A Non-cyclical Capital Adequacy Rule and the Aversion of Systemic Risk

Raphael Douady

2009-12-07

Abstract

We present in this note a method for computing the regulatory capital of financial institutions, along with the Basel Committee requirements, which avoids the pitfalls of the Value-at-Risk and, in particular, the fact that – as observed during 2008 crisis – it aggravates systemic risk rather than preventing it. The computation is based on stress testing, with the following principles: (i) market scenarios are defined by the regulator; (ii) sensitivities are estimated by each institution, as well as the impact of scenarios defined by the regulator and reported to it; (iii) the regulator not only counts the number of violations of the risk reporting but also their size; (iv) the regulatory capital is a multiple of the worst stress test, where the multiplier depends on the size and the frequency of the violations. By letting the institutions estimate their sensitivities to extreme market shifts, the regulator not only avoids a costly burden, but also keeps institutions responsible for their reporting. On the other hand, by keeping control on the list of stress tests involved in the computation of the capital, the regulator offers itself a very strong lever to prevent speculative bubbles, by making them costly in terms of capital requirements.

Reasons and Problems of the Current Capital-adequacy

The current Capital Adequacy rule, as stated by Basel II agreement, computes the "regulatory capital" of a financial institution as a multiple of its "Value-at-Risk" (VaR), itself the sum of several risk sources: market risk, credit risk, counterparty risk, operational risk, etc.

The recommendations for computing the "economic capital" follows the same guidelines.

The VaR depends on two parameters: this horizon h (usually 5 days) and a percentile q (usually 99%). It is the amount V such that the probability that the institution loss over the horizon h exceeds V is equal to $1 - q$. An abundant literature has been published on the various methods for computing the Value-at-Risk.

The required capital to operate on markets is $k \times$ VaR, where the multiplier k is a number between 3 and 10, which depends on the "quality" of the VaR computation. The assessment of the "quality" is the result of two verifications – and as the result, the worst of both:

- a qualitative assessment of the process
- a quantitative back-test counting the number of exceptions, i.e. where the loss exceeds the VaR, and comparing this number with the declared frequency $1 - q$.

The good thing in this setting is that the Regulator lets the institution compute its risks, assuming that it has a better knowledge of its details to better track the pitfalls of its own risk evaluation. It only acts as a *verifier* who checks afterwards that the risk has been correctly computed. If the risk was underestimated, the sanction is a higher ratio, hence a higher cost in regulatory capital for further operations.

It also represents a substantial economy for the Regulator, who leaves the burden of computing risks to the institutions. It is indeed a massive distribution of the burden across all the institutions, each of them taking care of its own risk computation. The verification task is less costly than the computation itself by orders of magnitude.

However, this approach leads to numerous problems, which became stringent throughout the 2008 crisis:

1) The rationale for the economic capital is to avoid bankruptcy, hence the loss should *never* exceed $k \times$ VaR. This trigger is surprisingly enough never tested.

2) The risk measure is 1-*dimensional* and neither tells the exact risk source, nor the market scenario it corresponds to. As a consequence, the Regulator cannot realistically require that the loss *never* exceeds the declared risk.

3) The most serious problem is *pro-cyclicality*: in a market downturn, the risk measure increases, leading most market participants to sell out positions in order to meet capital adequacy, adding to the market turmoil.

Proposed Solution for an Non Pro-cyclical Capital Adequacy Rule

Pro-cyclicality results from the fact that only a global risk measure is considered (see point 2 above) and from the *reactivity*, rather than *pro-activity* of the risk measure. Preventing a risk measure from being pro-cyclical is not an easy task. It requires the regulator to anticipate market crises, using its knowledge of the financial and economic situation. It also requires verifying that financial institutions have a correct monitoring of their extreme exposures. Indeed, any type of rule that forces institutions to act in the middle of the turmoil will mechanically aggravate the liquidity crisis and add to the turmoil.

General Principle

We here propose to include *stress tests* in the measure of capital adequacy in such a way that we respect the following 3 *golden rules*:

1) The Regulator defines which stress to apply to which indices. This will ensure that economic research is unbiased to anticipate potential market shifts.

2) The Institution computes itself the impact of stress scenarios on its activity. It is free to add other scenarios that the Regulator didn't think of for its particular case, either by stressing other risk factors or by increasing the stress size given by the Regulator.

3) The Regulator verifies that losses incurred by the Institution (if ever) do not exceed what could have been anticipated given the declared stress tests and the actual market moves. In other words, the Institution is responsible for correctly anticipating the impact of markets on its activity, but not for the moves of markets itself.

The required operating capital is proportional to the worst declared stress test (as of 2). The initial multiplier value is 1. In case of a violation, the impact on the multiplier depends on the amount of the violation. Minor violations have a minor impact, while large violations severely impact the multiplier. If there is no violation, the multiplier is progressively brought down back to the value 1.

We mean to exclude all notions of probability in this framework. Institutions should be responsible for the amounts *they declare, not for the* probability *of such or such event.*

Details of the Regulatory Capital Computation

Stress Tests Defined by the Regulator

Each period (e.g. month, but can be more frequent if necessary), the Regulator issues a list of market indices to stress with, for each index, a list of amounts by which it should be stressed. This means a table of the following kind:

Index	Stress++	Stress+	Stress0	Stress-	Stress--
S&P500	+20%	+10%	0%	–10%	–20%
TB Yield 10Y	+200bp	+100bp	0bp	–100bp	–200bp
BAA Credit Spread	+500bp	+200bp	–10bp	–100bp	–200bp
…					

Each entry of this table corresponds to a shift of an index I_i by an amount ΔI_{ij}. Ideally, each ΔI_{ij} corresponds to some percentile of the anticipated probability distribution of the index shift over the next coming month. For instance:

Stress++ = 99% percentile up
Stress+= 84% percentile up
Stress0 = Median
Stress– = 84% percentile down
Stress-- = 99% percentile down

We here gave an example with round figures, but the regulator is free to apply any quantitative model leading to the values of the shifts ΔI_{ij}. It is somehow recommended that these figures be rather stable through time, especially the most extreme ones, as capital requirements will be direct functions of them. In particular, rather than keeping them strictly constant for some time, then re-adjusting them suddenly, one should estimate them in the most anticipative manner, in order to temper down the probability of a large jump.

A sequence of 5 stresses for each index seems reasonable but this number can be subject to discussions. More important is the list of basic indices to be stressed. These should cover all asset classes (equity, fixed-income, credit, volatility, FX, emerging markets, commodities, real estate, etc.), as well as the most important drivers of majorly traded securities.

It is extremely important that the Regulator makes all efforts to anticipate the distribution of possible forward moves, and does not simply rely on the past volatility of each index.

Computation of Stress Scenarios by Institution

Each ΔI_{ij} must be seen by the institution as a full market scenario, not just a single shift of a single asset class. For instance, if the scenario represents a 15% increase of the oil price, the impact of such a scenario on other asset classes, such as energy related stocks or whatever market should be accounted for. The importance of this point will be clear when we shall describe how the Regulator will verify the accuracy of risk computations.

First, the Institution can, on a fully customary basis, decide to divide itself into "divisions" $D_1,\ldots,D_n$ which correspond to business units mostly exposed one dominant source of risk. Then, every reporting period (e.g. every week), each division D_k produces a Risk report that contains its own P/L estimate in case of scenario ΔI_{ij}, which we denote L_{ijk}.

This P/L estimate is supposed to be the *lower bound* of a confidence interval of the impact of the scenario. Specifying the probability to which this confidence interval corresponds to is not necessary. It is in the interest of the Institution not to overpass this lower estimate, or by a limited amount, as we shall see.

The institution is free to add more scenarios ΔI_{ij} either by adding other risk factors I_i or other shifts ΔI_{ij} for existing factors. It is in the interest of the institution to be as exhaustive as possible in the declaration of its risk sources, in order to avoid violations which may increase its multiplicative ratios. For instance, would one of the divisions be particularly exposed to Kazakhstan equities, a risk factor not listed by the Regulator, it is in the interest of the Institution to report a potential exposure to this risk factor, in order to avoid a violation of declared risks in case of a pure Kazakh crisis.

The regulatory capital C is computed as follows, where λ represents the multiplier:

$$C = \lambda \sum_k \max_{i,j} L_{ijk}$$

In other words, the Institution's capital is the sum of that of its divisions and, for each division, it is the maximum potential loss stemming from one of the declared scenarios.

Back-testing Stress Scenarios

The key point of this regulatory framework is the ability for the Regulator to back-test the accuracy and completeness of risk reports by financial institutions. Risk reports are established at dates t for a horizon h. Institutions provide, for each market scenario ΔI_{ij} the possible loss $L_{ijk}(t)$ of each division D_k. At the end of the period $t+h$, the Regulator observes the actual shift of each index $\Delta I_{i\text{act}}(t) = I_i(t+h) - I_i(t)$. By a simple linear interpolation, or extrapolation if this shift happens to exceed the range of published ΔI_{ij}, $j = 1...5$, the Regulator computes what should have been the impact of this shift on each division D_k:

$$L_{ik,\text{impact}}(t) = \text{Inter/Extrapolation of } L_{ijk} \text{ for } \Delta I_{i\text{act}}(t)$$

Note that, in case $\Delta I_{i\text{act}}(t)$ is outside the bounds of the 5 values ΔI_{ij}, this computed loss may exceed the maximum declared loss. Then the *Maximum Accepted Loss* of the division is computed:

$$MAL(D_k, t) = \max_i L_{ik,\text{impact}}(t)$$

The Institution *Maximum Accepted Loss* is the sum of that of its divisions:

$$MAL(t) = \sum_k MAL(D_k, t)$$

A violation is incurred when the Institution actual loss $L_{\text{act}}(t)$ over the period $[t, t+h]$ exceeds $MAL(t)$. In this case, the *violation ratio* is simply:

$$V(t) = \max(1, L_{act}(t) / MAL(t))$$

Computation of the Multiplier

The multiplier $\lambda(t)$ is re-computed at every reporting period, according to 2 rules:

1) If there is no violation, it is reduced in order to converge after a reasonable period of time to the value 1.
2) If there is a violation, it is increased by an amount depending on the severity of the violation.

The formula proposed here to compute $\lambda(t)$ from its previous value $\lambda(t-1)$ and from the violation ratio is quite simple. First we compute a natural dampening of $\lambda(t-1)$:

$$\bar{\lambda}(t) = \max\left(1, \min\left((1-\alpha)\lambda(t), \lambda(t) - \varepsilon\right)\right)$$

In this formula, $1-\alpha$ is a dampening factor, whose role is to tame down the multiplier. If the multiplier is already rather close to 1, the reduction is at least ε, without possibility to drop below 1. Parameters α and ε depend on the reporting frequency and should be set in such a way that, approximately after a year without violation, the multiplier is set back to 1. For instance, for a weekly reporting, $\alpha = 1\%$, $\varepsilon = 2\%$ ($\approx 1/52$).

In words, every week, if there is no violation, the multiplier is reduced by 1% of its value, the reduction not being less than 0.02 until the value 1 is reached. If $\lambda(t-1) > 2$, then $\bar{\lambda}(t) = 0.99\lambda(t-1)$, if $1.02 < \lambda(t-1) \leq 2$ then $\bar{\lambda}(t) = \lambda(t-1) - 0.02$ and if $1 \leq \lambda(t-1) \leq 1.02$, then $\bar{\lambda}(t) = 1$.

In case of a violation, i.e. $V(t) > 1$, the multiplier is simply multiplied by the violation ratio:

$$\lambda(t) = \bar{\lambda}(t)V(t)$$

This way, the penalty for violating the MAL is strictly proportional to the size of the violation.

Eliminating Pro-cyclicality

The key point to prevent cyclicality is the de-correlation between violations of risk measures and market events. This de-correlation will be achieved under two conditions:

1) The Regulator decides the stresses to apply, hence is in a position to smoothly impose deleveraging before it becomes an unsolvable problem. This is why the regulator must have an anticipative measure of factor risks and, in particular, of systemic risk.

2) Risk reporting is not a figure, but a function of markets hence violations are not due to markets swings but to misreporting of extreme risks. If institutions correctly report their extreme exposures, there is no reason why they would more violate their assessment during a crisis than during normal periods.

There is a chance that, even with such a setting, capital adequacy constraints still remain pro-cyclical if one of the following occurs:

a) The Regulator fails to anticipate systemic risks,

b) Institutions fail to correctly estimate their exposures to extreme market conditions

For these reasons, it is of utmost importance that the Regulator puts in place appropriate tools and analyses to cleverly monitor and update the list of imposed stress tests.

Other Risk Sources

It is in the interest of the Institution to foresee other risk sources within its risk reporting. These will add to the regulatory capital, but will avoid costly violations. The custom definition of "division" allows institutions to easily include extra risk sources in their reporting. For instance, in order to include Operational Risk, one can simply add a division supposed to entirely support this risk. The strength of this framework is to allow reporting not only a maximum amount, but an amount that may depend on external risk factors in a purely customary way. Let us here examine a few examples.

Operational Risk

The simplest way to include such a risk is to create a specific division D_{op}. Assuming operational risk cannot be related to any market factor, one will declare a fixed amount L_{op} which will be added to the overall MAL.

Operational losses, which are included in the actual loss $L_{act}(t)$, will be compared to this extra buffer in the MAL.

Counterparty Risk

Each major counterparty can be made a division. Losses stemming from the default of a counterparty are directly related to market events, in two ways: first by the amount of the engagement, second by the probability of default, which naturally depends on market conditions. It is in the interest of the Institution to estimate the reliability of its counterparty and to optimize the risk of violation and its cost in capital vs. its probability of occurrence. Using hedges, such as credit derivatives, or diversification strategies across several counterparties will be a direct consequence of this optimization.

Liquidity Risk

When the Institution deals with illiquid assets or even with assets exposed to liquidity risk, that is, assets that may suddenly become illiquid, such events materialize by a sharp price drop, which can be anticipated through nonlinear (optional) modelling. It is in the interest of the institution to anticipate such events and the market scenarios that may trigger them. If such scenarios are outside the bounds specified by the regulator, it might be optimal to include more extreme scenarios in the reporting, or to include scenarios rather specific to such or such asset class which were not in the Regulator's list.

Default Risk

The default risk on an asset is directly identifiable with a price drop. It is quite easy to anticipate as the impact of some market scenario and to include it into the risk computation.

Counterparty Risk: monitoring "too connected to fail"

One important teaching of the recent crisis is that market risk netting at the level of an entire institution is missing an important issue: counterparty risk. In fact, each major counterparty of the institution should be considered as a separate portfolio as, in case of default, the firm-wide netting will be destroyed.

For this reason, institutions should report their exposure to major counterparties in the same format as they do it for their own risk. In other words, a risk report of the same nature as that described above should be produced for each counterparty that accounts for a significant portion of the firm business. The “negative” part of the engagement with the counterparty (i.e. when the institution owes to the counterparty) is a simple market risk, to be aggregated with other market risks in the global market risk report. However, the “positive” part represents counterparty risk and should be subject to reporting in the form of stress tests as above.

The amount at risk, given by the worst case scenario, should be considered as a “loan” to the counterparty and be regulated as such with prudential rules. This being a risk figure and not a foreseeable amount with certainty, the Cooke ratio does not need to be applied in full. However, monitoring this potential risk with full knowledge of the market scenario in which it materializes will be crucial for the regulator to anticipate possible cascade effects. In fact, with such information, the Regulator will be able to run simulations and identify institutions that are “too connected to fail” as opposed to “too big to fail”.

A New Macroeconomic Lever

In the traditional Keynesian economy, the Government basically holds two macroeconomic levers in order to optimally monitor the country growth: short term interest rates and government spending. The third one – printing paper money – is not available to all governments (e.g. EU countries) and, generally speaking, is subject to constraints and must be used with extreme care.

Moreover, all these levers jointly act on the “financial world” and on the so-called “real economy”. Implicitly, the “financial world” is supposed to be in line with the “real economy”. Unfortunately, recent economic history has shown that the development of financial technology – both computerization and new financial instruments – allow strong discrepancies between the two. These are precisely those discrepancies which are targeted by regulation on capital adequacy.

The setting described here offers a new type of command lever to the Regulator. By allowing more or less regulatory capital to financial institutions, it can accurately monitor the general leverage of the system and discriminatively act on the financial world without touching the “real economy”, that is the industrial corporations, hence preventing too large discrepancies as observed during speculative bubbles.

Conclusion

The methodology presented here for capital adequacy rules is a strong framework for regulators to improve the current Basel II principles and, in particular, avoid pro-cyclicality without incurring uncontrollable costs for the Regulator to monitor the risks of financial institutions. This is achieved by applying the following principles:

1) The Regulator defines which market scenarios should be anticipated and tested. By its *anticipative* action, it avoids pro-cyclicality.

2) The Institutions provide their own risk estimate in each of these scenarios. The Regulator does not need to control the details of each institution, avoiding the corresponding costs.

3) The regulator checks that risks are correctly reported by comparing actual losses to an amount which depends on both the risk report of the institution and the actual market move. Institutions must correctly declare the potential impact of markets, but are not required to actually anticipate markets themselves. This task is left to the Regulator.

4) The Regulator penalizes institutions proportionally to the *amount* of the violation and not with respect to their frequency only. This puts a responsibility on the Institutions to provide *readable* risk figures and not abstract numbers potentially not in relation with actual losses.

Thanks to these rules, one can establish a healthy operating framework in which Institutions and Regulators keep confidence in each other. Institutions are naturally led to report risks as exhaustively as possible and to avoid "putting the dust under the carpet". Moreover, the extra safety gained by the approach will probably allow a reduction of capital for a number of healthy institutions that were penalized by the hazardous activity of others, which forced the Regulator to increase margins when this was not absolutely necessary.
Finally, the definition of stress scenarios used for computing the regulatory capital is a new command lever for preventing speculative bubbles, by accurately acting on the appropriate asset class and, therefore, avoiding systemic risk.