The relationship between risk capital and required returns in financial institutions: some preliminary results*

Alistair Milne† and Mario Onorato‡

April 2007

Abstract

We use arbitrage arguments to characterise the relationship between required shareholder returns and the exposure to downside risk, as measured by VaR at stated confidence level and time horizon. We show that skewness and diversification have a major impact on zero-NPV RAROC (return on capital) hurdles, implying that use of return on capital with a constant hurdle rate can lead to substantial loss of shareholder value. We propose an alternative performance measure consistent with standard valuation models and discuss implications for prudential regulation.[93 words]

Journal of Economic Literature number: G22, G22, G31

Keywords: Risk Management, Downside Risk, RAROC, RORAC, Economic Capital, Capital Allocation, Risk-Return Tradeoffs, Performance Measurement.

---

*The views expressed here are not necessarily those of the Bank of Finland or of Algorithmics. We are grateful for comments from Esa Jokivuolle, audiences at Cass Business School, the University of Lugano, Bocconi Univerity, the 2007 Chicago Enterprise Risk Management Workshop, and from two anonymous referees. Any remaining errors are our own responsibility.

†Corresponding author, Faculty of Finance, Cass Business School, City University, London; and Monetary Policy and Research Department, Bank of Finland, Helsinki. email: `amilne@city.ac.uk`

‡ Mario Onorato, Director, Algorithmics Inc. & Faculty of Finance, Cass Business School, City University, London. e-mail: `monorato@algorithmics.com`

1 Introduction

A majority of financial institutions in the developed world – including all the largest and most internationally active – now use models of risk capital for assessing risk-return tradeoffs.[1] This practice is often described as 'economic capital management' or 'capital allocation'. The best known capital based performance measure is 'return on economic capital' (RAROC) comparing expected revenues on a particular exposure with its contribution of the exposure to an institution wide downside (VaR) risk.[2]

RAROC and related measures are now the most widely used of all tools for assessing investment in financial assets;[3] but there is little available analysis of these measures from the perspective of standard financial economics.[4] This paper addresses this gap in the literature, providing a complete discussion of the relationship between the zero-NPV RAROC hurdle

[1]Consultancy companies are a good source of information on these methods, see in particular KPMG (2004) and PWC-EIU (2005). For a published collection of recent practitioner views see Dav (ed) (2006).

[2]Matten (2000, pp 146-166) describes RAROC and alongside several related performance measures. The various acronyms (RAROC, RORAC, RARORAC, etc.) are not applied by practitioners in a consistent manner. RAROC is the most common acronym for the the most commonly used measure, the one that we analyse; but we warn readers that this same measure is frequently referred to by other names and acronyms, and that RAROC is also sometimes applied to different performance measures.

[3]Smithson (2002), page 266, reports that 78% of the respondents to his 2002 Rutter Associates survey of credit portfolio managers, used RAROC to evaluate the performance of their portfolio of credit assets. PWC-EIU (2005), covering more than 200 medium sized and large banks and insurance companies worldwide, finds that more than half now conduct such capital allocation and most use the resulting return measures for various purposes, including business decision making, product pricing, and the determination of bonuses "economic capital is fast gaining critical mass within the industry". A more recent 2006 update of this survey (not yet available on the web) shows even greater adoption. These tools are not confined to banks and insurance companies: asset managers also make widespread use of return on risk capital as a performance measure when acting on behalf of both retail and institutional investors.

[4]While several papers consider performance measurement in financial institutions, as far as we are aware only one previous study (Crouhy et. al. (1999)) addresses the correspondence between risk capital and required returns, and does not cover all aspects of this relationship.

numerator (required returns) and denominator (risk capital).[5] The remainder of this introduction provides a short summary of the literature on capital and performance measurement in financial institutions and an overview of our own analysis.[6]

Several contributions discuss the private and social benefits of capital as a protection from default and develop models of the risk capital needed to reduce the probability of default to some desired level.[7] Previous studies also consider the use of performance measures, in both non-financial and financial companies;[8] and consider why conventional performance measures used by non-financial corporates are not always appropriate in financial institutions.[9]

An issue very close to that which we address is whether for performance measurement risk should be measured relative to the financial instution's own portfolio or relative to the market as a whole. Froot and Stein

[5]We avoid using the term 'economic capital', despite its widespread currency amongst practitioners. This usage assumes without justification that comparisons of expected return and downside risk provide a consistent measure of contribution to shareholder value. We instead prefer the term 'risk capital' to refer to the amount of capital required to protect the institution against insolvency to some target tolerance threshold.

[6]A fuller survey is provided by Schroeck (2002).

[7]The role of capital in default protection is discussed for example in Berger et. al. (1995). Gordy (2000) provides an overview of models of risk capital. As noted by Zaik et. al. (1996) models of risk capital can help optimise capital structure i.e. find the proportion of equity to assets that minimizes the bank's cost of funding, as well as for risk-return assessment.

[8]Zechner and Stoughton (2003) note that a key reason for using performance measures for enterprise management is to enable the delegation of decision making within a large organisation, when managers responsible for investment decisions have privileged information. They develop a model, drawing on the literature on capital decisions in non-financial companies, showing that a RAROC performance measure can overcome the information asymmetries between divisional managers responsible for portfolio decisions and central bank management. Their model however assumes that risks are normally distributed, a special assumption that allows a single measure of risk to be used for both required returns and downside risk.

[9]Merton and Perold (1993) emphasize that performance measurement in financial institutions must take account of the costs of leverage for banks. These are different from those for industrial companies because bank customers are often also their largest liability holders and as a consequence, a high credit rating is generally essential for banks to maintain their business activities, e.g. as dealers or customers in OTC markets, to underwrite securities or to compete effectively in the corporate banking and deposit markets. A further reasons for using capital allocation as a performance measure is that this method can be easily extended to the important non-funded off-balance sheet exposures of financial institutions, where the familiar method of internal rate of return cannot be applied.

(1995) point out that, faced with an increasing cost of raising external funds banks will behave in a risk-averse fashion towards risks that are diversifiable at a market level. Specifically, a business unit's contribution to the earnings volatility of the bank will be an important factor in the capital allocation and capital structure decisions and also in the decision to hedge earnings risk. Capital structure, hedging and capital budgeting are therefore inextricably linked together.[10]

Froot and Stein (1998) further develop this point, demonstrating in a two period model that the hurdle rate for bank investments can be calculated from a two factor pricing model, namely the covariance of the return on the tradable component with the market R_m and the correlation on the non tradable component of the new risk $\left(\mu_i^N \right)$ with the non-tradable risks of the existing portfolio:

$$\mu_i = \gamma \text{cov}\left(\mu_i, R_m\right) + \lambda \text{cov}\left(\mu_i, R_P\right)$$

where γ is the market unit price of risk for the (market) priced factor R_m and λ is the unit cost for volatility of the banks portfolio.

A different point, less closely related to our own discussion, is made by Stoughton and Zechner (2003), is the optimal method of capital allocation when there are several business units. They show that the appropriate measure of economic capital for performance measure should depend on each unit's incremental contribution to total portfolio Value at Risk (the "IVaR"). This can be defined in such a way that sum of the IVaRs is equal to the institution's overall VaR.

Finally the literature address the relationship between capital allocation and required regulatory capital. Misalignment of economic and regulatory capital is thought to have distorted business decisions and encouraged the use of financial transactions such as securitizations to reduce regulatory capital without altering bank exposure to downside risks.[11] A stated goal of the new Basel II accord on bank capital has been to achieve a closer alignment of regulatory capital with economic capital and so reduce these incentives.[12] Capital allocation is further promoted by a key further principle of the new

[10]See also Stulz (1998)

[11]See Jones (2000) for illustration of this practice of "Regulatory Capital Arbitrage".

[12]The Basel committe writes (Basel Committee (1999), page 11) "...during the 1990s the [1988 Basel] Accord became an accepted world standard, with well over 100 countries applying the Basel framework to their banking system. However, there also have been some less positive features. The regulatory capital requirement has been in conflict with increasingly sophisticated internal measures of economic capital....In addition the accord does not sufficiently recognise credit mitigation techniques such as collateral and guarantees. These are the principal reasons

capital regulations, the so called 'use test'.[13] A number of studies compare these various measure of capital, including Jokivuolle (2006) who reviews the approach of the new accord from the perspective of capital allocation; and Elizalde and Repullo (2006) who compare regulatory capital computed by the Basel II IRB risk curves with the capital chosen by banks (both with and without capital regulations) in the context of a simple dynamic model of banking risks.[14]

Our paper is organised as follows. Section 2 states our theoretical results. Proposition 2 shows that in order for RAROC to be applied with a single zero-NPV hurdle rate to different exposures, it is both necessary and sufficient that downside tail risk and the market valuation of risk bear a constant relationship to one-another. Proposition 3 provides a further intuition, indicating that RAROC can be applied if all different exposures have the same degree of skewness. These results are derived under the assumption of unlimited shareholder liability. Appendix 1 extends the analysis, showing that the same propositoins apply in the case of a financial institution with limited shareholder liability and with 100% deposit insurance (i.e. incorporating the assumptions of Merton (1974) and of Merton (1977)).

Section 3 illustrates that differences in skewness make a substantial quantitative difference to zero-NPV RAROC hurdle rates. We begin with two standard return distributions commonly applied to market risk, the arithmetic normal and log-normal, finding (Figure 1) that the required return on risk-capital is unaffected by portfolio variance in the case of the arithmetic normal, but that because of the right-skew it is increasing in the case of the log normal. This is consistent with the results of Crouhy et. al. (1999). We then examine the required return on risk capital for credit portfolios obeying the Vasicek asymptotic default distribution, the distribution underlying the Basel II pillar 1 risk curves, showing (from comparison of Figure 1 and Figure 2) that for good quality corporate exposures the required return on risk capital is about one quarter of that appropriate for equity market portfolios. Section 4 discusses the implications of our analysis for both financial institution management and regulators.

---

why the Basel committee decided to propose a more risk-sensitive framework in June, 1999."

[13]In order to qualify for the IRB method for credit risk and the AMA method for operational risk under Pillar 1 of the new Basel accord, the underlying systems must be applied by banks to their business decision making, not just used for regulatory compliance. A similar use test will apply for the recognition of advanced modelling methods in the forthcoming European Solvency II insurance regulations.

[14]Elizalde and Repullo (2006) use exactly the same model of credit risks as we apply in our section 3.3.

2 The zero-NPV RAROC hurdle: theoretical analysis

This section analyses the conditions under which returns required by shareholders on a bank exposure are proportional to its contribution to risk capital – a measure of the amount of additional capital required to protect the institution against insolvency to some desired target threshold probability – i.e. when a RAROC equation of the following form:

$$r_{RC} = \frac{\text{Expected Net Revenues}}{\text{Risk Capital}} \tag{1}$$

can be used with a constant (zero-NPV) hurdle rate in order to compute the contribution of an exposure to shareholder value.

The results reported here are based on standard arbitrage arguments used in asset and derivative pricing. The analysis will appear somewhat abstract, but this is in order to be as general as possible and demonstrate that the divergences between RAROC and shareholder value we document are not driven by the particular return distributions and valuation model used for illustration in Section 3.

The first subsection states our assumptions. The second subsection considers the case of unlimited shareholder liability, proceeding as follows. First we derive an expression for the market value of equity finance (the market value of risk capital) $\hat{E}(0)$ that must be supplied by shareholders in order to keep the probability of insolvency to a target level of p^* and an expression for the expected return on this market measure of risk capital $\hat{r}_{RC}$. We then establish (Proposition 1) that $\hat{r}_{RC}$ is the zero-NPV hurdle rate for equation (1), where r_{RC} the return on risk capital measured on an accounting rather than market value basis (any exposure earning a higher return on accounting risk capital than this threshold will have a market value that exceeds its accounting value (cost of acquisition) and so will create shareholder value; any exposure earning a lower return will destroy shareholder value).

We identify (in Proposition 2) necessary and sufficient conditions for the zero-NPV RAROC hurdle rate $\hat{r}_{RC}$ to be a constant for comparison of across a set of different investment opportunities. Finally we provide (in Proposition 3) a more intuitive sufficient condition for the use of a constant zero-NPV hurdle performance evaluation of different bank exposures, namely that all distributions are mean-preserving spreads of a representative distribution i.e. RAROC is a valid performance measure if all distributions have the same degree of skewness.

An appendix demonstrates that Propositions 2 and 3 continue to apply in the case of limited shareholder liability (exactly), and 100% creditor protection (to a close approximation). For concreteness the analysis is developed for the case of a bank, although it could equally well be applied to another financial intermediary such as an insurance company.

2.1 Notation and modelling assumptions

Consider a bank considering the choice of whether or not to invest in a loan asset or portfolio with a market value of $\hat{A}_i(0)$ at time $t = 0$. This asset is one of a number of potential assets i.e. $i \in (1, 2, \ldots, I)$. To avoid the need to discuss portfolio diversification, we make the further assumption that the bank invests only in a single asset (the implications of diversification are discussed in Section 3.2). We therefore drop the superscript i until Proposition 2, when we explicitly consider the comparison between assets. We use a 'hat' to distinguish market measures of assets ($\hat{A}(0)$) and also net worth (capital) ($\hat{E}(0)$) and return on risk capital ($\hat{r}_{RC}$) from their corresponding accounting measures ($A(0)$, $E(0)$ and r_{RC}).

If it proceeds, the bank finances this investment by issuing debt with a market (and in this case also accounting) value of $D\,(t)$. Since we assume shareholders are subject to unlimited liability i.e. this debt is risk-free and the bank market value balance sheet is:

$$\hat{A}(0) = \hat{E}(0) + D(0) \tag{2}$$

Consider now period $t = 1$ returns. Becuase of unlimited shareholder liabliity the debt is risk free so (with a risk free rate of interest is r_f) we have $D(1) = D(0)\,(1 + r_f)$

End period asset returns $A\,(1)$, net of all costs, are uncertain and continuously distributed. We will assume that all risks are tradeable in liquid markets and the returns on the asset are a (possibly non-linear) function of one aggregate priced market risk factor z.[15] $A\,(1) = R_A + \alpha(z) + \gamma\phi$ where ϕ is the specific asset risk. z and ϕ are independently distributed with the bivariate joint density function $f(z)g(\phi)$ with $\int_{-\infty}^{+\infty} zf(z) = \int_{-\infty}^{+\infty} \phi g(\phi) = 0$ and $\int_{-\infty}^{+\infty} z^2 f(z) = \int_{-\infty}^{+\infty} \phi^2 g(\phi) = 1$. It is convenient to write the random component of asset returns as $w = \alpha(z) + \gamma\phi$ with density function given by:

$$h(w) = \int_{\infty}^{\infty} f(\alpha^{-1}(w - \gamma\phi))g(\phi)\mathrm{d}\phi \tag{3}$$

with corresponding cumulative density $H(w) = \int_{\infty}^{w} h(w)\mathrm{d}w$.

[15] The assumption of a single pricing factor is inessential (we could instead work with an arbitrary number of n priced factors) but simplifies our exposition.

We assume markets are complete, implying (as discussed in any asset pricing textbook e.g. Cochrane (2005)) that the market value of the bank asset can be expressed as a function of the aggregate (priced) risk factor:

$$\hat{A}(0) = (1 + r_f)^{-1} \left(R_A + \int \alpha(z) q(z) \mathrm{d}z \right) < (1 + r_f)^{-1} R_A \tag{4}$$

where $q(z)$ is the risk-neutral probability density for z and $(1 + r_f)^{-1} q(z)$ is the linear pricing function (pricing kernel). If investors are risk averse $\int \alpha(z) q(z) \mathrm{d}z < 0$ and so as indicated by the inequality in (4) this valuation is less than the expected return discounted at the risk-free rate of interest.

As for accounting valuations, we suppose that the bank asset requires total funding of $L(0)$ (in the case of a lending operation $L(0)$ corresponds to the book acquisition value the loan portfolio, but the funding requirement can equal zero, or even be negative if the bank is undertaking business such as the writing of options which generates a positive cash flow). The usual accounting convention applies where the loan asset is valued at the cost of acquisition ($L(0)$) not its market value ($\hat{A}(0)$). The funding provided by equity holders (the book value of equity) is $E(0) = L(0) - D(0)$

Note that the investment creates value if $\hat{E}(0)) = \hat{A}(0) - D(0) > L(0) - D(0) = E(0)$ i.e. stated simply – an exposure creates value if its market value (expressed in terms of the either the underlying assets or the shareholder equity) exceeds its acquisition value. A marginal project, one that neither adds to nor subtracts from shareholder value, is one where $\hat{E}(0)) = E(0)$ or equivalently $\hat{A}(0) = L(0)$.

2.2 The case of unlimited shareholder liability

The first stage of the analysis is deriving the expected return on the bank's risk capital, measured on a market value rather than on an accounting basis. Note that the bank is insolvent if $A(1) = R_A + w < D(1)$. The bank chooses its capital structure (the amount of debt $D(1)$) so that the probability of insolvency is maintained at a target level p^*. This requires that

$$D(1) = R_A + H^{-1}(p^*) \tag{5}$$

and – in the assumed case of unlimited shareholder liability where the banks debt is risk-free – we have

$$D(0) = (1 + r_f)^{-1} \left(R_A + H^{-1}(p^*) \right) \tag{6}$$

and the risk capital of the bank, measured at market values, is then given by:

$$\hat{E}(0) = \hat{A}(0) - (1 + r_f)^{-1} \left(R_A + H^{-1}(p^*) \right) \tag{7}$$

while the return on risk capital (again measured on a market rather than accounting basis) for this investment is:

$$\hat{r}_{RC} = \frac{R_A - (1 + r_f)D(0)}{\hat{E}(0)} - 1 \tag{8}$$

where the hat once again denotes a measurement on a market not an accounting basis.

The next stage of the analysis is to establish the following proposition:

Proposition 1 $\hat{r}_{RC}$ *from equation (8) – the return on risk capital evaluated when the loan asset is measured at market values – is the appropriate zero-NPV hurdle rate for the corresponding performance measurement computed using (1). Any exposure earning a higher rate will create shareholder value while any exposure earning a lower rate will destroy shareholder value.*

Proof

Since the debt issue consistent with maintaining the target default probability p^*) $D(0)$ is given by (6), the book value of equity capital required to support the loan is given by: $E(0) = L(0) - (1 + r_f)^{-1} \left(R_A + H^{-1}(p^*)\right)$ and so the return on economic capital from equation (1) is given by:

$$r_{RC} = \frac{R_A - (1 + r_f)D(0)}{E(0)} - 1 \tag{9}$$

It is then apparent from comparing (8) and (9) that the investment creates value (i.e. $\hat{E}\,(0) > E\,(0)$ if and only $r_{RC} > \hat{r}_{RC}$.

QED □

The remainder of the subsection addresses the following question, under what conditions is $\hat{r}_{RC}$ constant, so that a single institution wide hurdle rate can be applied to r_{RC} in order to determine if an exposure creates value? i.e. when are risk capital and required accounting returns consistently aligned across business exposures?

We now consider the choice between different assets $i = (1, \ldots I)$. The different assets can have very different return distributions, so we now use the subscript i to distinguish the cumulative density function of returns (H_i^{-1}) and the impact of the aggregate risk factor on these returns ($\alpha_i(z)$).

The following proposition summarises necessary and sufficient conditions for applying a single zero NPV hurdle rate for return on risk capital ie for the equivalence of risk capital and required returns.

Proposition 2 *Consider a bank with unlimited shareholder liability. The rate of return on risk capital for a marginal bank asset takes the same value r^* for any marginal investment opportunity (one where $A_i(0) = L_i(0)$) drawn from a set of potential marginal investment opportunities indexed by $i \in (1, 2, \ldots, I)$ **if and only if** the ratio of $\theta^i = \int \alpha_i(z)q(z)dz/H_i^{-1}(p^*)$ for asset i is a constant for all i.*

Proof

Substituting for its component terms, the expression for the return on risk capital for a marginal (zero NPV) investment i.e. equation (8) can be rewritten:

$$\hat{r}_{RC} = \frac{-(1+r_f)H_i^{-1}(p^*)}{\left(\int \alpha_i(q(z))\mathrm{d}z - H_i^{-1}(p^*)\right)} - 1 = \frac{1+r_f}{1-\theta^i}$$

The left hand side of this equation remains constant for all investment opportunities if θ^i is a constant. Similarly if $\hat{r}_{RC}$ is constant so is θ^i. Both necessity and sufficiency are established.

QED □

Proposition 2 offers a simple intuition in capital based risk-performance measurement. RAROC works when a single risk measure can be used to represent both downside tail risk ($H_i^{-1}(p^*)$) and the compensation required by shareholders for bearing the risk of the exposure ($\int \alpha_1^i(q(z))\mathrm{d}z$). In general however this will not be the case and tail risk and the cost of risk must be distinbuished.

The following sufficient condition, while slightly less general than the previous proposition, provides further practical intuition:

Proposition 3 *A sufficient condition for Proposition 2 to apply is that the distribution of asset returns w^i for any given i can be expressed as a mean-preserving spread of a single underlying asset return distribution w^0.*

Proof. Consider a mean-preserving spread of the underlying asset return distribution, for convenience indexed by the parameter i, so that $w^i = iw^0 = i\alpha(z) + i\gamma\phi$.

The denominator of the first term in the RAROC performance measure r_{RC}, the risk capital required to protect against w^i can then be rewritten as:

$$
\begin{aligned}
RC^1 &= \hat{A}(0) - D(0) \\
&= (1+r_f)^{-1}\left(R_A + \int ki\alpha(q(z))\mathrm{d}z\right) - (1+r_f)^{-1}\left(R_A + iH^{-1}(p^*)\right) \\
&= (1+r_f)^{-1}\left(\int i\alpha(q(z))\mathrm{d}z - iH^{-1}(p^*)\right) \\
&= iRC^0
\end{aligned}
$$

i.e. risk capital increases in proportion to i.

The numerator of the first term of r_{RC} can be rewritten as:

$$
R_A - (1+r_f)D(0) = R_A - \left(R_A + H^{-1}(p^*)\right) = H^{-1}(p^*)
$$

and this also increases in proportion to i. Hence, the return on risk capital $\hat{r}_{RC}$ for a marginal investment opportunity is unaffected by a mean-preserving spread in asset returns.

QED □

The assumption of unlimited liability has been only a presentational device. As established in the Appendix these same Propositions 2 and 3 continue to apply even with limited shareholder liability and (to a close approximation) with implicit or explicit deposit insurance.

It should be apparent that the sufficient conditions for applying a RAROC hurdle given in Proposition 2 and 3 are very strong. Proposition 3 requires for example that the degree of skewness of all bank investments is the same. In practice banks must make choices for investment opportunities that differ greatly in their degree of skewness and there is therefore potential for substantial bias in business decision making from applying a single RAROC hurdle.

3 Illustrations of zero-NPV RAROC hurdles

The previous section has shown that the use of RAROC as a performance measure is only consistent with standard asset pricing theory under highly restrictive conditions, in particular that all return distributions exhibit the same degree of skewness. This section explores some of the practical consequences, quantifying the impact on the zero-NPV hurdle for return on risk capital of assuming different return distributions or altering exposure characteristics such as volatility of returns or probability of default.[16] Section

[16] Mathematica coding for all the Figures reported in the section is available from the authors.

3.1 compares two standard cases appropriate to the analysis of market risks, those of arithmetic and lognormal returns. This will show the link between our work is consistent with that of Crouhy, at. al. (1999), the paper closest to our own in the literature, who derive the "zero-NPV" RAROC hurdle for the case of log-normally distributed investment portfolio. They show (the final column of their Table 1, page 12) that in this case the RAROC hurdle (what they refer to as return on equity, i.e. the return required by the market) increases with the volatility of returns. This sub-section will also analyse the impact of diversification on the RAROC hurdle, showing that constant RAROC hurdle is biased against specialised institutions whose asset portfolio is not fully diversified against movements in market risk factors.

Section 3.2 analyses the determinants of the hurdle RAROC in a standard credit risk model, the asymptotic portfolio loss model of Vasicek underlying the Basel II pillar 1 risk curves and widely used in contexts such as CDO tranche pricing. This suggests that the RAROC hurdle rates applied when using Basel II measures of risk capital for loan credit portfolios should be much lower, much less than one half those applied to investments in marketable securities.

The figures of the zero-NPV RAROC hurdle for return on risk capital reported throughout this section are all based directly on the analysis is Section 2, computed using (equation (8)). For any given return distribution $A(1)$ and confidence threshold p for avoiding default, we compute the current market value $A(0)$ of the prospective investment and thus the market value of the initial equity $E(0) = A(0) - D(0)$ that must be provided by shareholders to reduce the default probability to p, yielding the required return on this risk capital.

We assume quadratic investor utility, so that the pricing function $q(z)$ used to compute $A(0)$ is the capital asset pricing model, in which the expected rate of return on the market value of the asset is given by:

$$r_A - r_f = \beta_{A,M} \left(r_M - r_f \right) = s_A \rho_A \Phi_M \tag{10}$$

and $\beta_{A,M}$ is the beta of the return on asset A with the market M and $r_M(t)$ is the market return at time t. This assumption, while convenient, is not especially restrictive. We could instead have adopted one of many other asset pricing models. While the quantitative differences between required returns and risk capital might differ from those we report here, the general conclusions would be unaffected. The calculations presented here in fact make use of the right-hand expression in (10), the reformulation more closely related to the Sharpe ratio, where ρ_A is the correlation of the asset return with the return on the single factor driving market returns, where s_A is the standard deviation of asset returns, and Φ_M is the market price of risk. [17]

Except where otherwise indicated we assume that all the portfolios (equity or credit) are fully diversified i.e. that $\rho_A = 1$ – an appropriate assumption when risk capital is measured by contribution to the default risk of a very large financial institution where the factors driving its returns may be assumed identical with those for the economy as a whole. We further assume that $\Phi_M = 1$, but this is only a scaling factor, assuming a larger value would raise all RAROC hurdles proportionately and not affect the differences in these hurdles which we report.

3.1 Arithmetic versus lognormal returns with full and partial diversification.

This subsection presents calculations of the RAROC hurdle for a marginal investment opportunity (the $\hat{r}_{RC}$ evaluated on a market value basis as in equation) while varying the standard deviation of returns on a market investment portfolio.

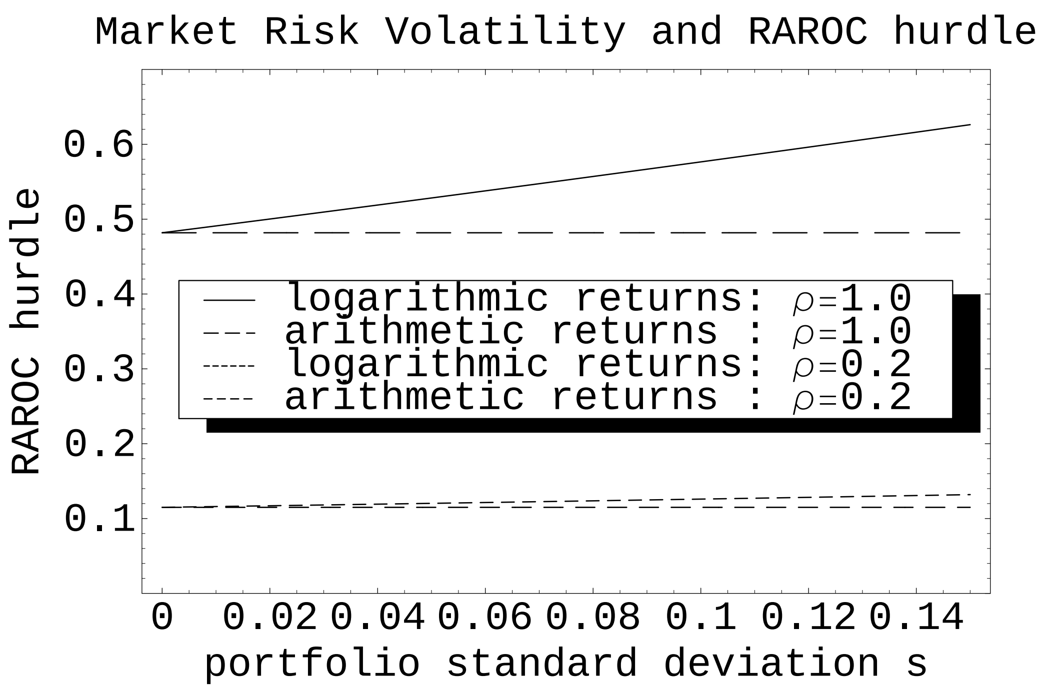
The results are shown in Figure 1. Consider first the upper pair of lines, for a fully diversified portfolio with correlation against the market of $\rho = 1$. The horizontal line is derived assuming an arithmetic normal distribution. This is as predicted by proposition 3, in this case an increase in the standard deviation of returns is an mean-preserving spread in the return distribution, and hence $\hat{r}_{RC}$ remains constant. In this case r_{RC} can be used as a valid peformance measure with a constant hurdle rate.

The lines that slope upwards are for the log-normal distribution of returns, previously analyzed by Crouhy et. al. (1999). This distribution or returns has a right hand skew. An increase in the standard deviation of returns results in a less than proportionate increase in downside tail risk. The denominator of the expression for return on economic capital rises less than proportionately to the increase in asset returns (the numerator). Hence the RAROC hurdle rises as the standard deviation of returns σ increases.[18] Comparing the two cases – arithmetic and log-normal distribution – the figure shows that increasing the volatility of returns from 0% to 14%, the required return on risk capital increases from 48% to about 63% at the 99.97% con-

[17] obtained using $\beta_{A,M} = \rho_A/(s_A s_M)$ and $(r_M - r_f)/s_M = \Phi_M$

[18] We have also rerun our calculations so as to replicate Table 1 on page 12 of Crouhy et. al. (1999). While the relationships are similar, the results differ slightly from theirs, for technical reasons. First we do not include the put option arising from deposit insurance. Secondly we use an exact rather than approximate conversion between continuous time returns and standard deviations (for the log-normal distribution) and discrete time returns and standard deviations.

Figure 1: RAROC hurdles for market risks†



† Computed assuming $\rho = 1$, $\Phi = 1$. In the case of the log-normal using exact discrete time expected returns and standard deviations.

fidence threshold for the log-normal distribution, whereas for the arithmetic normal it remains constant at 48%.

Figure 1 also reports the zero-NPV or required returns on risk capital for an undiversified portfolio with $\rho = 0.2$. The same contrast emerges between the arithmetic and lognormal distribution. With an arithmetic distribution, for which skewness is unrelated to volatility, the required return is constant; while in the case of the log-normal distribution the required return rises as volatility and hence the right skew of the distribution increases. Also now the required return on the partcially diversified portfolio is very much lower than for the fully diversified portfolio. The intuition here is simple - holding the standard deviation of returns constant, the same amount of equity capital is required to protect an undiversified portfolio as a fully diversified portfolio. However – for any given level of portfolio volatility – shareholders are exposed to much less systematic risk with the partially diversified portfolio than with the fully diversified portfolio, in the former case they are able to remove much of this volatility through diversification within their own holding of the market portfolio. Therefore investors have a very much lower required return on risk capital for the partially diversified institution, the lower lines in Figure 1.

3.2 An asymptotic credit portfolio distribution

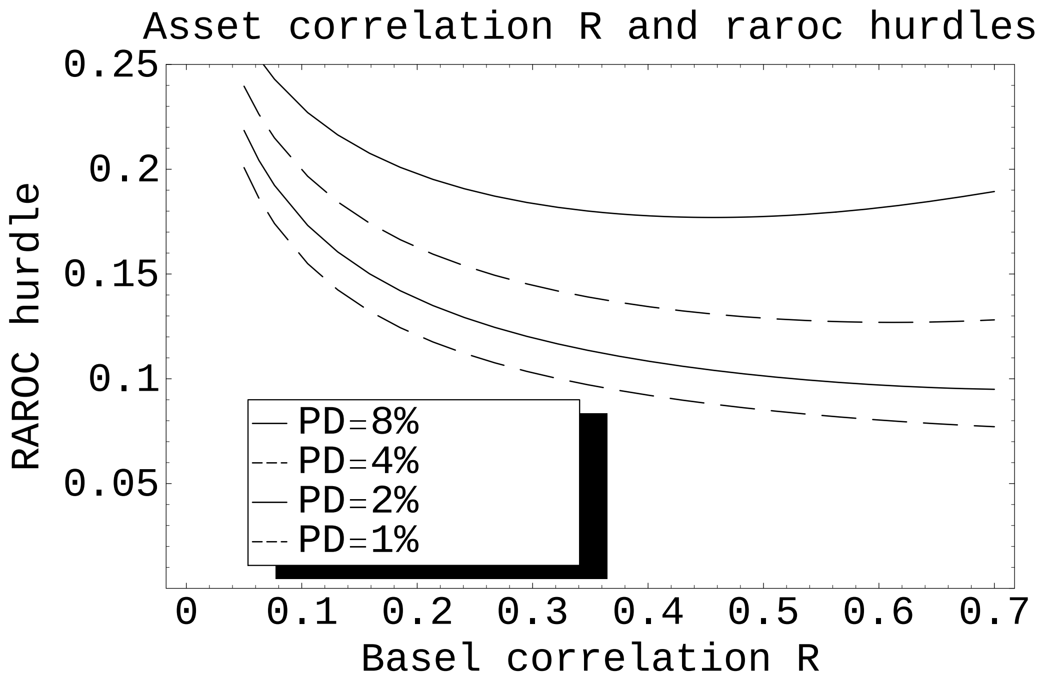
Figure 2 illustrates the RAROC zero-NPV hurdles for the standard credit portfolio model proposed by Vasicek (1987), an asymptotic model of the distribution of returns on a portfolio of defaultable claims and the model underlying the IRB risk-curves in pillar 1 of the Basel II accord.

This Vasicek model of defaultable losses reproduces many basic features of credit risk that cannot captured by either arithmetic or log-normal return distributions. The return distribution is leftward skewed, bounded above at the par value and bounded below at zero. In this model of risky credit portfolio returns, for most plausible parameter choices, the standard deviation of annual returns is relatively small. With the range of parameter values we have explored in Figure 2 the standard deviation of annual portfolio returns falls in the range 0% to 2%, much less than the range illustrated in Figure 1 appropriate for market investments such as equities.

Figure 2 is derived as follows. In the asymptotic Vasicek model of portfolio returns the end period portfolio return $A0$ (relative to a promised value of 1) is given by:

$$A(0) = 1 - \text{LGD} * N(\frac{N^{-1}(PD) + R^{0.5}X}{\sqrt{1-R}}) \tag{11}$$

Figure 2: RAROC hurdles and credit correlation†



† Computed assuming $\rho = 1$, $\Phi = 1$, $PD = 0.02$.

This return is conditional on the underlying normally distributed aggregate factor of X, the constant loss given default LGD, the probability of default PD, and the underlying asset correlation between any two credits R

We assume that returns on the market portfolio are proportional to the same risk factor X. We then use numerical integration over the range $X \in [-6, +6]$ to compute the correlation of credit portfolio returns with the returns on the market portfolio – and CAPM pricing to obtain the period 0 market value of the credit portfolio $\hat{A}(0)$ – and use the Basel risk curve formula (the right hand part of equation (11) with $X = N^{-1}(p^*)$) to compute the required risk capital $\hat{E}(0)$. The hurdle rates shown in the Figure are then computed directly from (8).

The two parameters varied in Figure 2 are the probability of default

PD and the underlying asset correlation R between two credit risky assets.[19] We show the zero-NPV RAROC hurdle for a range of these parameter values covering most bank credit portfolios (PD here ranges from 0.5% to 8%. R for most corporate loan portfolios is usually somewhere in the range 0.3-0.5; while for retail credit portfolio exposures it is much lower, typically in the range 0.01 to 0.15.)

Figure 2 indicate that the zero-NPV RAROC hurdle rate for a corporate loan portfolio with returns behaving according to the Vasicek model should be very much lower than the corresponding RAROC hurdle for market exposures shown in Figure 1. For good quality exposures ($PD < 4\%$) this hurdle is around 10-12% at a 99.97 % confidence threshold, compared with thresholds of over 40% for market exposures. This is a very large difference. Note that this difference is not due to the impact diversification, since the Vasicek model is an asymptotic model which assumes that the credit portfolio is already fully diversified. If the credit portfolio to be 'granular' i.e. not fully diversified then the zero-NPV hurdle rates would be even lower.

Figure 2 suggests that RAROC hurdles will be higher for retail credit portfolios (characterised by rather higher PD and much lower R than corporate portfolios) perhaps in the range of 20% to 30%. But comparison with Figure 1 indicates that fairly large discrepancies can still arise between the hurdle appropriate for market and for retail credit portfolios.

Why are these required RAROC hurdles for credit and market risks so hugely different? This is because of pronounced differences in the shape of the loss distributions. The credit portfolio return distribution computed using the Vasicek model have a very pronounced left skew. This is in contrast to the arithmetic and log-normal distributions used for Figure 1. This substantial difference in the skewness of returns means that the amount of shareholder equity i.e. the risk capital, required to protect a credit portfolio from default can be around five times larger as multiple of portfolio return volatility than is required to protect investment in an equity portfolio. A credit portfolio thus absorbs a much larger amount of risk capital than an equity portfolio, relative to the return required to compensate shareholders for accepting the portfolio risk (which under the CAPM assumption underlying these figures depends only on the volatility of returns and their correlation with market returns.)

[19] We do not report senstivity to LGD since this has almost no impact on the zero-NPV RAROC hurdle, the change in the spread of returns and of the correlation with the aggregate factor almost offsetting each other.

4 Conclusions and implications

This paper has examined the relationship between risk capital (the contribution of an exposure to default risk for a financial institution) and required shareholder returns. If required returns are proportional to risk capital then return on risk capiotal i.e. RAROC (equation (**??**)) can be used as a performance measure with a single hurdle rate for assessing the creation of shareholder value.

We establish (in Propositions 2 and 3) that the conditions for return on risk capital to be used as a valid measure of contribution to shareholder value with a constant hurdle rate are very strict. This is possible in only in two situations, if all bank asset returns distributions belong to a single 'family' with the same degree of skewness, or (rather implausibly) if changes in the skewness of the distribution are by coincidence exactly offset by changes in the degree of correlation with aggregate priced risk factors. Otherwise it is not possible to summarize risk using a single measure and it is instead necessary to have different measures for downside risk (risk capital) and for the cost to shareholders of carrying that risk (required return).

We further show (Figures 1 and 2) that differences in skewness, for standard distributions of market and credit portfolio returns, can make a very large difference to zero-NPV RAROC hurdle rates. We reproduce the finding of Crouhy et. al. (1999) that in the case of the log-normal distribution the zero-NPV RAROC hurdle rate rises with volatility and contrast this with the case of the arithmetic normal where the hurdle rate remains constant. We further show (i) that this hurdle rate should be much higher in large institutions with highly diversified portfolios than in smaller institutions with undiversified portfolios i.e. the fact that small institutions need more shareholder capital to protect them from default risk should not of itself create any difference in the assessment of risk-return tradeoffs; and (ii) that the required zero-NPV RAROC hurdle can be as much as four times as large for a market equity portfolio as for a high quality credit portfolio, even when both portfolios are highly diversified.

Our analysis casts considerable doubt on the dominant current practice for assessment of risk-return tradeoffs by financial institutions. The quantitative loss of shareholder value induced by use of RAROC merits further research, expecially in the context of:

- Alternative pricing models than the CAPM, imposing a relatively higher penalty on downside outcomes.
- Other return distributions, e.g. mark-to-market credit risk models and those avoiding the restrictive assumptions of the Gaussian copula.

- A dynamic model of capital structure and investment decisions that takes into account the loss of revenues and franchise value resulting from undercapitalisation (i.e. a model which endogenises the choice of p^* rather than as we do setting it at a somewhat arbitrary level.

Our analysis assumes that all risks can be priced against market risk factors. A major issues, for many general insurance companies but also for a number of bank exposures, is how to measure performance for risks not priced on financial markets.

How should capital allocation to be conducted, in order to overcome the weaknesses of RAROC, but to adequately reflect the concerns of a financial institution in maintaining an adequate credit rating and protecting its solvency? One possibility would be to impose exposure specific hurdle rates for r_{RC} , for example our modelling framework can be used to compute lower hurdles for credit portfolios than for market portfolios. Finer distinctions could be introduced. Alternatively, in order to keep a single hurdle rate, it would be possible to change the denominator - the amount of risk-capital – by just the amount required to offset the differences in required return on risk capital.

However neither of these approaches seems adequate to capture the distinction between market wide and institutional specific risks discussed by Froot and Stein (1995, 1998). We suggest instead that this be undertaken in two stages. The first stage captures shareholder required returns in a RAROC type measure but computing this through a risk adjustment of the numerator rather than the denominator, according to:

$$r^*_{RC} = \frac{\text{Expected Net Revenues} - \text{Market Cost of Risk}}{\text{Risk Capital}} \tag{12}$$

where the new term – the market cost of risk – can be calculated using some standard asset pricing framework. It is the difference between the expected value of portfolio returns and their certainty equivalent value. It can also be interpreted as the cost of hedging or insuring risk on the market.[20]

The second stage is to capture institution specific balance sheet constraints by an appropriate choice of the hurdle rate for applying (12). This

[20]For example using the CAPM this cost of risk is $s_A \rho_A \Phi_M$ although it could instead be based on other pricing models that allow for greater aversion to downside tail risks.

should be raised above zero to the point $\hat{r}^*_{RC}$ at which the accepted investments (those for which r^*_{RC} exceeds $\hat{r}^*_{RC}$) exactly utilise all the risk capital available on the balance sheet. This then maximises the creation of shareholder value subject to the prudential constraint that risk-capital absorbed by individual exposures does not exceed the total risk capital on the balance sheet. In this case $r^*_{RC}\times$ risk capital can still be used as a measure of shareholder value added (not $(r^*_{RC} - \hat{r}^*_{RC})\times$ risk capital, because the hurdle rate is a shadow price not a real resource cost.) The magnitude of $\hat{r}^*_{RC}$also provides an indication of the shortage of risk capital and this can be used to make a case for retaining additional earnings or raising additional capital. If the institution is unconstrained then $\hat{r}^*_{RC} = 0$.

Our work has some related messages for bank regulators. We note that the contribution of an individual exposure to its *regulatory* capital requirement is only a business concern to a financial institution if it has – or is danger of having – insufficient capital to meet the overall regulatory requirement. Most banks have a very substantial buffers of capital over and above their regulatory capital requirements. This in turn implies that healthy financial institutions should not be concerned with the level of capital that regulators require to back a particular exposure. Just as there is no reason for shareholders to require returns based on consumption of risk capital, nor should they require returns based on consumption of regulatory capital.[21] Moreover it does not matter if regulatory capital and the financial institution's own measure of risk capital do not correspond.[22] This in turn suggests that (a) the practice of 'regulatory capital arbitrage', using securitisation as a method for reducing regulatory capital requriements, is value-destroying rather value-enhancing for the unconstrained financial institution, creating no additional value creating investment opportunities and large fee payments to investment banks; and (b) the goal of 'aligning' economic and regulatory

[21]It is possible that an institution might facing a binding regulatory capital constraint but not a binding risk-capital constraint, in the situation where risk-capital is lower than actual captial which is in turn lower than regulatory capital. In this case it would be appropriate to use a variation of equation (12), but with the contribution to regulatory capital requirement as the denominator, and with a hurdle rate chosen to make efficient use of the banks actual capital against this requirement.

[22]As has been pointed out to us there is a problem of comparing not just apples and pears, but apples, pears, and oranges i.e. risk capital, required return, and regulatory capital.

capital in the Basel II accord is misplaced since higher regulatory capital has no impact on the market pricing of risks. It also suggests that care is required in the application of the 'use test' in the new risk-sensitive regulations: regulators should not normally expect to see financial institutions take direct account of either risk capital or regulatory capital requirements in business performance measurement.[23]

Widespread adoption of an analytical framework such as that offered in the present paper will yield considerable private and social economic benefits. Risk-return tradeoffs will be assessed in a way which is much more supportive of shareholder value creation. Safety and soundness in the financial system will be promoted, since higher levels of risk capital will no longer (and incorrectly) be percieved as having any direct implications for shareholder value. Downside risk and the returns required by shareholders for accepting risky exposures are indeed apples and pears that must not be confused.

[23]This should be done *only* when the bank is under pressing risk or regulatory capital constraints.

A Appendix: Extension to case of limited liability

A.1 Limited liability with risky deposits

Unlimited liability is a simple and transparent special case. But our results also apply to the case of limited liability with risky financial institution debt.

Proposition 4 *In the case of shareholder limited liability, provided that debt holders are fully liable for any losses not borne by shareholders and the bank is able to pre-commit its asset choice at the time it issues debt, then Proposition 1 continues to hold.*

Proof

End-period debt $D(1)$ is determined as before by p^* through equation (**??**). The credit riskiness of debt now reduces the current market value by an amount V_{put}, the present market value of the put option on banks assets written by deposit holders:

$$D(0) = (1 + r_f)^{-1} D(1) - V_{\text{put}} = (1 + r_c)^{-1} D(0) \tag{13}$$

Here r_c is the implied interest rate on credit risky debt.

The present value of equity (the required risk capital) then becomes:

$$\hat{E}(0) = \hat{A}(0) - D(0) + V_{\text{put}} \tag{14}$$

while expressing the expected absolute return on equity is given as:

$$R_E = R_A - (1 + r_c) D(0) + (1 + r_f) V_{put} \tag{15}$$

and we have as before $\hat{r}_{EC} = R_E / \hat{E}(0$

Substituting (13) into the right hand side of (14) and (15), all terms in V_{put} cancel. Neither the expected return nor the amount of required risk capital is affected by the presence of risky debt. It follows immediately that the return on risk capital for each project $\hat{r}_{EC}$ is also unaffected by limited liability in this case with risky debt and so Proposition 1 continues to apply.

QED □

This result reflects the completeness of markets. Debt holders must be compensated for bearing default risk and since this compensation is paid by equity holders, the outcome is that neither the market value of equity nor the expected return on equity is affected by limited liability. The pre-commitment assumption is needed because otherwise the bank could increase the value of V_{put} after raising debt finance, hence transferring wealth from debt to equity holders (the agency cost of debt).

A.2 Limited liability with 100% protection for creditors

Creditor protection, most obviously the provision of deposit protection through either an explicit insurance scheme or an implicit safety net for failed institutions, increases both the return to equity holders and the absolute return to equity holders. In this case (focussing on the case of a purely deposit financed bank with 100% deposit insurance), while we have not been able to prove Proposition 1, a version of Proposition 2 still applies:

Proposition 5 *In the case of shareholder limited liability with 100% insured bank debt, then under the assumption of Proposition 2 (that all risk can be described as a mean-preserving spread of a single underlying distribution) the zero-NPV RAROC hurdle is a monotonically decreasing function of the spread of returns i, falling between the unlimited liability threshold r^* and the risk-free rate with limiting values $\lim_{i\downarrow 0}\hat{r}_{RC} = r^*$ and $\lim_{i\uparrow\infty}\hat{r}_{RC} = r_f$,returnprudentialcapitaleqn*

Proof

Let $r^* = -H^{-1}(p^*) / \left(\hat{A}(0) - (1+r_f)^{-1}(R_A + H^{-1}(p^*))\right) - 1 > r_f$, be the return on risk capital for the zero-NPV project under unlimited liability. With 100% deposit insurance we have:

$$\hat{r}_{RC} = \frac{-H^{-1}(p^*) + (1+r_f)V^i_{\text{put}}}{\hat{A}(0) - (1+r_f)^{-1}(R_A + H^{-1}(p^*)) + V^i_{\text{put}}} - 1 \quad (16)$$

$$= \frac{1 - (1+r_f)V^i_{\text{put}}/H^{-1}(p^*)}{(1+r^*)^{-1} - V^i_{\text{put}}/H^{-1}(p^*)} - 1 \quad (17)$$

$$= (1+r_f)\frac{(1+r_f)^{-1} - V^i_{\text{put}}/H^{-1}(p^*)}{(1+r^*)^{-1} - V^i_{\text{put}}/H^{-1}(p^*)} - 1 \quad (18)$$

This establishes that when $i = 0$ i.e. the volatility of return equals zero and so $V_{\text{put}}^i = 0$, $\hat{r}_{RC} = r^*$. V_{put}^i is an increasing function of i and hence also $\hat{r}_{RC}$ is a decreasing function of i. Finally since $\lim_{i\uparrow\infty} V_{\text{put}}^i = \infty$, $\lim_{i\uparrow\infty} \hat{r}_{RC} = r_f$

QED □

100% deposit insurance does make a difference to our results. But it should be realised that for a safe bank, one say holding capital to maintain its annual default probability to a target of 0.1% or less, the value of the the deposit insurance put option V_{put}^i held by equity holders is very small and will make little difference to the RAROC hurdle.

References

1. Basel Committee on Banking Supervision (1999a) ”The New Basel Accord: An Explanatory Note”, available as http://www.bis.org/publ/bcbsca01.pdf

2. Basel Committee on Banking Supervision (1999b) A new capital adequacy framework, Basel: Switzerland.

3. Berger A. N., Herring R.J. and Szego G. P. (1995) The role of capital in financial institutions, Journal of Banking and Finance19, pp. 393-430.

4. Crouhy, Turnbull and Wakeman (1999) Measuring Risk-adjusted performance, Journal of Risk, Vol 2 no 1, Fall, pp. 5-35.

5. Cochrane, John H (2005) Asset Pricing: Revised, Princeton University Press

6. Dev, Ashish (2006) Economic Capital: A Practitioner's Guide, Risk Books

7. Elizalde, Abel and Rafael Repullo (2006) Economic and Regulatory Capital: What is the Difference?

8. Froot, K. and Stein, J. (1998) Risk Management, Capital Budgeting and Capital Structure Policy for financial institutions: An Integrated Approach, Journal of Financial Economics, 47: 55-82.

9. Gordy, M. B. (2000), A Comparative Anatomy of Credit Risk Models, Journal of Banking and Finance, Vol. 24, No. 1-2, pp 119 –149.

10. Jokivuolle, Esa (2006) Aligning Regulatory with Economic Capital, chapter 21 of Dev, Ashish; Economic Capital: A Practitioner's Guide, Risk Books, pp 455-479

11. Jokipii, Terhi and Alistair Milne, (2007) Understanding European Banks Capital Buffer Fluctuations, Journal of Banking and Finance, forthcoming

12. KPMG (2004) Basel II- A closer look: Managing Economic Capital. downloadable from several national KPMG websites including www.us.kpmg.com

13. Matten, Chris (2001) Managing Bank Capital, 2nd edition, Wiley

14. Merton, R (1974), 'On the Pricing of Corporate Debt: The Risk Structure of Interest Rates', Journal of Finance, 25, May, pages 449-70.

15. Merton, R (1977), 'An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees: An Application of Modern Option Pricing Theory', Journal of Banking and Finance, 1, pages 3-11.

16. Merton, R. and Perold, A. (1993) Theory of risk capital in financial firms, Journal of Applied Corporate Finance, 6: 16-32.

17. PWC-Economic Intelligence Unit (2005) Effective capital management: Economic Capital as an Industry Standard? which can be found via http://searchpwc.com/extweb/pubsrc

18. Schroeck, Gerhard (2000) Risk Management and Value Creation in financial institutions, Wiley.

19. Smithson, Charles W. (2002), Credit Portfolio Management, Wiley

20. Stern, Joel M. (2000) EVA and Strategic Performance Measurement, Global Finance 2000, New York.

21. Stewart, G. Bennett (1994) EVA: Fact and Fantasy, Stern Stewart & Co, Journal of Applied Corporate Finance, volume 7 no 2.

22. Stoughton, M. Neal and Zechner, Josef (2003) Optimal Capital Allocation using RAROC and EVA, Working paper.

23. Stulz, Rene M. (1996) Rethinking risk management, Journal of Applied Corporate Finance, vol. 9 number 3, pp. 8-24.

24. Vasicek, O. (1987) Probabilty of Loss on a Loan Portfolio, Working Paper, KMV, published in Risk, december 2002, under title 'Loan Portfolio Value'.

25. Wilson, Thomas C. (1992) RAROC remodeled, Risk Magazine, Vol. 5 number 8, September 1992, pp. 112-119.

26. Zaik E., Walter J., Kelling G. and James C. (1996) RAROC at Bank of America: from theory to practice, Journal of Applied Corporate Finance.