THE COST OF CAPITAL FOR FINANCIAL FIRMS

By C. J. Exley and A. D. Smith

[Presented to the Institute of Actuaries, 23 January 2006]

ABSTRACT

Most businesses have assets financed by capital providers. The cost of capital is a measure of the returns required by those capital providers. Its main use is to set a target for the profits, which must be achieved on the firm's assets in order to satisfy equity and bond holders.

This paper describes the classical theory of the cost of capital, and then applies it to the special case of banking and insurance firms. We develop implications for product pricing, performance measurement and capital structure optimisation.

KEYWORDS

Cost of Capital; Capital Asset Pricing Model; Franchise Value; Market Consistent Valuation; Embedded Value; Franchise Insurance Premium; Frictional Costs; Taxation; Performance Measurement

CONTACT ADDRESSES

C. Jon Exley, Barclays Capital, 5 The North Colonnade, Canary Wharf, London E14 4BB, U.K. E-mail: Jon.Exley@barclayscapital.com
Andrew D. Smith, Deloitte & Touche LLP, Stonecutter Court, 1 Stonecutter Street, London EC4A 4TR, U.K. E-mail: andrewdsmith8@deloitte.co.uk

Charlie [Munger] and I have not the faintest idea what our cost of capital is and we think the whole concept is fairly crazy, frankly Warren Buffett, May 2003

1. Introduction

1.1 *Why does the Cost of Capital Matter?*

1.1.1 Some people think that capital incurs a cost just by being there. A business must meet capital costs in addition to other expenses of doing business. Others talk confidently about the cost of capital, as if they have already measured it, comprehensively taken account of all accounting effects, risks, regulations, taxes and other complications, survived a rigorous audit process, and now use cost of capital as a guiding principle for running every aspect of their lives. The rest of us sit in meetings feeling left out, making a mental note to research the ingenious secret. If that sounds familiar, then this paper is for you.

© Institute of Actuaries and Faculty of Actuaries

1.1.2 Some firms have made significant investment in advanced quantitative systems for risk management and economic capital quantification. We believe that a similar degree of effort and science can profitably be applied assessing the commercial cost of meeting economic capital requirements.

1.1.3 Cost of capital has become an important deduction from accounting profit in performance measurement and reporting. Thousands of financial sector employees find that a cost of capital calculation affects their bonuses. Business units are bought, sold, closed down or expanded on the basis of whether they meet their cost of capital. Apparently, you make better commercial decisions if you take account of capital costs.

1.1.4 Many analysts' corporate valuation models feature the cost of capital as an input. Ideally, a cost of capital used in this way should reflect, not only the risk and potential of a firm, but also conditions in the capital market. To the extent that one believes that analysts' opinions affect security prices, these calculations of a firm's cost of capital potentially affect its share price and its ability to raise finance when required. At the very least, the way in which analysts do their cost of capital calculations is important if management behaviour reflects attempts to second-guess analysts' cost of capital estimates.

1.1.5 If capital has a real cost, then that cost should feed through into the prices which consumers pay for financial products. Accurate assessment of capital cost means a competitive advantage in accurate product pricing. In some jurisdictions, financial firms have to satisfy price regulators that customers are not being overcharged. Capital costs in a disclosed pricing basis can justify greater pricing freedom, while demonstrating that shareholder returns remain fair.

1.1.6 Some regulators are considering whether financial firms should demonstrate financial resources, not only to meet customer liabilities in one year with a high probability, but also to meet the cost of any capital tied up over the period for which the liabilities exist. If that were the case, then the higher the cost of capital, the more the regulator would expect firms to hold.

1.1.7 If the cost of capital increases, you might expect firms to find ways to use less capital or use it more efficiently, just as motorists seek more efficient cars when the price of fuel rises.

1.1.8 All of these issues increase pressure on managements to articulate their capital strategy and assessment of capital costs.

1.2 *What is the Cost of Capital?*

1.2.1 Let us start where the literature starts. Modigliani & Miller (1958) consider a simple company with a set of assets, for example representing plant and inventory. The assets are financed partly out of borrowed money (debt) and partly by shareholders (equity).

1.2.2 The lenders may be banks providing overdraft facilities, or may be subscribers to an issue of commercial paper or corporate bonds. The shareholders and lenders are, in aggregate, known as *capital providers*.

1.2.3 In the past, companies used to set budgets to ensure positive operating results. In other words, the price charged for goods must cover the cost of the labour and materials to make the goods. A company was considered broadly healthy, provided that its operating result was positive or that its dividends were maintained.

1.2.4 More recently, companies have focused on *adding value*. This is defined as covering, not only the operating cost, but also providing an acceptable return to capital providers. The cost of debt is interpreted as the interest payable, while some economists refer to a *cost of equity* as the profit which shareholders require. The weighted average cost of capital, as a money amount, is the total cost of equity and cost of debt.

1.2.5 References to 'cost of equity' are controversial and sometimes confusing. The cost of equity is not a cash flow. Neither accounting standard setters nor taxation authorities recognise cost of equity as a cost. Even solvency regulators are not yet requiring firms to provide for future equity costs. Cost of equity does not accumulate arrears if you miss a payment, and cannot trigger default or bankruptcy. Nevertheless, describing required returns on equity as a cost has fallen into common usage, and we follow that here.

1.2.6 The usual approach for measuring the cost of equity is the concept of opportunity cost. This means the return which a shareholder could have obtained with alternative investments of similar risk.

1.3 *Customers as Capital Providers*

1.3.1 How then can we apply the cost of capital ideas to financial firms? Many of the ideas are applicable, but only after some conversion in terminology. The main additional complication is that financial firms, typically, have significant customer liabilities, either in the form of bank deposits accepted or provisions for insurance claims and associated expenses. These customer commitments are often highly regulated. There may also be additional borrowings, usually subordinated to (i.e. lower priority than) the customer commitments, sometimes raised at a parent level and then injected as equity into a financial subsidiary.

1.3.2 In classical corporate finance parlance, all of these are sources of capital, and should be taken into account in a cost of capital calculation. On the other hand, when financial firms talk about capital, they usually mean equity and subordinated debt, excluding liabilities to retail customers. Confusion can easily arise from these conflicting uses of the word 'capital'.

1.3.3 It is difficult to assign a cost of debt to customer liabilities. Bank deposits may appear to be a cheap source of capital, if customers demand an annual interest rate below that of the capital markets. However, a focus on credited rates overlooks the high front-end costs in terms of marketing and commission for attracting the customers in the first place. Although accountants may treat marketing as an expense rather than as an investment, shareholders are equally entitled to demand that expenditure on marketing is justified by future returns.

1.3.4 Insurance liabilities are also problematic. We might value them using a discounted cash flow calculation, but where does the discount rate come from? Whatever discount rate is chosen, this will turn out to be the cost of debt. For example, if reserves are computed on a prudent statutory basis, then the cost of debt will appear equal to the statutory discount rate. However, it would be a mistake to interpret a low discount rate as an indication that customers supply cheap capital.

1.3.5 As a result of difficulties in debt cost measurement, the weighted average cost of capital (WACC) and associated measures are challenging to apply to financial firms if customers are treated as capital providers. Instead, in a financial context, the stated WACC usually averages capital costs only over equity and subordinated bondholders, excluding customer liabilities from the equation.

1.4 *Reported Cost of Capital*

1.4.1 Many financial companies report financial results net of cost of capital, to disclose some measure of value added which is used for performance measurement. For example, Lloyds TSB provides the following commentary in their 2004 financial statements:

> "A common approach is applied across the Group to assess the creation of shareholder value. This is measured by economic profit (the profit attributable to shareholders, less a notional charge for the equity invested in the business). The focus on economic profit allows the Group to compare the returns being made on capital employed in each business."

1.4.2 The cost of capital used is a simple 9% of average shareholders' equity, as shown in Table 1.

1.4.3 Some insurers now use the language of cost of capital to describe their embedded value calculation. Here, the cost of capital is applied, not as a retrospective performance measure, but as a discount rate for assessing the

Table 1. Lloyds TSB calculation of economic profit 1998 to 2004

£m	1998	1999	2000	2001	2002	2003	2004
Average equity	9,822	10,189	12,556	12,338	10,672	8,460	10,878
Shareholder profit	2,073	2,439	2,654	2,229	1,781	3,254	2,421
Less cost of equity	884	917	1,130	1,110	960	761	979
Economic profit	1,189	1,522	1,524	1,119	821	2,493	1,442

Source: Lloyds TSB Annual Report and Accounts, 1999 to 2004

Table 2. Derivation of WACC in Aviva's EEV presentation

	Cost	Weight
Cost of equity	9.5%	70%
Cost of debt	4.1%	30%
Unadjusted WACC	7.9%	
Adjustment for risks counted elsewhere	–0.1%	
Adjusted WACC	7.8%	

Source: Aviva analysts' presentation, June 2005

Table 3. Derivation of cost of equity in Aviva's EEV presentation

Risk free rate	5.1%
Equity risk premium (ERP)	3.0%
Beta (β)	1.48
Cost of equity = risk free + β* ERP	9.5%

value of shareholders' interest in a block of business. A typical example is provided in Aviva's June 2005 EEV presentation. They derive their WACC as a weighted average of equity and debt cost, as in Table 2.

1.4.4 The cost of equity is calculated using the Capital Asset Pricing Model (CAPM), with the assumptions given in Table 3.

1.4.5 The embedded value method also requires assumptions for the expected returns on assets. In Aviva's case, the same equity risk premium of 3% is used for the projection and the WACC calculation. However, expected returns must reflect the beta of the equities concerned. For example, an insurer may well have a higher beta than the average of the shares in which it invests. In that case, the risk discount rate will exceed the assumed earned rate of return on equities.

1.4.6 Readers may wonder why the cost of debt comes out below the risk free rate. The explanation is that debt interest cost results in a reduction in corporation tax. The 4.1% debt cost is net of tax. In comparison, the risk free rate of 5.1% is a gross-of-tax interest rate.

1.4.7 Other insurers have adopted a so-called 'market consistent' bottom up approach to embedded value calculations. In the market consistent framework, asset risk premiums are irrelevant, as, in theory, they cancel out between projection and discounting. The question arises whether any further adjustment is required to allow for the cost of capital.

1.4.8 The market consistent approach does allow for the price of pure market risks — such as those arising from equity or interest rate movements. However, there are a number of other more subtle effects, including double taxation of investment income on retained capital, risks of firm failure and also agency costs. These are collectively known as 'frictional costs'. See

Cumberworth *et al.* (2000), Hancock *et al.* (2001), or Ng *et al.* (2003) for more details on these items.

1.4.9 Firms are beginning to get to grips with some of the more subtle frictional costs. O'Keeffe *et al.* (2005) provides a description of recent developments in this area. A recent disclosure by Friends Provident describes the additional allowances as follows:

> "We have allowed for the frictional costs of holding this locked-in shareholder capital, being the tangible costs of holding capital on a market-consistent basis."

1.4.10 Yet other firms measure returns on capital relative to some hurdle rate, to assess relative attractiveness of different parts of the business. For example, Fortis tabulated the following:

> "Economic capital has been developed within Fortis in order to provide a consistent and comparable measure of risk across all risk types and geographies. It is a measure of the 'Value at Risk' to a confidence interval of 99.97% and a horizon of one year.
>
> The Return on Risk Adjusted Capital (RORAC) is a performance yardstick which establishes a consistent relationship between the risks and returns of Fortis' various activities. RORAC is calculated by dividing the risk-weighted return by the economic capital. The risk-weighted return is itself determined on the basis of the net operating results, with provisions for credit risks being replaced by estimated, cycle-neutral expected losses."

1.5 *Remainder of this Paper*

1.5.1 This introductory section has laid out some initial notions of equity cost and how firms currently use it. The remainder of this paper is more forward looking, and makes some suggestions for improving the application of cost of capital techniques in financial firms.

Table 4. Fortis: economic capital and return on risk adjusted capital

	Economic capital (€ milliard)		RORAC	
	2003	2004	2003	2004
Network banking	7.4	7.3	13%	19%
Merchant banking	3.0	2.9	17%	13%
Investment services	0.4	0.4	82%	85%
Total bank	10.8	10.5	16%	20%
Fortis AG	0.8	0.9	12%	34%
Fortis ASR	1.2	1.5	32%	47%
Fortis insurance international	0.5	0.4	23%	26%
Total insurance	2.4	2.8	31%	37%
General sector	0.9	0.5		
Fortis total	14.1	13.8	18%	22%

Source: Fortis Report and Accounts, 2004

1.5.2 Section 2 considers what we know about shareholder required returns. These are taken from the body of academic literature known as 'asset pricing theory'. These theories seek to explain market returns in terms of risks underlying different investments. Required market returns form the theoretical basis of most attempts in practice to assess capital costs.

1.5.3 Section 3 raises the main theme of this paper, that is return biases. While asset pricing theory informs our understanding of required returns on market value, many applications require hurdle returns to be set in terms of accounting quantities, such as return on equity. The difference in denominators, combined with differences in risk, can give rise to biases which distort the apparent attractiveness of different investment opportunities. We investigate possible adjustments to relate a market value return to accounting returns, including an allowance for franchise value (which we define), and the use of economic, rather than accounting, capital measures.

1.5.4 Section 4 investigates alternative models of corporate valuation allowing for capital costs. We compare dividend discount models to market consistent methods. We also discuss how financial firms create value, the role of frictional costs, and use these ideas to motivate the calibration of valuation models. Finally, we consider the possibility of ruin and its effect on corporate valuations.

1.5.5 Section 5 considers what capital costs can tell us about the ideal amount of capital for a financial firm to hold. The optimum balances frictional costs of overcapitalisation against financial distress costs of undercapitalisation. We consider reasons why economic capital is not the same as the amount of capital which a firm might wish to hold. Finally, we are able to identify a cost of economic capital in terms of default spreads, and their effect on corporate valuations.

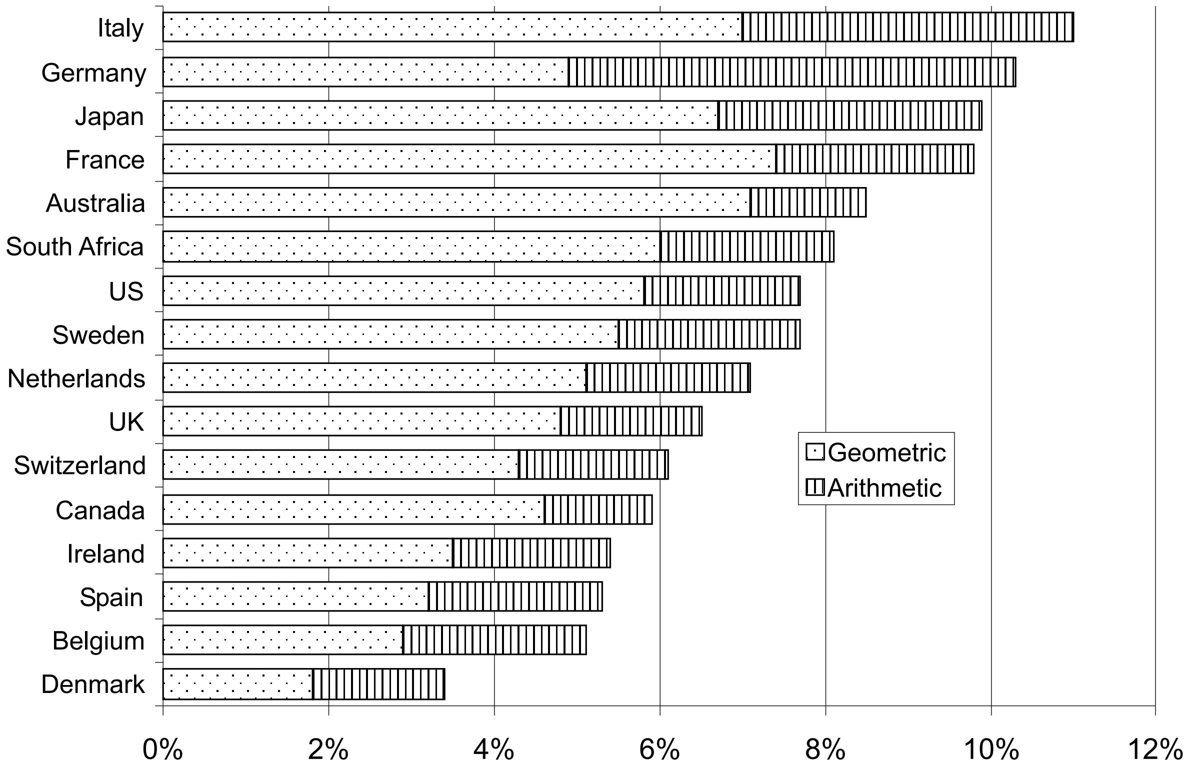
1.5.6 Section 6 concludes with a proposed holistic approach to financial management, applying consistent science to financial reporting, risk management and product pricing.

2. What Returns do Shareholders Require?

2.1 *Long-Term Equity Returns*

2.1.1 For many years, Barclays capital has published an annual study of historic returns on United Kingdom equities and gilts. Other assets have been added over time. Dimson *et al.* (2002) provide international comparisons. Figure 1 shows the historic excess returns on equities (above cash returns) for 16 countries over the period 1900 to 2000, according to Dimson *et al.*

2.1.2 As a first approximation, we could assume that these average historic returns represent returns which shareholders will require in future. This would imply, for example, that shareholders in Italy are much more demanding than those in Spain.



Source: Dimson *et al.*

Figure 1. Mean equity returns in excess of T-bills, 1900 to 2000

2.1.3 If we believe that there is a fixed cost of equity for each country, then we can change a firm's average cost of capital by changing the weight between equity and debt. We can minimise the cost of capital by financing as much as possible by debt.

2.2 *Asset Pricing Theory*

2.2.1 Asset pricing theory is the study of expected returns on different asset classes, and how they relate to risk. At the simplest level, analysts may plot an *efficient frontier* of asset risks and returns. These are discussed further by Cumberworth *et al.* (2000). An efficient frontier does not tell you the required return at a given level of risk. It merely gives an upper bound, that is the highest possible required return at a given level of risk.

2.2.2 The first, and still the most popular, model for relating risk to return is the CAPM. The CAPM splits risks into two components: systematic and non-systematic risks. The systematic component, traditionally denoted by β, determines required expected returns. The other components of risk do not affect expected returns.

2.2.3 Historic betas do not tell the whole story. If two portfolios have the same beta, then they appear to have the same exposure to changes in stock market levels. However, they might still have different exposures to movements in yield curves, real estate markets, credit spreads, currencies or implied volatilities. The Casualty Actuarial Society, in the United States of America, commissioned a report to define best practice on the cost of capital. The output was a survey of many possible models (Cummins *et al.*, 2000), including some which allow for other market exposures.

2.2.4 Some financial firms use stochastic scenario generators, for example in order to assess economic capital or for quantifying the cost of options or guarantees. In most cases, these scenario generators will already include a relationship between investment risks and returns. These same models can then be used for assessment of capital costs.

2.3 *Expected or Actual Returns?*

2.3.1 Traditional returns on equity models work out the long-term returns which shareholders require. If the risk free rate is 4.5% and the equity risk premium is 6%, then this implies that a share with a beta of 80% should return 9.3%.

2.3.2 This is a blunt tool which works well in a year where there is no adverse stock market performance, but life insurance companies are at an immediate disadvantage in years where the stock market performs poorly. This is because the long-run average return is a poor estimate of the actual return in a given year, and yet the asset performance of a financial firm (unlike an industrial company) depends directly on the actual market return.

2.3.3 Hancock *et al.* (2001) describe an alternative approach for insurers, based on replicating portfolios. The idea is to measure shareholder returns relative to the actual return on a replicating portfolio, rather than using some estimate of long-term returns. The replicating portfolio relative to actual assets can be considered as a leveraged investment fund, while the insurance operations can then be examined after stripping out the effect of market moves. This idea is shown schematically in Figure 2, taken from Hancock *et al.* (2001).

2.3.4 A financial firm can build a replicating portfolio model of assets and liabilities. This quantifies the fair value exposures to many market factors, based on a bottom up model of how the business actually works, rather than on historic correlations alone. Performance measurement then involves comparison of actual shareholder returns by business unit relative to the corresponding replicating portfolio return.

2.3.5 Performance measurement relative to actual, rather then to expected, returns is not a new idea. Investment manager mandates typically involve the construction of a benchmark portfolio and a margin by which the manager seeks to exceed the benchmark. More recently, so-called liability driven investment benchmarks for pension funds in the U.K. have, in turn, based this benchmark portfolio on assets replicating pension scheme liabilities. We suggest that these approaches are relevant, not only to investment

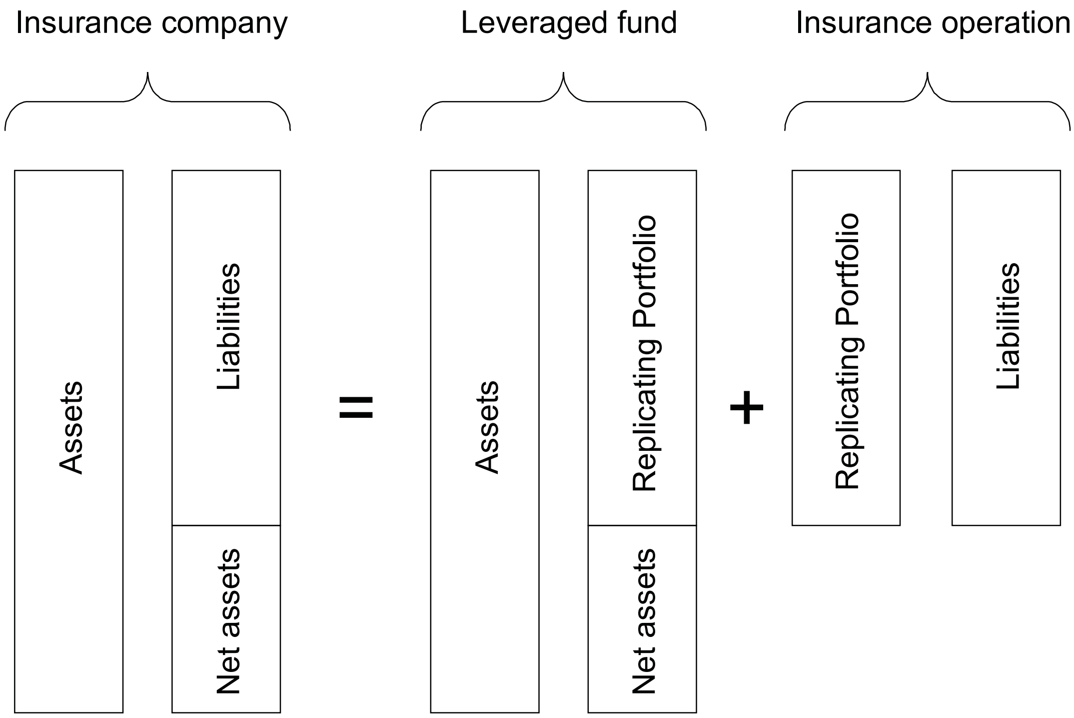


Figure 2. The replicating portfolio approach to measuring required shareholder returns (Hancock *et al.*, 2001)

management contracts and pension fund benchmarks, but also to the assessment of financial firms' performance.

2.3.6 The replicating portfolio method is more informative than the traditional beta approach, because it:

— eliminates any effect of market moves, whether beneficial or detrimental, by capturing market moves in the replicating portfolio;
— is objective rather than subjective;
— removes the need to estimate long-term parameters, such as the equity risk premium; and
— can price correctly for interest rates, credit, and other sensitivities, as well as for stock market movements.

2.4 *Introducing Franchise Value*

2.4.1 However, the replicating portfolio described above does not quite capture the whole story. There may be more to the market valuation of a financial firm than merely the value of its net assets. Instead, we can consider the market capitalisation of a financial firm to be made up of two components: net assets and franchise value, as shown in Figure 3.

2.4.2 Terminology varies. In this paper we treat franchise value as being any element of market capitalisation in excess of statutory net assets. In particular, we regard life insurance Value in Force as part of the franchise value.

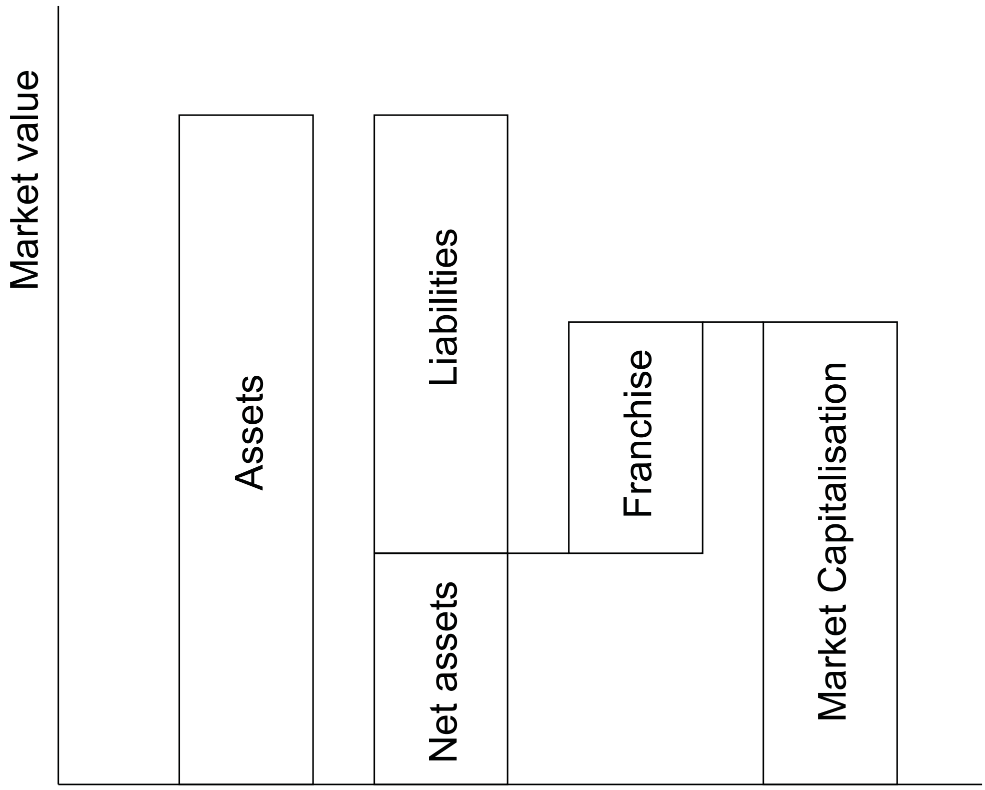


Figure 3. Components of value for a financial firm (after Hancock *et al.*, 2001)

2.5 *Franchise Insurance Premium*

2.5.1 There is an inconsistency in the way in which WACC is conventionally calculated. The equity portion is an expected return over all possible outcomes. However, the stated cost of debt is usually the actual interest rate charged. From the debt holders' perspective, the interest rate is not the expected return; it is the maximum return, because there is some possibility of default. WACC is, in practice, a weighted average of the mean equity return and the best case debt return.

2.5.2 It would be tempting to fix the inconsistency by allowing for some default probability in the debt cost calculation. However, for ex post performance measurement, this argument makes little sense, because in the scenarios where we are measuring performance, the full maximum debt cost will have been incurred. The default event can also happen, but, in that case, all management performance measures are likely to have been abandoned!

2.5.3 Instead, the resolution of the paradox involves adjusting the required return on equity consistently with the cost of debt calculation. That means using only that part of the distribution conditional on the firm having survived.

2.5.4 Although that may sound mysterious, parallel concepts exist in insurance. For example, an annuitant who survives a year probably obtains a better than the risk free return, made possible by the cessation of annuity payments to those who died. A benchmark annuity return for application to the living would incorporate this survivorship effect. To take another example, a catastrophe reinsurer expects to collect premiums for being exposed to risks, even in a year where none of the risks materialise — and their budgeting process reflects this.

2.5.5 We can interpret the additional required return as a form of insurance premium against the death of a firm. To put the investor in a financial firm into the same position as a holder of the replicating portfolio, the financial firm investor needs also to purchase insurance against the loss of franchise value should the firm fail.

2.5.6 In the case of an annuity or catastrophe reinsurance contract, well established procedures exist for determining the sum at risk and the associated risk premium. For an insurance investor, we argue that the relevant sum at risk is the franchise value, with the appropriate risk premium being related to the spread on credit default swaps. For this reason, we think of the adjustment as a franchise insurance premium.

2.5.7 Well established procedures exist for calculating replicating portfolios which mirror the balance sheet sensitivities to market moves. We propose that replicating portfolios should also replicate with regard to credit risk. To replicate the loss of franchise value in the event of failure, the replicating portfolio might include a sold credit default swap. In that case, the credit spread on the swap forms part of the replicating portfolio return, and there is no need for a separate manual addition of a franchise insurance premium for performance measurement purposes. Section 4 considers this idea in more detail, including a calculation of the notional swap amount for inclusion in the replicating portfolio.

2.6 *Return Denominators*

2.6.1 All our theories of required returns are based on market value returns. We know a lot — or at least, have many theories — for market value returns on different asset classes.

2.6.2 However, applications in performance measurement often relate to accounting numbers rather than to observable prices. For example, most WACC calculations use as weights the relative accounting values of equity and debt, rather than the market values.

2.6.3 Sometimes the use of accounting values is forced upon us. This is particularly the case for performance measures at the business unit or more granular levels, where separate market values may not be readily available. In other cases, the avoidance of market values may be deliberate, for example in order to avoid volatility in performance measures linked to market prices.

2.6.4 There is no reason why accounting returns should equal market value returns. In the next section we consider the reasons for differences, or biases, between accounting and market value returns, and suggest some corrections to ensure that accounting-based performance measures make sense.

2.7 *Required Returns at Business Unit Level*

2.7.1 Performance measures are straightforward to define at a group level. Whether we measure performance on market capitalisation or accounting net assets, this information is readily available at an aggregate level.

2.7.2 However, management decisions usually require a greater degree of granularity. Allocating profit by business unit is a routine accounting exercise. We can also evaluate replicating portfolios based on the assets and liabilities of each business unit. It is less obvious how to allocate market capitalisation or balance sheet equity.

2.7.3 The most popular algorithm for allocating capital between business units involves the measurement of each unit's risk exposure — for example in terms of modelled 'Value at Risk'. The hypothesis is that capital is required to absorb risks; therefore relative risk becomes an attractive way of allocating capital between units for performance measurement.

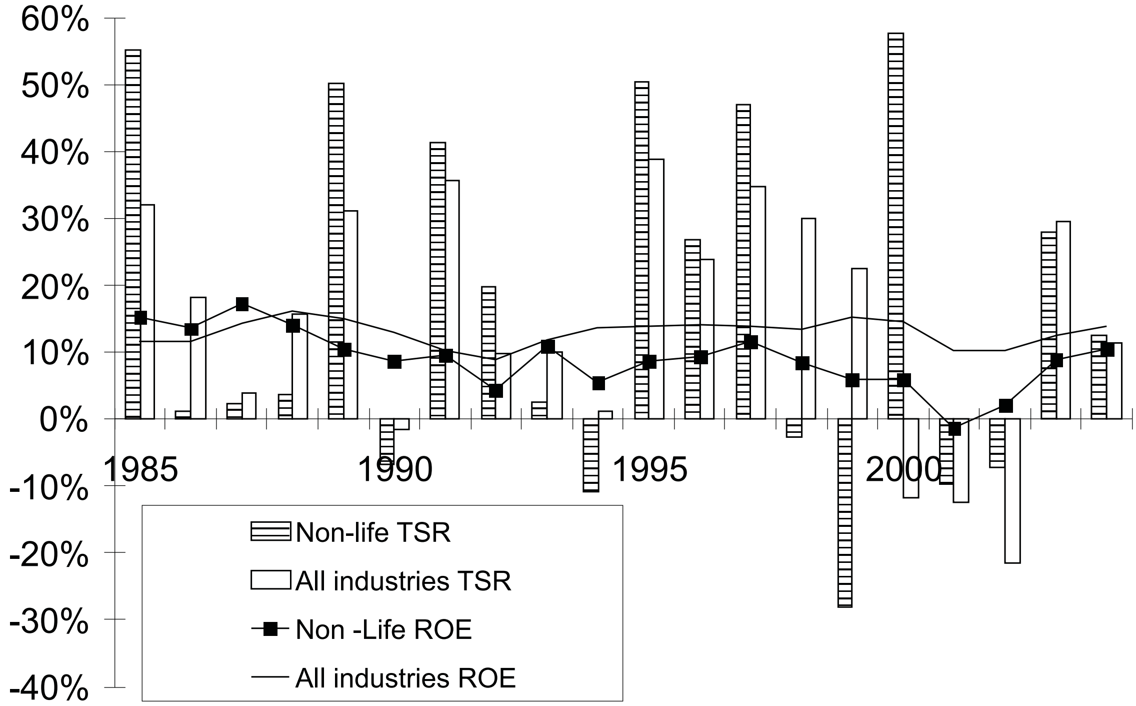
2.7.4 An alternative, and potentially superior, approach is, instead, to allocate market capitalisation between business units. A quick way to do this is to use analysts' rules of thumb, such as the price earnings ratio, to apportion value. Comparisons with the market prices of peer group firms, particularly specialist mono-line firms, can give further information.

2.7.5 A more intensive approach to value apportionment makes use of appraisal values. This is a discounted profit valuation of the whole business (see Burrows & Whitehead, 1987; Burrows & Lang, 1997). The assumptions are chosen to replicate the known market value — for example by flexing new business growth until the present value comes into line with the market capitalisation. The present value of business unit profits then provides an indication of how the market capitalisation splits between business units. The breakdown is not unique, so the process of setting appraisal value assumptions can be the subject of heated negotiation. While many insurers have adopted appraisal value techniques, banks have shown less enthusiasm.

3. Mending Bias in Accounting Measures of Return

3.1 *Historic Experience*

3.1.1 Figure 4 shows the historic return on equity (ROE), based on U.S. figures, for banks, insurers and industry as a whole. Dr. Hartwig at Insurance Information Institute has kindly provided the return on equity data; the total shareholder returns are taken from Datastream indices.



Source: Insurance Information Institute

Figure 4. Historic return on equity and total shareholder returns for U.S. non-life insurers and all industries

3.1.2 The question which analysts face is whether the low ROE from property/casualty insurers implies that they are under-performing as businesses, or are poor investments. It is clear that non-life insurers have under-performed industry as a whole over this period, as measured by ROE, and, in particular, in every year since 1990 the ROE for non-life insurers has been below the average for all industries. On the other hand, $100 invested in 1990 would have returned more invested in the non-life sector than in industry as a whole. This discrepancy between ROE and market returns is called the 'ROE bias'.

3.1.3 There is an established literature, starting from Soloman & Laya (1967), investigating biases in accounting measures of return. They consider whether some of the ROE differences between different industries are due to

Table 5. Total shareholder returns for U.S. non-life insurers and all industries, 1990 to 2004

	Value at 31/12/1990	Value at 31/12/2004	Average ROE
Non-life	$100	$582	7.3%
All industries	$100	$539	12.7%

Source: Datastream

imperfections in measurement rather than to differences in management performance. However, ROE biases are still little understood among practitioners, and we think that this needs to change, especially for financial firms. For example, the recent *Financial Times*/PriceWaterhouseCoopers text on cost of capital (Ogier *et al.*, 2004) makes no mention of ROE biases, nor, indeed, of frictional costs.

3.1.4 The classic example in Brealey & Myers, Section 12.5 (1983) compares pharmaceutical and chemical companies. Chemical companies have mostly fixed assets, consisting of bricks and mortar — or plant and machinery. Pharmaceutical companies have many intangible assets, such as intellectual property (research and development) and patent rights.

3.1.5 Financial statements typically capitalise fixed assets, but do not recognise intangible assets. Therefore, the pharmaceutical company will produce a higher return on equity number than the chemical company, even if the underlying cash flows are the same. This shows that the definition of equity is not absolute, and may not be applied consistently, even within a single organisation.

3.1.6 Thus, the low ROE for U.S. non-life insurers can be explained by the relatively low barriers to entry, yet the need for significant capital to support non-life risk. Non-life insurers, therefore, tend to have relatively low franchise value and relatively high equity, relative to annual profits. As shareholders pay little more than net assets to acquire shares in a non-life insurer, a relatively low ROE can still result in an acceptable market return to shareholders. There is no reason to doubt that the bias can be sustained indefinitely.

3.1.7 Of course, equity is one factor of production which happens to be important for financial firms. Another factor of production is iron — the entire workforce may be supported by chair legs made of it. Car manufacturers and construction firms have a low return on iron (ROI), because their businesses are iron intensive. By this measure, financial firms are highly efficient, because their ferrous needs are actually quite modest. This does not mean, however, that society, as a whole, benefits from closing down manufacturers and reallocating resources to the financial services industry. After all, someone has to assemble the cars and the buildings so that the insurers can insure them.

3.1.8 Brealey & Myers (1983) outline the theoretical solutions to ROE bias in terms of discounted cash flow valuations for assets and liabilities, using methodologies consistent with the market value of the firm. However, they then dismiss these tools for practical use, arguing that the cost of preparing the discounted cash flow projections would exceed any benefit of improved management.

3.1.9 For financial firms the situation is different. Discounted cash flow valuations are already available for most assets and liabilities, for with-profits insurance funds, and these are increasingly prepared under market consistent methods, as part of principles and practice of financial management. Therefore, the kind of improvements in performance measurement which elude pharmaceutical companies can still benefit financial firms.

3.1.10 We were flattered to notice that our original drafting of this section featured in a specimen examination paper for Communication, subject CA3. Candidates were invited to explain the nature of the problem for the benefit of a works production manager. The question, and model solution, are available at www.actuaries.org.uk.

3.2 *The Illusion of Equity Efficiency*

3.2.1 A focus on ROE, rather than on total shareholder returns, creates two particular biases:
— a failure to recognise activities which build franchise value; and
— an artificial incentive to reduce reported equity.

3.2.2 The desire to reduce ‘expensive’ balance sheet equity has been a major plank of financial strategy, not only for banks, but also for other financial institutions. The market in credit derivatives and asset securitisation has been helpful for institutions seeking a leaner, more highly geared, balance sheet. Similar patterns are seen elsewhere, for example airlines, which increasingly lease rather than buy their aircraft.

3.2.3 The case for (and against) straightforward leverage of a balance sheet is well rehearsed in the world of classical finance literature. However, in the world of accounting numbers, the benefits of a financial engineering solution which reduces balance sheet equity may be very misleading, since the accounting cost of debt (in the form of leasing arrangements, for example) can appear low compared with an artificially inflated ROE (due to the understated balance sheet equity in the denominator). Indeed, in the airline industry it is interesting to consider whether apparently more efficient capital usage, through reduced balance sheet equity, has produced the better shareholder returns suggested by the accounting measures, or has simply increased the financial distress costs identified by classical finance.

3.3 *Return Biases and Portfolio Selection*

3.3.1 There is a helpful analogy which relates the group function in a corporation to the role of an active investment manager. Just as an investment manager seeks investments with a high return and low risk, so the corporate centre of a financial group allocates capital to businesses which produce a high return on capital.

3.3.2 Return biases often arise because of inconsistencies in accounting treatment. For example, consider a three-year-old portfolio of poorly performing mortgages, compared with a portfolio of high interest credit card receivables. The mortgage portfolio might have a return on equity of -10%, while the credit cards show a return on equity of 30%. A RORAC analysis suggests the allocation of more equity to the credit card business at the expense of the mortgage business.

3.3.3 This strategy involves disposing of mortgage assets and acquiring new credit card assets at face value. We would love to return bad mortgages to the shop, like unwanted Christmas presents, and trade them in for something which suits us better. Unfortunately, financial markets are less forgiving than Marks and Spencer. You can buy and sell blocks of business at market value, for example using securitisations, but historic cost transactions are harder to arrange. This is especially the case in our example, where the market value of the mortgage book is below face value, and the credit card receivables trade above face value. Our RORAC methodology has told us nothing more than to buy assets cheap and sell them dear.

3.3.4 Ironically, the transaction may still be positive for management bonuses. The reason is that the sale and acquisition of portfolios may be treated as an exceptional item, and therefore taken below the line of the income statement. It is interesting to consider whether the reduced flexibility under International Financial Reporting Standards (IFRS) relating to below-the-line items will dampen the enthusiasm for RORAC driven transactions.

3.3.5 If we revalue the portfolios at their securitisation value, the ROE differential goes away. The high ROE on the credit card portfolio is an artefact of its prudent valuation, just as the low ROE on the low quality mortgage portfolio is an artefact of its less than prudent valuation. Decisions to buy or sell portfolios turn out to be driven by ROE bias, and not by shareholder value.

3.3.6 The same issues arise in insurance. Businesses with low ROE often correspond to life products with embedded options and guarantees; the corresponding liabilities may be undervalued on the balance sheet. Conversely, non-life run-off portfolios can appear to have a high ROE, because of the prudence implicit in undiscounted reserves. However, it would be naïve to assume that capital can easily be reallocated from life to non-life, as this would require charitable third parties prepared to take on the life liabilities at their understated value, while happy to pay an undiscounted price to cede non-life liabilities.

3.4 *Return Biases and Monopoly Regulation*

3.4.1 Accounting biases also have interesting consequences for price regulation. Some competition regulators argue that a loading for profit of, say, \$10, is evidence of market failure, because shareholders get \$10 for nothing. Their solution is the regulation of the price of financial services to eliminate the \$10 margin, thus protecting consumers from exploitative financial firms. Most recently, we have seen such intervention in lending to small and medium sized enterprises.

3.4.2 To charge the \$10, the firm must persuade the regulator that \$10 represents a fair return to shareholders. The difficulty is that a CAPM return applied to capital produces a lower number. However, we now know that such an application of CAPM creates a bias which may require correction.

3.4.3 A firm might legitimately argue that its ability to command premium prices reflects past investments in customer service and branding. If the observable franchise value reflects a stream of $10 profits, then the franchise correction to the ROE bias justifies the inclusion of the $10 profit margin.

3.4.4 Suppose that the regulator disputes the use of franchise value to remove ROE bias. The predictable effect of such an attitude is to discourage firms' investment in the brand or service which develops the franchise value. This results in reduced customer loyalty and greater customer switching between financial service providers, and — here is the irony — the switching generates additional administrative and acquisition costs of, say, $20, which are recognised in the regulatory formula, thereby permitting firms to pass them on to consumers. A naïve monopoly regulators' impact is to confiscate $10 from consumers, $10 from shareholders and spend the $20 on avoidable administration.

3.4.5 We conclude that, while firms need to understand ROE biases for accurate pricing, so too must price regulators.

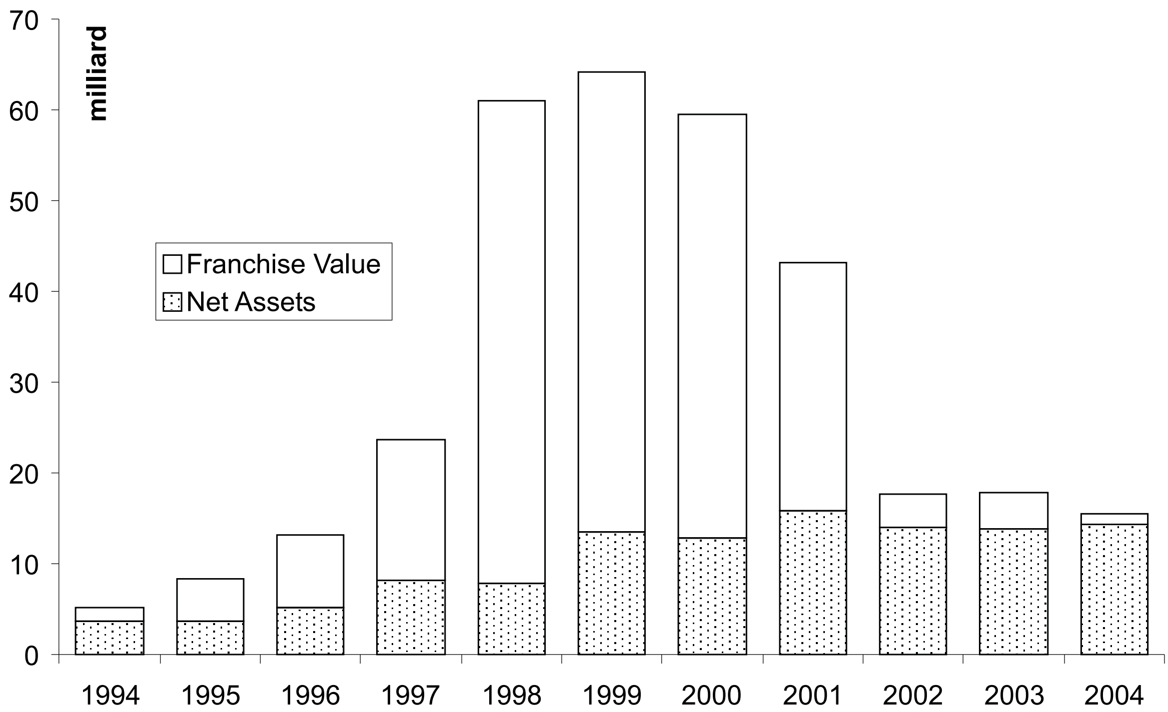
3.5 *Fixing the Bias*

3.5.1 On a market value basis, the shareholder return is any gain in the market value of the share plus dividends paid (or less any new equity subscribed). We can also write the total shareholder return as a weighted average of the return on equity and the franchise growth rate, for the following reason:

$$
\begin{aligned}
&(\text{Equity} + \text{Franchise}) \times \text{Total shareholder return} \\
&= \text{Increase in net assets} + \text{Increase in franchise value} + \text{Dividend} \\
&= \text{Increase in franchise value} + \text{Retained profit} + \text{Dividend} \\
&= \text{Franchise} \times \text{Franchise growth rate} + \text{Equity} \times \text{Return on equity}.
\end{aligned}
$$

3.5.2 This identity holds in theory, but validation in practice is a fearsome challenge, especially when working from published financial statements. It is difficult to find a single company for which the stated dividend plus the increase in equity actually equals the stated profit. The reasons are many and varied, sometimes relating to prior year adjustments or restatements due to share splits, scrip dividends or other capital transactions. Current accounting standards provide a number of loopholes, where firms can report changes to balance sheets without any effect on the income statement. IFRS should alleviate this problem, although not entirely eliminate it.

3.5.3 It would be convenient for performance measurement if the return



Source: Aegon historical data, www.aegon.com

Figure 5. Aegon net assets and franchise value, year ends 1994 to 2004

on equity corresponded directly to the total shareholder return. This would happen if franchise growth is the same as return on equity.

3.5.4 It so happens that many firms currently trade somewhere near their net assets. However, this has not always been the case. Figure 5 shows recent changes in net assets and franchise value for Aegon.

3.5.5 Recent cross-sectional analysis by Scotti (2005) examined factors which explain franchise to net asset (or, equivalently, market capitalisation to net assets, the so-called market-to-book) ratios in U.S. insurers. It applied a number of adjustments to translate accounting assets and liabilities to a fair value basis. Its analysis indicates that three factors — underwriting profitability, growth and size — together explain two thirds of the observed variability in franchise to equity ratios.

3.5.6 Where franchise value is material, we need to restate our weighted average to make the return on equity the subject of the equation:

$$\begin{aligned}\text{Return on equity} &= \text{Total shareholder return}\\ &\quad + \text{Franchise/Equity} * (\text{Total shareholder return} - \text{Franchise growth rate}).\end{aligned}$$

3.5.7 To set a return on equity target, we therefore need to estimate the franchise/equity ratio, as well as the franchise growth rate. This process is illustrated in Figure 6.

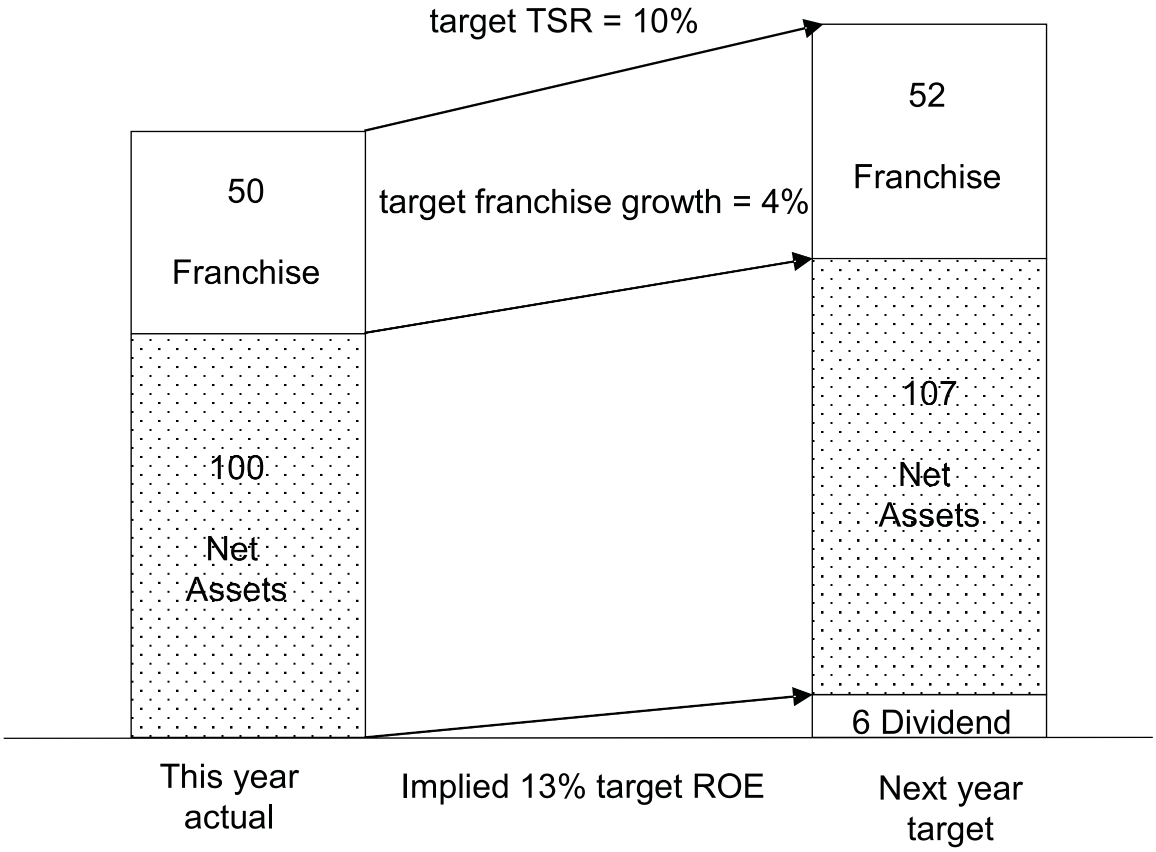
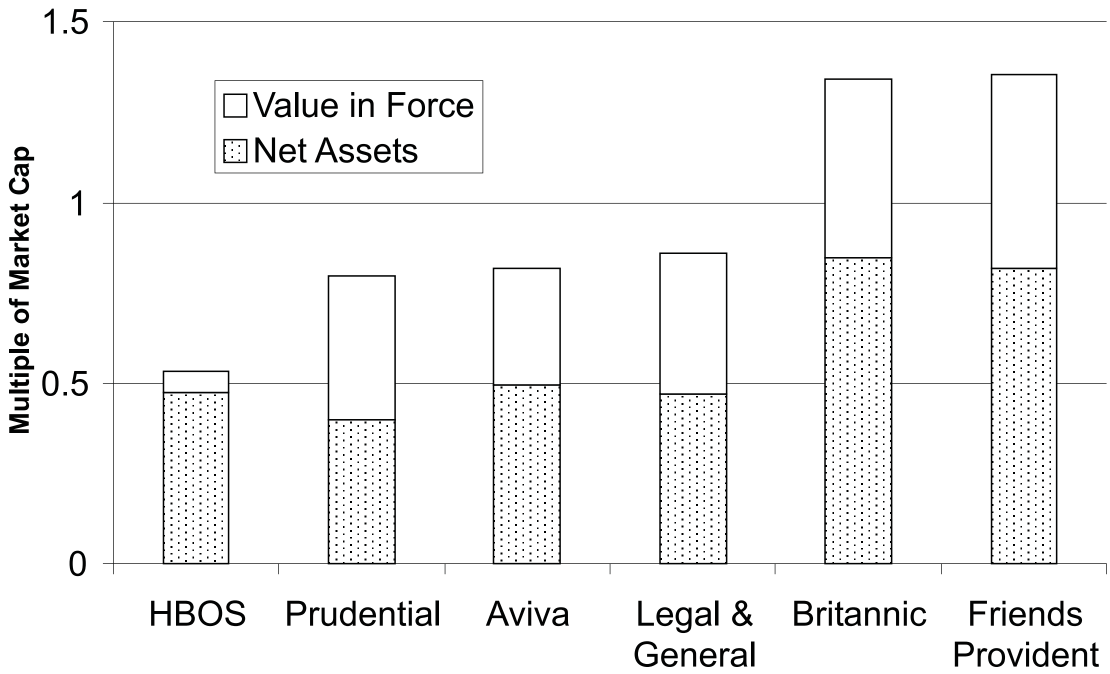


Figure 6. Grossing up the return on equity target to allow for franchise growth

3.6 *Embedded Value*

3.6.1 One of the most substantial innovations to address the ROE bias is the emergence of embedded value reporting in Europe. The idea is to calculate, explicitly, the value of future margins in a block of in-force insurance business, allowing for capital costs, then treat this value as an additional asset, called 'Value in Force' (VIF). Embedded value refers to the sum of net assets and VIF. An entire parallel accounting system is developed, including balance sheet and income statements, all assuming that this value of future margins is treated as an asset. There has been a significant change in embedded value methodology over the last year, with the arrival of European embedded values, which reflect more explicitly the cost of options and guarantees as well as risk costs. The figures presented here are based on year end 2004 figures, and are all on the more traditional basis. For a fuller description of the old and new methodologies, see O'Keeffe *et al.* (2005).

3.6.2 We can see from Figure 7 that the value in force explains some of the gap between net assets and market capitalisation. For this reason, ROE numbers based on embedded value accounting are often close to market total



Source: Report and accounts, Datastream

Figure 7. Net assets and value in force at 31 December 2004 for selected U.K. financial firms

shareholder return (TSR) numbers. A firm which increases embedded value ROE is therefore likely also to improve TSR.

3.6.3 Embedded value reporting has a particular strength in relation to multi-year contracts. For example, let us consider a ten-year term assurance product. An insurer may price this product allowing for the cost of capital over the ten years of the product.

3.6.4 Market consistent valuation tools, for example as used in the U.K. with-profits realistic reporting regime (Sheldon & Smith, 2004), would revalue the policy at the end of the first year, with no allowance for capital cost. As a result, all the margins emerge in the first year. The first year result shows a great return on capital, while older policies show a return on capital close to zero, because all the margins in the valuation basis were squeezed out in the first year. This gives rise to an ROE bias; we might get the impression that policies are only attractive in their first year.

3.6.5 Embedded value reporting resolves this issue by explicitly valuing the cost of the capital required to support the business as it runs off. In that case, the unwinding of the embedded value discount rate provides the profit needed to meet the cost of capital in each future year.

3.6.6 To remove the ROE bias entirely, a different sort of embedded value is required, which reconciles to market capitalisation. This includes making an allowance for new business and using profitability assumptions consistent with the market view rather than based on internal management views, as well as allowing for all capital likely to be held on a balance sheet rather than a statutory minimum.

3.6.7 Insurers seeking to explain embedded value techniques would find it helpful if these tools were not unique to the life insurance industry. It is interesting to consider why banks, non-life insurers and other financial firms have not developed their own versions of embedded value reporting, especially in relation to the valuation of bank loan portfolios. Here are three possible reasons:

— There is a perception that the need for embedded value arises only because the statutory reporting regime is artificially prudent. Some would argue that, if you value the balance sheet accurately in the first place, then there is no need for embedded value adjustments.
— Overdrafts and credit cards are open-ended facilities. Customers have flexibility over their loan balances, within authorised lending limits. Therefore, the projection of future margins is more speculative than is the case for an insurance policy with fixed contractual cash flows. Indeed, the fundamental embedded value distinction between in force and new business is difficult to apply. If a customer starts to use a long-standing overdraft facility, is this in-force or new business? Similar issues apply to some degree with personal lines insurance — it is ambiguous whether a renewed motor insurance policy represents new business or a low customer lapse rate.
— Banks, encouraged by regulators, have poured resources into the calculation of economic capital. Approaches abound for managing businesses and measuring value based on economic capital calculations. Arguably, the enthusiasm for economic capital methodologies has crowded out alternative value measures.

3.7 *Economic Capital*

3.7.1 There are two popular approaches to the allowance for risk in performance measurement:

— modify the required return to reflect risk, for example using replicating portfolios; or
— modify the equity on which a cost of capital is charged.

3.7.2 The first of these is called 'contingent claim pricing' or 'asset pricing theory', while the second route is called 'economic capital' or 'RORAC' — return on risk adjusted capital. *Economic capital* is the amount of equity required such that a firm has a specified probability of meeting a defined liability over a defined time horizon. The probability is often set by reference to historic defaults for corporate bonds with a particular credit grade. Economic capital is often measured on a fair value net assets basis, ignoring the impact of regulatory or accounting constraints.

3.7.3 The methodology recognises that the required profit for each business unit depends on its risk. Capitalising operating units to an agreed rating standard is an attempt to standardise the risk. The implicit assumption is that two units capitalised to the same rating standard would merit the same required return. In other words, the cost of capital is assumed to be a function of the aspired credit grade. Risk from a shareholder perspective is equated to the risk of default.

3.7.4 Economic capital can be looked at from a bottom-up and a top-down manner. Top-down economic capital can mean different things to different people — for example, owing to a different choice of time horizon (measuring daily, yearly or longer-term risk) or to different approaches to the constraints imposed by regulatory regimes (with pure approaches pretending that these regimes do not exist). Economic capital can help investors to understand the level of risk capital and diversification which a company has. It can be useful for regulators and rating agencies as well.

3.7.5 The divergence of approaches leads to the following common criticisms. Firstly, it is impossible to compare economic capital from one organisation to another, because no two organisations use the same metric. This undermines claims that 'economic' capital measures are free of arbitrary conventions or are more fundamental than accounting or regulatory capital measures. Secondly, the lack of uniform standards and independent audit procedures leave users of economic capital numbers uncertain as to how they have been calculated or how reliable they are.

3.7.6 For pricing, a bottom-up approach can be used, which incorporates both market consistent pricing and economic capital in respect of non-hedgeable risk, from which a frictional cost of risk capital is obtained. To the extent that a mismatch risk is taken, additional capital is calculated and charged for. We develop this further in Section 4.

3.7.7 Some businesses develop more sophisticated approaches, where capital is tiered according to risk, each tier implying a different capital cost. Hare & King (2005) recommended four tiers, corresponding to an internal capital assessment, additional regulatory capital, respectability capital and excess capital, as shown in Figure 8.

3.7.8 The introduction of economic capital models has been a catalyst for other developments, which are useful in their own right, including:

— an understanding of the need to manage risk as well as to increase reported profit;
— the discipline of implementing risk modelling capabilities in each business unit;
— a focus on value created for shareholders, and its link to financial planning; and
— an input to distance-to-default models for quantifying credit risk.

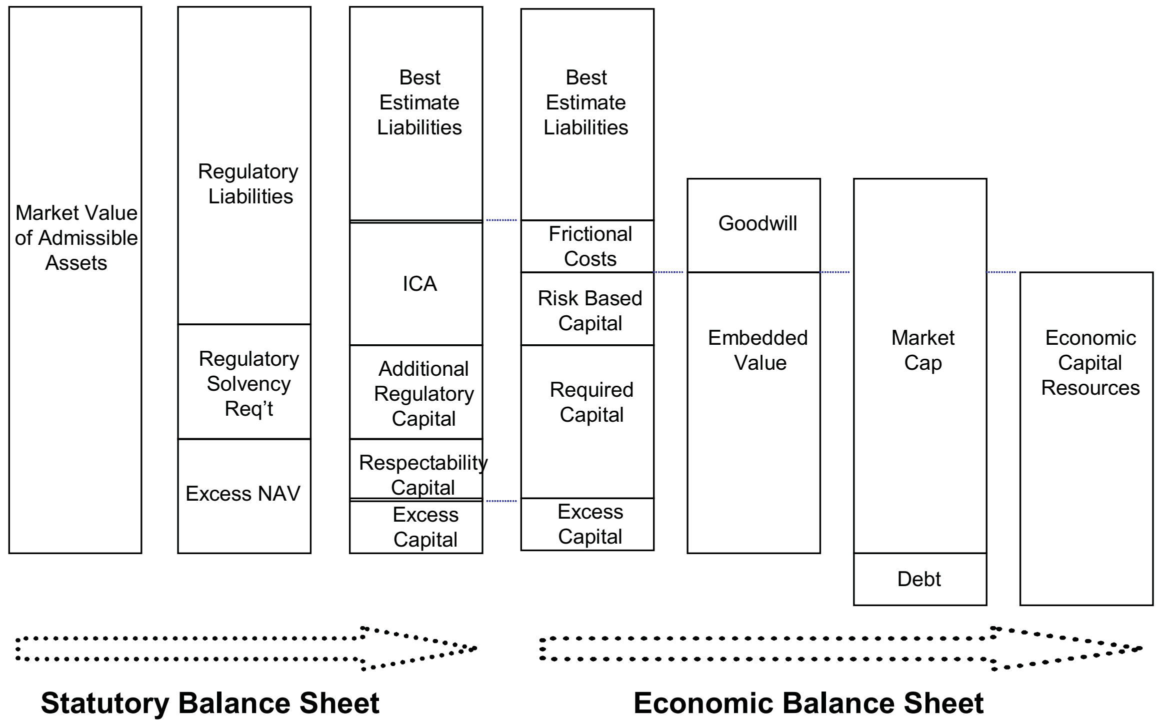


Figure 8. Tiers of capital in a life insurer (based on Hare & King, 2005)

3.7.9 The question arises of whether RORAC is helpful in removing the known biases in ROE. We suggest that the answer in many cases is no. In particular, the RORAC concept of capitalising to a given credit rating is inconsistent with asset pricing theory — which has long abandoned the notion that extreme percentiles are the main determinant of expected market returns. Secondly, RORAC makes no attempt to allow for differences in franchise value.

3.7.10 For example, many diverse financial businesses find that their highest RORAC units are transaction based, such as current accounts or share dealing services. These units are allocated little capital, as they carry only a modest amount of operational risk. On the other hand, these businesses are relationship based, and so can comprise significant franchise value.

3.7.11 On paper, it might seem a simple matter to reallocate capital from more capital intensive businesses, such as mortgages or life insurance savings products, to transaction-based units. However, a paper re-allocation of capital is not sufficient to generate the extra transaction volume. Extending the franchise generally involves expenditure in marketing, product development, commission or acquisition of competitors. Once again, the apparent attractiveness of transaction-based units turns out to be an artefact of ROE bias, and not a profound statement about shareholder value.

4. Corporate Valuation and the Cost of Capital

4.1 *Valuation Supported*

4.1.1 For a given profit target, we can ask: "How would the market value a company which is expected to meet its profit targets?"

4.1.2 Let us suppose, first, that a company's target is to earn its own cost of capital, measured on its accounting net assets. Such a target could provide adequate compensation for a shareholder buying shares at net asset value. In other words, it could support a share price equal to net assets.

4.1.3 Let us suppose, instead, that a company seeks only to break even, that is, make zero profit. Shareholders, therefore, have no prospect of return, either in dividends or in growth, but, nevertheless, the company keeps going indefinitely. This is clearly worse than a company which earns only its cost of capital. Shares in such a company would arguably be worthless, as the recovery of net assets is postponed to an indefinite point in the future. To put this differently, a target merely to make a profit would support a share price of zero. This is clearly inadequate, from a shareholder perspective, if they have paid a positive price for the shares.

4.1.4 Consider an example of a company with €500 of assets and €400 of liabilities, and so €100 of equity. This does not mean that the company trades at €100 in the market. Maybe the company trades at €150. It makes little sense to assume that shareholders require a return only on the €100. A combination of previous management actions and a favourable trading environment has persuaded the market that a return on €150 is achievable, and it is up to current management to deliver a return to support the total market value. It is misleading to regard everything above a return on €100 as 'value added'. The market gives a clear signal, in the market price, of the future returns required.

4.1.5 While some managers thrive on challenging targets, not all fall into this category. There is not much enthusiasm for targets based on market capitalisation. Instead, many managers argue for a reduction in the capital on which a return is required, even below the level of accounting net assets.

4.1.6 In insurance, there is a history of locking in adjustments, which implicitly charge a cost of capital only on that part of net assets required as a legal minimum — in our example this might be €60. Banks have gone further in their use of economic capital, which seeks to identify the capital required to achieve a given probability of financial distress — which might, in our example, give an economic capital of €40. Both of these approaches would support a share price significantly below net assets — so probably represent a poor deal for shareholders.

4.1.7 These values, supported by different capital definitions, are illustrated in Figure 9.

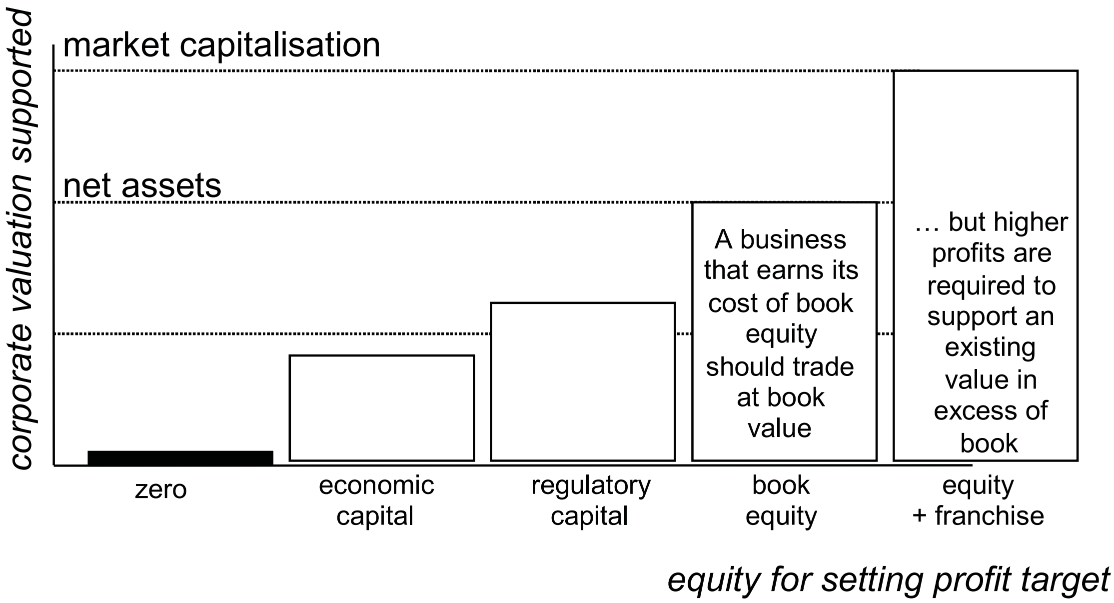


Figure 9. Corporate valuation and profit targets

4.2 *Discounted Cash Flow Valuations*

4.2.1 The cost of equity is defined in terms of the most fundamental of economic calculations — discounted cash flow, in this case the present value of dividends. It is often difficult to establish the dividend impact of a transaction. For example, an insurance contract may have a known distribution of claims, but to model the dividend impact, we would need to know whether a future claim reduces dividends or retained profit. It would be more convenient to study the value impact by looking at the profitability of a contract, rather than at its dividend impact.

4.2.2 Fortunately, there are some simple algebraic rearrangements of the dividend discount model which enable us to measure profits instead of dividends. In this section we describe these ideas.

4.2.3 We need some notation. All these variables are measured at annual intervals. We assume that dividends are paid at the year end, immediately before the accounts are drawn up:

— E_t is the shareholders' equity on the balance sheet at time t; this is the difference between the assets at time t and the liabilities;
— F_t is the franchise value at time t, that is the excess of the market capitalisation over the shareholders' equity;
— D_t is the dividend paid to shareholders at time t; and
— COE is the cost of equity.

4.2.4 By definition of the cost of equity, we know that the market capitalisation is the present value of the dividends. This means that:

$$E_0 + F_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1 + COE)^t}.$$

4.2.5 The profit during the year can be expressed as:
— distributed profit, that is dividend; plus
— retained profit, that is increase in equity.

4.2.6 To make the accounting work, we need to treat new equity raised as a negative dividend. We can then rewrite the corporate valuation to refer to profit (the term in square brackets):

$$E_0 + F_0 = \sum_{t=1}^{\infty} \frac{[D_t + E_t - E_{t-1}] - COE.E_{t-1}}{(1 + COE)^t} - \sum_{t=1}^{\infty} \frac{E_t}{(1 + COE)^t} + \sum_{t=1}^{\infty} \frac{E_{t-1}}{(1 + COE)^{t-1}}.$$

4.2.7 The last two sums cancel, except for the term E_0. This implies that:

$$F_0 = \sum_{t=1}^{\infty} \frac{ROE_t - COE}{(1 + COE)^t} E_{t-1}.$$

4.2.8 We have, therefore, expressed the franchise value as the present value of profit, minus an additional term $COE.E_{t-1}$. This term is the product of the cost of equity and the equity on the balance sheet. It reflects the profit margin required to satisfy a valuation equal to the net assets.

4.2.9 The expression $[ROE_t - COE]E_{t-1}$ is sometimes called the *economic value added*, or (as that phrase has been trademarked by Stern Stewart and Co.), *economic profit*. The accounting equity represents original capital subscribed plus cumulative retained earnings, that is the historic shareholder investment. To the extent that franchise value is positive, it represents the gain to shareholders over and above their historic investment. Our identity has, therefore, demonstrated that franchise value is the present value of economic profit discounted at the cost of equity.

4.2.10 In Section 3 we repeatedly cautioned against biases which arise when a market value cost of equity is applied to an accounting equity number. Yet, here we have precisely that product appearing in a valuation identity. The reader might conclude that it is correct, after all, to apply a market return to accounting equity in order to establish economic profit.

4.2.11 On the other hand, we could argue that economic profit turns out to be precisely the accounting bias, that is the difference between the accounting return ROE and market return COE. The fact that franchise value is the present value of accounting biases emphasises the importance of accounting biases. If we use COE naïvely as a hurdle rate for ROE, we are unlikely to gain insight into franchise value.

4.2.12 The relationship between franchise value and economic value added does not depend on finding a theoretically pure or market consistent set of accounting procedures. Of course, a change in accounting practice would affect reported equity, and so would alter both franchise value and future reported profit. However, the identity that:

$$\text{Franchise value} = \text{Present value of economic value added}$$

holds, irrespective of the chosen accounting conventions. We cannot use this identify to justify one accounting framework against another, or to argue that one definition of profit is more economic than another.

4.2.13 To take two extreme examples: suppose that we could find an accounting standard where net assets were equal to equity market capitalisation. In that case, the observed return on equity is the market return on the shares, which is a form of cost of capital. Furthermore, since the market capitalisation and the balance sheet equity remain in step, we find that the franchise value is zero. To take the other extreme, we could account on a cash flow basis with all assets and liabilities valued at zero. In that case, the 'economic profit' is defined as the dividend payment, the net assets are zero, and the franchise value is the present value of dividends.

4.3 *Market Consistent Valuations*

4.3.1 Financial firms usually have assets or liabilities which also trade in financial markets. Ensuring that these assets and liabilities are priced consistently with capital markets removes some of the subjectivity which accompanies dividend discount models. Other subjectivity remains, particularly in relation to expenses, margins and new business volumes.

4.3.2 We define a firm's asset returns using the linked method, as is common in the fund management industry. The return over a short period is constructed as investment income plus any capital gains on asset revaluation. In this construction, we define the values in relation to their balance sheet accounting treatment. For assets such as loans, we need to translate banking terminology into investment management terminology, as shown in Table 6.

Table 6. Conversion of bank loan terminology into investment terminology

Bank loan terminology	Investment terminology
Loans advanced	Investments bought
Loans redeemed	Investments sold
Interest received	Investment income
Increase (decrease) in bad debt provisions	Capital loss (gain)

4.3.3 We define liability returns as the returns obtained over a period by a firm's creditors — that is depositors, policyholders and bondholders. This reflects the fact that a financial firm's liabilities are someone else's assets. We consider all creditors in aggregate. Our return calculation reflects opening and closing balance sheets, as well as intermediate cash flows. In this calculation, inward cash flows, such as funds deposited or insurance premiums, are treated as investment purchases by the creditors. Conversely, outward cash flows, such as deposits redeemed or insurance claims paid, are considered to be the proceeds of investment sales.

4.3.4 Now let us suppose the following:

— the financial firm has no expenses and pays no tax;
— investors can also invest directly in the firm's assets;
— investors can invest directly in the firm's liabilities;
— the asset and liability valuations on the balance sheet are at market prices;
— external investors earn the same asset and liability returns as the firm; and
— the firm never goes bankrupt.

4.3.5 In this case, we could compare a financial firm to a geared investment trust. The investment trust raises debt and equity in order to acquire assets. Most investment trusts trade at close to net asset value, or slightly below. Franchise value is negligible.

4.3.6 The absence of franchise value comes from an arbitrage argument. To replicate the firm's assets, we can either buy those assets directly, or we can acquire the firm and simultaneously buy the liabilities in the market. Therefore, the price which we pay for the firm must be the difference between its assets and liabilities.

4.3.7 All of this thought experiment is counterfactual. Financial firms are not free of costs and taxes. Balance sheets are not always on a market value basis. Bankruptcy can happen. Firms' assets and liabilities are not all traded on financial markets, although some of them are. Franchise value exists. We now explore the implications of non-traded assets and liabilities.

4.4 *Margins and how they Persist*

4.4.1 It is worth considering the reason why retail borrowers pay higher interest rates than capital markets, and why investors in insurance policies may obtain lower returns than other investors in the market. The reasons for these margins explain why financial intermediaries have a role in society, and why shareholders get compensated for financing these intermediaries.

4.4.2 A different way of asking this question is to consider why retail customers do not simply buy and sell their insurance policies, savings and debt on a huge public exchange. The answer is that operating such an exchange would be costly; so costly, in fact, that less transparent bilateral transactions with financial firms are economically more efficient. To understand the role of financial firms, we need to explain why their cost structure is efficient compared with the public exchange alternative. We investigate several reasons:

- customisation and transaction costs;
- fiscal reasons;
- market failures;
- regulated sales process;
- depositor protection;
- relationships and service; and
- underwriting and information asymmetries.

4.4.3 Personal loans, deposits and insurance policies are customised products which are special to one individual. Customers will pay to find a product which is suitable for them. Financial intermediaries offer an environment which makes that search easy, and allows deals to be consolidated in smaller quantities than would be possible in capital markets. This explains, from the customer perspective, why companies may earn margins from meeting customer needs.

4.4.4 Products delivered by financial intermediaries may be taxed or subsidised differently from those delivered directly via capital markets. For example, in many countries benefits paid under insurance policies are tax exempt. Income tax may be rebated in relation to pension savings or mortgage interest. Smaller effects may include a financial firm's ability longer to defer tax on capital gains, or to take advantage of tax management economies of scale, which would be too costly for customers to implement on their own.

4.4.5 From the intermediaries' perspective, it is worth considering why retail profit margins do not get competed out of the market. One reason could be lack of competition. Perhaps, a few large players have sufficient economies of scale or control over distribution channels to exclude smaller competitors. Price fixing cartels may be in operation.

4.4.6 Customers may pay margins in order to benefit from a regulated environment. In many cases, financial intermediaries operate within a regulatory environment which seeks to find the most suitable product for the consumer, with a wide variety available under one roof, and rights of redress if, with hindsight, advice was inappropriate. Industry or government backed funds often provide an additional layer of protection if the original contracting firm is unable to meet its obligations. In comparison, capital markets usually force buyers to perform their own research, and stand or fall by their own decisions.

4.4.7 Although we measure accounting returns on asset value, the accounting asset value does not capture all the resources which shareholders may have invested. Airlines also invest in their people and distribution capabilities. A producer must market his products to potential customers. A hotelier seeks to build a good reputation. All of these firms are spending shareholder resources in order to gain some intangible benefits, even though the accounting treatment immediately writes off the expenditure. Likewise, many financial service customers value their relationships with banks or insurers. Competition is not purely on price. Retail financial services rely to a degree on trust; this trust may take many years to establish. Financial services will invest in infrastructure, customer service and brand in order to retain profitable customers.

4.4.8 The underwriting of loans and insurance policies involves the processing of personal and confidential information. Privacy concerns prevent individual policies from trading in open markets. Part of the profit margin which financial intermediaries achieve can be considered a fee for the safe and responsible handling of personal data.

4.4.9 These are the economic factors which we can expect to explain the franchise value which we see in financial markets.

4.5 *Asset Swaps*

4.5.1 We have considered how financial firms are valued when their assets and liabilities are traded. We now seek to construct a similar arbitrage argument for non-traded assets. To do this, we need to assume that outside investors can access a firm's assets, but only via a derivative contract, which we describe, for ease of explanation, as an asset swap. Strictly speaking, the contract may be better described as a total return swap, but we adopt the term 'asset swap' for convenience, when introducing the concept of a 'liability swap' counterpart, below.

4.5.2 The return on assets is a vital lever of shareholder value for most enterprises. An airline seeks to gain the maximum revenue from a limited number of airplanes. A producer seeks the maximum production from his farm or factory. A retailer maximises traffic through a supermarket. An hotelier seeks to maximise occupancy. All of these businesses try to create wealth by working their assets as hard as possible.

4.5.3 The accounting return on assets is usually measured as income less expenditure, making an allowance also for depreciation on the assets. The accounting return does not try to incorporate past intangible investments in items such as staff training, marketing or distribution networks.

4.5.4 Therefore, even when capital markets are competitive, we would not expect accounted returns on balance sheet assets to correspond directly to returns available in competitive markets, unless some allowance is made for investments in intangibles.

4.5.5 In the same way, mortgage lenders seek to maximise the return on their loan advances. For example, a mortgage lender may seek to lend money at some margin over base interest rates to prudent customers who will repay in full.

4.5.6 In the case of mortgage lending, we do have some measure to relate accounting returns to market returns. The comparison arises because many lenders now securitise their asset portfolios, and sell them on into capital markets. Interestingly, the market value of a securitised mortgage portfolio is usually higher than the face value of the mortgages. For example, a lender may be able to lend £95m face value of loans, and re-sell them via securitisation for £100m.

4.5.7 An alternative route to compare market returns is to enter an *asset swap*. Under the terms of such a swap, the lender pays to a third party the return on a mortgage portfolio, in return for a series of floating payments, for example LIBOR + 1%. The 1% is a market number which would be determined by the asset swap market at the time when the swap is entered. In some sense, the 1% compensates the lender for past investments in intangible assets, such as underwriting capability, which allow the lender to write mortgages profitably. The other swap party is then able to gain exposure to the return on mortgage portfolios, paying an additional 1% to avoid the overhead of building its own underwriting and distribution capability.

4.5.8 The present value of the 1% stream for a block of business roughly corresponds to the 5% difference between loan face value and securitised value. Conversely, if we re-invent the asset swap based on the securitised value, the 1% spread above LIBOR disappears. After all, the point of securitisation is to give investors direct access to mortgage returns. When investors have an option of direct investment in a securitisation, there is no reason to forgo 1% return to achieve the same exposure via an asset swap. We note that, in real markets, the swap based on market value is more likely to trade than the swap based on face value. However, as the price of the one is easily transformed into the price of the other, we continue here to consider the more exotic book value hedge, which will make our subsequent arbitrage arguments more transparent.

4.5.9 Asset returns for insurers are much more challenging, as the majority of insurers' assets will be direct investments held at market value. Investment markets are highly competitive, and it is open to question whether insurers, generally, have any competitive advantage in selecting undervalued investments. In theory, one could imagine an asset swap where an investor pays LIBOR in exchange for the return on an insurer's asset portfolio. However, it is unlikely that the insurer would gain any spread above LIBOR, because the swap counterparty has the alternative of simply buying the underlying assets.

4.5.10 This introduces an important difference between insurers and other firms. Most firms would expect to gain a positive spread to LIBOR if they swapped their asset return in the market, while insurers would not expect to see such a spread. This conclusion changes if assets are not accounted at market value — for example, if assets are held at historic cost, which happens to be below current market values, then an asset swap for the accounting returns could well deliver a spread above LIBOR.

4.5.11 There are several more subtle effects which affect asset swap margins. There may be small margins for traded assets, where there are differences in tax or liquidity. An example would be gilts going special on repurchase, when there is a squeeze on a particular stock.

4.5.12 The margin is positive if the firm enhances returns by particular skills which are not generally available. Examples would include loan underwriting and active investment management. Equally, it is possible to be less than averagely skilful, in which case the market determines a negative swap spread. A negative swap spread could also occur if the underlying assets were priced to produce profit margins for third parties.

4.5.13 The asset swap margin is unaffected by market risk premiums. If an asset earns a long-term risk premium relative to LIBOR, then the asset swap markets do not reflect that risk premium in the form of a spread above LIBOR. Instead, the risk premium is reflected in a positive expectation of swap payments, for the party paying LIBOR, and receiving the asset return. For this reason, risk premium estimates do not feature in valuations based on the arbitrage construction.

4.6 *Liability Swaps*

4.6.1 We have considered a derivative which exchanges asset returns for LIBOR plus or minus a spread. In the same way, we could consider derivatives based on liability returns. The derivatives which we have in mind exchange a liability return for LIBOR minus a spread. We use this convention — that the spread is measured below LIBOR — to be consistent with our asset swap convention that a positive spread means positive value for the financial firm's shareholders.

4.6.2 These derivatives also exist for some liabilities, although in a very limited form. Most of the examples in issue relate to catastrophe bonds, that is bonds whose coupons are forfeit in the event, either of a market loss of a certain size or the loss to a particular insurer. See Ahmed *et al.* (1997) for some descriptions of how they work. Scotti (2005) provides a chart of historic spreads to LIBOR on the few actively traded insurance linked securities, which mostly relate to non-life insurance. See Blake *et al.* (2006) for a discussion of similar products in relation to mortality risk.

4.6.3 Some liabilities are traded in the market, most obviously corporate debt in issue. In that case, we would expect the liability spread to be zero, or close to zero, enforced by arbitrage. Tax and liquidity may be possible reasons for deviations from zero.

4.6.4 If liabilities are valued on a market consistent basis, and there is no new business, then the liability swap margin should still be zero. Where a firm is open to new business, the liability swap margin approximates to the value of one year's new business, divided by total liabilities.

4.6.5 Where customers are granted below market returns, the liability swap margin reflects the discount to market returns which is passed on to customers. For example, a unit linked fund with a 1% management charge, or a deposit account paying 1% below LIBOR, would both imply liability swap spreads of 1%.

4.6.6 A positive liability swap spread could arise because of specialist skills not available to direct investors, such as skill in underwriting insurance risks.

4.6.7 To specify the liability swap fully, we must also define how it behaves in the event of insurer failure, where the closing accounting liability value may not be calculated, or at least not honoured in full. In that case, we assume that the liability return for the swap is computed from the opening liabilities and closing assets for the period. This reflects the fact that, in a ruin situation, control of a firm's assets passes from shareholders to creditors. This means that the parties to a liability swap participate fully in actual liability cash flows when ruin occurs, as opposed to the promised cash flows, which may be higher.

4.6.8 This definition of liability return is natural in relation to total return swaps for corporate bonds. However, it generates some additional complications in relation to non-traded liabilities.

4.6.9 For example, let us suppose that insurance liabilities are carried at market consistent value, based on gilt yields, with no deduction for insurer default. In that case, the liability swap spread is positive, because the actual return on the liabilities is worse than gilts in the default event. On the other hand, if liabilities are valued at a lower value which allows for the possibility of default, then the liability swap spread is zero.

4.7 *The Hedge Construction*

4.7.1 We are now in a position to apply a hedge construction to value a financial firm. We use the notation given in Table 7.

4.7.2 The income statement over the first period appears as in Table 8.

4.7.3 We leave to the reader the accounting exercise of verifying that our asset and liability returns do pick up all the relevant elements of the

Table 7. Notation used for hedge construction

A_t	Balance sheet assets at time t ($t = 0, 1$)
L_t	Balance sheet liabilities
F_t	Franchise value
R_A	Actual asset return
R_L	Actual liability return
R_f	Risk free rate, in this case LIBOR
k_A	Asset-related expenses, as a proportion of A_0
k_L	Liability-related expenses as a proportion of L_0
k_T	Tax paid, as a proportion of pre-tax profit
m_A	Margin above LIBOR on asset swap
m_L	Margin below LIBOR on liability swap

Table 8. First period income statement

Asset return	$A_0 R_A$
Less liability return	$-L_0 R_L$
Less asset expenses	$-k_A A_0$
Less liability expenses	$-k_L L_0$
= Pre tax profit	$A_0(R_A - k_A) - L_0(R_L + k_L)$
Less tax	$-k_T A_0(R_A - k_A) + k_T L_0(R_L + k_L)$
= Post tax profit	$(1 - k_T)\{A_0(R_A - k_A) - L_0(R_L + k_L)\}$
Less dividend	$A_1 - L_1 - A_0 + L_0$ $-(1 - k_T)\{A_0(R_A - k_A) - L_0(R_L + k_L)\}$
= Retained earnings	$A_1 - L_1 - A_0 + L_0$

profit calculation. This is an illustrative simplified example. The income statement could be made many times more complex than our example here, particularly in regard to taxation.

4.7.4 An investor in the firm starts with initial wealth $A_0 - L_0 + F_0$, equal to the market capitalisation. At the end of the year, that investor has the new market capitalisation plus the dividend. The total investor wealth at time one is then calculated as in Table 9.

4.7.5 It is now time to invoke the use of asset and liability swaps. We have assumed that we can swap the asset return R_A for $R_f + m_A$ and the liability return for $R_f - m_L$. In that case, we can construct the hedged wealth at time one, as in Table 10.

4.7.6 We now make the heroic assumption that we know in advance the one-period franchise value F_1. In that case, the hedged wealth at time one is

Table 9. Total investor wealth at time one

Opening market capitalisation	$A_0 - L_0 + F_0$
Closing market capitalisation	$A_1 - L_1 + F_1$
Dividend	$(1 - k_T)\{A_0(R_A - k_A) - L_0(R_L + k_L)\} - A_1 + L_1 + A_0 - L_0$
Total wealth at time one	$A_0 - L_0 + (1 - k_T)\{A_0(R_A - k_A) - L_0(R_L + k_L)\} + F_1$

Table 10. Development of hedged investor wealth over first period

Opening market capitalisation	$A_0 - L_0 + F_0$
Unhedged wealth at time one	$A_0 - L_0 + (1 - k_T)\{A_0(R_A - k_A) - L_0(R_L + k_L)\} + F_1$
Hedge	$(1 - k_T)\{A_0(R_f + m_A - k_A) - L_0(R_f - m_L + k_L)\}$
Hedged wealth at time one	$A_0 - L_0 + (1 - k_T)\{A_0(R_f + m_A - k_A) - L_0(R_f - m_L + k_L)\} + F_1$

deterministic, and we have created a risk free asset. To prevent arbitrage, it must earn the risk free rate. We equate the initial value rolled up at the risk free rate to the final value:

$$(1+R_f)\{A_0 - L_0 + F_0\} = A_0 - L_0 + (1-k_T)\{A_0(R_f + m_A - k_A) \\ - L_0(R_f - m_L + k_L)\} + F_1.$$

4.7.7 We can re-arrange this to give an inductive expression for the franchise value:

$$(1+R_f)F_0 = (1-k_T)\{A_0(m_A - k_A) + L_0(m_L - k_L)\} - k_T R_f(A_0 - L_0) + F_1.$$

4.7.8 We can interpret the right hand side as follows. The first two terms are the post-tax margins on assets and liabilities, after deduction of expenses. The penultimate term is a frictional cost of holding capital. In this case, the effect is double taxation, since, even if capital is invested in cash at the risk free rate, the interest received is subject to corporate tax.

4.7.9 We can apply this result inductively. For example, if we assume that the firm lasts 1,000 years, then the franchise value is zero after 1,000 years. We can then use the relation in annual steps to deduce F_{999} and so on, until, finally, we have an estimate of the warranted franchise value F_0 at time zero.

4.7.10 These ideas are well known in the finance literature. The allowance for tax in the general case derives from Modigliani & Miller (1963), which forms the basis for the WACC methodology. Myers & Cohn (1987) provided formulae equivalent to those above in the special case of non-life insurance.

4.7.11 We can relax the assumption that F_1 is known in advance. For example, it suffices to know the forward value of time one franchise. Alternatively, we might think that we know the mean value, combined with an appropriate risky discount rate. Although future growth rates are uncertain, the growth rate is often a balancing item in setting the modelled capitalisation to its market value. This gives a market implied growth rate which, despite fluctuations over time, is still more robustly defined than it might, at first, appear.

4.8 *The Mechanics of Failure*

4.8.1 We now consider a final enhancement to our model — an allowance for possible financial failure. One possible model of bankruptcy events is illustrated in Figure 10.

4.8.2 In default, a number of events are assumed to take place:

— net assets become negative (this is the usual trigger for default);
— control of the firm's assets passes from shareholders to creditors;
— shareholders also lose any franchise value which may have built up; and
— a tax loss occurs, resulting in an irrecoverable tax credit.

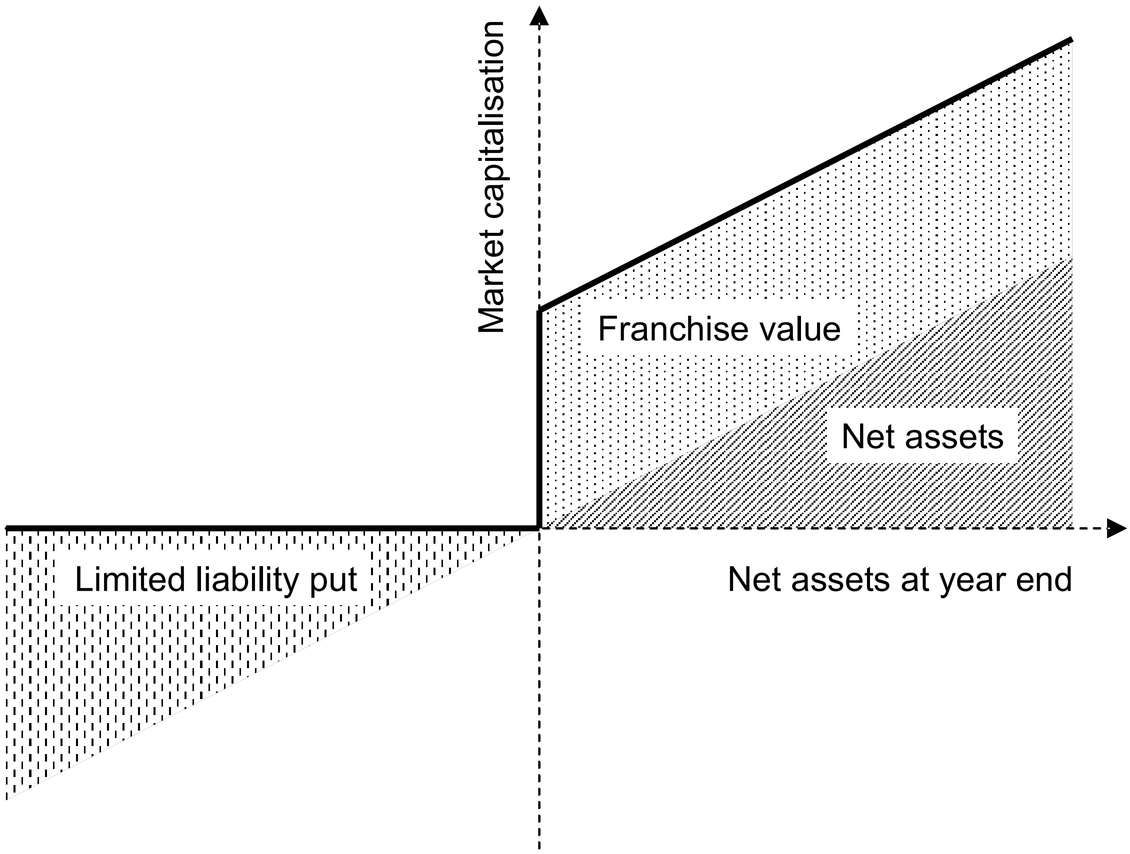


Figure 10. Market capitalisation as a function of net assets

4.8.3 The limited liability put option describes the excess of liabilities over assets in the default event. This is the shortfall in meeting creditor obligations. Sometimes the limited liability put is considered as an asset of shareholders, as it is shareholders who have the option to put the firm's assets up and walk away from the liabilities.

4.8.4 Regulators require financial firms to put up a certain amount of capital in order to transact business. It is possible that a firm could have positive net assets, but still insufficient to meet regulatory requirements. In that case, we assume that the firm is able to raise sufficient capital to support ongoing business. In this model, a state of zero net assets is assumed to represent the point of no return, either because the markets will not grant any further capital or because the regulator does not let you try.

4.8.5 Many real life default events are more complicated than our simple model permits. For example, firms may try to stave off ruin using financing arrangements to put VIF, or even future business margins, onto the balance sheet. This is a possible motivation for using embedded value accounting to determine net assets in our model — although you would need to be convinced that the Value in Force is realisable in a transaction.

4.8.6 The emergency use of transactions to boost net assets, supports analysts' arguments for restating financial firms' accounts in market consistent terms. However, a firm's deathbed conversion to market consistency may be partial in effect, with the greatest zeal being applied to areas where assets would increase or liabilities decrease.

4.8.7 More complex capital structures may provide for several tiers of capital, capable of defaulting separately. There may also be multiple accounting bases in use, for separate determination of solvency, tax or GAAP reporting. Segregated funds bring complications of their own; for example, a proprietary with-profits life office can fail because of inadequate total assets, but it could also fail if shareholder equity is exhausted, even when the life fund itself has a healthy surplus. See, for example, Doherty (1997), Froot & Stein (1998), Hancock *et al.* (2001), Merton & Perold (1999) for more details in these areas.

4.9 *Revised Hedge Valuation*

4.9.1 To complete our hedge construction, we need to introduce a final derivative which will enable us to price the risk of failure. We choose a credit default swap (CDS).

4.9.2 A CDS usually offers an option to sell a corporate bond issued by a specified entity at its par value, usually including accrued interest. The option to sell the bond only comes into effect if a specified default event occurs — and, in that case, the option is almost always in the money. The option premium is expressed as an annual rate calculated on the bond par value. The annual premiums continue until the entity defaults or the swap matures, whichever is the earlier.

4.9.3 If the corporate bond becomes worthless on default, then the credit default swap is worth the full par value. More usually, a bond falls in value, but does not become totally worthless. The bond market value, expressed as a proportion of par, is known as the *recovery rate*. The value of a CDS on default is then the par value multiplied by (one minus the recovery rate). There is an arbitrage relationship which keeps CDS premiums close to the yield premium on a corporate bond, expressed relative to risk free rates.

4.9.4 To construct our hedging argument, we need the price of a more exotic derivative — the zero recovery CDS. A zero recovery CDS pays a fixed amount if the default event occurs, and zero otherwise. Although we cannot observe prices of zero recovery CDSs exactly, we can estimate theoretical prices from the premiums on conventional CDSs, grossing up for an assumed recovery rate. We denote the zero recovery swap rate by s.

4.9.5 The CDS premium is not a good forecast of actual default frequencies. Default expectations are difficult to estimate, chiefly because of uncertainty regarding the frequency of the rare liquidity crises causing an avalanche of defaults. Current default spreads are higher than recent default frequencies in many economies. There is no contradiction here; lenders can demand a risk premium just as equity investors can; but we cannot deduce the risk premium (nor unbiased default probabilities) from CDS spreads. Given the difficulties of estimation, we consider the lack of risk premium inputs as a strength, rather than as a weakness, of our proposed hedge construction.

4.9.6 The arbitrage construction is somewhat subtle. More details are provided below. The result is the following inductive formula for the initial franchise value F_0:

$$(1+R_f)F_0 = A_0(1-k_T)(m_A-k_A) + L_0(1-k_T)(m_L-k_L) + (1-s)F_1 - (R_f+s)k_T(A_0-L_0).$$

4.9.7 We note that this reduces to the default free value if the default spread $s = 0$. Otherwise, the default spread appears in two places. Firstly, it appears in a survival factor $1-s$ for discounting future franchise value. Secondly, it appears in the final double taxation term, this time reflecting the irrecoverable tax credit against the losses which precede failure.

4.10 *More Details of the Hedge with Default Risk*

4.10.1 The share payoff is calculated as previously in the case of no default, but we now recognise a total loss of equity on default, as in Table 11.

4.10.2 Possible hedging instruments include cash, the asset swap, the liability swap and the credit default swap. All of these are one-year instruments. We note that, in the event of default, the liability swap is

Table 11. Full equity pay out distribution including default event

Structure	Value at time zero	Value at time one if no default	Value at time one on default
Share + dividend	$A_0 - L_0 + F_0$	$A_0[1+(1-k_T)(R_A-k_A)]$ $-L_0[1+(1-k_T)(R_L+k_L)]$ $+F_1$	0

Table 12. Full pay out distribution for available hedging instruments

Structure	Value at time zero	Value at time one if no default	Value at time one if default
Cash	$(1+R_f)^{-1}$	1	1
Asset swap	0	$R_A - R_f - m_A$	$R_A - R_f - m_A$
Liability swap	0	$R_L - R_f + m_L$	$A_0 L_0^{-1}(1+R_A-k_A)$ $-(1+R_f+k_L-m_L)$
Credit default swap	0	$-s$	$1-s$

Table 13. Build up of hedge for full equity pay out distribution

Structure	Value at time one if no default	Value at time one if default
$A_0(1-k_T)$ $\times$ Asset swap	$A_0(1-k_T)\times(R_A-R_f-m_A)$	$A_0(1-k_T)\times(R_A-R_f-m_A)$
$-L_0(1-k_T)$ $\times$ Liability swap	$-L_0(1-k_T)\times(R_L-R_f+m_L)$	$L_0(1-k_T)(1+R_f+k_L-m_L)$ $-A_0(1-k_T)(1+R_Ak_A)$
$-[k_T(A_0-L_0)+F_1]$ $\times$ Credit def swap	$[k_T(A_0-L_0)+F_1]s$	$-[k_T(A_0-L_0)+F_1](1-s)$
$A_0\{1+(1-k_T)$ $\times(R_f+m_A-k_A)\}$ $-L_0\{1+(1-k_T)$ $\times(R_f-m_L+k_L)\}$ $+F_1(1-s)$ $-[k_T(A_0-L_0)]s$ in cash	$A_0\{1+(1-k_T)\times(R_f+m_A-k_A)\}$ $-L_0\{1+(1-k_T)(R_f-m_L+k_L)\}$ $+F_1(1-s)$ $-[k_T(A_0-L_0)]s$	$A_0\{1+(1-k_T)\times(R_f+m_A-k_A)\}$ $-L_0\{1+(1-k_T)(R_f-m_L+k_L)\}$ $+F_1(1-s)$ $-[k_T(A_0-L_0)]s$
Total	$A_0\{1+(1-k_T)\times(R_A-k_A)\}$ $-L_0\{1+(1-k_T)(R_L+k_L)\}$ $+F_1$	0

computed using the return on the assets, as creditors in this situation gain control of the firm's assets.

4.10.3 We now develop a linear combination of hedging instruments which has the same value as the share in every outcome. We start by estimating the quantity of asset and liability swaps required, by inspection of the non-default time one value.

4.10.4 By inspection of the difference between default and non-default cases, comparing the share to the hedge, we can determine the quantity of credit default swap.

4.10.5 Finally, we determine a quantity of cash which, when added to the three swaps, replicates the share price plus dividends in both default and no-default scenarios, as shown in Table 13.

4.10.6 The only component of the share price with non-zero initial capital is the cash. Discounting at the cash return to time zero, we obtain the initial share price:

$$\begin{aligned}(1+R_f)(A_0-L_0+F_0) = {} & A_0\big\{1+(1-k_T)(R_f+m_A-k_A)\big\}\\ & -L_0\big\{1+(1-k_T)(R_f-m_L+k_L)\big\}\\ & +(1-s)F_1-sk_T(A_0-L_0).\end{aligned}$$

4.10.7 Subtracting the initial net assets, we find the initial franchise value, which was the result we set out to prove:

$$\begin{aligned}(1+R_f)F_0 = {} & A_0(1-k_T)(m_A-k_A)+L_0(1-k_T)(m_L-k_L)\\ & +(1-s)F_1-(R_f+s)k_T(A_0-L_0).\end{aligned}$$

4.10.8 A special case of interest is the geometric growth scenario. In this case we do not know F_1, except that $F_1 = (1+g)F_0$. On substitution, we can then deduce the value of F_0:

$$(R_f + s - g + sg)F_0 = A_0(1-k_T)(m_A - k_A) + L_0(1-k_T)(m_L - k_L) - (R_f + s)k_T(A_0 - L_0).$$

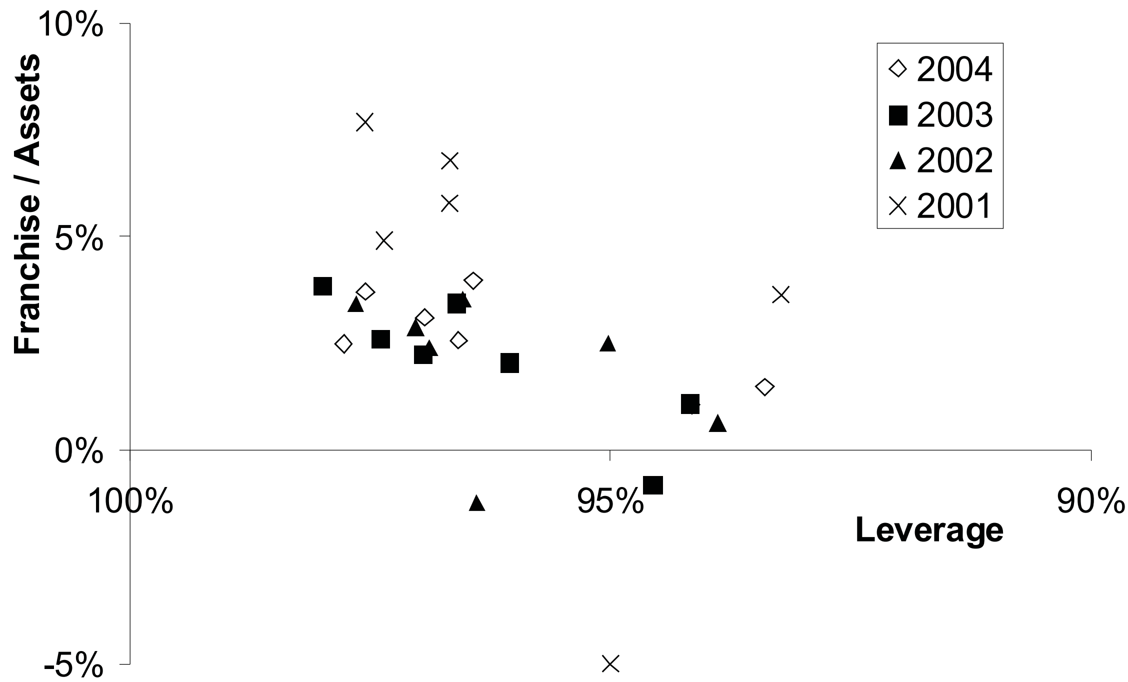
5. Capital Structure and Leverage

5.1 *Leverage*

5.1.1 The leverage of a financial firm, or, indeed, any firm, is defined as the ratio of non-equity liabilities to assets. In the case of financial firms, we include in the liabilities all liabilities to customers, in addition to any bondholders.

5.1.2 Figure 11 shows the leverage and franchise value for a range of U.K. active financial firms at the year ends 2001 to 2004. We show leverage decreasing from left to right, so that equity as a proportion of assets increases from left to right. Equivalently, net assets increase from left to right.

5.1.3 In this section we consider factors determining an optimal leverage.



Source: Report and accounts, Datastream

Figure 11. Franchise value and liabilities as a proportion of total assets for selected U.K. financial firms

Changes in leverage occur by transactions such as dividend payments, share buybacks or new issues.

5.1.4 Let us take divided payments as an example. This also applies to new equity raised, if we treat new equity raised as a negative dividend. The change in shareholder value on dividend payments is the change in market capitalisation plus the dividend itself, that is:

$$\text{Value gain} = \text{Increase in net assets} + \text{Increase in franchise value} \\ + \text{Dividend paid.}$$

5.1.5 As the dividend is paid out of net assets, the increase in net assets is equal to minus the dividend. We then have the identity:

$$\text{Value gain} = \text{Increase in franchise value.}$$

5.1.6 It is now clear that shareholder wealth is maximised by maximising the franchise value. It does not make sense to maximise the market capitalisation, because each extra pound of equity raised will increase the market capitalisation. The trick is to stop raising capital when the marginal pound increases the market capitalisation by less than a pound. This point is a maximum of the franchise value.

5.2 *RORAC and Embedded Value Approaches to Optimal Leverage*

5.2.1 The RORAC methodology supposes that each business unit has a return on capital associated with that unit. Value is created by allocating capital to the areas where RORAC is highest. Provided that some business unit creates value in RORAC terms, this maximum is above the cost of capital. The value is proportional to the amount of capital allocated. Value is linear and increasing in capital. The optimal capital is as much as you can raise; the optimum leverage is zero.

5.2.2 The embedded value approach works in the reverse direction. In an insurance context, the discount rate usually exceeds the assumed return on assets, so that any assets held in excess of liabilities are implicitly valued at a discount to market value, the discount being called a *locking in adjustment*. Applied to value in force, the locking in adjustment often only reflects statutory capital requirements, but, in a whole firm valuation context, the logical approach is to count all the capital which the firm expects to hold. Similarly, in a banking context, the cost of equity is higher than the cost of debt, so an embedded value approach recognises an increase in value whenever expensive equity is substituted for cheap debt. As with the RORAC approach, franchise value is linear in leverage, but this time the slope is positive, and so the optimal leverage is 100%, or as high as can be achieved, subject to regulatory capital constraints.

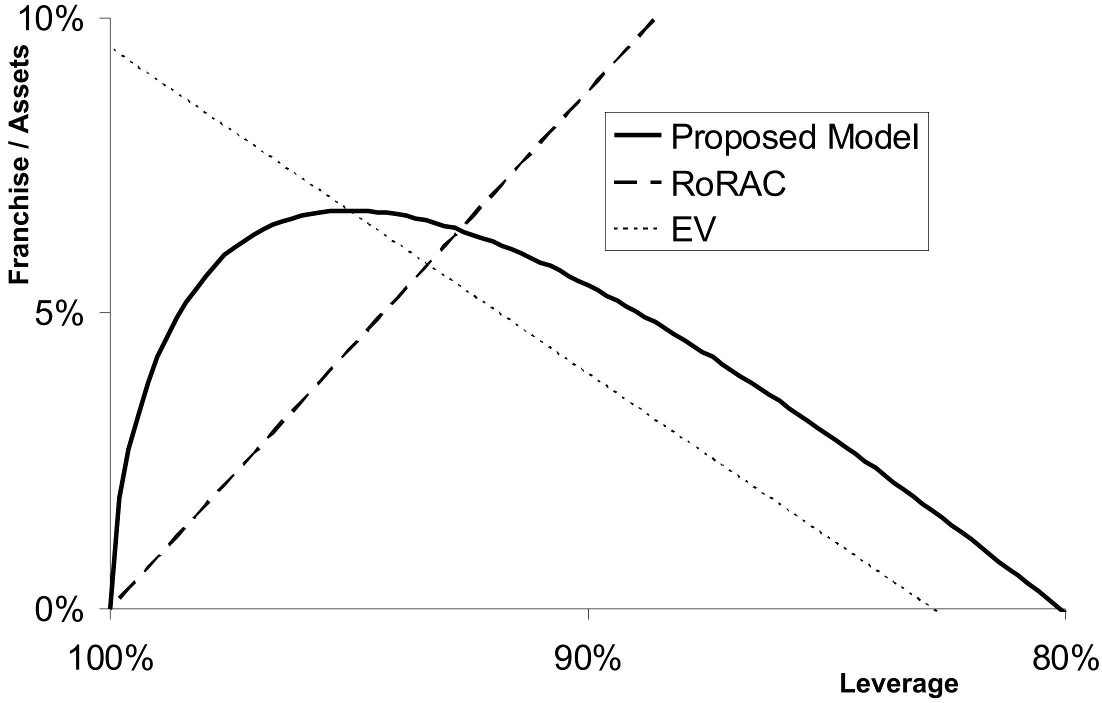


Figure 12. Oil and water — comparing the RORAC and EV framework, plus proposed solution

5.2.3 RORAC and market consistent embedded value have been two big ideas on insurance management from the consulting industry in the last ten years. It can be seen in Figure 12 how they blend together, like oil and water, to make a framework for running a business.

5.3 *Hedge-Based Valuation and Optimal Capital Structure*

5.3.1 We now consider the hedge-based valuation, using the valuation formula from Section 4.9. We will see that, at the high equity (low leverage) end, the tax and agency costs reduce the merited franchise value. On the other hand, when equity is low, then the risk of franchise loss reduces the franchise value. There is some optimal leverage where the franchise value is maximised. This curve is similar to the figure described in Hancock *et al.* (2001).

5.3.2 In contrast to the RORAC or EV approaches, we have established an optimal amount of capital within a single modelling framework. We avoid the need, seen in economic capital exercises, to pick an arbitrary percentile or time horizon in order to calculate capital needs. Thus, we have recovered a financial approach, which simultaneously delivers a pricing basis and an optimal capital structure.

5.3.3 A chief challenge with this approach is to establish which parameters might change with leverage and which are more likely to be constant. We now consider these points.

5.4 *Margin Sensitivity to Leverage*

5.4.1 In a bank setting, it is natural to assume that the assets are fixed, and net assets vary by considering alternative liabilities. Insurers are more likely to treat their liabilities as fixed, and vary the quantity of assets.

5.4.2 It is plausible to believe that asset margins are unaffected by the firm's capital position. That is because the seller of an asset does not care about the credit quality of the buyer, once the sale has taken place. For similar reasons, retail loan customers care little for the credit quality of their lender.

5.4.3 Liabilities are a different matter. If a firm owes me money, I am certainly concerned about whether I get paid back. This credit sensitivity is fundamental to both banking and insurance operations. We would surely expect weaker financial firms to offer better rates on deposits or lower insurance premiums.

5.4.4 However, our definition of the liability swap already includes the effect of default risk. If we assume that the liability swap rate is constant as net assets vary, this still implies that terms become more generous for policyholders, in such a way that the market value of the contractual revision offsets the effect of the credit risk.

5.4.5 There are, of course, reasons why customers might price risk differently from the market. Policyholder and consumer protection schemes insulate creditors from some of the losses, but customers are still better off getting paid in the first place than having to face the delay and uncertainty of a compensation scheme. Risk-based scheme levies increase the extent to which shareholders bear the cost of the limited liability put. On the other hand, there is an argument that customers price credit risk more severely than capital markets, on the basis that it is costly for customers to diversify their credit exposure to financial service providers.

5.4.6 For the remainder of this section we take the middle road, and suppose that the effect of diversification costs and compensation funds broadly cancel out. In that case, we can hold the liability swap margin m_L fixed, as we change capital structure.

5.4.7 This approach is different to that described by Hancock *et al.* (2001), who essentially assumed that contractual liability terms are unchanged as leverage increases. Their approach, therefore, produces a limited liability put option effect, which becomes more valuable as leverage increases. Our approach does not produce these shareholder gains, as, instead, we assume that creditors adjust the terms of trade to pass any default cost back to shareholders.

5.5 *Effect of Expenses*

5.5.1 We investigate a series of sensitivities and their effect on franchise value. Figure 13 shows the effect of a reduction in asset expenses.

5.5.2 There is an increase in franchise value at all levels of leverage.

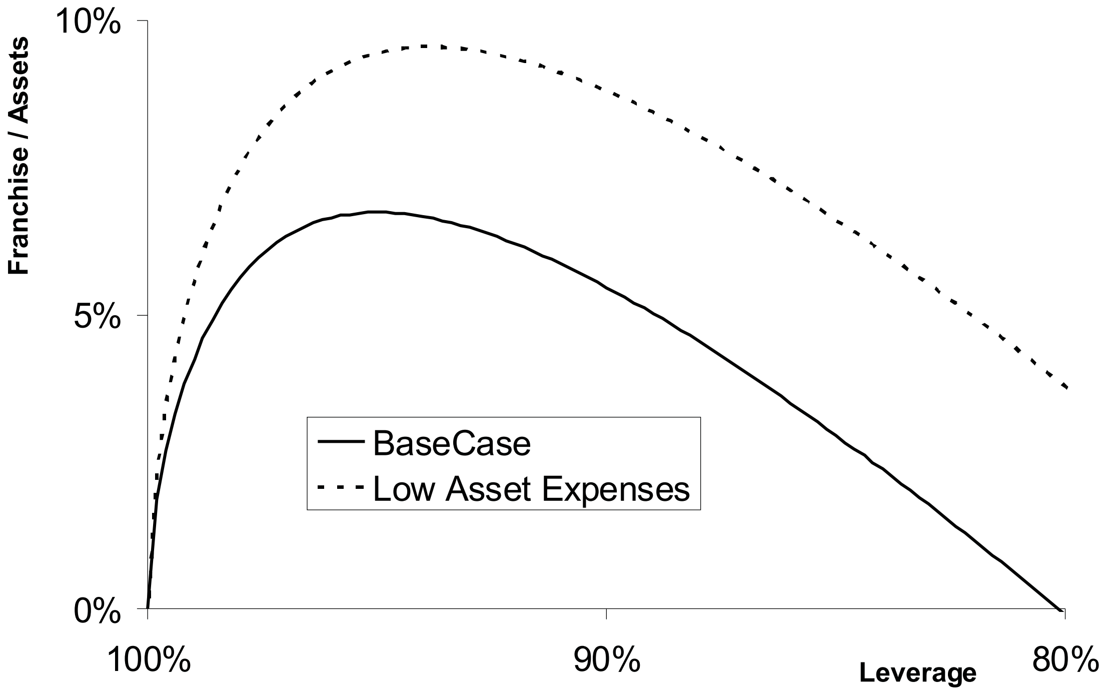


Figure 13. Franchise value effect of a reduction in asset expenses

This is to be expected — a reduction in expenses can only leave shareholders better off. However, the amount of the improvement is greater at lower levels of leverage, that is to the right of the figure. This happens because, at lower leverage levels, the firm is likely to survive longer, and therefore shareholders benefit from the expense reduction over a longer period.

5.5.3 The optimal level of leverage is that which maximises the franchise value. We see that the optimal leverage decreases, or optimal net assets increase, when expenses reduce. This happens because the shareholder rationale for holding capital is to preserve franchise value. If franchise value increases, for example following an expense reduction, then shareholders will be more anxious to preserve their higher franchise value than they were before. As a result, shareholders would, rationally, require the firm to raise more capital, or retain dividends, following a reduction in expenses.

5.5.4 Observed management behaviour could be quite different to this. Management may well argue that an expense reduction has improved cash flow, some of which they can prudently distribute to shareholders without increasing the firm's risk. In this situation, a franchise value model can generate interesting debate among management, as it tries to articulate the rationale for an increased dividend.

5.6 *Effect of Growth Rate*

5.6.1 Figure 14 shows the effect of a higher growth rate assumption. Franchise value increases for all degrees of leverage. The effect is less when

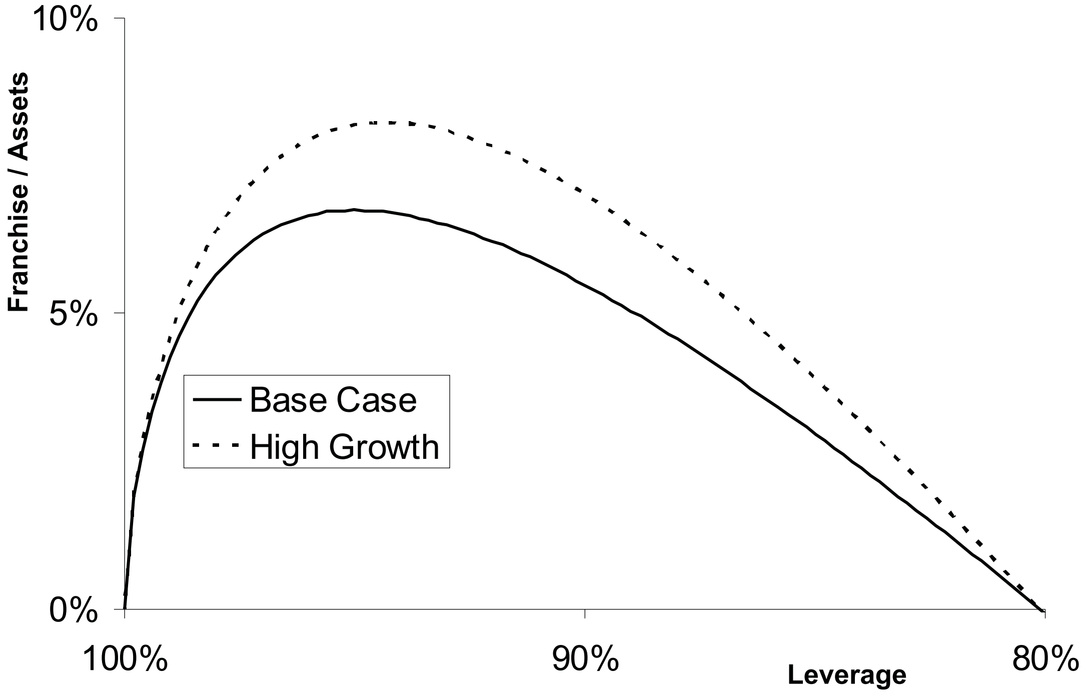


Figure 14. Effect of higher growth rate

leverage is very high. This is because of a high ruin probability, so that it is unlikely that the firm survives long enough to benefit from the improved growth.

5.6.2 At low levels of leverage, there is also little impact. This is because, at those levels, the capital costs offset profits from the core business. High growth is of little value if the growing business generates no profit.

5.6.3 We see a small increase in the optimal level of net assets, or, equivalently, a small fall in optimal leverage. This is driven by the increased franchise value, which is, therefore, more worth preserving from a shareholder perspective.

5.7 *Effect of Tax*

5.7.1 Figure 15 shows the effect of tax on franchise value. We see that a lower tax regime creates value at all levels of net assets. However, the effect is most significant for large values of net assets. The reason is that the double taxation of the return on net assets is one of the costs which the valuation considers. A reduction in the rate of tax on that investment therefore creates a greater benefit when there are more free assets whose return can be taxed.

5.7.2 In this case, also, the optimum level of free assets increases significantly. This reflects the fact that the double taxation of net asset investment return is one of the more important deterrents to the holding of excess capital. With that deterrent removed, firms optimally increase their level of free assets.

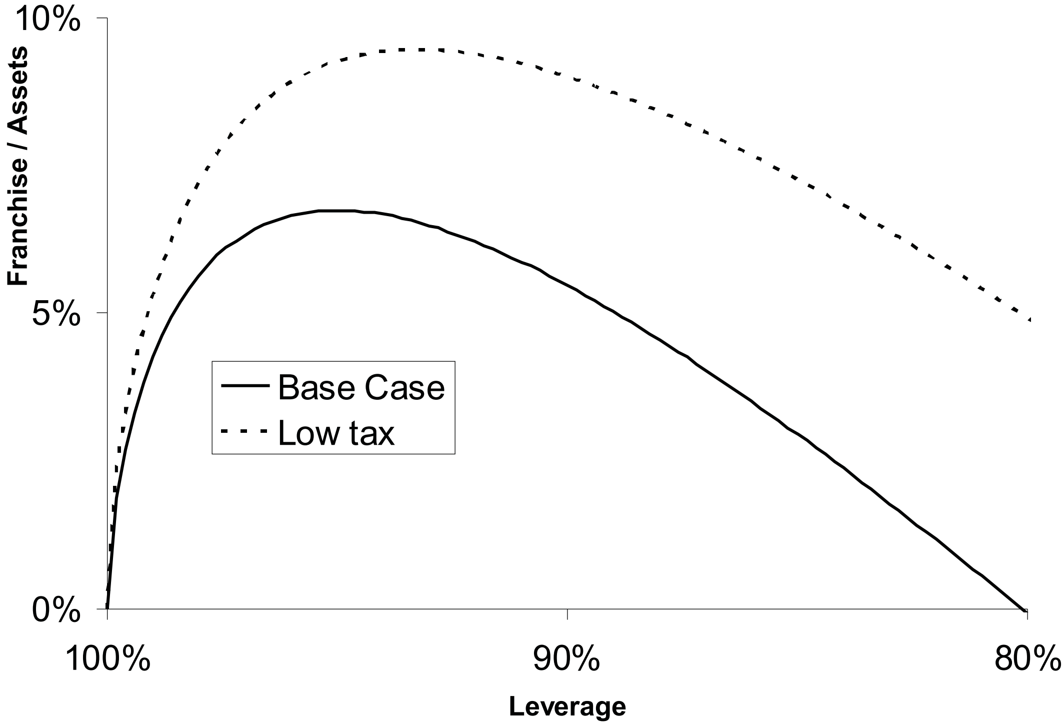


Figure 15. Effect of tax on franchise value

5.8 *Agency Costs*

5.8.1 Agency costs arise from the risk of poor stewardship of shareholder assets in the hands of corporate managers. For example, managers may be tempted to squander shareholder resources on ambitious acquisitions, which destroy value, but enhance the status of the management team concerned. Jensen & Meckling (1976) show how these costs can arise from conflicts of interest between shareholders and managers. From a shareholder perspective, agency costs act as a form of tax on assets entrusted to third party managers.

5.8.2 There are good reasons for model agency costs to be proportional to net assets. While times of financial hardship are bad for credit risk, they can increase pressure on management to reduce expenses, even when this involves tough decisions. Conversely, a firm with more than enough free assets may believe that the change of needing further capital is negligible. That would remove an important incentive to control costs.

5.8.3 It may appear circular to expect managers to quantify their own conspiracy against shareholders when setting premiums. We argue that markets allow for such costs when pricing insurance shares; therefore, it is natural to use this information also in pricing decisions.

5.8.4 Agency costs are, at least in principle, measurable empirically. The difficulty is stripping them out from other costs. One approach is to assume that, at current levels of leverage, costs are consistent with observed history,

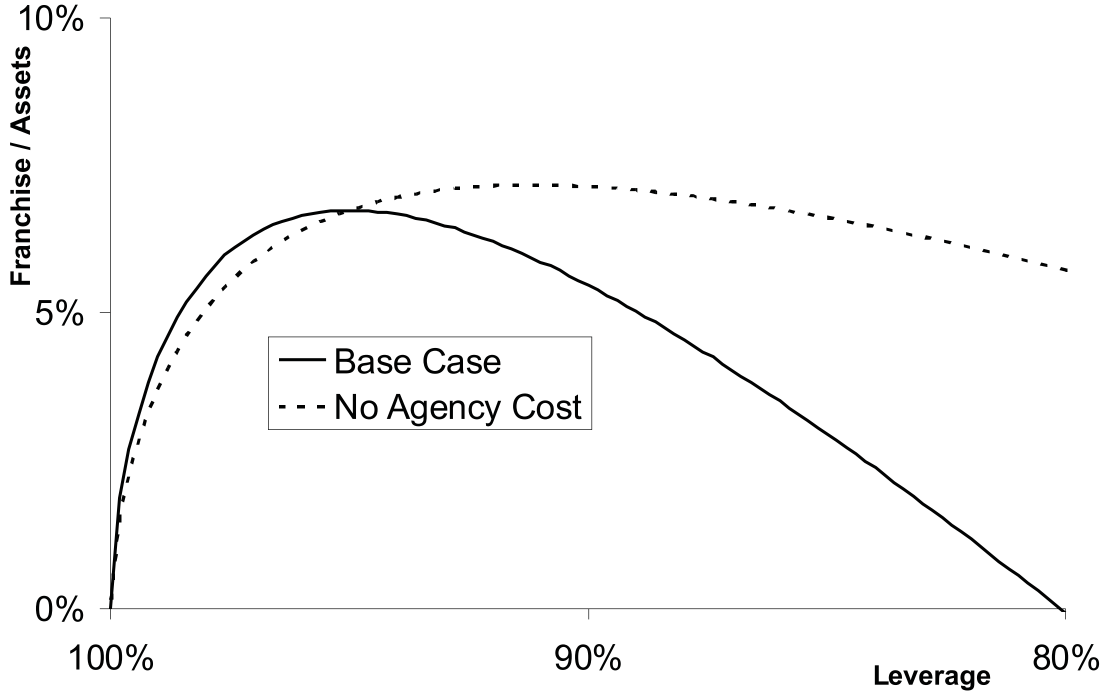


Figure 16. Effect of agency costs on franchise value

and require no agency cost addition. On the other hand, if extra equity capital is raised, then there is a proportional increase in agency costs, and a proportional fall if leverage increases.

5.8.5 We model agency costs for alternative structures by assuming that an additional cost k_E is incurred for each unit of net assets in excess of the base case $A_0 - L_0$. At the base case, expenses are unaffected. At low levels of net assets relative to the base case, our formula reflects an improved operating performance arising from expense reduction.

5.8.6 Figure 16 shows the effect of an agency cost reduction. Agency costs have a dramatic effect in reducing the level of free assets which a firm optimally holds. This is because agency costs operate as a form of tax on free assets.

5.9 *Economic Capital and the Credit Default Spread*

5.9.1 We now consider how the credit default spread varies as a function of leverage. In the embedded value framework, this is taken as constant, not depending on leverage.

5.9.2 It is more natural to assume that the credit default spread increases as leverage increases, and this has been reflected in the examples which we have considered so far. One possible approach to calibration is to observe the leverage of a cross-section of firms, to investigate how the leverage affects the price of their issued debt.

5.9.3 Such analysis gives a broad indication of the relationship between leverage and credit spreads. However, it suffers from several shortcomings. Firstly, the issued bonds will generally have lower priority than customer obligations. Secondly, the price of bonds allows for a partial recovery, while our particular version of credit default swap is based on zero recovery. Finally, there is, at best, a broad relationship between net asset ratios and default risk; our empirical analysis would fail to take account of the different levels of variability in net assets between firms.

5.9.4 An alternative approach would be to build a stochastic model of business risks faced, and to use this to determine the likely credit default spread with different levels of starting assets.

5.9.5 Many firms have already built such models as part of their economic capital investigations. As a result, firms already have estimates of the volatility of net assets, or indeed of economic capital, which, in the case of a normal distribution, is proportional to the volatility. A pragmatic approach, then, is to assume that the default spread is a function of the distance to default, that is the net assets relative to economic capital. This approach is already familiar for the modelling of credit risk, and is embedded into KMV's widely used model. Bohn & Crosbie (2003) give more details of this approach.

5.9.6 Figure 17 shows the effect of economic capital on franchise value. We can see that a reduction in economic capital increases franchise value. Conversely, an increase in economic capital represents a cost to shareholders, which we are now able to quantify.

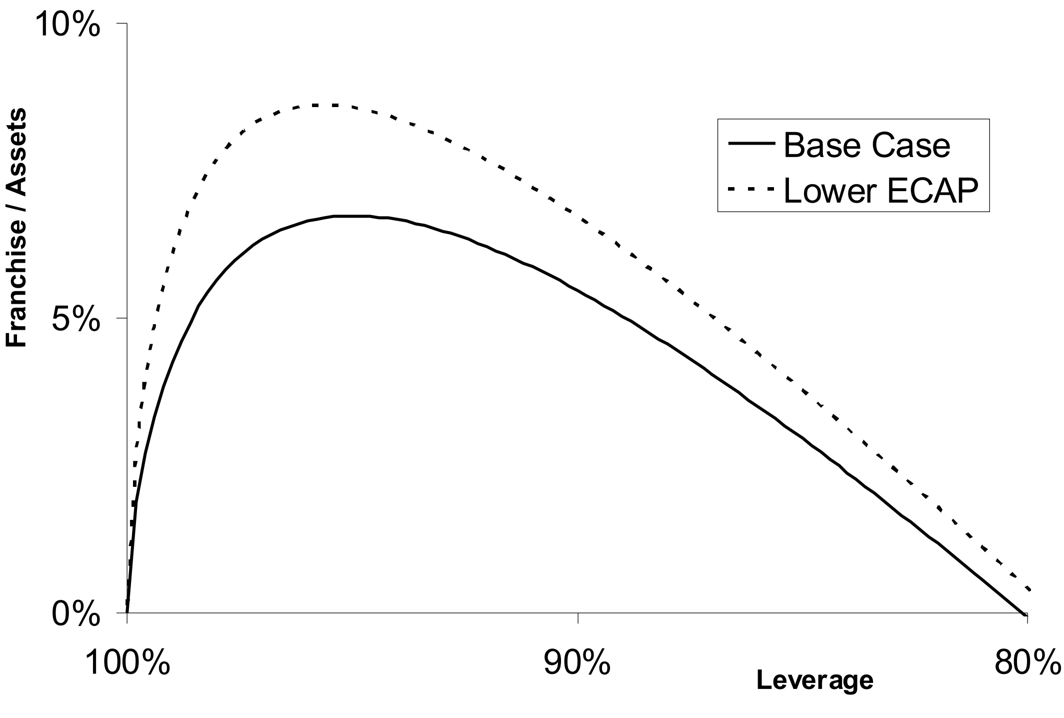


Figure 17. Effect of economic capital on franchise value

5.9.7 As with all our previous examples, the effect on franchise value at high leverage is negligible. The effect is also modest at low levels of leverage. Here, the reason is that actual capital is so large relative to economic capital that the calculated level of economic capital becomes irrelevant.

5.9.8 However, unlike the previous examples, we now see that the optimal level of capital has fallen as economic capital fell. This is because, with a lower level of risk, we can afford to hold less actual capital.

5.10 *The Cost of Economic Capital*

5.10.1 We have now come full circle. We started with a survey of all the business decisions where cost of capital has a role to play. We now have a tool to quantify the cost of risk capital in financial decision making.

5.10.2 There are many ways in which financial firms can reduce exposures to risk. For example, firms can buy reinsurance or derivative hedges from external providers. Asset securitisation is another popular tool for managing default risk. In addition to external transactions, firms can adjust internal processes to reduce risk, such as restrictions on lending criteria or insurance underwriting.

5.10.3 In all these cases, the cost is a reduction in expected profit, because of margins paid to third parties. The benefit is a reduction in risk, usually measured as a reduction in economic capital. Therefore, decisions rest critically on the assumed cost saving when economic capital is reduced.

5.10.4 RORAC theory prices the capital reduction using an expected shareholder return. The most important inputs are the firm's beta and the assumed equity risk premium. If a firm believes in a large equity risk premium, it will hedge more market risk, securitise more credit risk, buy more reinsurance and operate tight underwriting criteria.

5.10.5 Our hedge-based valuation gives an alternative view. We quantify the cost of economic capital by its effect on the franchise value. That effect is determined by the frictional costs of capital, by CDS spreads and by the franchise value at risk. Risk premiums are irrelevant, because risk premiums always cancel out in hedge constructions.

5.10.6 Economic capital is most important for firms where capital is an active constraint. In these cases, our analysis shows that franchise value is sensitive to economic capital. Economic capital is less costly for highly leveraged firms, especially those with low franchise value. In that case, our analysis suggests that the optimal shareholder strategy is the pursuit of profit with comparatively less regard for risk.

5.10.7 Economic capital is also less costly for firms with very large capital resources, relative to risks run. That is because capital is unlikely to be a constraint on business decisions, nor to affect shareholder cash flows.

6. Conclusions

6.1 *No Big Secret, but Many Definitions*

6.1.1 After much deliberation, we have come to this conclusion: the only secret about the cost of capital is that there is no secret. Different people mean different things when they refer to the cost of capital. Clear definitions are a prerequisite for a productive discussion.

6.1.2 Firms are right to be concerned when they fail to find a universal capital definition which makes sense for all their businesses. The capital needed to secure a given credit rating is not the same as the capital on which shareholders require a return, and is different again from the optimal capital to hold on the balance sheet. The cost of capital which analysts use to value a firm is not the same as the cost of capital for use in performance measurement or product pricing. Neither of these is the same as the CAPM implied cost of capital, which you can find at the push of a button from your market data supplier.

6.1.3 When analysts talk about the cost of capital, they usually mean the return which shareholders require on an investment, given a level of risk. This percentage return applies to the whole market capitalisation, rather than merely to the balance sheet net assets. The expected return is difficult to observe, and most firms rely on theoretical models, such as the CAPM, to construct expected return estimates.

6.2 *Financial Firms*

6.2.1 Although all firms have a cost of capital, application of the concept to financial firms is challenging, because:

- liabilities to customers represent a form of debt capital (and these customers may be particularly credit sensitive);
- rather than balance sheet assets being invested in plant and machinery, they may be invested in financial instruments, whose values may be marked to market, and their growth unsuited to comparison against CAPM expected returns; and
- the difference between balance sheet net assets and total equity market capitalisation may be material, reflecting significant franchise value, which can be identified in terms of customer relationships.

6.2.2 The good news is that known biases in the cost of capital methods for other industries are most easily remedied for financial firms. This is because, provided that each of the above issues is addressed carefully, the assets and liabilities of financial firms can more easily be represented in terms of replicating portfolios, including information from securitisation and other financial transactions. Thanks to the replication approach, CAPM implied capital costs, based on expected returns, are irrelevant in this calculation.

6.3 *The Importance of Frictional Costs*

6.3.1 Apart from required shareholder returns, there are cash flow costs associated with the actual level of capital on the balance sheet. These are collectively known as frictional costs, including double taxation and agency costs. Measurement of frictional costs is an empirical question. As yet, theoretical models of frictional costs have been slow to develop, but it is plausible that frictional costs are related to the accounting balance sheet net assets.

6.3.2 Frictional costs may be offset by the benefits of holding capital, in particular the reduction in the risk of failure. This normally benefits both customers, who experience lower default risk, and shareholders, who increase the probability of retaining franchise value.

6.4 *Economic Capital*

6.4.1 Financial firms often identify their 'economic capital', usually defined as the difference between current net assets and some low percentile of an estimated future net asset probability distribution. Economic capital also generates costs: firstly, the cost of computing it; plus any increases in the compensation which managers negotiate with reference to economic capital arguments; and, finally, the effect of economic capital on the probability of ruin. The economic capital calculation is one important factor in determining how much actual capital a firm chooses to hold. However, the other factors discussed above, including the frictional costs of holding capital and the impact of franchise value, are, potentially, just as important.

6.4.2 Capital held affects performance measurement, because of its effect on the possibility of failure. A firm, whose actual capital is a small multiple of economic capital, must set demanding performance targets in the events where it survives, in order to offset the total shareholder loss which accompanies failure. The effect is most marked for firms with a large franchise value, as shareholders then have most to lose.

6.5 *Using the Cost of Capital in Practical Applications*

6.5.1 Hurdle rates used in transactions, product pricing or embedded value reporting may, at first sight, be hard to explain. We can rationalise observed rates as a combination of CAPM implied returns, plus a loading to reflect the possibility that a firm fails before the cash flows are realised, plus a further loading to offset optimism in cash flows. The optimism may be partly due to a failure to model frictional costs.

6.5.2 Market consistent reporting provides an improved substitute for the CAPM component, but the need remains to allow for possible firm failure, frictional costs and other cash flow optimism.

6.5.3 This paper proposes empirical investigations, which may be used to quantify these effects, and suggests that financial firms are an ideal setting for cost of capital methodologies in developing pricing tools and target setting.

Acknowledgements

This paper is based on another paper of the same title, presented to the Institute and Faculty of Actuaries Finance and Investment Conference in 2003.

We are grateful to many friends, colleagues and clients who have encouraged us to develop our models in this area. Particular thanks are due to Richard Baddon, Paul Coulthard, Rolf van den Heever, Martin Lees, Tim Medcalf, Shyam Mehta, Ian Moran, Norbert Schnadt, Tim Sheldon, Guillaume Valois, Elliot Varnell, David Walczak, Andrew Wallace-Barnett, Martin White and two anonymous scrutineers. We would also like to thank participants at a pricing for risk workshop in May 2003, discussants at the 2003 Finance and Investment Conference, participants in the 2004 Bowles symposium in Atlanta, Georgia, and contributors to a discussion at the Polish Society of Actuaries, Warsaw, in 2005. Any remaining errors are our own.

References

The following list includes, not only the works referred to in the paper, but other publications which would be of use to readers.

Ahmed, M.A., Bulmer, J.R., Clark, P.K., Collins, J.A., Fulcher, G., Lualdi, M., Robinson, E.J.A., Sayers, J.E., Spence, A.C. & Walker, S.R. (1997). The securitisation of insurance risk. General Insurance Convention.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203-228.

Blake, D., Cairns, A.J.G. & Dowd, K. (2006). Living with mortality: longevity bonds and other mortality-linked securities. *British Actuarial Journal* (to appear).

Bohn, J. & Crosbie, P. (2003). Modelling default risk. Moody's KMV. Downloadable at http://www.moodyskmv.com/research/files/wp/ModelingDefaultRisk.pdf

Borch, K. (1969). The optimal reinsurance treaty. *The ASTIN Bulletin*, **V**(2), 16-20.

Borch, K. (1982). Additive insurance premiums: a note. *Journal of Finance*, **37**, 1295-1298.

Boulton, R., Roberts, D., Smith, A.D. & White, M. (2002). Investment strategies for general insurers. General Insurance Convention, Institute of Actuaries.

Brealey, R.A. & Myers, S. (1983). *Principles of corporate finance*, 8th edition (2005). McGraw Hill.

Burrows, R.P. & Lang, J. (1997). Risk discount rates for actuarial appraisal values of life insurance companies. *7th AFIR International Colloquium*, **1**.

Burrows, R.P. & Whitehead, G.H. (1987). The determination of life office appraisal values. *Journal of the Institute of Actuaries*, **114**, 411-465.

Butsic, R., Cummins, D., Derrig, R. & Phillips, R. (2000). The risk premium project. Phase I and II Report. CAS Website, http://casact.org/cotor/rppreport.pdf

Christofides, S. & Smith, A.D. (2001). DFA The value of risk. CAS Spring Forum, Casualty Actuarial Society. http://casact.org/pubs/forum/01spforum/01spf153.pdf

CUMBERWORTH, M., HITCHCOX, A., MCCONNELL, W. & SMITH, A. (2000). Corporate decisions in general insurance: beyond the frontier. *British Actuarial Journal*, **6**, 259-296.

CUMMINS, J.D. (1990). Multi-period discounted cash flow models in property-liability insurance. *Journal of Risk and Insurance*.

CUMMINS, J.D., PHILLIPS, R.D., BUTSIC, R.P. & DERRIG, R.A. (2000). The risk premium project. Phase I and II Report. Casualty Actuarial Society. http://www.casact.org/pubs/forum/00fforum/00ff165.pdf

DIMSON, E., MARSH, P. & STANTON, M. (2002). *Triumph of the optimists. 101 years of global investment returns.* Princeton.

DOHERTY, N. (1997). Financial innovation in the management of catastrophe risk. *Proceedings of the 28th Astin Colloquium/7th AFIR Colloquium*. Institute of Actuaries of Australia.

DUFFIE, D. (1996). *Dynamic asset pricing theory.* Princeton University Press.

EXLEY, C.J. & SMITH, A.D. (2003). The cost of capital for financial firms. Finance and Investment Conference.

FAMA, E.F. & FRENCH, K.R. (1988). The corporate cost of capital and the return on corporate investment. CRSP working paper. http://ssrn.com/abstract=75999

FROOT, K. & STEIN, J. (1998). Risk management, capital budgeting and capital structure policy for financial institutions, an integrated approach. *Journal of Financial Economics*, **47**, 55-82.

HANCOCK, J., HUBER, P. & KOCH, P. (2001). *The economics of insurance — how insurers create value for shareholders.* Swiss Re Technical Publishing.

HARE, D. & KING, D. (2005). Capital allocation and performance measurement: a case study. Life Convention.

JENSEN, M. & MECKLING, W. (1976). Theory of the firm: managerial behaviour, agency costs and ownership structure. *Journal of Financial Economics*, **3**, 305-340.

MORAN, I.R., SMITH, A.D. & WALCZAK, D. (2003). Why financial firms can charge for diversifiable risk. Bowles Symposium, Society of Actuaries.

MERTON, R. & PEROLD, A. (1999). Theory of risk capital in financial firms. In Chew, D. (ed.), *The new corporate finance: where theory meets practice*. Boston: Irwin — McGraw Hill.

MILLER, M. (1999). Does M&M apply to banks? In *Merton Miller on Derivatives*, Princeton.

MODIGLIANI, F. & MILLER, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, **48**, 261-297.

MODIGLIANI, F. & MILLER, M. (1963). Corporate income taxes and the cost of capital. *American Economic Review*, 433-443.

MYERS, S.C. & COHN, R.A. (1987). *A discounted cash flow approach to property-liability insurance rate regulation, fair rate of return in property-liability insurance.* Kluwer-Nijhoff.

NG, H.M. & VARNELL, E.M. (2003). Frictional costs. Paper presented to the Staple Inn Actuarial Society.

O'KEEFFE, P.J.L., DESAI, A.J., FOROUGHI, K., HIBBETT, G.J., MAXWELL, A.F., SHARP, A.C., TAVERNER, N.H., WARD, M.B. & WILLIS, F.J.P. (2005). Current developments in embedded value reporting. *British Actuarial Journal* (to appear).

OGIER, T., RUGMAN, J. & SPICER, L. (2004). *The real cost of capital — a business field guide to better financial decisions.* Prentice Hall.

PEROLD, A.F. (2001). Capital allocation in financial firms (2001). Harvard Business School No. 98-072. http://ssrn.com/abstract=267282

SCOTTI, V. (2005). Insurers' cost of capital and economic value creation: principles and practical implications. *Sigma 2005*, **3**. Swiss Re publications.

SHELDON, T.J. & SMITH, A.D. (2004). Realistic valuation of life insurance business. *British Actuarial Journal*, **10**, 543-626.

SHERRIS, M. (2002). Economic valuation: something old, something new. *Australian Actuarial Journal*, **9**(4).

SOLOMON, E. & LAYA, J.C. (1967). Measurement of company profitability: some systematic errors in the accounting rate of return. *Financial research and management decisions.* (A.A. Rominchek, ed.) John Wiley, New York.

TAYLOR, G. (1994). Fair premium rating methods and the relations between them. *Journal of Risk and Insurance*, **61(4)**, 592-615.

WANG, S. (2002). A universal framework for pricing financial and insurance risks. SCOR working paper.

WANG, S., YOUNG, V. & PANJER, H. (1997). Axiomatic characterisation of insurance prices. *Insurance Mathematics and Economics*, **21**, 173-182.

Printed by Bell & Bain Ltd., Glasgow