PART I

Micro- and macro-prudential regulation

1

The road from micro-prudential to macro-prudential regulation

ESTER FAIA AND ISABEL SCHNABEL[1]

Introduction

One of the most important lessons from the 2007–2009 crisis was that micro-prudential regulation and supervision are insufficient to stabilize financial systems. Despite a highly sophisticated micro-prudential framework, the world experienced the most severe crisis since the Great Depression. This showed quite plainly that a focus of regulation on individual intermediaries is not enough to prevent the breakdown of the financial system, and that instead more attention has to be paid to the evolution of systemic risk. Nowadays, both academics and regulators agree that the micro-prudential focus should be complemented by a macro-prudential perspective.

The emergence of the largely micro-prudential Basel framework goes back to the 1980s. After several decades of financial calm with strongly regulated and rather closed economies worldwide after the Second World War, the 1980s and 1990s saw a wave of deregulation in many spheres, including financial markets. At the same time, globalization proceeded rapidly and evoked the need for harmonized financial regulation to create a global "level playing field." This gave rise to the Basel process, which became the global regulatory framework for banking. This framework was largely micro-prudential in nature – it tried to ensure the safety and soundness of individual institutions, putting more emphasis on the goal

[1] Ester Faia is Professor of Monetary and Fiscal Policy at Goethe University Frankfurt, and Isabel Schnabel is Professor of Financial Economics at Johannes Gutenberg University Mainz. The authors would like to thank Carmelo Salleo, Michalis Haliassos, as well as conference participants at the SAFE Workshop on Financial Regulation for useful comments and discussions.

of depositor protection than on system stability. At the same time, there was an overall consensus that financial markets were by and large efficient and that they were capable of selecting stable equilibria in which individual incentives could, in most cases, be disciplined by "the market." The regulator only had the task of limiting the scope of activities for individual intermediaries, in particular by asking for minimum capital holdings.

As the regulatory framework came increasingly under pressure due to regulatory arbitrage and financial innovation, regulation was modified in a stepwise fashion, raising the degree of complexity substantially and trying to capture risks at an ever-increasing granularity (Haldane 2012). The major innovation was the reliance on banks' internal models – first with respect to market risk, then, under Basel II, also with respect to credit risk. The formulation of value-at-risk models was left in the hands of individual intermediaries, the underlying idea being that intermediaries and financial markets were able to elaborate the optimal management of risks by themselves. This approach ignored that banks maximize profits rather than social welfare and that their behavior may entail externalities on the financial system and the real economy that a profit-maximizing bank would not take into account. Hence, the reliance on banks' internal models is inherently micro-prudential.

Consequently, unweighted capital ratios often fell to levels of no more than 2 percent and banks built up substantial off-balance sheet risks in the form of special purpose vehicles investing in structured products, such as mortgage-backed securities, while relying on short-term funding and liquidity guarantees from their sponsors.[2] In the financial crisis, the initial failures of particularly weak banks were magnified by macroeconomic repercussions in the financial system – for example, through fire sale externalities. These amplification mechanisms were not the result of an exogenous shock, but they were caused by banks' endogenous responses to such shocks. Hence, aggregate risk proved to be much more than the sum of individual risk positions.

Although the severity of the crisis, the timing, and the degree of interconnectedness of financial institutions were surprising for many observers, the underlying mechanisms were already well understood long before the crisis. Hellwig (1995) wrote about them as early as 1995. Crocket (2000) started an intensive debate on the importance of the macro-prudential dimension of financial stability at the Bank for

[2] For detailed accounts of the financial crisis, see Brunnermeier (2009) and Hellwig (2009).

International Settlements (BIS), which was continued by Borio (2003) and others. Daníelsson et al. (2001) criticized the Basel II Accord for not taking into account the endogeneity of risk and its inherent pro-cyclicality. However, these criticisms were not taken up by policy makers.

This chapter presents our vision of an appropriate framework for macro-prudential regulation. We proceed by first identifying the transmission channels of aggregate risk, secondly by defining the different dimensions and targets of macro-prudential regulation, and thirdly by outlining the institutional framework implemented in Europe. Finally, we discuss some important issues concerning policy implementation, such as the debate on rules versus discretion and cross-border coordination, as well as possible conflicts or overlaps with other macro-policies, such as monetary policy.

Transmission channels of systemic risk

The financial system is subject to various types of shocks, including idiosyncratic and aggregate shocks. Both types of shocks can lead to systemic stress through various channels (see Figure 1.1 for an illustration). The most basic manifestation of systemic risk is a situation in which an *exogenous aggregate shock* hits many banks at the same time because they are exposed to the same type of risk. For example, the recent crisis was triggered by a sharp drop in real estate prices in the United States. Since many banks were exposed to U.S. mortgages, the price decline generated a loss of values on the asset side of many financial intermediaries at the same time.

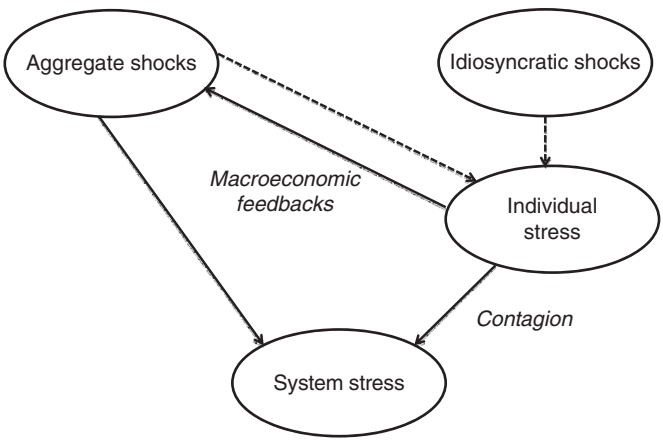

Figure 1.1 Transmission channels of systemic risk

Systemic stress can also arise from *idiosyncratic shocks* if individual stress is transmitted to other financial institutions (*contagion*). Contagion may work through three different channels: information, interconnections among banks, and macroeconomic feedbacks. Information contagion (as in Chen 1999) occurs when the failure of one institution signals the fragilities of other financial institutions. Such effects appear to have played an important role in interbank markets during the crisis.

The second channel works through banks' interconnectedness, such as lending exposures on interbank markets or inter-linkages through CDS and other insurance contracts, which lead to domino effects and default cascades. Consider the case of interbank exposures: as a single bank is hit by an idiosyncratic shock, it may be unable to meet its obligations on the interbank market. The inability of one bank to honor its debt would transmit losses to other banks, which then potentially become insolvent (as in Allen and Gale 2000). This chain of insolvencies is produced by network externalities and can induce widespread defaults even as a result of a shock to a single bank.

The third type of contagion is caused by macroeconomic feedbacks. In this case, individual stress induces banks to adjust their behavior in a way that has repercussions for the macro-economy. These *endogenous* aggregate shocks in turn affect all financial institutions, with the potential of generating vicious cycles between individual stress and the macro-economy. Such cycles may be amplified by regulation, making evident how the prescriptions of micro- and macro-prudential regulation can diverge.

For example, Basel rules prescribe that banks increase their capital or reduce their assets when they incur losses, and capital requirements are binding. Due to risk-weighting, the same reaction is required when asset risk increases, as happens in a recession. The underlying rationale is that banks shall be able to cope with increased potential losses in portfolio values. While such behavior seems prudent from the point of view of the individual intermediary, it might produce disruptive consequences at an aggregate level. Reducing assets in the middle of a recession may induce a credit crunch, exacerbating the downturn (see Angeloni and Faia 2013). A macro-prudential regulator concerned with the stabilization of aggregate credit would therefore prescribe countercyclical capital buffers, such as those featured in the Spanish dynamic provisioning or in the Basel III countercyclical capital buffers.

In the above example, the amplification works through banks' capital positions. Another type of feedback is related to bank liquidity. When a

bank suffers an (idiosyncratic) liquidity shock, it may be forced to sell assets in order to produce the liquidity needed to honor its obligations. Again, this seems like a reasonable response from a microeconomic perspective. However, the fire sale may produce a fall in the market price of that asset, especially in times of market distress, leading to liquidity spirals (Brunnermeier and Pedersen 2009). In the presence of mark-to-market accounting procedures, the fall in this asset's price will inflict portfolio losses on all other banks who have invested in the same asset. But even in the absence of mark-to-market accounting it limits other banks' possibilities of generating liquidity by selling assets.[3]

Banks do not take such externalities into account when choosing their capital structure (equity versus debt) or the maturity structure of their debt. This may lead to both excessive leverage and maturity transformation, yielding a rationale for regulation, such as increased capital requirements and the introduction of liquidity requirements under Basel III.

Summing up, systemic stress can be generated either by an exogenous aggregate shock or it can propagate endogenously through bank runs, interconnections, or macroeconomic feedbacks. Micro-prudential regulation has generally neglected the possibility of endogenous self-propagation of risk and its time-varying nature. It traditionally focuses on preventing individual stress (i.e., on the dashed arrows in Figure 1.1). In contrast, macro-prudential regulation concerns contagion effects, exposures to macroeconomic risk at the system level, and macroeconomic feedback effects (as shown by the solid arrows in Figure 1.1).

The two dimensions of macro-prudential regulation

The literature distinguishes between two dimensions of macro-prudential regulation (see Borio 2003): the cross-sectional and the time series dimension. The cross-sectional dimension refers to the varying levels of systemic risk emanating from financial institutions at a given point in time. This dimension captures the strength of contagion effects as described above. The time series dimension is concerned with the evolution of systemic risk over time and is hence closely related to the evolution of macroeconomic prices and quantities (such as credit), caused by exogenous shocks or driven by the endogenous dynamics of

[3] This transmission channel results from a classical pecuniary externality. Such externalities cause distortions only in the presence of other frictions (see Hanson et al. 2011).

the financial system. Macro-prudential regulators have to be concerned with both dimensions if they want to limit systemic risk.

Appropriate definition of the policy objective, as well as instruments, is surely one of the most important aspects of macro-prudential regulation. The two dimensions manifest themselves in different objectives and instruments, some of which will be outlined in the following sections.[4] Objectives will capture the cost of aggregate externalities to be minimized: this goes beyond the mere aggregation of individual risk, capturing the endogenous risk propagation mentioned above. An obvious challenge is the measurement of such propagation mechanisms. In contrast to objectives, which may apply at the aggregate level, instruments (such as capital or liquidity requirements) should be bank-specific: due to bank heterogeneity, aggregate instruments, applied equally to all institutions, are likely to induce distortions.

The cross-sectional dimension of macro-prudential regulation

The central idea of the cross-sectional dimension is that macro-prudential regulation should be calibrated in a way that captures individual contributions to systemic risk. This necessarily implies that systemic banks should be regulated more strictly than non-systemic banks. This is a departure from traditional (micro-prudential) regulatory practices, stressing the importance of a level playing field (although the playing field was not so level after all, given implicit bail-out guarantees). An example under Basel III is the capital surcharge for systemically important financial institutions (SIFIs).

In order to measure systemic risk, supervisors may either use an indicator approach or a quantitative measure based on banks' return distributions. In the more common indicator approach, banks are categorized according to different determinants of systemic relevance. The most frequently used criterion is bank size, which is easy to measure and is seen as a proxy for many other determinants of systemic relevance.

A second criterion is interconnectedness. Theory-based measures of interconnectedness try to capture externalities within network models. While taking into account the role of interconnections, such measures are also able to identify systemically important institutions or key spreaders of risk. In this context there is a distinction between static ex-ante

[4] An extensive discussion of macro-prudential instruments is beyond the scope of this chapter.

measures versus dynamic ex-post measures. Ex-ante metrics include network centrality or input–output metrics. Centrality indices, which were first developed in graph theory, produce rankings that identify the most important nodes. The variety of centrality measures is identified by the way importance is defined. The idea behind input–output metrics is as follows: if a bank is hit by a shock, it will transmit it to the rest of the system according to the coefficient of the input–output matrix, which in a banking context represents a transformation of the interbank exposure matrix. One advantage, compared to centrality measures, is that systemically important banks emerge also in more sparse systems, which do not necessarily feature a vertex to which many banks are connected. Both centrality and input–output metrics are by their nature static as the matrix of connections is taken at a certain point in time: they can only signal the critical nodes, i.e., those institutions that are potential spreaders of risk. However, they lack predictive power as they do not indicate how risk can spread within the system.

A second type of measure is represented by the Shapley value and other measures borrowed from the literature on cooperative game theory. Given a multivariate distribution of shocks across banks, those metrics capture the contribution of each bank to the diffusion of risk by considering all possible permutations in which the same bank enters the network matrix. Those metrics capture the ex-post diffusion of risk as they indicate the contribution of each bank to total losses in the system after the shock has been spread across the network. Due to their dynamic nature they are more apt for use in crisis management. Although such network measures are well-developed theoretically, their use is often hampered by a lack of data. At the current stage, their practical application is therefore limited.

As an alternative to the indicator approach, a number of measures have been developed to capture the systemic risk of financial institutions on the basis of joint return distributions. Examples are the change in the conditional value at risk (ΔCoVaR) developed by Adrian and Brunnermeier (2011), the marginal expected short-fall (MES) by Acharya et al. (2012), the capital short-fall (SRisk) by Brownlees and Engle (2012), and the change in the conditional joint probability of default (ΔCoJPoD) developed by Radev (2012). All such measures try to capture tail risk conditional on some distress event.

Their main advantage is that they are based on data that is readily available for publicly traded firms. In Europe, many banks are not traded, which limits the usefulness of these measures. Moreover, the measures

rely on the assumption that market prices have some predictive power of distress. However, as was seen before the recent crisis, markets tend to understate risks in boom times. Furthermore, the different measures of systemic risk yield widely varying results regarding the systemic relevance of different financial institutions. Hence, further research is needed in this area. The most problematic issue is the strong pro-cyclicality of such measures. For example, ΔCoVaR rose dramatically over the course of the financial crisis (see Barth and Schnabel 2013). Hence, linking capital requirements to this type of variable would introduce an additional pro-cyclical element into financial regulation. Thus, the cross-sectional and the time series dimension may be contradictory. This speaks for using through-the-cycle concepts, which are purged from cyclical factors, to capture the cross-sectional aspect of systemic risk.

The goals of macro-prudential regulation would be to make systemic banks safer, to make it less attractive for financial institutions to become systemic, and to reduce the competitive distortions caused by implicit government guarantees for systemic banks. In order to achieve these goals, various regulatory instruments can be linked to the described systemic risk measures, including capital and liquidity requirements or bank taxes. In order to achieve the desired incentive effects, it is crucial to link regulation to a bank's contribution to systemic risk rather than burdening all institutions to a similar degree. This suggests, for example, that banks' contribution to the Single Resolution Fund should be calibrated to banks' systemic risk.

In addition, changes in the financial infrastructure can help to reduce contagion effects and remove distorted incentives from implicit government guarantees. Important examples are the introduction of central counterparties (CCPs) for derivatives trading and the implementation of bank resolution procedures, as envisaged in the Single Resolution Mechanism (SRM) of the European Banking Union. Disclosure requirements can also be useful – for example, concerning the interconnectedness in interbank markets.

The time series dimension of macro-prudential regulation

The time series dimension of macro-prudential regulation focuses on two objectives: dampening the financial cycle, and preventing the emergence of bubbles in certain market segments.

In order to dampen the financial cycle, macro-prudential regulation should be adjusted over time to the evolution of macroeconomic