Petroski, Henry, 1985, *To Engineer is Human: The Role of Failure in Successful Design* (New York: St. Martin's Press).

Reese, Stuart and Laurent Guy, 1990, Fixed Income Portfolio Performance: Myth Versus Reality (unpublished manuscript).

Smith, Clifford W., Jr., and Charles W. Smithson, eds., 1990, *The Handbook of Financial Engineering* (New York: Harper Business).

Staking, Kim, 1989, Interest Rate Sensitivity and the Value of Surplus: Duration Mismatch in the Property-Liability Insurance Industry, Unpublished dissertation, University of Pennsylvania, Philadelphia, PA.

Stengel, Robert F., 1986, *Stochastic Optimal Control: Theory and Application* (New York: John Wiley and Sons).

Tversky, Amos, and Daniel Kahneman, 1982, Belief in the Law of Small Numbers, in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment under Uncertainty: Heuristic and Biases* (Cambridge: Cambridge University Press).

Weyrick, Robert C., 1975, *Fundamentals of Automatic Control* (New York: McGraw-Hill).

Williams, Arthur III, 1983, Performance Evaluation in Fixed Income Securities, in Frank J. Fabozzi and Irving M. Pollack, eds., *The Handbook of Fixed Income Securities* (Homewood, IL: Dow Jones-Irwin).

Williams, Arthur III and Noreen M. Conwell, 1983, Fixed Income Indices, in Frank J. Fabozzi and Irving M. Pollack, eds., *The Handbook of Fixed Income Securities* (Homewood, IL: Dow Jones-Irwin).

Ziemba, W. T., and R. G. Vickson, 1975, *Stochastic Optimization Models in Finance* (New York: Academic Press).

6 ESTIMATING DIVISIONAL COST OF CAPITAL FOR INSURANCE COMPANIES

Franklin Allen

Professor of Finance and Economics
Vice Dean and Director
Wharton Doctoral Programs
The Wharton School
University of Pennsylvania

1. Introduction

In a competitive market, what should the rate of return on insurance companies' equity be? What premiums should insurance companies charge? Traditionally, the answers to these questions have been based on actuarial and accounting concepts.[1] More recently financial models of the insurance firm have been developed. Ferrari (1968) suggested a descriptive model which allowed an algebraic expression for the rate of return on equity as a function of the premiums charged to be derived. Combining this with the capital asset pricing model (CAPM) meant that an equilibrium value for the return on equity

---

I am grateful to David Cummins for helpful discussions and comments. Financial support from the NSF is acknowledged.

[1] See, for example, Cummins and Chang (1983).

and the corresponding level of premiums could be found. This model is known as the insurance CAPM.[2] The development of other asset pricing models in finance has also led to insurance counterparts. Thus the Arbitrage Pricing Theory, and option pricing models have been used to derive the return on equity and the level of premiums.[3] Although the approaches based on asset pricing models have advantages compared to those based on traditional actuarial and accounting concepts, they are not ideal. One of the most important problems is that they are not well suited for finding premiums when an insurance company has multiple divisions. The difficulty is that it is not clear how earnings on reserves should be allocated among the various divisions of the firm.

The other financial method of pricing insurance is the Discounted Cash Flow approach. There are two versions of this. Myers and Cohn (1987) have suggested an adaptation of the adjusted present value method for calculating the price of insurance. The premium is found by discounting the expected cash flows associated with the insurance at the appropriate discount rates. The other method is the NCCI (1987) approach which uses an internal rate of return methodology. This involves finding the discount rate such that the net discounted cash flow is zero. The fair premium is the one such that this discount rate is equal to the opportunity cost of capital. Cummins (1990c) shows that these models are essentially the same if properly applied and the choice between them is a question of which is easier to use. Unfortunately, as with the approaches based on asset pricing models, neither of these methods is well suited to finding insurance premiums when a company has multiple divisions. The problem is again how to deal with reserves and allocate the earnings from these to the different divisions.

Existing financial methods of pricing insurance take the structure of the insurance firm as given. It is argued below that a more fruitful approach is to start with the question of why insurance firms have the particular structure that they do. In an ideal world there would be no need for insurance companies to be involved in financial markets. If contracts were costlessly enforceable, it would not be necessary for

[2] See Cooper (1974), Biger and Kahane (1978), Fairley (1979) and Hill (1979).

[3] For surveys of this literature see D'Arcy and Doherty (1988) and Cummins (1990a; 1990b).

premiums to be paid in advance. Instead, premiums could be paid by those that did not suffer losses at the time when payouts were necessary. In this case, the insurance company would simply be a conduit for premiums from those who did not suffer any loss to those who did.

In practice, of course, this type of insurance company could not survive because contracts are not costlessly enforceable. It would be very difficult, if not impossible, to force the people who did not suffer a loss to make payments. For this reason, insurance companies require payment for insurance at the beginning of the period rather than at the end. It is this problem which causes insurance companies to be involved in financial markets since they invest the premiums they receive until payouts are necessary.

Although the case where payments are made at the end of the period is unrealistic, it is nevertheless important as a benchmark. It means the insurance aspects of the problem can be separated from the investment aspects. Among other things, this allows the level of premiums to be determined when an insurance company has multiple divisions. It also allows a theory of the competitive rate of return on insurance companies' equity to be developed.

Section 2 starts with the benchmark case where there are perfect contracts and capital markets. Section 3 considers what happens if contracting possibilities are imperfect but capital markets are perfect. In Section 4, imperfect capital markets are introduced into the analysis. Section 5 focuses on the role of shareholders, Section 6 introduces risky insurance liabilities and Section 7 considers the effect of taxes. Finally, Section 8 contains conclusions.

2. Perfect Contracts and Capital Markets

If contracts are costlessly enforceable, insurance firms will not need to collect premiums until losses are incurred. This benchmark case is analyzed first. Frictionless capital markets and advance collection of premiums are then introduced.

Perfect Contracts

Initially, consider a simple scenario. There are two dates $t = 0, 1$. The number of consumers is very large. They all have the same initial wealth and the same opportunities. They are risk averse with identical utility functions $U(W_1)$ where $U' > 0$, $U'' < 0$ and W_1 is wealth at date 1.

At date 0, each faces a probability π ($0<\pi<1$) of a loss L at date 1 so the expected loss is EL = πL. This loss is observable to an insurance company. The risks consumers face are independent so that the standard error of their average loss is zero. Every consumer is made better off by insuring against the risk and guaranteeing a level of consumption of W_1 - EL.

The loss can be thought of as a property loss such as a house burning down. An alternative interpretation is that households are the relevant unit and one of the members of the household dies with a resultant loss in earning power. The model is thus applicable to both property and to life insurance.

As far as the insurance industry is concerned, there are no barriers to entry, the market is competitive and there are no costs in setting up and running an insurance company. This implies that profits will be zero in equilibrium. The prices referred to below are equilibrium prices.

There is *perfect* contracting in the sense that all contracts are costlessly enforceable. As a result, insurance companies can ensure that everybody receives W_1 - EL at date 1 by issuing contracts of the following form. At date 0, before consumers know whether they will suffer a loss or not, they sign a contract promising to pay EL at date 1 if they do not suffer a loss. In return, the insurance company will promise to pay L-EL to all those consumers who do suffer a loss at date 1. In this way the risk associated with the loss L can be entirely eliminated and consumers' welfare is maximized. Competition among the insurance companies ensures that payments are set at this level and their profits are zero since revenues are $(1-\pi)$EL and costs are π(L-EL) = $(1-\pi)$EL.

This equilibrium will be referred to as the *benchmark* equilibrium. Its assumptions essentially correspond to the Arrow-Debreu framework and the allocation that results is therefore Pareto efficient.

In the simple case presented, the only issue in pricing the insurance is the expected loss. There is no need for a cost of capital because all payments are made at the same time.

The insurance contracts with payments at date 1 do not, of course, correspond to actual insurance contracts where payments are required before the insurance starts. The reason that advance payments are necessary in practice, is obviously that contracts are not costlessly enforceable. It would be very difficult for insurance companies to make those consumers who do not suffer a loss pay EL at date 1. This problem is overcome by requiring all consumers to pay a premium at date 0 and

then making payouts to those consumers suffering a loss at date 1. The complication that this introduces is that the premiums can be invested in financial markets between dates 0 and 1. Section 3 examines the effect of relaxing the assumption that contracts are perfectly enforceable. However, before doing this it is helpful to consider what happens if capital markets are introduced into the analysis and premiums are paid at date 0.

Perfect Capital Markets

Suppose that in addition to perfect contracting, capital markets are *perfect*. In other words, all agents have equal access so both firms and consumers face the same interest rates and investment opportunities. There is no difference between borrowing and lending rates, all agents have the same information, the market is perfectly competitive, there are no transaction costs and so forth. In addition, markets are complete so that there are full risk sharing possibilities. There are no taxes.

In these circumstances, there are a number of different types of insurance contract which allow consumers to eliminate the risk associated with the loss L. One alternative is the benchmark case where insurance contracts are signed at date 0 but all payments are made at date 1. In addition to arranging insurance at date 0, consumers use their initial wealth to purchase an optimal portfolio Z*. The ex post return on this portfolio is denoted R* and the expected return is ER*. At date 1, consumers use the proceeds from their investment to pay out EL if they did not suffer a loss. If they did suffer a loss they receive L-EL from the insurance company. Their total wealth at date 1 is therefore R*-EL whatever happens. Consumers thus have insurance against the loss L they face and so do not have to bear any of this risk but they do bear an optimal amount of investment risk.[4]

The existence of perfect capital markets means that an alternative arrangement is for premiums to be paid at date 0 and for the insurance companies to invest them. The interest and dividends earned ensures a premium lower than EL will be charged. For example, suppose the premiums were invested in a risk free asset yielding r_F. In this case the premium at date 0, denoted p_0, could be set at

[4] See Chapter 4 of Sharpe (1970) for an analysis of optimal risk bearing by investors.

$$p_0 = \frac{EL}{1 + r_F} \tag{1}$$

and the company would still have sufficient funds to cover its liabilities. However, notice that the allocation of resources would be exactly the same as before. The only difference would be that policyholders would reduce their holdings of the risk free asset by $EL/(1+r_F)$ between dates 0 and 1 and would pay the premium a period earlier.

This is just an application of the well-known Modigliani and Miller (1958) result from corporate finance. The theorem asserts that with perfect and complete capital markets the value of a firm does not depend on its capital structure because investors' opportunity sets are not affected by its capital structure. If the firm takes on more debt, for example, shareholders will not be any better off because if they had wanted levered equity they could have borrowed on their own account.

In the insurance context, policyholders can offset any action of the insurance company in terms of the timing of the premium by adjusting their holding of the risk free asset appropriately. No matter what the spread of payments between dates 0 and 1 the policyholder's opportunity set and hence the allocation of resources will not be affected. What happens if an insurance company were to require a premium at date 0 and instead of investing it in the risk free asset invested it in a risky asset with random rate of return r and mean rate of return Er? The insurance company could charge a premium

$$p_0 = \frac{EL}{1 + Er} \tag{2}$$

and on average have enough to cover its liabilities. The "on average" is important here since some of the time the realization of r will be such that they have more funds than they require to meet their liabilities and other times there will be a shortfall. In order to ensure they can always meet their liabilities, companies can use contracts which require policyholders to make extra payments when there is a shortfall. Any surplus at date 1 can be paid out to policyholders. To see how this might work, suppose the funds an insurance company has available at date 1 as a result of investing the premiums at date 0 are αEL. If $\alpha < 1$, those who did not suffer a loss would be required to pay $(1-\alpha)EL$ each and those who did suffer a loss would receive $L-(1-\alpha)EL$. If $\alpha > 1$, those who do not suffer a loss receive a refund of $(\alpha-1)EL$ and those who do suffer a loss receive $L-(1-\alpha)EL > L-EL$.

Are the policyholders any better off because of the reduction in the average premium? The Modigliani-Miller analysis is again applicable here. Since markets are complete, all that would happen is that the policyholders would alter their portfolios to offset the position the insurance company takes. For example, if an insurance company with N policyholders invests a total of X in a security, each policyholder would reduce his or her investment in that security by $x = X/N$. Overall, the allocation of resources would be the same as in the benchmark equilibrium when all insurance payments are made at date 1. Policyholders would consume the same amount for every possible realization of r and the insurance company would make zero profits. This discussion gives the following result.

The Insurance Modigliani and Miller Theorem

If insurance contracts are perfectly enforceable, capital markets are perfect and complete and there are no taxes, an insurance company's investment strategy does not affect its policyholders' welfare.

One immediate corollary of this result is that even when there is only one type of loss being insured there is more than one level of premium that is optimal. An equivalent way of putting this is that the cost of capital at which firms should discount expected liabilities to arrive at the breakeven premium is not unique. It depends on the investment strategy the firm has decided to pursue but does not affect policyholders welfare since they will simply take offsetting positions. When there are multiple divisions in the insurance company, the same result will be true. There will be no sense in which there is a unique optimal premium or a well-defined divisional cost of capital.

3. Imperfect Contracts and Perfect Capital Markets

In this section the case where payments at date 1 cannot be enforced but capital markets are perfect is considered. In the previous section there were many different policies that an insurance company could follow. One of these policies stands out in the current context. It guarantees that there will be no shortfall of funds at date 1 so the fact that payments can only be collected at date 0 does not affect its feasibility. Suppose the insurance company charges a premium

$$p_0 = \frac{EL}{1 + r_F} \tag{3}$$

at date 0 and invests all its funds in the risk free asset. At date 1 it pays out L to all those who suffer a loss and nothing to those who do not. In this case there is no risk that it cannot meet its liabilities. Moreover, the allocation of resources is the same as the benchmark case when payments at date 1 are costlessly enforceable. The policyholders simply adjust their holdings of the risk free asset to offset the fact that the insurance company is effectively investing in the risk free asset on their behalf. They bear the same degree of capital market risk as they do in the benchmark equilibrium where all contracts are fully enforceable.

The cost of capital the insurance company should use in this case in determining its premium is the risk free rate. This is true if it has one division or many divisions.

The other important point here is that investing in the risk free asset is the best thing an insurance company can do for its policyholders. Investing in risky assets may allow it to lower its premium but this will not affect the welfare of its policyholders. To see this, consider what would happen if an insurance company were to invest the premiums in risky assets with expected rate of return $Er > r_F$ and with lower bound on the return of $0 < 1+r_l < 1+r_F$. To ensure that it can meet its liabilities at date 1 it must charge a premium of

$$p_0 = \frac{EL}{1 + r_l} > \frac{EL}{1 + r_F}. \tag{4}$$

Except when the realized rate of return is r_l, the insurance company will be able to refund part of the dividends to its policyholders. The average effective cost of the insurance in date 0 dollars, denoted c_0, will be

$$c_0 = \frac{EL}{1 + Er} < \frac{EL}{1 + r_F}. \tag{5}$$

The important point here is that even though the average effective premium is lowered, none of the firm's policyholders are made better off. They will simply adjust their portfolios to offset the investments made by the insurance company. Thus with perfect and complete capital markets

the insurance company cannot do any better than invest in the risk free asset and use the risk free rate as its divisional cost of capital.

4. Imperfect Contracts and Capital Markets

The next obvious case to consider is what happens when there are imperfect capital markets and all policyholders cannot simply adjust their portfolios to offset the investment strategy of insurance companies. The imperfections that are likely to lead to this are when policyholders can only borrow at rates in excess of r_F and short sales entail transaction costs; in extreme cases they will not be able to borrow or short sell at all.

For those policyholders who hold sufficient quantities of the risk free asset in their portfolios, the analysis in the previous section remains valid. They will reduce their investment in the risk free asset to offset the effect of paying a premium at date 0. In equilibrium, insurance companies will offer this clientele of consumers a policy with a premium and effective cost of

$$p_0 = c_0 = \frac{EL}{1 + r_F} \tag{6}$$

and will invest the proceeds in the risk free asset.

What about consumers who do not hold sufficient quantities of the risk free asset to do this? They will not be able to adjust their portfolios to perfectly offset the insurance companies' actions if the only policy that is offered corresponds to investing premiums in the risk free asset. It will be necessary to offer these people a different policy. Suppose these consumers all hold an optimal portfolio Z*, with expected rate of return ER* and a lower bound on the return of $1+R_l > 0$. In order to guarantee it can meet its liabilities the insurance company must charge a premium of

$$p_0 = \frac{EL}{1 + R_l}. \tag{7}$$

As in the example in the previous section, it will also refund any return on the investment portfolio above $1+R_l$ to consumers. Hence the expected cost of insurance in date 0 dollars is

$$c_0 = \frac{EL}{1 + ER^*}. \quad (8)$$

In this case, consumers are made strictly better off by insurance companies offering this policy in addition to the risk free policy. They can now adjust their portfolio to offset the investments of their insurance company and can have the overall portfolio that is optimal for them. Consumers would make this adjustment by reducing their holdings of the optimal portfolio by p_0 since this is the amount the insurance company is effectively investing on their behalf.

What happens if consumers have different optimal portfolios because, for example, their degree of risk aversion or level of wealth differs? In equilibrium, competition ensures each group of consumers will be offered a policy tailored for its particular preferences. Premiums will be set to guarantee that the company can meet its liabilities and the investment policy of the firm will mirror the optimal portfolio of the clientele the policy is designed for. The expected cost of the insurance will depend on the expected return of the consumers' optimal portfolio.

The analysis thus indicates that when there are multiple groups the same type of insurance will be priced differently both in terms of the premium and the expected cost even though the risk of loss that is being insured is the same. Even when there is one division there will therefore be many values for the discount rate, or cost of capital, to be applied to expected liabilities. Essentially it is policyholders that bear the investment risk and the appropriate discount rate is their opportunity cost of capital.[5] The premium has to be paid at date 0 because of the enforceability problem. However, the insurance company is effectively acting on each policyholder's behalf and is investing the premium in assets the policyholder would have invested in if there were no enforceability problem.

In the scheme described, the premium is set at a sufficiently high level to ensure that there will be no shortfall. In practice, the situation where it is perhaps easiest to do this is life insurance. Here the policies typically last for many years and the probability of a claim is low initially. This means that even with moderate premiums it is relatively easy to build up a surplus. When the policy eventually expires the surplus can be returned to the policyholder.

[5] It is the policyholders' opportunity cost of capital that is relevant here rather than the equityholders' since the policyholders are providing the funds to be invested. The case where equityholders provide funds is discussed below.

5. The Role of Insurance Company Equityholders

In the analysis of Section 4, it is assumed that there is a lower bound $1+R_l > 0$ on the returns to consumers' optimal portfolios. By setting a premium of $p_0 = EL/(1+R_l)$ it is possible to guarantee that the company can meet its liabilities. If $1+R_l$ is small this premium could be large and might exceed the available wealth of the consumers the insurance is targeted at. If $1+R_l = 0$ it will clearly be impossible to charge a premium which will guarantee that liabilities can be met. How can a company meet its obligations in these situations?

Insurance companies have been modelled so far as competitive firms which make zero profits no matter what happens. The owners or equityholders of the companies have not played a role at all. When there is a possibility of a shortfall this is no longer the case. The equityholders will have to bear this risk and will receive an appropriate return to compensate them.

How would such equityholder guarantees work in practice? One possibility is for there to be unlimited liability and for the equityholders to guarantee to make good any shortfall that occurs at date 1. In Sections 2 and 3 it was argued that it would not be possible in practice to write contracts with policyholders requiring those that do not suffer a loss to make a payment at date 1. In contrast, with insurance company equityholders this type of contract is observed. In essence this is the way that Lloyd's of London operates. People above a certain threshold of wealth apply to be "names" and pledge all their wealth to make good any shortfall when claims come due. Since the wealth threshold is set at a high level the costs of enforcing payment at date 1 are relatively low.[6] Thus the system where equityholders' payments occur at date 1 is of interest in its own right. As in Section 2 it is also of interest as a benchmark.

[6] Recent problems in the Lloyd's insurance market have led to resistance by some names to paying the amount they have been called on for. This suggests the costs of enforcing ex post payments are increasing.

The people that bear the residual liability in a Lloyd's type of system receive income to compensate them. Suppose the residual risk is represented by the random variable ε with mean $E\varepsilon$ and variance σ_ε^2. The income received by the people that bear this risk has two components. The first is to compensate them for the expected residual risk and is just $E\varepsilon$. The second is to compensate them for the remainder of the risk. It is shown in the Appendix, which contains an analysis of the market for residual risk, that for small risks this component will only depend on σ_ε^2. If the price that is paid for each unit of variance borne is Θ, the amount that is given at date 1 to compensate the person for bearing the risk is

$$y_1 = E\varepsilon + \Theta\sigma_\varepsilon^2. \tag{9}$$

It is also shown in the Appendix that the demand for residual risk by individuals is Θ/a where a is the person's coefficient of absolute risk aversion. The equilibrium value of Θ is the value such that the demand for bearing residual risk is equal to supply. If the risk borne is large, the income required to compensate for the residual risk will also depend on the third and higher moments of ε. In the analysis below it is assumed that only mean and variance are of any importance.

In equilibrium, the price Θ ensures that people are willing to bear the risk. It will depend on the extent to which the risk can be diversified away. When everybody is well diversified, as in the CAPM, Θ will be zero if the risk is diversifiable but nonzero if it is correlated with market risk. In practice, it appears that most investors are not very well diversified. Blume, Crockett and Friend (1974) found that the average amount of diversification is equivalent to having an equally weighted portfolio with two stocks. Blume and Friend (1978) provide more detailed evidence of this lack of diversification; they find that a large proportion of investors have only one or two stocks and very few have more than ten. One possible explanation of this lack of diversification is that there is a fixed set up cost of participating in a market.[7] Whatever the reason, if investors are not well diversified Θ will be positive.

The theory presented suggests that mutual insurance companies are feasible when $(1+R_l)$ is sufficiently large that the guaranteeing

[7] See Allen and Gale (1991) for a discussion of the evidence for lack of diversification and possible reasons for it.

premium can be afforded by many consumers. However, if $1+R_l = 0$ or is small so the guaranteeing premium $EL/(1+R_l)$ would be too large to be affordable, equityholders will be necessary for the insurance company. Consider how the required payment for the residual risk borne by equityholders affects the pricing of insurance. Policyholders must compensate the equityholders for bearing the residual liability.

In the current framework the reason residual risk arises is that the insurance company is acting as the policyholder's agent in investing the premium between dates 0 and 1. If a cost must be paid for this residual risk it may be that the policyholder's optimal portfolio is changed so that less risk is borne. In some cases it may be that they prefer a policy where the insurance company bears no risk and always invests the premiums in the risk free asset. For example, if markets are complete there will be no advantage from insurance company equityholders bearing the risk. However, in general, when markets are incomplete both equityholders and policyholders will bear risk.

The other decision concerns the level of the premium. The higher the premium the lower the equityholders' residual risk and hence the less compensation that must be paid by policyholders. The level of the premium will be determined by the ability of the policyholder to tie up wealth in the hands of the insurance company relative to the cost of compensating equityholders for the residual risk they bear. If there are constraints on borrowing and short sales, the ability of policyholders to pay large premiums may be severely limited and this may often be the determinant of the level of premiums.

Suppose the mean of the residual risk borne by the equityholders given the optimal premium and the optimal investment strategy for the policyholders' funds is μ and the variance is σ^2. Then the policyholders must pay $\mu+\Theta\sigma^2$ at date 1. The average cost of the insurance in date 0 dollars will be

$$c_0 = \frac{EL}{1 + ER^*} + \frac{\mu + \Theta\sigma^2}{1 + r_F}. \tag{10}$$

Notice that this assumes that the part of the premium to cover liabilities, $EL/(1+ER^*)$, is being held on the policyholders' behalf and is invested at expected return ER^*. The part of the premium to cover the cost of the residual risk, $(\mu+\Theta\sigma^2)/(1+r_F)$, is held on the equityholders' behalf. For simplicity, it is assumed these equityholders do not face any capital market imperfections and it is therefore optimal (i.e. as good as any other strategy) to invest it in the risk free asset.

As far as the level of the premium is concerned, it will depend on the amount of residual risk that it is optimal for equityholders to bear. In general, the higher the premium the less residual risk equityholders will bear and the smaller $\mu+\Theta\sigma^2$ will be.

The cost of capital (i.e. the value liabilities should be discounted at) corresponding to c_0 in (10) is the value of γ such that

$$\frac{EL}{1+\gamma} = \frac{EL}{1+ER^*} + \frac{\mu+\Theta\sigma^2}{1+r_F}. \tag{11}$$

As before, there is not necessarily a single cost of capital even when there is one type of loss being insured. Policies are designed for particular clienteles and the cost of capital used in pricing the insurance will be clientele specific. When there are multiple divisions this will be true for each division.

Apart from a Lloyd's-type system of equityholder payments at date 1, the other way to guarantee that an insurance company can meet its liabilities is for the shareholders to put up the necessary wealth at date 0 to guarantee payments at date 1. Similarly to Section 2, this avoids the problem of securing payment from them after the event. Conceptually, the surplus is being held on behalf of the equityholders. If it is not required at date 1 to meet insurance liabilities it is returned to them. These funds should therefore be invested on their behalf. As in Section 2, if the equityholders do not face any capital market imperfections or incompleteness a Modigliani-Miller type of analysis is again relevant. The precise way these funds are invested will be irrelevant as far as the shareholders' welfare is concerned because they will be able to take offsetting positions; if the insurance company buys x of a security per shareholder, each shareholder reduces their holding of it by x. Given the purpose of the funds is to assure liabilities at date 1 can be met, investing the surplus in the risk free asset is prudent. If equityholders face capital market imperfections then investing in risky assets on their behalf may simply result in the necessity of putting up more funds at date 0 to guarantee liabilities at date 1. For simplicity, it will be assumed below that all the surplus held on behalf of equityholders is invested in the risk free asset.

The crucial point here is that the funds are held by the insurance company to overcome problems of collection at date 1. In terms of

compensating equityholders for bearing the residual risk, the analysis for the Lloyd's type of situation is still valid. The equityholders in the insurance company will require compensation in the same way and the premium and cost of capital will be calculated as above. Thus the return on the surplus collected at date 0 to guarantee payment of liabilities at date 1 should not be factored into the premium at all. The cost that is important for the premium is that incurred in bearing residual risk.

The expected return that shareholders receive is a combination of the expected return earned on surplus held by the insurance company on their behalf and the compensation for bearing the residual risk. For example, if shareholders put up a guaranteeing surplus of G (and these funds are invested in the risk free asset) the expected fair rate of return r_E would be given by

$$r_E = \frac{(1+r_f)G + \mu + \Theta\sigma^2}{G} - 1. \tag{12}$$

An implication of this analysis is that the issue of how to allocate earnings on surplus to particular types of policy or divisions does not arise. What is important in determining the cost of insurance in different divisions is the residual liability that equityholders bear. What is important in determining the fair rate of return to equityholders is that they are appropriately compensated for the residual risk they bear. For example, suppose that a company has two divisions. The residual risk in division i (= 1, 2) has mean μ_i and variance σ_i^2. If the expected loss being insured in division i is EL_i and the expected return on the portfolio the premiums are invested in is ER_i^*, the average cost of the insurance would be

$$c_{0i} = \frac{EL_i}{1+ER_i^*} + \frac{\mu + \Theta\sigma_i^2}{1+r_F}. \tag{13}$$

The fair rate of return for shareholders who put up a total G of guaranteeing surplus is

$$r_E = \frac{(1 + r_F)G + \mu_1 + \Theta\sigma_1^2 + \mu_2 + \Theta\sigma_2^2}{G} - 1. \tag{14}$$

The extension to the case where there are more than two policies or divisions is straightforward.

6. Risky Insurance Liabilities

An important assumption of the analysis above is that the risks that are insured can be pooled so that overall the risk is eliminated. In many situations this may be an appropriate assumption. However, in others it may not. Even after pooling the policies, there may be residual risk. How does this affect the analysis of the cost of capital?

It is again helpful to address this question by considering the benchmark situation where contracts are costlessly enforceable. It is then possible to have all payments at date 1 so that the insurance component can be separated from the financial market component in the usual way. The loss can again be represented by the random variable L with mean EL but now the residual risk that cannot be pooled away has variance σ_L^2 per policyholder. Equityholders in the insurance company will be needed to ensure that the residual risk that remains after pooling is covered. Initially, it is assumed they have unlimited liability or the funds they commit at date 0 are sufficient to cover the highest realization of the residual risk. This is again like the Lloyd's-type system in Section 4. As there, suppose there is a price Θ per unit of variance at date 1. The date 1 payment in this case for policyholders who do not suffer a loss will be

$$p_1 = EL + \Theta\sigma_L^2 \tag{15}$$

rather than $p_1 = EL$. Those policyholders that do suffer a loss will receive $L - (EL + \Theta\sigma_L^2)$ rather than $L - EL$.

The next step is to suppose that premiums are collected at date 0. In this case the analysis will be the same as above but with $EL + \Theta\sigma_L^2$ replacing EL. It will be assumed initially that there are perfect capital markets so all premiums are invested in the risk free asset. The date 0 premium will be

$$p_0 = \frac{EL + \Theta\sigma_L^2}{1 + r_F}. \tag{16}$$

The cost of capital (i.e. the rate for discounting the expected liabilities) alone will be the value of γ such that

$$\frac{EL}{1 + \gamma} = \frac{EL + \Theta\sigma_L^2}{1 + r_F}. \tag{17}$$

For different divisions the residual risk born by equityholders will differ and so σ_L^2 and hence the cost of capital will differ across divisions as in the previous section.

In the case where there are capital market imperfections and policyholders do not hold a sufficient amount of the risk free asset to offset the premium but instead hold a portfolio of risky assets, insurance companies will offer a range of policies depending on the clientele that is sought. Suppose that the portfolio the premiums are invested in has expected rate of return ER*, and its returns are independent of the insurance losses then the expected cost of insurance is

$$c_0 = \frac{EL}{1 + ER^*} + \frac{\Theta\sigma_L^2}{1 + r_F}. \tag{18}$$

The corresponding cost of capital is the value of γ such that

$$\frac{EL}{1 + \gamma} = \frac{EL}{1 + ER^*} + \frac{\Theta\sigma_L^2}{1 + r_F}. \tag{19}$$

Again there will be multiple costs of capital within each division. Across divisions σ_L^2 will vary so for a given value of ER* the cost of capital will also differ.

If portfolio returns and aggregate losses are correlated there will be an extra term in (18) to account for this. When losses are positively correlated with returns the expected cost will be higher than in (18) but when they are negatively correlated it will be lower.

As far as the premium is concerned, this will depend on the value of R_l as before and the amount of residual financial risk that is optimal. If there is to be no residual financial risk then the premium must be sufficient to cover all liabilities and payments to the equityholders, as in Section 3, so

$$p_0 = \frac{EL}{1 + ER_l} + \frac{\Theta \sigma_L^2}{1 + r_F}. \tag{20}$$

If $1+R_l = 0$ or is small then it will not be possible to have a guaranteeing premium and in addition to the residual risk from the insurance, there may also be residual risk from the investment of the premium. In this case the average cost of the insurance will be

$$c_0 = \frac{EL}{1 + ER^*} + \frac{\Theta \sigma_L^2 + \mu + \Theta \sigma^2}{1 + r_F} \tag{21}$$

and the premium will depend on the amount of residual financial risk that is optimal.

An assumption of the analysis above is that equityholders either have unlimited liability or the funds that they pledge at date 0 are sufficient to cover the highest realization of the residual risk. If this is not the case there is a real possibility of bankruptcy of the insurance company. There is then another residual holder, namely the fund which in most states guarantees the liabilities of insurance companies. This extra level of risk bearing could be added to the analysis. The precise way in which premiums and the cost of insurance will be affected depends on how the guarantee is implemented.

7. Taxes

An important assumption in the derivation of the Insurance Modigliani and Miller theorem is that there are no taxes. The introduction of taxes adds a number of dimensions to the analysis. If there is a difference between the way in which investment income is taxed when the investments are held directly and when they are held by the insurance company the Insurance Modigliani and Miller theorem will no longer hold. Typically, insurance companies are taxed less heavily in

this respect than individuals and there may be opportunities for tax arbitrage. In this case, premiums will be higher than they need to be for pure insurance purposes to increase the investment component. Once again policies will be tailored to clienteles, but here the important thing will be tax status rather than the investors' optimal portfolios. In a competitive insurance market, it will be policyholders who obtain the benefits from the tax avoidance opportunities provided by insurance policies.

The main impact in terms of the effect on premiums and the cost of capital will be the fact that investment income and residual risk premia will be taxed. To illustrate, consider the case where capital markets are perfect, premiums are invested in the risk free asset and there is some residual risk from the liabilities. Here the counterpart of (16) if equityholders are to receive $\Theta\sigma^2$ after corporate taxes at date 1 is

$$p_0 = \frac{EL + \dfrac{\Theta \sigma_L^2}{1 - t}}{1 + (1 - t) r_F} \tag{22}$$

where t is the tax rate on insurance companies' income. The after-tax cost of capital will, similarly to (17), be the value of γ such that

$$\frac{EL}{1 + \gamma} = \frac{EL + \dfrac{\Theta \sigma_L^2}{1 - t}}{1 + (1 - t) r_F}. \tag{23}$$

As before σ_L^2 will vary across divisions and so the divisional cost of capital will differ. Other cases can be similarly analyzed.

8. Conclusions

Traditional financial analyses of the divisional cost of capital which take the structure of the insurance firm as given have not adequately dealt with the issue of how to allocate the earnings from surplus to different divisions. This paper starts with the benchmark case where contracts are fully enforceable. This allows the insurance market component of companies' activities to be separated from the financial market component and leads to the notion of there being a market for residual risk. It is the price of residual risk and the amount of residual risk in each division that determines the compensation insurance companies' equityholders receive. As a result the allocation of earnings from a firm's surplus to the various divisions does not arise and it is possible to calculate a divisional cost of capital.

One important issue that does arise is how the scheme for estimating the divisional cost of capital described here could be implemented in practice. The main problem in this respect is finding an empirical estimate for the price of residual risk. This could be done in two ways. First of all data from markets where residual risk is directly guaranteed such as Lloyd's of London could be used. Alternatively, the excess premium earned by insurance company equityholders could be estimated and compared with the residual risk that they bear. Once the price of the residual risk has been found, using it to estimate the divisional cost of capital is relatively straightforward.

APPENDIX

The Market for Residual Risk

Consider a simple case where an investor has wealth W_1 at date 1 and von Neumann Morgenstern utility function $U(W_1)$. The person's coefficient of absolute risk aversion is defined to be

$$a(W) = -\frac{U''(W)}{U'(W)}. \qquad \text{(A1)}$$

Suppose the residual risk ε is normally distributed with mean $E\varepsilon$ and variance σ_ε^2. If the investor assumes a part of the residual risk $X\varepsilon$, then he or she receives an amount $X(E\varepsilon + \Theta\sigma_\varepsilon^2)$ at date 1 in return. The person's expected utility is

$$EU = EU[W_1 - X\varepsilon + X(E\varepsilon + \Theta\sigma_\varepsilon^2)]. \qquad \text{(A2)}$$

Using a Taylor's series expansion it can straightforwardly be shown[8] that for small $X\varepsilon$ this can be written in the form

$$EU \approx U(W_1 - 0.5aX^2\Theta_\varepsilon^2 + X\Theta\sigma_\varepsilon^2). \qquad \text{(A3)}$$

Choosing X to maximize this gives the demand for residual risk as

$$X \approx \frac{\Theta}{a}. \qquad \text{(A4)}$$

As might be expected intuitively, the less risk averse the person and the higher the price, the greater is the demand to bear the residual risk.

The price of residual risk Θ will be determined by the total demand for bearing it from individuals and the total supply of residual risk from insurance companies.

When the risk borne by each person is large, (A3) will no longer hold and it will be necessary to use more terms in the Taylor's series expansion. The risk premium individuals require to bear the residual risk will then depend on the third and higher moments of ε as well as the mean and the variance.

---

[8] See Pratt (1964).

REFERENCES

Allen, F. and D. Gale (1991). "Limited Market Participation and Volatility of Asset Prices," Rodney L. White Center Working Paper 2-92, University of Pennsylvania.

Biger, N. and Y. Kahane (1978). "Risk Considerations in Insurance Ratemaking," *Journal of Risk and Insurance* 45, 121-132.

Blume, M., J. Crockett and I. Friend (1974). "Stock Ownership in the United States: Characteristics and Trends," *Survey of Current Business*, 16-40.

Blume, M. and I. Friend (1978). *The Changing Role of the Individual Investor: A Twentieth Century Fund Report*, New York: Wiley.

Cooper, R. W. (1974). *Investment Return and Property-Liability Insurance Ratemaking* (Philadelphia: S.S. Huebner Foundation, University of Pennsylvania).

Cummins, J. D. (1990a). "Asset Pricing Models and Insurance Ratemaking," *Astin Bulletin* 20, 125-166.

Cummins, J. D. (1990b). "Financial Pricing of Property and Liability Insurance," working paper, University of Pennsylvania.

Cummins, J. D. (1990c). "Multi-Period Discounted Cash Flow Ratemaking Models in Property-Liability Insurance," *Journal of Risk and Insurance* 57, 79-109.

Cummins, J. D. and L. Chang (1983). "An Analysis of the New Jersey Formula for Including Investment Income in Property-Liability Insurance Ratemaking," *Journal of Insurance Regulation* 1, 555-573.

D'Arcy, S. and N. A. Doherty (1988). *Financial Theory of Insurance Pricing* (Philadelphia: S.S. Huebner Foundation, University of Pennsylvania).

Fairley, W. (1979). "Investment Income and Profit Margins in Property-Liability Insurance," *Bell Journal of Economics* 10, 192-210.

Ferrari, J. R. (1968). "A Note on the Basic Relationship of Underwriting, Investments, Leverage and Exposure to Total Return on Owners' Equity," *Proceedings of the Casualty Actuarial Society* 55, 295-302.

Hill, R. (1979). "Profit Regulation in Property Liability Insurance," *Bell Journal of Economics* 10, 172-191.

Modigliani, F. and M. H. Miller (1958). "The Cost of Capital, Corporation Finance, and the Theory of Investment," *American Economic Review* 48, 261-297.

Myers, S. and R. Cohn (1987). "Insurance Rate Regulation and the Capital Asset Pricing Model," in J. D. Cummins and S. E.

Harrington, eds., *Fair Rate of Return in Property-Liability Insurance* (Norwell, MA: Kluwer Academic Publishers).

National Council on Compensation Insurance, (1987). *The Impact of the Tax Reform Act of 1986 on Property-Liability Insurance Profits* (New York: The Council).

Pratt, J. W. (1964). "Risk Aversion in the Small and in the Large," *Econometrica* 32, 122-136.

Sharpe, W. F. (1970). *Portfolio Theory and Capital Markets* (New York: McGraw Hill).