Building a Coherent Risk Measurement and Capital Optimisation Model for Financial Firms

Tim Shepheard-Walwyn and Robert Litterman

I. Introduction

Risk-based capital allocation methodologies and regulatory capital requirements have assumed a central importance in the management of banks and other financial firms since the introduction of the Basle Committee's Capital Accord in 1988. However, as firms have progressively developed more sophisticated techniques for measuring and managing risk, and as regulators have begun to utilise the output of internal models as a basis for setting capital requirements for market risk, it is becoming increasingly clear that the risk as measured by these models is significantly less than the amount of equity capital that the firms themselves choose to hold.[1]

In this paper, we therefore consider how risk measures, based on internal models of this type, might be integrated into a firm's own methodology for allocating risk capital to its individual business units and for determining its optimal capital structure. We also consider the implications of these developments for the future approach to determining regulatory capital requirements.

Tim Shepheard-Walwyn is managing director, Corporate Risk Control, UBS AG. Robert Litterman is managing director, Asset Management Division, Goldman Sachs.

II. Why Do Financial Firms Need Internal Risk Measurement and Risk-Based Capital Allocation Methodologies?

The core challenge for the management of any firm that depends on external equity financing is to maximise shareholder value. To do this, the firm has to be able to show at the margin that its return on investment exceeds its marginal cost of capital. In the context of a nonfinancial firm, this statement is broadly uncontentious. If the expected return on an investment can be predicted, and its cost is known, the only outstanding issue is the marginal cost of capital, which can be derived from market prices for the firm's debt and equity.

In the case of banks and other financial firms, however, this seemingly simple requirement raises significant difficulties. In the first place, the nature of risk in financial markets means that, without further information about the firm's risk profile and hedging strategies, even the straightforward requirement to be able to quantify the expected return on an investment poses problems. Second, the funding activities of financial firms do not provide useful signals about the marginal cost of capital. This is because, for the majority of large and well-capitalised financial firms, the marginal cost of funds is indifferent to day-to-day changes in the degree of leverage or risk in their balance sheets. This, in turn, leads to a third problem, which is how to determine the amount of capital that the firm should apply to any particular investment. For a nonfinancial company, the amount of capital tied up in an investment can be more or less equated to the cost of its investment. However, in the case of a financial firm, where risk positions often require no funding at all, this relationship does not hold either.

It therefore follows that a financial firm that wants to maximise shareholder value cannot use the relatively straightforward capital pricing tools that are available to nonfinancial firms, and must seek an alternative shadow pricing tool to determine whether an investment adds to or detracts from shareholder value. This is the purpose that is served by allocating risk capital to the business areas within a financial firm.

III. Risk Measurement, Shadow Pricing, and the Role of the Sharpe Ratio

Since the objective of maximising shareholder value can be achieved either by increasing the return for a given level of risk, or alternatively by reducing the risk for a given rate of return, the internal shadow pricing process needs to be structured in a way that will assist management in achieving this objective. In other words, the shadow pricing tool has to have as its objective the maximisation of the firmwide Sharpe Ratio, since the Sharpe Ratio is simply the expression of return in relation to risk. Seen in these terms, we can draw a number of important conclusions that will assist us in determining how we should build our shadow pricing process.

First, and importantly, the shadow pricing process should operate in a manner that is independent of the level of equity capital in the firm. This follows because, where the perceived risk of bankruptcy is negligible, as is the case for most large financial firms, the Sharpe Ratio is independent of the amount of equity within a firm (see appendix). Thus, for any given set of assets, the amount of equity the firm has does not alter the amount of risk inherent in the assets, it merely determines the proportion of the risk that is assumed by its individual equity holders. Consequently, for any given level of equity, shareholder value can always be enhanced either by increasing the ex post rate of return for the given level of risk, or more importantly for a bank, which has little scope for significantly enhancing the earnings on its loan portfolio, by reducing the variance of those earnings through improved portfolio management.

Second, if the purpose of the process is to maximise the firm's Sharpe Ratio by encouraging risk-optimising behaviour, it has to capture all the important components of a firm's earnings volatility. The Sharpe Ratio that is relevant to the investor is simply the excess return on the firm's equity relative to the volatility of that return.

In ex post terms, this can be expressed as:

$$Sharpe\ Ratio_t = \frac{R_{pt} - R_{ft}}{\sigma_{pt}},$$

where

R_{pt} is the observed firmwide return on the investment in time t,

R_{ft} is the return on the risk-free rate at time t, and

σ_t is the standard deviation of R_{pt} measured at time t.

Management's objective at time t is therefore to maximise the expected Sharpe Ratio over the future period $t+1$. In order to do this, management has to be able to predict R_{pt+1} and σ_{t+1}. This means that we need to be able to understand both the components of $E(R_{pt+1})$ and the determinants of its variance, σ_{t+1}.

In a simple model of the firm, we can express $E(R_{pt+1})$ as follows:

$$E(R_{pt+1}) = E(\Delta P_{t+1} + Y_{t+1} - C_{t+1}),$$

where

$E(R_{pt+1})$ is the forecast value of earnings in time $t+1$,

ΔP_{t+1} is the change in the value of the firm's portfolio of assets in time $t+1$,

Y_{t+1} is the value of the firm's new business revenues in time $t+1$, and

C_{t+1} is the costs that the firm incurs in time $t+1$.

We can express $Var(R_{pt+1})$ as σ^2_{t+1}, so that by definition:

$$\sigma^2_{t+1} = \sigma^2\Delta P_{t+1} + \sigma^2 Y_{t+1} + \sigma^2 C_{t+1} + 2(Cov(\Delta P_{t+1}, Y_{t+1}) - Cov(\Delta P_{t+1}, C_{t+1}) - Cov(Y_{t+1}, C_{t+1})).$$

Because this is a forward-looking process, the firm cannot rely solely on observed historical values. It needs to be able to estimate their likely values in the future. The firm must therefore understand the dynamics of each of ΔP_{t+1}, Y_{t+1}, and C_{t+1}, and in particular the elements that contribute significantly to both their variance and covariance. These are the risk drivers of the business, which need to be identified and modeled if the firm is to have an effective shadow pricing process for its risk.

As a result of this approach, it is possible to think in terms of a generic risk pricing approach for maximising shareholder value, using generally agreed-upon risk pricing tools that could be applicable to all financial firms. Just as value at risk measures for market risk have become a common currency for comparing and analysing market risk between firms, a similar approach to other risk factors could readily be developed out of this model.

IV. Determining the Optimal Capital Structure for the Firm

As we have explained, there is no causal link between the level of gearing that a firm chooses and its Sharpe Ratio. However, this is subject to one important caveat, which is that the amount of equity capital that a firm holds has to be large enough to enable it to survive the "normal" variability of its earnings. This means that at the minimum, a firm will need to have some multiple of its expected earnings volatility—$(\sigma_{t+1})k$, where k is a fixed multiplier—as equity capital. Failure to maintain such an amount should lead to a risk premium on the firm's equity, which would make the cost of capital prohibitive. In most cases, though, management will choose to operate in some excess of this minimum level.

The question we therefore need to address here is how much equity capital in excess of $(\sigma_{t+1})k$ will a well-managed firm choose to hold, and how should it reach that decision?

Although by definition the amount of equity that the firm chooses will itself be a multiple of $E(\sigma_{t+1})k$,[2] the methodology for deciding how to set that amount needs to be significantly different from the methodology by which the shadow pricing amount σ_{t+1} is determined. This is so for three reasons. First, financial markets are prone to the characteristics of fat tails, which means that it is dangerous to rely solely on the properties of statistical distributions to predict either the frequency or the size of extreme events. Given that one of the responsibilities of the management of a financial firm is to ensure the continuity of that firm in the long term—which will in turn help to ensure that the perceived risk of bankruptcy is kept to a minimum—the firm needs to be able to analyse the nature of these rare events and ensure that the capital and balance-sheet structure are robust enough to withstand these occurrences and still be able to continue in business thereafter.

Thus, while in the case of certain risk factors the potential stress or extreme loss that the firm faces and needs to protect against may indeed be best estimated by an extension of the statistical measures used to calculate σ_{t+1}, in other cases the results of scenario analysis may yield numbers well in excess of the statistical measure. (The 1987 market crash, for example, was a 27 standard deviation event—well outside the scope of any value-at-risk measure.) As a result, statistical techniques that are applicable to a risk pricing process need to be supplemented with effective scenario and stress analysis techniques in order for management to assess the potential scale of the firm's exposure to such extreme events.

The second consideration in managing the firm's capital is how to optimise the firm's equity structure in an imperfect world. In theory, in the absence of any significant risk of bankruptcy, the market should be indifferent between different levels of leverage for firms with the same Sharpe Ratio, but it is not clear that this is the case. In particular, highly capitalised banks, which should have lower target returns on equity to compensate for their lower risk premia, appear to remain under pressure to provide similar returns on equity to more thinly capitalised firms.

Third, management has the additional requirement to ensure that it complies with regulatory capital requirements, set by reference to regulatory measures of risk, which often do not correspond with internal risk measures and in many cases conflict with them.

This means that one of the principal strategic considerations for management is to optimise the capital structure, bearing in mind the three different considerations of protecting the firm against catastrophic loss, meeting shareholder expectations, and complying with external regulatory requirements.

The essential requirement for this optimisation exercise is to ensure that the two following conditions are always met:

$$(\sigma_{t+1})k_i \leq Total\ Capital_i , \qquad \text{(Condition 1)}$$

where

$(\sigma_{t+1})k_i$ is the minimum level of capitalisation at which firm i can raise capital funds in the market for its given level of risk, and $Total\ Capital_i$ is the amount of capital that the firm actually holds

and

$$Regulatory\ Capital_i \leq Total\ Capital_i , \qquad \text{(Condition 2)}$$

where

$Regulatory\ Capital_i$ is the amount of capital that firm i is required to hold under the existing regulatory capital regime.

This formulation shows clearly why in a shadow pricing approach to risk, based on the calculation of σ_{t+1}, the amount of capital at risk and therefore being charged to the business is always likely to be less than the total capital of the firm.

Furthermore, from the perspective of the firm, the preferable relationship between these three considerations would also be such that

$$(\sigma_{t+1})k_w < Regulatory\ Capital_w < Optimal\ Capital_w , \qquad \text{(Condition 3)}$$

where

$Optimal\ Capital_w$ is the amount of capital that the firm would choose for itself in the absence of a regulatory constraint.

Where this condition can be met, the firm can concentrate solely on optimising its capital structure and maximising shareholder value without having to factor considerations about the impact of a regulatory capital regime into its optimisation exercise.

For completeness, we can also note here that the further necessary condition should exist from the regulatory perspective for any regulatory capital regime to be appropriately represented as risk-based, which is

$$(\sigma_{t+1})k_i \leq Regulatory\ Capital_i , \qquad \text{(Condition 4)}$$

so that the risk-based regulatory capital requirement is at least consistent with the market's assessment of the minimum amount of capital a firm should have in order to protect against the risk inherent in its business. This, in turn, by combining Conditions 2 and 4, leads us to the minimum requirement for a satisfactory regulatory capital regime that

$$(\sigma_{t+1})k_w \leq Regulatory\ Capital_i \leq Total\ Capital_i . \qquad \text{(Condition 5)}$$

We return to this issue, and in particular the relationship between the regulatory requirements and optimal capital structure for the firm in more detail in Section VI.

V. Risk Measurement—The Challenge of Normalisation

Now that we have distinguished between the different purposes of risk measurement for shadow pricing of risk and for the determination of the optimal capital structure, we can move on to consider the challenges of building an effective risk measurement system. The objective here is to enable management to assess the different risks that a firm faces in a broadly similar fashion, and to understand their interrelationships. This requires both a common measurement framework and a methodology for ensuring that the risk process covers all the material risks that may impact the shadow pricing process or the decisions about the capital structure.

At the outset, a firm has to have a clear understanding of the meaning of risk if it is to develop an effective risk measurement methodology. For the purposes of this paper, we can define the risk in a firm on an ex post basis as the observed volatility of the firm's earnings over time around a mean value. The firm's risk measures are thus the firm's best estimates of that volatility, which management can then use to make choices between different business strategies and investment decisions and to determine the firm's capital structure.

In order to achieve this, it is necessary to distinguish between the three measures of expected, unexpected, and stress loss as follows.

The *expected loss* associated with a risk factor is simply the expected value of the firm's exposure to that risk factor. It is important to recognise that expected loss is not itself a risk measure but is rather the best estimate of the economic cost of the firm's exposure to a risk. The clearest example of this at present is the treatment of credit risk, where banks know that over the credit cycle they will incur losses with a high probability, but only account for those losses as they occur. This introduces a measure of excessive volatility into the firm's reported earnings, which is not a true measure of the "risk," given that the defaults are predictable with a high degree of confidence. The true risk is only that part of the loss that diverges from the expected value.

Having established the expected loss associated with a risk, it is then possible to measure the variance of that cost in order to establish the extent to which it contributes to the overall variance of the firm's earnings, which we term the *unexpected loss* associated with the risk factor. Both VaR for market risk and the credit risk measures produced by CreditMetrics and CreditRisk+ are examples of measures of unexpected loss that can be used in an internal risk pricing process of the type discussed in Section III. However, comparison of these two approaches also points up the significance of adopting different time horizons in measuring different risks.

VaR measures for market risk are typically either a one-day or ten-day measure of risk. By contrast, the modeling of default risk, which is still at an early stage of development, typically utilises an annual observation period, since default frequencies change over a much longer time horizon than market prices. As a result of these different time horizons, a ten-day 99 percent confidence interval for market risk would imply that the VaR limit could be expected to be exceeded once every three years. An annually based VaR of 97.5 percent for credit risk, however, would be expected to be exceeded only once every forty years. Aggregating the two measures into a single measure of the firm's risk—even assuming for the moment that the firm's market and credit risk were independent—would not provide a satisfactory indication of the aggregate risk that the firm faces.

A further problem with the estimation of unexpected losses is the availability of reliable data for the different risk factors that a firm faces. Significant progress has been made on measuring market risk because of the availability of daily data for prices and for revenues within firms, and more recently progress has also been made on modeling credit risk, although here the data quality problem is proving more challenging. In the case of other risk factors such as liquidity, legal, and operational risks, however, the analysis is likely to have to rely on firms' own internal data, and very little work has yet been undertaken to examine the statistical properties of those risks. Moreover, meaningful estimates of the covariances between risk factors will only be possible once reliable estimates can be made of unexpected loss on a stand-alone basis.

In addition to the need to develop expected and unexpected loss measures, which are particularly relevant to the firm's risk pricing methodology, the firm also has to have a methodology for determining the extreme or *stress loss* that it might face over the longer term horizon as a result of its exposure to a risk factor in order to make meaningful decisions about its capital structure and risk limits systems. A number of risk measures and limits, such as the concentration limits that banking regulators use to limit the proportion of a bank's capital that can be at risk to any one counterparty, are derived explicitly or implicitly from this type of measure. The methodology that a firm may choose for calculating the potential stress loss associated with a particular risk will vary from risk factor to risk factor, but will typically consist of a form of scenario simulation, which envisions the type of situation where the firm could potentially be put at risk from a particular risk factor, or a combination of factors, and then assesses the firm's capital resources and limits structures by reference to the results of this exercise.

Given that the purpose of measuring risk is to estimate the exposure of the firm to earnings variability from its principal risk drivers, the firm also needs to have a factor model that identifies the key risk factors to which it is exposed and measures their impact on the volatility of the earnings stream. The issue we now need to address is, What are these risk drivers and how can they be measured effectively?

In order to establish a starting point for this exercise, we can use the 1994 Basle Committee paper on risk management for derivatives, which identified six risks that firms face—market risk, credit risk, settlement risk, liquidity risk, legal risk, and operational risk. If we relate this list back to the shadow pricing equation in Section III, we can readily see how much still remains to be done in establishing an effective internal risk pricing process.

As we discussed in Section III, firms have started this process by analysing their trading exposure to *market risk*, which is where the data are most readily available. It is interesting to note, however, that even in the context of market risk, few firms are yet able to measure their overall revenue exposure from areas such as corporate finance or funds management to movements in market variables, even though these may be significantly more powerful factors in determining the quality of their earnings in the medium term, not least because the time horizons are different.

In a manner similar to their work on market risk, firms have turned their attention more recently to the issues associated with the measurement of the unexpected loss associated with *credit risk*. Work in this area derives from two parallel initiatives. On the one hand, there has been increasing interest, stimulated in considerable part by the Basle Committee's model-based approach to capital requirements for market risk, in developing models of the specific risk in the trading book. On the other hand, there has been an increasing effort to develop reliable models for measuring the default risk in the banking book.

The third category of risk identified in the 1994 paper in the context of derivative products was *settlement risk*. In practice, settlement risk is a special case of credit risk, since it arises from the failure of a counterparty to perform on a contract. Its particular characteristic is that it arises on a daily basis as transactions—particularly in foreign exchange and payments business—are settled, and the magnitude of the daily exposure between different financial institutions in relation to settlement risk is many times larger than for other risk factors. The primary challenge for a financial firm is therefore to be able to capture and monitor its settlement risk in a timely manner. Once this has been done, the same methodology for measuring expected and unexpected loss can be applied to settlement risk as for other types of credit risk.

To date, the techniques for measuring *liquidity risk* have tended to focus on the potential stress loss associated with the risk, whether in the form of the cash capital measure used by the U.S. securities firms or the funding gap analysis undertaken by bank treasuries. Both are attempts to quantify what might occur in extreme cases if the firm's funding sources dried up. While this is clearly a prudent and desirable part of corporate financial management, it should also be possible to apply the framework of expected and unexpected loss to liquidity risk by measuring the extent to which the liquidity risk inherent in the business gives rise to costs in hedging out that risk through the corporate treasury function.

In a similar way to the approach to liquidity risk, the focus to date in analysing the impact of *legal risk* and other aspects of *operational risk* has been in seeking to prevent the serious problems that have given rise to the well-publicised losses, such as those of Hammersmith and Fulham in the context of legal risk, or those of Barings and Daiwa Bank in the context of operational risk more generally. As with liquidity risk, however, the issue that has yet to be addressed in the context of internal risk pricing is how these risk factors contribute to the earnings volatility of the firm, since operational risk can be seen as a general term that applies to all the risk factors that influence the volatility of the firm's cost structure as opposed to its revenue structure. It is therefore necessary for the firm to classify and analyse more precisely the nature of these risk factors before any meaningful attempt can be made to fit them into a firmwide risk model of the type envisaged by this paper.

As the foregoing analysis indicates, a considerable amount of further work clearly still remains to be undertaken in the development of risk modeling in financial firms. Nevertheless, despite the evident gaps in the development of a full risk model, this does not preclude proceeding to implement a risk pricing methodology for those risks that can be measured. This is because with risk pricing there is no presumption that the risk measures should add to the total capital of the firm, and thus there is no danger of misallocating capital to the wrong business, which can occur if a risk-based capital allocation model is used with an incomplete risk model. Given this fact, the integrity of the risk measure for the particular risk factor is the primary consideration, and the need for a strict normalisation of risk measures—so that the measures for each risk factor can be aggregated on a consistent "apples for apples" basis—assumes a lesser importance as an immediate objective.

VI. Risk Allocation Methodologies and Regulatory Capital Requirements—A Synthesis?

Having outlined the components of an integrated approach to risk pricing and capital optimisation within financial firms, we can now consider the implications of this analysis for the structure of a satisfactory regulatory capital framework. In this context, we do not seek to analyse the different rationales for capital regulation, but simply note that it is now widely accepted that any regulatory capital requirement should be risk-based and should be consistent with firms' own internal risk measurement methodologies, so that a firm that carries more risk is subject to a higher capital requirement than one that carries less risk.

As we have explained, the core objective of a firm's own internal risk pricing mechanism should be to enhance shareholder value by encouraging behaviour that will improve the firm's overall Sharpe Ratio. In normal circumstances, this will be separate from the process of determining the optimal capital structure for the firm. The difference between the two is that the risk pricing exercise is based on a measure of unexpected loss and is designed to operate at the margin, at the level of the individual business decision. The decision on the capital structure should, by contrast, be based on an assessment of stress loss scenarios and be independent of activity at the margin, leading to the minimum capital condition that, identified in Section III, that

$$(\sigma_{t+1})k_i \leq Total\ Capital_i . \qquad \text{(Condition 1)}$$

In Section III, we also derived the following minimum condition, which we believe should be satisfied in order to characterise a regulatory capital regime as adequately risk-based

$$(\sigma_{t+1})k_i \leq Regulatory\ Capital_i \leq Total\ Capital_i , \qquad \text{(Condition 5)}$$

and we identified the desirable condition for a well-managed and well-capitalised firm that

$$(\sigma_{t+1})k_w < Regulatory\ Capital_w < Optimal\ Capital_w . \qquad \text{(Condition 3)}$$

We can now assess how these requirements compare under three alternative approaches to setting regulatory capital requirements, which can be summarised as follows:

- the fixed ratio approach (Basle 1988/CAD/SEC net capital rule)
- the internal measures approach (Basle market risk 1997/Derivatives Policy Group proposals)
- the precommitment approach.

The *fixed ratio approach* calculates the required regulatory capital for a financial firm by reference to a regulatory model of the "riskiness" of the firm's balance sheet. The problem associated with any regime of this sort, which seeks to impose an arbitrary measure of the riskiness of a firm's business on a transaction-by-transaction basis, is that there is no mechanism for testing it against the true risk in the firm, which will by definition vary from firm to firm. As a result, the only part of Conditions 3 and 5 that this approach can satisfy a priori is that

$$Regulatory\ Capital_i \leq Total\ Capital_i ,$$

which is achieved by regulatory requirement. But Condition 1 is violated because we cannot be sure that

$$(\sigma_{t+1})k_i \leq Regulatory\ Capital_i$$

and equally, there is no way of ensuring for a well-managed firm that Condition 3 can be met because there is no mechanism for ensuring that

$$Regulatory\ Capital_w < Optimal\ Capital_w .$$

Given these flaws, it is difficult to see how a fixed ratio regime could realistically be adapted to meet our conditions for an optimal capital structure.

By comparison with the fixed ratio approach, the *internal models approach* is clearly preferable from the viewpoint of the well-managed firm, since it seeks to equate regulatory capital to

$$(\sigma_{t+1})m,$$

where m is the regulatory multiplier.

If we assume that m is set at a level that is higher than k (the minimum capital requirement for a viable firm) but at a level that is still economic, it is likely that the well-managed firm will be able to live with this regime, provided it has a sufficient margin of capital between $(\sigma_{t+1})m_w$ and $Optimal\ Capital_w$.

However, it is questionable whether such a "full models" regime is genuinely optimal, or could be introduced quickly, since neither the industry nor the regulators are yet able to define the model that determines σ_{t+1} for the whole firm. Consequently, a decision to use a full models approach for regulatory capital purposes would commit both regulators and financial firms to a significant investment of resources, with an indeterminate end date, and would at the same time provide no assurance that the outcome was superior to a simpler and less resource-intensive approach.

The *precommitment approach*, by contrast with either the fixed ratio or internal models approach, has the attraction of simplicity and synergy with the firm's own processes since it allows firms to determine their own capital requirement for the risks they face. If the regulators are able to ascertain that the firm's internal procedures are such as to ensure that

$$(\sigma_{t+1})k_i \leq Total\ Capital_i$$

with sufficient margin to satisfy the regulatory needs for capital, then precommitment in its most complete sense has the simple result that

$$(\sigma_{t+1})k \leq Total\ Capital_i \equiv Regulatory\ Capital_i,$$

which satisfies the requirements of our three conditions.

However it is questionable whether a full precommitment approach, as outlined, can be defined as a regulatory capital regime at all. It would probably be better described as an internal controls regime, since in substance it would mean that the regulator would review the methodology whereby the firm undertook its risk pricing and capital structuring decisions and would either approve them—allowing precommitment—or impose a capital requirement if they were not satisfied with the process. In addition, the regulatory authority would be susceptible to criticism, in the event that a problem was encountered at a firm that had been allowed to employ the precommitment approach, that it had unnecessarily foregone an important regulatory technique.

Given the evident problems of a move that is as radical as the precommitment proposal, we therefore believe that it is worthwhile to consider a fourth approach, which we refer to as the *base plus* approach. Under this approach, the regulator would determine directly on a firm-by-firm basis the regulatory capital requirement for the forthcoming period as an absolute amount, say R_{t+1}, based on some relatively simple rules such as a multiple of the firm's costs or revenues in the previous year, and modified to take account of the risk profile of the firm. The basis for setting this requirement should be clearly defined, and would need to be sufficient to ensure that the condition for the well-managed firm was met such that

$$(\sigma_{t+1})k_w < Regulatory\ Capital_w < Optimal\ Capital_w.$$

However, in order to prevent the firm from exploiting this fixed capital requirement by changing its risk profile after the capital requirement was set, the firm would also be required to supplement its regulatory capital by a precommitment amount that should be sufficient to cover the amount that its risk profile changed during the reference period.

The advantage of this approach would be that it would be simple from the firm's perspective, it would require relatively little detailed assessment by the regulator of the firm's own internal models regime, and would not be conditional on the firm having modeled every material risk before it took effect. At the same time, it could have incentives built in, since the more confident the regulator was about the quality of the firm's internal controls the lower could R_{t+1} be set, while still leaving the regulator the ultimate authority to ensure that all firms were capitalised at a level sufficiently in excess of $(\sigma_{t+1})k$ to protect the overall system against the risk of extreme systemic events. From the perspective of the firms, the fact that additional capital was required at the level of changes in $(\sigma_{t+1})k$ and not based on a higher multiplier would ensure that the regulatory regime remained in line with the requirements of the internal risk pricing, so avoiding the risk of regulatory arbitrage arising from inappropriate capital rules.

VII. Conclusion

It is becoming increasingly clear that the regulatory capital requirements for both banks and securities firms are not appropriately aligned either with the risk that those firms are taking or with the way in which those firms manage their own risks in order to maximise shareholder value and optimise their capital structures. In this paper, we have argued that this process has two elements. Internal risk measures such as value at risk can be used by financial firms as a means of enhancing shareholder value by targeting directly the firmwide Sharpe Ratio rather than through the indirect mechanism of internal capital allocation. However, we argue that these measures of unexpected loss need to be supplemented by techniques such as scenario analysis when assessing the firm's potential exposure to stress loss and thus determining the firm's optimal capital structure.

In light of these considerations, we do not believe that any of the current proposed regulatory capital regimes, which we characterise as the regulatory ratio approach, internal models approach, and the precommitment approach, are consistent with this account of risk pricing and capital optimisation within firms. By contrast, we believe that our proposal for a base plus approach to regulatory capital would be consistent with both regulatory objectives and firms' own internal processes, and as such would provide a sound basis for a regulatory capital regime for financial firms in the twenty-first century.

Appendix: The Scale Independence of the Sharpe Ratio

1. Definitions:

   I Arbitrary Amount of Investment

   F Financing Amount of Investment I

   C Capital Allocated to Investment I

   Such that:

   $I = F + C$.

   (This is merely a restatement of an accounting fact that assets = liabilities.)

   Further:

   $Exp(P)$ Expected Profits from Investment I net of direct and allocated indirect costs before funding

   $Exp(P_{net})$ Expected Net Profits, that is, profits after funding costs

   $Exp(R)$ Expected Return (percent) on (arbitrary amount) Capital Allocated to Investment I,

   where:

   $$Exp(R) = \frac{Exp(P_{net})}{C}.$$

   Vol_P Volatility of Profits

   Vol_R Volatility of Return on Equity

   r_f the Default Free Interest Rate

   In its simplest form, the Sharpe Ratio is defined as the excess return of an investment over the standard deviation of the excess return. If we assume that interest rates are fixed over the time horizon of the investment, then the volatilities of returns and of excess returns are the same.

2. First Result:

   Many activities in banking effectively require little or no investment at the outset (if regulatory capital requirements are neglected for a moment), such as swaps and futures. For this reason, we choose to start with an absolute revenue-based Sharpe Ratio and extend it to a relative (percent) measure in a second step.

   The excess profits over the risk-free rate of interest for capital and after any refinancing costs are given by:

   $$Exp(P) - r_f F - r_f C,$$

   and the Sharpe Ratio therefore by

   $$\frac{Exp(P) - r_f F - r_f C}{Vol_P} = \frac{Exp(P) - r_f(F + C)}{Vol_P}$$

   $$= \frac{Exp(P) - r_f I}{Vol_P} = \frac{Exp(P_{net}) - r_f C}{Vol_P}.$$

   The Sharpe Ratio of the Expected Revenues is thus given by the profits net of the costs for full (that is, 100 percent) refinancing over the volatility of earnings.

3. Second Result:

   If return is measured as the ratio of absolute return to allocated capital (which can be an arbitrary amount), then the following result holds for volatilities:

   $$Vol(Return) = Vol\left(\frac{P}{C}\right) = \frac{1}{C}Vol(P).$$

   This simple result obviously guarantees that the Sharpe Ratio does not change its value since both the numerator and the denominator are scaled by the same amount. A closer examination of the above formula, however, gives some intuition for this result

   $$\frac{\frac{Exp(P) - r_f I}{C}}{Vol(R)} = \frac{\frac{Exp(P) - r_f I}{C}}{Vol\left(\frac{P}{C}\right)} = \frac{\frac{Exp(P)}{C} - r_f\left(\frac{F}{C} + 1\right)}{\frac{1}{C}Vol(P)}.$$

   Apart from the fact that the C cancels out, one can see that the higher the leverage the higher the expected return on the one hand, but the higher also the volatility of the returns, which leaves the Sharpe Ratio unchanged.

4. Conclusion:

   As long as the institution can refinance itself at approximately the risk-free rate, or its refinancing rate is indifferent to changes in volatility over the relevant range, the amount of capital that it allocates to the business will not affect its Sharpe Ratio. This can be seen by solving the Sharpe Ratio backwards for some (arbitrary) capital allocation C:

$$\frac{Exp(R) - r_f}{Vol(R)} = \frac{\frac{Exp(P_{net})}{C} - r_f}{Vol\left(\frac{P}{C}\right)} = \frac{\frac{Exp(P)}{C} - r_f\left(\frac{F}{C} + \frac{C}{C}\right)}{\frac{1}{C}Vol(P)}$$

$$= \frac{\frac{Exp(P)}{C} - r_f\frac{I}{C}}{\frac{1}{C}Vol(P)} = \frac{Exp(P) - r_f I}{Vol(P)}.$$

Of course, this whole relationship changes as soon as the marginal cost of funding becomes a function of the credit quality of the institution. In that case, the costs of funding become an increasing function of the volatility of the profits (or returns) and, as a consequence, the Sharpe Ratio drops.

It is for this reason that the absolute level of capital in banks is held at some multiple of the volatility of the earnings, since this ensures that the cost of funding at the margin remains independent of day-to-day changes in the risk profile of the firm.

Endnotes

The authors are grateful to Marcel Rohner of Swiss Bank Corporation for his contribution to the development of this paper and for providing the appendix.

1. This is borne out by the experience of the recent precommitment pilot study and by the value at risk returns provided by members of the Derivatives Policy Group in the United States to the Securities and Exchange Commission.

2. Strictly, we should denote our risk term as $E(\sigma_{t+1})_t$—that is, expected value at time t of the standard deviation of earnings at time $t+1$. For ease of notation, however, we adopt the term σ_{t+1} for the rest of this paper.

References

Froot, Kenneth A., and Jeremy C. Stein. 1997. "Risk Management, Capital Budgeting and Capital Structure Policy for Financial Institutions." Unpublished paper, February.

McQuown, J. A., and S. Kealhofer. 1997. "A Comment on the Formation of Bank Stock Prices." KMV Corporation, April.