*Paul Wright**

THE CONCEPT OF TOO BIG TO FAIL: ISSUES FOR CONSIDERATION[1]

INTRODUCTION

A great deal of international debate is currently focused on the treatment of firms that are judged to be systemically important. The recent crisis has reminded us that large and complex firms can fail and that such failures can have very serious consequences for financial stability.

Because the consequences of failure can be so serious, there has been a temptation in the past for governments to rescue such firms, typically using public money to take capital stakes in them. But this cannot be a sustainable policy in the long term. In the first place it is prohibitively expensive. The Bank of England has estimated that official support to the banking systems of the UK and US in the period 2007–09 accounted for around three quarters of GDP[2]. These are staggering figures even if (as has already been the case to a considerable degree) much of the cost is eventually recovered. Closely linked to this is the fact that such rescues are highly unpopular politically. And less visible, but no less important, is the pernicious effect that the expectation of such actions has on markets and market discipline. If creditors and market counterparties believe that the firm they

* Senior Director, Institute of International Finance, Washington DC.

1 This article is an expanded version of a presentation made at a conference held by the Bank Guarantee Fund in Warsaw on May 21 2010. The views expressed should be taken as those of the author and not necessarily those of the IIF.

2 Bank of England Financial Stability Report, June 2009.

are dealing with will be rescued in the event of failure, they will not undertake proper due diligence and they will not price risk correctly. That creates moral hazard which stops markets from working as they should and is highly corrosive.

Policy makers are therefore faced with a dilemma. Firms will fail and exit the market from time to time. They should not be rescued using public money. But the systemic consequences of disorderly failure are often judged to be unacceptable. How can policy makers respond to this?

PRELIMINARY CONSIDERATIONS

Before turning to possible solutions, it is worth setting out some preliminary considerations which need to be borne in mind as we face up to this dilemma.

The first important point is that systemic risk – which lies at the heart of this debate – is a very elusive concept that is hard to pin down. One official report recently described it as follows:

'A risk of disruption to financial services that is (i) caused by an impairment of all or parts of the financial system and (ii) has the potential to have serious negative consequences for the real economy. Fundamental to the definition is the notion of negative externalities from a disruption or failure in a financial institution, market or instrument. All types of financial intermediaries, markets and infrastructure can potentially be systemically important to some degree'[3].

This definition highlights a number of important features of systemic risk.

- ❖ It is varied in form. It can arise from the failure of a single large bank (or non-bank financial institution); the simultaneous failure of a number of small banks or the failure of key element of market wide infrastructure. One particularly lively debate which is currently under way is whether insurance firms can be a source of systemic risk[4].
- ❖ It can mutate and spread quickly. It is noteworthy for example that the recent crisis which had as one of its principal origins retail mortgage underwriting practices in the banking sector, necessitated liquidity and de facto solvency support to be extended to non-bank intermediaries, especially in the US[5].

---

3 Guidance to Assess the Systemic Importance of Financial Institutions, Markets and Instruments: Initial Considerations. Joint FSB/IMF/BIS Report, October 2009.

4 The consensus, to the extent that there is one, suggests that typically they are not. See for example the report of the Geneva Association 'Systemic Risk in Insurance: An analysis of Insurance and Financial Stability', March 2010.

5 See comments by Paul Tucker, Deputy Governor of the Bank of England 'Shadow Banking, Capital Markets and Financial Stability' remarks at the BGC Partners Seminar, London, January 2010.

- Systemic risk is highly time and condition dependent. History provides multiple examples of firms and groups of firms which would not in any normal circumstances be regarded as systemic but which have been treated as such in circumstances in which confidence is particularly fragile, usually at times of more general systemic tensions.

Because systemic risk is such a complex and elusive phenomenon, the policy response to it needs to be carefully thought through. One clear overall message emerges from this, namely that it is very misleading to take a simplistic or uni-dimensional view of the issue. Some recent debate has fallen into this trap – in particular a tendency to equate size or geographic scope with systemic risk[6].

This leads on to the second point, which is about the need for balance in this debate. A great deal of emphasis has (understandably) been placed on the risks attaching to large, complex and global firms. But there has been a tendency to lose sight of the other side of the equation – the benefits that such firms bring to the global economy. There is great superficial appeal to the idea that the failure of large/complex/global firms has catastrophic consequences and that the only realistic response is therefore to force such firms to be less large, complex or global in the first place – that is break them up or restrict their activities. This argument is flawed on a number of levels. But even if it conceded that such firms *may* be harder to manage and supervise and that their failure *may* (other things equal) pose particular systemic challenges, the simplistic solution of 'breaking them up' or requiring global groups to operate through tightly ring fenced national structures would impose enormous costs on the global economy.

A recent study by the IIF drew on a wide array of case studies to demonstrate that large, globally active firms provide services to the globally economy that would simply be unavailable if they were replaced by firms engaged in a more limited range of activities with more limited geographic reach[7]. Such firms match savings and investments on a global level; they support the banking needs of international corporations; they facilitate the growth of regional companies; they make markets on a global scale and they provide payments services across the globe. They also spread good practice and expertise which has been developed in one market to others, to the benefit of their customers and they act as drivers of economic and financial development in emerging markets. It is obviously the case that such companies pose particular challenges in the areas of management, supervision and (as we will see) resolution and it is imperative that considerable effort and resources are invested in doing better in these areas. But simplistic approaches to

---

6 For a much fuller discussion of these issues see 'Systemic Risk and Systemically Important Firms: an Integrated Approach', Institute of International Finance, May 2010.

7 Institute of International Finance, May 2010 (op. cit.).

making them 'less systemic' would impose very great costs on the global economy and would create complex unintended outcomes.[8]

We can be reasonably confident that policy makers will avoid the trap of taking an over simplistic view of the problem and its solution. The remainder of this article looks at what is being done to address the problem of 'too big to fail' and to hint at what more needs to be done in the future.

REDUCING THE PROBABILITY AND IMPACT OF FAILURE

Two concepts which underline much conventional risk management are the *probability* of failure and the *impact* of failure. In other words, how likely is it that an undesirable event will occur? And if it does, how much will it matter? An approach based on this simple but powerful distinction underlies much of the development of regulatory policy that is currently under way.

Financial regulation typically seeks to reduce the likelihood that firms will fail to some acceptable level. It does not aim to eliminate the possibility of their doing so – that would be impossible and it would undermine essential market mechanisms. The most basic function of regulators is to ensure that firms have enough financial resources – capital and liquidity – to enable them to support the risks inherent in their business. The Basel Committee on Banking Supervision is currently examining ways of strengthening these requirements. The details are complex but the objective – of making firms more resilient – is clear and one which the industry generally supports.

One question which has arisen in this context is whether systemically important firms should hold more capital and liquidity in relation to the size of their balance sheets than other, non-systemic ones. The rationale for this can be expressed in a number of ways. Perhaps the simplest would be that the failure of these firms results in disproportionate costs (for society as well as their owners) so the likelihood of their failing should accordingly be made very remote[9]. One variant on this argument is that systemically important firms might be required to hold 'contingent' capital. This term has been used to mean different things but in this context it means that firms would be required to issue debt instruments which would convert into capital on the basis of some pre-determined trigger if the firm

---

[8] It is very tempting here to quote the American journalist and sage H.L. Mencken who famously said that 'for every complex problem there is an answer that is clear, simple …. and wrong'.

[9] The economic version of this argument is that the failure of systemic firms creates externalities which the firm, in the normal course of its operations, has no incentive to recognize or internalize.

became distressed. This would provide additional financial resources which would help it recover[10].

At the time of writing, the official sector has still to come up with firm proposals on these issues[11]. Until they do, it is hard to know how the industry will (or should) react. Some points however are clear even at this stage. First of all it is quite unrealistic to imagine that it would be possible to create a definitive list of firms that are 'systemic' – and hence eligible for some kind of capital surcharge. As discussed above, systemic risk is a complex and elusive phenomenon and any such list would be highly misleading. Even if it were possible to create such a list, policy makers would be very unwise to do so. The fact that the firms on the list were seen as being 'systemic' would tend to create a presumption that, in the event of their failure, they would be treated in a certain way. And as soon as such a view gained currency, it would lead to the kind of market distortions and moral hazard outlined above.

Policy makers know this of course and will not seek to create a simple list. They recognize that the most that can be established is that some firms exhibit more or less 'sytemic-ness' based on a range of indicators such as size, interconnectedness with other parts of the financial system and whether the firm provides services of a type or scale that could not easily be provided by others in the event of its failure. On this basis 'more systemic' firms might be required to hold more capital (for example) than less systemic ones. Such an approach would reduce – but not eliminate – the problem of moral hazard outlined above. Great care would still be needed to avoid the creation of perverse incentives and competitive distortions however, even if a more carefully calibrated approach to systemic risk surcharges were to be adopted.

A second aspect of reducing the likelihood of failure is to increase the intensity with which large/complex/global firms are supervised. There is compelling logic to the idea that such firms pose particular business and management challenges. Both managements and supervisors need to have a good grasp of these (something which has not always been evident in the past). Arriving at this understanding is all the more challenging the greater the complexity of the firm. This kind of thinking underlines the proposals for so called recovery and resolution plans or 'living wills'. These broad headings actually cover quite a diverse range of things, including:

---

[10] In some other variants of the idea, the conversion of contingent capital at, or close to the point of, failure would provide additional equity resources to the receiver to enable it to be liquidated in an orderly way without recourse to public funds.

[11] See for example 'Strengthening the Resilience of the Banking System', Basel Committee for Banking Supervision, December 2009.

- Recovery plans – outlining the mechanisms by which the firm would shed risk and rebuild its financial resources in the event of a (non-terminal) crisis, with a view to returning it to a healthy state; and
- Resolution plans – which are really about providing the information and wherewithal for the firm to be wound down in an orderly way in the event of its failure[12].

The logic of all this is unassailable. Putting such plans in place and keeping them workable and relevant will be a challenge for all concerned. Perhaps most important for firms will be the extent to which regulators see such requirements as a pretext for insisting on changes in business models or structures. Prior to the crisis these were not matters that regulators have typically involved themselves in. Regulators have already begun insisting that firms develop recovery plans and are examining the need for resolution plans. Doing this on a sustained and routine basis and on any significant scale, however, would involve quite significant burdens for supervisors and the broader implications, including the moral hazard resulting from any perceived transfer of responsibility from managements to supervisors, would need to be thought through carefully[13].

The other part of the conventional risk management equation – reducing the impact of failure – underlies the current emphasis on making markets and infrastructure more resilient. The drive to have credit default swaps cleared through central counterparties (CCPs), both to improve management of counterparty risks and to improve market transparency, is an example of this. Such measures are not a panacea – it is necessary to think about the potential systemic risks posed by the CCPs themselves and to ensure that they are regulated accordingly for example. But the recent crisis illustrated rather vividly that some markets and structures were more fallible than had been assumed and rectifying this is a key part of reducing systemic risk.

BETTER RESOLUTION ARRANGEMENTS

It is worth taking stock at this point. It is highly likely that within a reasonably short space of time, the official sector will have put in place a range of regulatory and other measures which will make it less likely that firms will fail and mitigate

---

[12] For an excellent discussion of recovery and resolution plans see 'Crisis: Cause, Containment and Cure' by Thomas Huertas. Palgrave MacMillan 2010.

[13] While supervisors have always challenged firms on whether their controls, management and financial resources are equal to the risks inherent in their business models, there has typically been a reluctance – other than in the most extreme cases – to challenge business models themselves.

the impact of their doing so. But the risk of failure will not have disappeared and the potentially systemic impact of failure will not have been eliminated altogether. Such outcomes would be neither feasible nor desirable. That would leave policy makers with three possible directions in which they could go. They could:

- ❖ Conclude that they have done everything that feasibly can be done and acquiesce in the residual risks to the financial system. This would be consistent with a conclusion that, with appropriate ongoing supervisory vigilance, the optimal balance has been achieved between risk and efficiency;
- ❖ Decide that the residual risks are still unacceptable and seek to minimize these at source by forcing firms to be less systemic – for example by placing restrictions on their size, scope of their operations (in terms of products and services) or their geographic reach; or
- ❖ Decide that the residual risks are still unacceptable and seek to minimize them further through improving the arrangements for resolving failing firms.

As noted above, the second option is highly problematic and would entail significant costs. There is therefore a powerful argument for going down the third route (which would also allow policy makers to feel more comfortable with the risk/efficiency trade off). The recent crisis demonstrated clearly that conventional insolvency arrangements are not equal to the task of resolving complex financial institutions – particularly where their activities are global in scope.

These issues were explored at length in a recent publication by the Institute of International Finance[14]. That report emphasized the following points:

- ❖ All firms, regardless of their size or complexity, should be able to exit the market in an orderly manner – that is without creating systemic trauma;
- ❖ There should be no presumption that taxpayers' money will be used to support failing firms;
- ❖ All major financial systems need to have in place special resolution regimes which facilitate the orderly winding down of financial firms. Such special regimes are necessitated by the specific challenges – especially with regard to complexity, speed, the need to minimize the destruction of value which invariably accompanies conventional insolvency and the need to protect depositors – that are more acute in the resolution of financial firms than in the resolution of commercial ones;
- ❖ The special resolution regimes need to be harmonized to the maximum possible extent internationally. At the very least they must not conflict – for example as a result of conflicting national depositor preference arrangements;
- ❖ But even a harmonization of national arrangements will not be sufficient. An ambitious global approach to resolution is needed if there is to be any hope of resolving globally active firms in an efficient and equitable way.

---

[14] A Global Approach to Resolving Failing Financial Firms: an Industry Perspective. Institute of International Finance, May 2010.

The issues here are complex and there is no scope for going into them in great detail in this article. Readers are referred to the IIF report on this subject. It is, however, worth highlighting a number of the key issues that need to be resolved if progress is to be made in this difficult area.

The first important point is the need to restore market discipline. It is understandable that, confronted with the prospect of widespread systemic trauma, governments have often followed the line of least resistance and used taxpayers' money to recapitalize banks and keep them in business. Not only is this a deeply unpopular thing to do but, as noted above, it sends the signal that certain types of firms will always be 'bailed out' – with very damaging consequences for market discipline. The costs of failure should not be borne by governments but, in the first instance, by the owners of the firm – that is the shareholders – and if the firm is taking too many risks, the shareholders should exercise discipline through the board to restrain them.

In reality, the systemic trauma that comes about from the failure of firms results from the fact that the equity capital in the firm is insufficient to absorb losses or to contain them in a way which avoids systemic damage. This means there needs to be a second line of defense. There is quite widespread agreement that this needs to be the unsecured creditors of the firm – those who took a view on the firm's risk by investing in its debt – in accordance with the normal hierarchy of creditor seniority. On this view, the debt markets would supplement the discipline exercised by the shareholders by demanding an appropriate risk premium. This very basic premise of corporate finance was to some extent eclipsed by reactions to the crisis and there is a need to reaffirm the principle that the risk implications of firms' activities need to be felt also by its unsecured creditors.

There is a variety of ways in which general unsecured creditors could be called upon to bear their share of losses. The idea of so called 'gone concern' contingent capital sees debt instruments being converted into equity near to the point of failure (an analogy is sometimes made here to 'pre-packed' corporate reorganizations)[15]. A similar idea which has gained a considerable amount of currency is that of the 'bail-in'[16]. Here, debt holders would see the value of their holdings reduced (ie subjected to 'haircuts') as part of the winding down process.

All of these ideas need to be examined further. Would investors have an appetite for gone concern contingent capital instruments? Exactly which categories of obligation would be subject to haircuts in a bail-in? Would traditional hierarchies of claims be respected? What would trigger the activation of these mechanisms and

---

[15] This is quite distinct from the idea of 'going concern' contingent capital in which debt is converted into equity but at a much earlier stage, as part of the recovery phase to restore the firm's capital base to enable it to continue as a going concern.

[16] Paul Calello and Wilson Ervin 'From Bail-out to Bail-in', "The Economist January" 28, 2010.

what legal and contractual changes would be needed to give effect to them? These are complex issues to which there are not ready answers. But the broad principle – that a range of creditors need to be on the hook to bear losses instead of the public – is well established.

A second set of issues arises out of the fact that the failure of any sizeable firm will undoubtedly entail costs. Imagine for example a firm which has a prominent role in the payments system. It is unlikely that the firm would be able precipitately to withdraw from these activities in the event of wind down. Similar issues (with even greater complexity) may arise with respect to any extensive operations in repo and derivatives markets. It is quite possible that parts of the firms' activities which are legitimately judged to have a systemic dimension would need to be transferred to a temporary 'bridge bank', even if there is a good prospect of the business line eventually being sold off. This would require financing – if only in the provision of working capital. The principle that taxpayers should not foot the bill for failures leaves only one place to go for this, namely the industry itself.

There is a remarkably wide measure of agreement about the principle that the industry as a whole should by responsible for residual losses once the equity and debt resources available from the failing firm are exhausted. There is however a lively debate about whether this should require the creation of a standing fund in advance of any failures (the *'ex ante'* approach) or whether the industry should be made liable to pick up any costs after the event (the *'ex post'* approach). Proponents of the ex ante approach argue that it would provide incentives to firms to monitor the risk taking activities of their competitors on an ongoing basis and that an ex post arrangement would penalize survivors who would be required to pay for the failures of less responsible competitors. Supporters of the ex post approach believe that the existence of a standing fund would provide a temptation to bail out failing firms in their entirety and obviate the need for the difficult decisions involved in forcing an orderly winding down. The European Union in its proposals on bank resolution funds[17] envisages the creation of an ex ante fund. The Dodd Frank Wall Street Reform and Consumer Protection Act in the United States makes provision for ex post financing.

Perhaps most difficult of all is the international dimension to the resolution issue. Many large firms operate on a global basis. They have product lines and management structures which extend across national boundaries. Insolvency regimes – and even special resolution regimes – tend however to be entirely national in their focus and modus operandi. This creates enormous tensions when global firms fail. It is frequently the case for example that in the period immediately ahead of failure, assets move quickly around the group so that when the music stops and the firm is officially declared insolvent, the geographical distribution of

[17] European Commission: Communication on Bank Resolution Funds, May 2010.

assets may be completely out of line with past experience and with the legitimate expectations of creditors.

These problems may be compounded by the fact that legal requirements applicable in traditional resolution and bankruptcy cases have in some cases created conditions which make it difficult or impossible for administrators or receivers to cooperate to achieve results that are optimal or equitable from an international or group-wide perspective.

For as long as these problems remain unresolved they will be insuperable barriers to the orderly resolution of global firms. There seem to three broad approaches that might be followed:

- Put in place financing mechanisms such as bail-ins which ensure: a) that all categories of creditors understand their potential liability in the event of failure; and b) that there will under most circumstances be sufficient resources to ensure that legitimate creditor claims can be met, without recourse to public funds (on anything other than a temporary basis). Ideally, 'bail in' arrangements should operate to treat all claimants equally, regardless of their geographical location. It is too early to say whether the current 'bail-in' and other proposals being discussed would deliver this outcome.
- Force (currently global) financial groups to adjust their structures and business models to conform more closely than at present to the (national) limitations of insolvency and resolution regimes. That would, in effect, involve breaking up global groups or requiring their operations in each market to be tightly ring fenced, in either case creating substantial management and efficiency constraints. As noted above, this would make it very difficult for the financial services industry to continue to provide many valuable services on which global businesses rely.
- To put in place a truly international framework for the resolution of global financial firms. Making progress in this will be difficult and it will take time. But if the only alternative is to delude ourselves into believing that global firms can really be made national again without severe costs, there is no realistic choice but to go down this route. The difficulty of achieving a fully effective international framework is not an acceptable argument against making the effort. As an interim matter, there is a great deal of progress that can and needs to be made in reforming national resolution regimes, not least to increase the scope for greater international coordination and cooperation among administrators and receivers.

CONCLUSION

There is a wide measure of agreement between the official sector and the financial services industry that we can no longer countenance a situation in which governments feel that they have no choice but to support large financial firms when they get into difficulty. That means that ways must be found to ensure that such firms can exit the market in a way which does not cause unacceptable systemic damage.

Improvements to regulation and supervision together with measures to make markets and infrastructure more resilient will go a long way towards reducing the incidence and impact of failures. Mechanisms such as bail-ins and gone concern contingent capital offer the prospect of bolstering the resources available to failing firms, permitting failures to occur in an orderly way and, critically, improving market discipline. There are many unresolved technical issues surrounding these ideas; these need to be resolved quickly so that the real value of these promising ideas can be fully evaluated. It is imperative that solutions to these problems are capable of addressing the very complex challenges posed in resolving globally active groups. All of this calls for high levels of ambition by the official sector and the industry alike.

What is clear however is that simplistic solutions will not do. Systemic risk is a complex problem and whether we like it or not, it may require complex solutions. We must not lose sight of the benefits that large global firms bring by rushing to find solutions to the problems that some of them have created in the past.