JIA 113 (1986) 499-508

THE OPTIMALITY OF THE NET SINGLE PREMIUM IN LIFE INSURANCE

By COLIN M. RAMSAY Ph.D., A.S.A.*

(of the Department of Actuarial Science in the University of Nebraska—Lincoln)

INTRODUCTION

Vadiveloo *et al* (1982) introduced a criterion for choosing, among the various net risk premium payment plans, the so-called optimum one. They considered a life insurance situation where benefits were payable at the end of the year of death. Of course, premiums cease after death! This optimum criterion was the minimization of the discounted net profit variance. Their model can be mathematically described as follows. Consider an insured life aged exactly x having a death benefit with present value, if death occurs at time t, denoted by $B(t)$. If the present value of the total premiums paid up to time t is $R(t)$ and the random variable T denotes the time of the death of x, then the profit at death has present value $Z(T)$ where

$$Z(T)=R(T)-B(T).$$

The optimum net premium payment plan, $R^*(t)$, is the one that minimizes $\mathrm{Var}[Z(T)]$ subject to $E[Z(T)]=0$.

Two cases were considered:

(1) $B(t)$ as a non-increasing, deterministic function of time. It was proved that the optimum net premium plan was a single premium, paid at $t=0$, of amount $E[B(T)]$.

(2) $B(t)$ as a non-decreasing function of time. This is typical for annuity-type benefits. The optimum net premium payment plan was to let $R(t)=B(t)$ for all $t\geqslant 0$, a savings account in other words.

The model used by Vadiveloo *et al* was quite restrictive in the sense that they considered both non-stochastic benefits and non-stochastic interest rates. In modern insurance product design, there is a move to allow for stochastic variations in the value of the benefit to be paid. An increasingly popular practice is to 'index-link' or 'unit-link' the maturity value of the insurance benefit to the performance of some stochastic index or unit such as the rate of inflation, the Consumer Price Index, the yield rate of a group of stocks, etc. See Wilkie (1981) for the treatment of linked financial contracts. However, the vast majority of insurance contracts are still written using non-stochastic interest and benefit assumptions.

* The essence of this research was completed during the summer of 1985 while the author was visiting The City University, London.

In the sequel a life insurance situation is considered where, without any loss of generality, the insurance death benefits are assumed to be paid immediately upon death. The size of these benefits is considered stochastic and can depend on the time of death, the evolution of some 'index' over the life of the contract, and/or other factors such as interest rate and inflation rate fluctuations. It is proved that, under certain conditions, the net single premium remains optimal (using the minimum variance of the discounted profit at death criterion), and this optimal premium produces the smallest expected underwriting gain at death. In practice these results will serve to highlight the conflict between the need for greater stability (as represented by the minimization of the discounted profit variance) and the desire for larger under-writing profits at death. Thus, for example, an insurance company interested in greater stability will prefer the net single premium plan or a plan where most of the premium payment is completed during the early life of the policy.

2. THE MODEL

Consider an insurance company which is about to insure a group of N similar lives all aged exactly x, each effecting the same type of insurance policy. The following seven assumptions will be made about the nature of the insurance environment:

(1) If T_i denotes the time of death of the i^{th} life, $i = 1, 2, \ldots, N$, then the T_i's are independent and identically distributed (i.i.d.) random variables with known distribution function $F(t)$, $0 \leq t < \infty$. The T_i's are independent of market conditions.

(2) Without loss of generality, the benefits are paid immediately upon the death of each insured life. The size of the death benefit paid, if death occurs at age $x+t$, is $Y(t)$. $Y(t)$ is considered to be a non-negative stochastic process. For 'traditional' life insurance products such as whole life or ordinary term insurance policies, $Y(t)$ is usually deterministic. However, over the past decade or so, insurance companies have been selling 'linked' policies. Such policies have no set (or fixed) insurance benefits, instead these benefits depend on the performance of certain 'units' or 'indices' which fluctuate in a random manner and thus yield random returns. The resulting $Y(t)$ is therefore a stochastic process.

(3) The rate of interest used to calculate the premium is a non-negative stochastic process. Let $V(t)$ be the discounted value of 1 to be paid at the end of t years, then $V(t)$ is a decreasing stochastic process. In most life insurance premium calculations this rate of interest is a deterministic function of time. However, since we assumed $Y(t)$ is stochastic, it seems natural to extend this assumption to $V(t)$.

(4) Premiums are non-negative and are not paid after death. Let $P(t)$ denote the total monetary premiums paid in $[0, t)$ i.e., excluding any interest earned. $P(t)$ will be called a premium payment plan if it is deterministic and is completely specified, *a priori*, for all $t \geqslant 0$. The present value of the aggregate premiums paid up to time t is a stochastic process and will be denoted by $R(t)$, i.e.

$$R(t) = \int_0^t V(s)\mathrm{d}P(s),\ t \geqslant 0. \tag{2.1}$$

Both $P(t)$ and $R(t)$ are of course non-negative and non-decreasing.

(5) The present value of the death benefit $Y(t)$ is denoted by $B(t)$ where

$$B(t) = V(t)Y(t),\ t \geqslant 0. \tag{2.2}$$

$B(t)$ is considered to be a non-negative and non-increasing stochastic process. The assumption that $B(t)$ is non-increasing is not a restrictive one. In fact a large class of insurance policies have death benefits which adhere to this condition. Among them are policies where (a) the sum assured is fixed or decreasing over time, and (b) the death benefit increases at a fixed or variable rate. For policies in (b), it is standard actuarial practice to discount such benefits, for the purpose of obtaining the present value, by assuming the discounting rate of interest exceeds the rate of increase of the death benefits.

(6) The variance of $B(T_i)$ does not depend on $P(T_i)$. This assumption is necessary for the purpose of separating the variances of $B(T_i)$ and $R(T_i)$, and it will simplify the subsequent analysis. Policies which include a return of premiums at death will violate this assumption.

(7) All stochastic processes have finite first and second moments with respect to t as well as T_i.

Overall these assumptions are not too restrictive in practice and are valid for a large class of insurance plans. However, in some insurance schemes the benefits are paid not only at death, but are paid if the insured life survives the duration of the policy (e.g., an endowment assurance). To include these types of insurance policies, we will assume throughout the rest of this paper that, if there exists a non-random point n beyond which the possibility of a benefit being paid is zero (i.e., $B(t)=0$ for $t>n$), then premiums are not paid beyond this point (i.e., $\mathrm{d}P(t)=0$ for $t>n$). Hence T_i can be viewed as the time at which a *benefit* is paid to the i^{th} policy (with the appropriate change in its distribution function). For example, if there is an n year endowment insurance, the new distribution function of T_i is $F_n(t)$, given by

$$F_n(t) = \begin{cases} F(t), & 0 \leq t < n \\ 1, & t \geqslant n. \end{cases}$$

This adjustment to the representation of $B(t)$ and $P(t)$ for values of t beyond n will ensure that the random variables T are not defective, i.e., $Pr[0 \leqslant T_i < \infty] = 1$.

Finally, having placed constraints on $B(t)$ and $P(t)$, we now consider a criterion for choosing the optimal premium payment plan. Let Φ be the class of all functions $P(t)$ satisfying $E[Z(T_i)]=0$, where

$$Z(t) = R(t) - B(t),\ t \geqslant 0. \tag{2.3}$$

Formally, Φ is defined as

$\Phi = \{P(t), t \geqslant 0$: $E[Z(T_i)] = 0$, $i = 1, 2, \ldots, N$ and $P(t)$ is a non-decreasing, non-negative, deterministic function of $t\}$

Any $P(t) \varepsilon \Phi$ is called a net premium payment plan. Note that the expectation of $Z(T_i)$ is taken with respect to both T_i and the various possible realizations of the stochastic process $Z(t)$, i.e.,

$$E[Z(T_i)] = \int_0^\infty E[Z(t)]\, dF(t). \tag{2.4}$$

A net premium payment plan $P^*(t)$ is said to be optimal if and only if $P^*(t) \varepsilon \Phi$ and $P^*(t)$ minimizes the variance of

$$\sum_{i=1}^{n} Z[T_i].$$

Quantities associated with this optimum plan will have an asterisk (*) appended to them. Since the T_i's are i.i.d., $P^*(t)$ will be optimal if and only if, for $i = 1, 2, \ldots, N$,

$$\mathrm{Var}[Z^*(T_i)] \leqslant \mathrm{Var}[Z(T_i)]. \tag{2.5}$$

In view of this, the subscript on T will be dropped in the subsequent analysis when there is little likelihood that confusion will be caused.

3. THE OPTIMUM PREMIUM

In their paper, Vadiveloo *et al* considered only the cases where $B(t)$ and $R(t)$ are deterministic functions. Using the model described in Section 2 will result in our considering a larger class of insurance policy types. However, the crux of the proof used to establish the optimality of the net single premium still remains the same, i.e., that $\mathrm{Cov}[B(T), R(T)]$ is negative. Vadiveloo *et al* proved that $\mathrm{Cov}[B(T), R(T)]$ is negative if $B(t)$ and $R(t)$ are non-increasing and non-decreasing functions respectively. Is this covariance negative for $B(T)$ and $R(T)$ described in Section 2? To answer this question the nature of the covariance of $B(T)$ and $R(T)$ must be more closely investigated.

In order to find the $P^*(t) \varepsilon \Phi$ which minimizes $\mathrm{Var}[Z(T)]$, we must first expand $\mathrm{Var}[Z(T)]$ in terms of $\mathrm{Var}[B(T)]$, $\mathrm{Var}[R(T)]$ and $\mathrm{Cov}[B(T), R(T)]$. To this end, let A be the net single premium for the insurance $B(T)$, i.e., for all $P(t) \varepsilon \Phi$, set

$$A = E\,[R(T)] = \int_0^\infty E\,[R(t)]\mathrm{d}F(t) \tag{3.1}$$

$$= \mathrm{E}\,[B(T)] = \int_0^\infty E\,[B(t)]\mathrm{d}F(t)$$

This leads to

$$\begin{aligned}\mathrm{Var}[Z(T)] &= E[(R(T)-A+A-B(T))^2] \\ &= \mathrm{Var}[R(T)] + \mathrm{Var}[B(T)] - 2\mathrm{Cov}[B(T), R(T)]. \qquad (3.2)\end{aligned}$$

But by assumption (6), $P(T)$ does not affect $\mathrm{Var}[B(T)]$, therefore $P^*(t)$ must minimize $\mathrm{Var}[R(T)]-2\mathrm{Cov}[B(T), R(T)]$. This covariance can be written as

$$\begin{aligned}\mathrm{Cov}[B(T), R(T)] &= E[(B(T)-A)(R(T)-A)] \\ &= \int_0^\infty E[(B(t)-A)(R(t)-A)]\mathrm{d}F(t). \qquad (3.3)\end{aligned}$$

Clearly, under the condition of a negative covariance in (3.3), $\mathrm{Var}[Z(T)]$ is minimized when $R^*(t)=A$ for all $t\geqslant 0$ because both $\mathrm{Var}[Z^*(T)]$ and $\mathrm{Cov}[B(T), R^*(T)]$ are zero. This gives $P^*(t)=A$ for all $t\geqslant 0$ and

$$\mathrm{Var}[Z^*(T)]=\mathrm{Var}[B(T)].$$

If, however, this covariance is non-negative, it might be possible that $\mathrm{Var}[Z(T)]$ can be made less than $\mathrm{Var}[B(T)]$ by choosing a net premium payment plan different from the net single premium. Thus it is important that $\mathrm{Cov}[B(T), R(T)]$ be negative or zero in order to establish the optimality of the net single premium. Since $B(t)$ and $R(t)$ move in opposite directions as t increases, then one will expect a negative co-variance to exist between them. Clearly if, for each $t\geqslant 0$, $B(t)$ and $R(t)$ are deterministic or are uncorrelated, the method of Vadiveloo *et al*'s Proposition A1 proves this covariance is non-negative. Thus, intuitively, one would expect a negative co-variance if both $B(t)$ and $R(t)$ are stochastic and there is not much random fluctuation between them at each value of t. In other words, the covariance will tend to be negative if market conditions (e.g. interest rate and stock price levels) do not fluctuate too wildly. See the Appendix for a 'probability theoretic' justification of this assertion.

So far we have considered only the class of net premiums. In practice, companies have to obtain gross premiums because of expense and profit considerations. The natural question to ask is whether or not the conditions described above will ensure that the gross single premium remains optimal. To investigate this further, assume the expenses incurred in $(t, t+dt)$ are expressed as percentages of the gross premium and death benefit, the percentages being denoted by $\rho(t)$ and $\beta(t)$ respectively. $H(t)$ is the discounted value of all other expenses paid in $(0, t)$.

Let $G(t)$ be the total monetary amount of gross premiums paid up to time t and $R_1(t)$ be discounted value of this less expenses related to premiums, i.e.,

$$R_1(t) = \int_0^t (1-\rho(s))\, V(s)\, \mathrm{d}G(s), \qquad t\geqslant 0.$$

If death occurs at time t, the present value of expenses not related to premiums is $X(t)$ where

$$X(t) = H(t) + \beta(t)\, B(t),\ t \geqslant 0.$$

If a profit margin with present value L is included, the gross premium $G(t)$ must be chosen in such a manner that it satisfies the following equation:

$$E[R_1(T)] = L + E[X(T) + B(T)].$$

If $X(t) + B(t)$ is a non-increasing stochastic process, the single premium equal $G^*(t)$ is optimal, where

$$G^*(t) = [L + A + E[X(T)]]/[1 - \rho(0)]$$

It must be pointed out that $\rho(t)$ will depend on the type of gross premium payment plan. For example, if premiums are level and are paid at the start of each year for, say, 20 years, it usually is the case that $\rho(0) \geqslant 1$. However, in the case of single premium this cannot be the case. Next we consider a property of the net single premium not considered by Vadiveloo *et al*, i.e., the expected underwriting gain at death.

4. THE UNDERWRITING GAIN

The underwriting gain at death, as a random variable, was introduced in the context of a paradox by Jewell (1980). It was shown by Chan and Shiu (1982) to have a positive expectation if interest rates are positive and constant, and premiums are of the level annual variety. The underwriting gain at death, $U(T)$, is the accumulated value of the profit on the net premium immediately after the payment of the death claim. Let $S(t)$ be the accumulated value of 1 to the end of t years, i.e.,

$$S(t) = 1/V(t),$$

then

$$U(T) = S(T)\, Z(T), \tag{4.1}$$

where T is the random time of death. It will be proved that under certain conditions, the net single premium $P^*(t) = A$ produces a non-negative expected underwriting gain at death and that this underwriting gain is the lowest among all net premium payment plans.

If $U^*(T)$ is the underwriting gain at death under the optimum plan $P^*(t)$, it satisfies

$$\begin{aligned} E[U^*(T)] &= E[S(T)\, Z^*(T)] \\ &= E[S(T)\,(A - B(T))] \\ &= -\mathrm{Cov}[B(T), S(T)]. \end{aligned}$$

However, since $B(t)$ and $S(t)$ move in opposite directions as t increases, the arguments used in the Appendix suggest that $\mathrm{Cov}[B(T), S(T)]$ will be negative if the market conditions do not fluctuate too much. This covariance is obviously negative if $B(t)$ and $S(t)$ are non-random for each $t \geqslant 0$. Therefore under conditions of low 'market' volatility,

$$E[U^*(T)] \geqslant 0. \tag{4.2}$$

Consider any other premium payment plan $P(t) \varepsilon \Phi$ and let $U(T)$ be its resulting underwriting gain at death. Clearly

$$U(T) - U^*(T) = S(T)(R(T) - A)$$

and

$$E[U(T) - U^*(T)] = \mathrm{Cov}[S(T), R(T)]$$

Since both $S(t)$ and $R(t)$ are increasing functions of t, one would expect their covariance to be positive if there is little market volatility. This leads to

$$E[\mathrm{U}(T) - U^*(T)] \geqslant 0. \tag{4.3}$$

Hence the expected underwriting gain at death is minimized under the net single premium plan (under stable market conditions).

5. COMMENTS

The object of this research has been to extend the results of Vadiveloo *et al* to cases where factors are stochastic. In so doing, it was demonstrated that their results can be extended successfully. In practice this means that the optimal net single premium puts a high probability on the discounted profit per life being close to zero. Due to the law of large numbers, this will be especially true for a large block of insured lives. These authors also suggested that if paying a single premium is unreasonable, then in order to maintain a smaller profit variance, it is desirable to collect most of the premium payments as early as possible in the life of the policy (see their Theorem 1). Even though this result was not proved in this paper, it seems clear that a similar result can be established if market conditions are not too erratic. Thus the insurance company selling 'linked' insurance contracts will, in the interest of variance reduction, design contracts with most of the premiums collected early in the life of the policy. However, in order to market this contract, the company will have to pay close attention to the preferences of the prospective policy holders.

The perspective thus far has been from the point of view of the insurer. However, to the insured life, the problem of choosing an optimum premium payment plan requires optimum criteria that are different from those we have considered so far. The insured life is usually interested, not in minimizing the profit variance, but in the types of premium payment plans and their respective periodic costs. Several authors have studied the effects of the various methods of paying premiums. Meyer and Power (1973), using the concept of opportunity cost, developed a model for the optimal number of payments the insured should choose (for non-life insurance). Their method can be used in life insurance situations with appropriate modifications. Polk (1974) introduced the flexible premium life insurance concept which permits the utmost freedom for policyholders to choose premium and benefit scales which best suit their needs. Since mortality tables are constructed with integral ages, the method of calculating the non-annual premiums is important. Skipper (1980) compared the relative cost rankings between premiums paid annually and those paid more frequently, and showed that the mode of premium payment has a substantial effect on the rankings of at least some insurers. Overall, it is clear that the interests of the insurer and the insured are, in large measure, conflicting. In general, the choice of an optimality criterion for the insured will be a difficult task and will depend almost totally on the particular individual's utility with respect to the various premium plans.

Finally, if one considers a casualty insurance portfolio then no random variable similar to T exists. Therefore the optimality criteria considered in this paper cannot be used. However, Ramsay (1986) considered an optimality criterion based on deviations of the risk reserve from its expected path. Since the process does not terminate at a random time, the net single premium is not optimal.

REFERENCES

(1) CHAN, F. Y. & SHIU, E. S. W. (1982) Underwriting gain as a random variable. *ARCH: Society of Actuaries*, **2**, 5–9.

(2) GRIMMETT, G. R. & STIRZAKER, D. R. (1982) *Probability and Random Processes*. Oxford University Press, New York.

(3) JEWELL. W. S. (1982) Models in insurance: paradigms, puzzles, communications and revolutions. *Trans. 21st I.C.A.* Vol. S, 87–141.

(4) MEYER, R. L. & POWER, F. B. (1973) Total insurance costs and the frequency of premium payments. *Journal of Risk and Insurance*. Vol. XL, 599–605.

(5) POLK, K. E. (1974) Variable premium life insurance. *Transactions of the Society of Actuaries*. Vol. XXVI, 449–466.

(6) RAMSAY, C. M. (1986) An optimum and equitable net risk premium payment plan. *Journal of Risk and Insurance*. Vol. LIII, 294–300.

(7) SKIPPER, H. (1980) The effect of premium payment frequency on life insurance cost rankings. *Journal of Risk and Insurance*. Vol. XLVII, 291–303.

(8) VADIVELOO, J., MEHROTA, K. & KAMINSKY, K. (1982) Some results on optimum premium payment plans. *Scand. Act. J.* 155–167.

(9) WILKIE, A. D. (1981) Indexing long-term financial contracts. *J.I.A.* **108**, 299–360.

APPENDIX

Let us now take a more detailed look at the form of Cov$[B(T), R(T)]$. This will require a slight indulgence in probability theoretic concepts. An elementary reference is the text by Grimmett and Stirzaker (1982). Formally we denote the probability space on which both $B(t)$ and $R(t)$ are defined, by (Ω, τ, μ). For any fixed $w \varepsilon \Omega$ there is a corresponding collection $[B(t;w)\colon t \geqslant 0]$ and $[R(t;w)\colon t \geqslant 0]$ called the realization or sample path of B and R respectively at w. For our purposes Ω can be regarded as the totality of all possible market conditions that can affect either $Y(t)$ and/or $V(t)$ (and as a result $B(t)$ and/or $R(t)$) for all $t \geqslant 0$. Thus each $w \varepsilon \Omega$ represents one possible evolution of the market and $d\mu$ represents the probability of an w. τ represents the sigma-algebra generated by the subsets of Ω. It is reasonable to assume that the times of death and the market conditions are mutually independent.

The covariance can now be written as

$$\begin{aligned}\mathrm{Cov}[B(T), R(T)] &= E[E[(B(t;w)-A)(R(t;w)-A)] \mid w] \\ &= \int_{\Omega}\int_{0}^{\infty} (B(t;w)-A)(R(t;w)-A)\,\mathrm{d}F(t)\,\mathrm{d}\mu. \qquad \text{(A.1)}\end{aligned}$$

Once $w \varepsilon \Omega$ is specified, $B(t;w)$ and $R(t;w)$ are non-random functions of t. Let $b(w)$ and $r(w)$ be conditional expectations, i.e.,

$$b(w) = \int_0^\infty B(t;w)\mathrm{d}F(t)$$

and

$$r(w) = \int_0^\infty R(t;w)\mathrm{d}F(t)$$

The equation (A.1) can now be rewritten as

$$\mathrm{Cov}[B(T), R(T)] =$$

$$\int_{\Omega}\int_{0}^{\infty} (B(t;w)-b(w))(R(t;w)-r(w))\mathrm{d}F(t)\,d\mu + \int_{\Omega} (b(w)-A)(r(w)-A)\mathrm{d}\mu. \qquad \text{(A.2)}$$

From assumption (5), $B(t)$ is a non-increasing function of t, so $B(t;w)$ is a non-increasing deterministic function of t for each $w \varepsilon \Omega$. Similarly $R(t;w)$ is a non-decreasing deterministic function of t for any given w. Using a method similar to Vadiveloo *et al* Proposition A1 yields, for each w,

$$\mathrm{Cov}[B(T;w), R(T;w)] = \int_0^\infty (B(t;w)-b(w))(R(t;w)-b(w))\mathrm{d}F(t) \leqslant 0$$

Thus, the first expression on the right hand side of (A.2) is negative.

The determination of the sign and the size of the second expression on the right hand side of (A.2) is by no means a simple matter. In fact, except for extremely simple or artificial examples, it will be almost impossible to evaluate this expression. The major difficulties will lie in the determination of the elements in Ω and the probability measure μ. However we know that if market conditions are completely predictable (i.e., $V(t)$ and $Y(t)$ are completely specified and non-random) then $b(w)$ and $r(w)$ will be identical to A (since Ω will be a singleton set), and

$$\int_{\Omega} (b(w)-A)(r(w)-A)\mathrm{d}\mu = 0.$$

So if Ω is such that the market conditions do not deviate too widely from its expected path, the above integral should be small in absolute value compared to the absolute value of the first term on the right hand side of (A.2). As a consequence one would expect $\mathrm{Cov}[B(T), R(T)]$ to be negative. Once this covariance is negative, the net single premium will be the optimal premium. Clearly each of the following conditions on $B(t;w)$ and $R(t;w)$ is sufficient to ensure the expression (A.1) is non-positive: For each $t \geq 0$

(i) $B(t;w)$ is non-random;
(ii) $R(t;w)$ is non-random;
(iii) $\mathrm{Cov}[B(t;w), R(t;w)]$ is non-positive.