise. And these criticisms are correct up to a point. Ways of dealing with it include trying to arrange regulation so that its effects only bite some of the time, when additional restraint is really needed, so that the costs and benefits to the regulated are not too far out of line,[61] and trying to limit the potential extent of disintermediation.[62]

The next criticism of effective regulation is that it will often be inefficient and lead to higher costs. The financial intermediaries within the boundary are often the most efficient. If their costs are raised, e.g. by higher CARs, then they will have to respond by raising the spread between interest rates on liabilities and on assets. The interest rates charged to borrowers will rise. When regulation is really needed, in asset bubbles, the outlook is generally optimistic. Everything looks good. As Alan Greenspan noted, no one can easily distinguish between an unsustainable asset bubble

---

[60] A major problem with the strategy of originate to distribute was that the distribution was often phony, originate and pretend to distribute to associated conduits, SIVs, etc., which were often artificially beyond the boundary, but where the risk and balance sheet burden flowed back to the main bank as soon as the market turned sour. While forcing all banks to retain some residual proportion of securitised products may well be desirable, in order to encourage properly diligent monitoring, the banks that got into worst trouble with CDOs and RMBS were those that retained, or were forced to take back onto their books, too much of such products.

[61] For example the provision of deposit insurance to bank depositors should allow banks to obtain retail funding more cheaply.

[62] A key component of time-varying regulatory controls could be the imposition of time-varying upper limits on loan to value ratios for residential mortgages. Such limits can be easily avoided by having a market for second mortgages, or by booking such mortgages abroad. But this could be deterred by making residential mortgage debts only legally recoverable if financed by a first mortgage issued by a bank sited in the country, i.e. including subsidiaries but not branches of foreign banks.

and a beneficial change in fundamentals. A regulatory initiative that has the effect of artificially raising interest rates, or tightening other borrowing conditions, e.g. LTVs, to borrowers at the height of the boom will be extremely unpopular to borrowers, banks and politicians. Moreover, in so far as the regulation succeeds in averting a future bust, it may then also have appeared to have been unnecessary!

It takes a lot of courage to take away the punch-bowl just as the party gets going. Even if regulators had sufficient instruments, (which they do not now have), to restrain cycles in credit expansion and asset prices, would they have the courage to use them, in the face of uncertainty in need and probable vilification in practice? One partial answer to such a 'time inconsistency' problem is to put more reliance on procedural rules, i.e. to state publicly in advance that regulation will be tightened in certain specified conditions, (e.g. when housing prices, according to index X, rise at an annual rate faster than Y; when overall bank credit in the country grows faster than Z; when lending by bank I grows faster than annual rate J, etc., etc.). Preferably there should be a ladder of responses, not a single trigger point. The FDIC Improvement Act of 1991 in the USA is an example of a proper regulatory procedure. We shall outline our own proposals in this respect in Section 4, the next Section.

The more effective regulation is, the greater the incentive to find ways around it. With time and considerable money at stake, those within the regulatory boundary will find ways around any new regulation. The obvious danger is that the resultant dialectic between the regulator and the regulated will lead to increasing complexity, as the regulated find loop-holes which the regulators then move (slowly) to close. Basel I metamorphosis into Basel II. So the process becomes ever more complex, almost certainly without becoming less porous.

How can one halt the onward march of this dialectic? This is not an easy task. One approach, as already noted, is to limit the periods in which regulation is effectively biting to those few in which it is essential, so that the overall costs, and hence the incentive to avoid, such regulation is lessened. Another, and perhaps more important, solution is to place the boundary at a point where flows across the boundary, substitution between claims on intermediaries within, and without, the boundary are likely to be relatively low. As described earlier, it is such flows that cause the main problems.

A key issue here relates to hedge funds. The aim should be to leave such funds outside the regulatory net, unless they become so large (or so connected with a key market) that their failure would be systemically catastrophic. But if ordinary people should begin to switch en masse between hedge funds and bank deposits, that would no longer be feasible. The authorities should require that all hedge funds operating in their own country impose high minimum limits on inward investment, or have available lock-up conditions on invested funds (so that outflows during crises can be constrained). It is bad enough that pension funds are already investing in hedge funds. The need is to insure that hedge funds will continue to be allowed to fail without public support.[63]

[63] Maybe the need is not so much to regulate hedge funds, but to limit the extent to which pension funds and life insurance companies can invest in them.

Moreover, should larger banks be more toughly regulated than smaller banks, or non-bank intermediaries, such as money market funds, this too would lead to boundary problems, with outwards flows to the less regulated in good times offset by a reverse rush during panics.

The other main issue is the incentive for intermediaries caught within the regulatory boundary to establish associated entities outside, to which business can be transferred. This is an obvious response for the regulated. So it was surprising, at least in retrospect, that regulators/supervisors appeared to have been often less than fully apprised, in 2006/7, of the development, and implications, of the chain of associate entities that banks had set up for this purpose. Some of these entities were legally separate, but remained reputationally connected. In that case how far will, or can, the bank within the regulatory system allow its, legally separate, associate outside to fail? If the answer should be that it may not feel able to do so, then the risks have <u>not</u> really been transferred off the balance sheet.

These problems of setting, and policing, the regulatory boundary are real and severe. There are no easy answers. But perhaps the first step towards resolving such problems effectively is to be aware of them. A guiding principle would be to design the interface between the regulated and the unregulated in such a way that the resulting incentive to shift business into unregulated channels, because of regulation, was so low that it never became systemic. Perhaps one conclusion from this is that regulation should be designed only to bite occasionally. If so, the time when it should bite is, surely, during periods of optimism, risk-seeking and rapid credit expansion, rather than at present when regulation tends to bite hardest just when the regulated are in any case most risk averse.