STOCHASTIC MODELLING OF PENSION SCHEME DYNAMICS

by

S HABERMAN

Actuarial Research Paper No. 106

Department of Actuarial Science and Statistics
City University
London

February 1998

ISBN 1 901615 24 3
“Any opinions expressed in this paper are my/our own and not necessarily those of my/our employer or anyone else I/we have discussed them with. You must not copy this paper or quote it without my/our permission”.
ABSTRACT

The paper describes mathematical models proposed for investigating the behaviour of defined benefit pension schemes in the presence of stochastic investment returns.

The paper describes two general approaches to modelling and controlling the funding of such pensions schemes: by adjusting contributions in response to emerging surpluses or deficits, and by using an optimal control theory approach to optimize particular performance criteria.

Interest here is particularly focused on the efficacy of the various control variables at the disposal of the actuary, including the choice of amortization period, the delay in fixing contributions, the frequency of valuation and the choice of funding method for eliminating the unfunded liability. The paper closes with suggestions for topics for further investigation.

Key words

Pension scheme modelling; stochastic investment returns; control.
1. **INTRODUCTION**

Defined benefit pension schemes, which are common in a number of countries including the UK, USA, Canada and the Netherlands, are arrangements for a group of members where the benefits promised in the event of various contingencies are defined by a formula while the contributions (to be paid by the employer and possibly the member) are to be determined by the actuary as part of the regular valuation process. (In recent years, there has been a shift towards defined contribution schemes: see Turner and Beller (1992) for a discussion and towards the development of hybrid schemes: see Khorsanee (1995) for an analysis).

The fund associated with such a scheme can be regarded as a reservoir into which income from contributions and investment earnings (including the proceeds of sales and maturities and unrealised capital growth) flow and out of which benefit payments on the contingencies of age retirement, disability, death and withdrawal and so on would be made.

The most financially significant benefit is age retirement. For this contingency, the benefit would be in the form of a pension, payable while the member is alive. There may also be a lump sum benefit payable at retirement and an entitlement to a reversionary pension payable to a surviving spouse. The defined benefit formula for the basic annual pension to be member would be normally of the form:

\[
\text{Annual Pension} = K \times (\text{number of years of membership}) \times (\text{earnings averaged over the h years before retirement})
\]

where K (the accrual rate) and h are specified in the scheme rules. In contrast, the annual contribution formula would be of the form

\[
\text{Annual Contribution} = c \times (\text{current pensionable earnings})
\]

where c is not specified in the scheme rules but is determined by the "funding method" used by the actuary at each valuation. The valuations take place at regular intervals and the actuary values the prospective liabilities (i.e. benefit promises) allowing for the value of the future contributions which are expected to be paid at the assumed rate c, and compares this result with the value of the assets currently held in the fund. These calculations (and more detailed analyses) are used to determine c, which is held fixed for the period up to the next valuation.

This financing arrangement depends critically on the presence of regular valuations at which assets, prospective liabilities and future contributions are compared. At each valuation, the actuary is required to make detailed assumptions about the demographic and economic future of the pension scheme. These assumptions may constitute "best
estimates" of the various parameters but are not long term predictions as the actuary will have the opportunity to revise these estimates at the next and subsequent valuations.

The methods and assumptions available to the actuary in these routine valuations are not prescribed in UK, US or Canadian practice. In the event of a surplus or a deficiency being revealed at the next valuation, the contribution rate would be adjusted for the future. The financial status of the scheme would then be reviewed at the next valuation. In the UK, it is common for these valuations to be annual, although legislation requires a valuation to be performed at least every 3½ years.

In recent experience, one of the principal sources of surplus or deficiency has been the rate of investment return on pension scheme assets. It has become customary for actuaries to analyse the behaviour of the finances of pension schemes using simulation. A recent development in the literature has been the use of simple, analytic models of pension scheme dynamics. These models can lead to definitive results which, providing the underlying assumptions and simplifications are reasonable, then can provide insight and identify avenues worthy of further investigation based on more complex models, perhaps with the assistance of simulation. Some of the results obtained have helped to elucidate the effects of varying investment returns on the moments of the contribution rate and level of assets and to identify the effects of the various control variables at the disposal of the actuary e.g. the spread or amortization period (for dealing with valuation surpluses and deficiencies), delay in fixing contributions, frequency of valuation and the choice of funding method. The purpose of the paper is to review the models proposed and the results obtained, and to offer suggestions for further work in this area.

2. TYPES OF FUNDING METHOD

Pension funding methods are used to determine the pattern over time of the contributions to be paid. In common practice, there are several different types of funding method used. These can be categorised in a number of ways. Thus, in the UK, the split between accrued benefit methods and projected benefit methods is widely recognised. We shall use a different categorisation based on the mathematical structure of the fundamental equations. Therefore, we shall consider individual and aggregate funding methods.

In pension funding, the normal cost is used to describe the (stable) level of contribution which would apply if all the valuation assumptions made were to be borne out in actual experience. The actuarial liability is used to describe the mathematical reserve held: for some pension funding methods it can be thought of as the difference between the present value of benefit promises expected to be paid out of the pension scheme and the present value of normal costs expected to be paid to the pension scheme.
With individual funding methods (e.g. Projected Unit Credit and Entry Age Normal), the normal cost (NC) and the actuarial liability (AL) are calculated separately for each member and then summed to give the totals for the population under consideration. With aggregate funding methods (e.g. Aggregate and Attained Age Normal), there may not be explicit determination of a normal cost or actuarial liability; instead the group of members is considered as an entity, ab initio.

Let C(t) and F(t) be the overall contribution and fund level at time t. We consider the case where F(t) is measured in terms of the market value of the underlying assets.

For an individual funding method,

\[ C(t) = \sum NC(x, t) + ADJ(t) = NC(t) + ADJ(t) \]  \hspace{1cm} (1)

where the summation is taken over all members and where NC(x,t) is the normal cost for a member aged x at time t, NC(t) is the total normal cost at time t and ADJ(t) is an adjustment to the contribution rate at time t represented by the liquidation of the unfunded liability at time t, UL(t). UL(t) is defined by

\[ UL(t) = \sum AL(x, t) - F(t) = AL(t) - F(t) \]

where the summation is taken over all members and where AL(x,t) is the actuarial liability for a member aged x at time t and AL(t) is the total actuarial liability, in respect of all members, at time t.

For an aggregate method, the overall contribution is directly related to the difference between the present value of future benefits and the fund.

Specifically,

\[ C(t) = (PVB(t) - F(t))/S(t)/PVS(t) \]  \hspace{1cm} (2)

where S(t) is the total salaries of active members at time t, PVB(t) is the present value of future benefits (of all members including pensioners) at time t and PVS(t) is the present value of future salaries (of active members) at time t. Here, the difference, PVB(t) - F(t), is spread over the remaining period of membership of current members, effectively by an annuity which allows for expected earnings progression and which has expected present value $\frac{PVS(t)}{S(t)}$. 
For individual funding methods, there are a number of choices for the ADJ(t) term. The most commonly used are the spread method (UK) and the amortization of losses method (Canada and USA).

Under the spread method, \( ADJ(t) = k \cdot UL(t) \) \( \quad (3) \)

where \( k \) is the reciprocal of a compound interest annuity value calculated at the valuation rate of interest, \( i \), i.e. 
\[ k^{-1} = \sum_{j=0}^{M-1} (1 + i)^{-j}. \]

So the unfunded liability is spread over \( M \) years, where \( M \) would be chosen by the actuary. It should be noted that this definition of \( ADJ(t) \) uses the same fraction of the unfunded liability regardless of the sign of the latter. So, surpluses and deficiencies would be treated in a comparable manner - this would not always be the case in practice (see, for example, Winklevooss (1993)). Typical values of \( M \) would be 20-25 years, corresponding approximately to the average remaining period of membership of current members. \( k \) is the fraction of \( UL(t) \) that makes up \( ADJ(t) \) and can be thought of as a penalty rate of interest that is being charged on the unfunded liability, \( UL(t) \).

For the amortization of losses method, we introduce the actuarial loss experienced during the interval between \( t-1 \) and \( t \), \( l(t) \), which is defined as the difference between \( UL(t) \) and the value of the unfunded liability if all the actuarial assumptions had been realized during the year \( t-1,t \). Then \( ADJ(t) \) is defined as the total of the interval valuation losses arising during the last \( m \) years (i.e. between \( t-m \) and \( t \)) divided by the present value of an annuity for a term of \( m \) years calculated at the valuation rate of interest (i.e. spread over an \( m \) year period). Thus,

\[ ADJ(t) = \frac{\sum_{j=0}^{m-1} l(t-j)}{\bar{a}_{m \mid}}. \] \( \quad (4) \)

As Dufresne (1989a) has shown, under these conditions, \( UL(t) \) satisfies a recurrence relation

\[ UL(t) = (1+i) (UL(t-1) - ADJ(t-1)) + l(t) \]

which can be solved to give

\[ UL(t) = \sum_{j=m}^{m-1} \lambda_i \cdot l(t-j) \text{ where } \lambda_i = \frac{\bar{a}_{m-j}}{\bar{a}_{m \mid}}. \] \( \quad (5) \)

Then \( F(t) = AL(t) - UL(t) \) and \( C(t) = NC(t) + ADJ(t) \).

Here, \( m \) would be chosen by the actuary and would typically lie in the range 5-15 years.
3. **CONTROL OF FUNDING**

We describe two general approaches to the control of pension funding.

We first consider the "classical" approach of adjusting contributions for the spread method and amortization of losses method, and consider the existence of optimal choices for the spread period M or the amortization period m.

We then consider a more holistic approach based on the use of optimal control methods to choose the contributions in order to optimize a range of performance objectives.

A third approach exists in the actuarial literature viz control through the choice of valuation methods and assumptions for the scheme assets and liabilities. This area is not discussed for reasons of space: interested readers are referred to Dyson and Exley (1995) for a discussion of smoothed asset valuation methods, and Wise (1987a, b) and Sherris (1992) for a discussion of portfolio selection and matching.

4. **ADJUSTMENT OF CONTRIBUTIONS**

4.1 **DISCRETE TIME MODELS**

At any discrete time \( t \) (for integer values \( t=0, 1, 2 \ldots \)) a valuation is carried out to estimate \( C(t) \) and \( F(t) \), based only on the scheme membership at time \( t \).

In the mathematical discussion, we initially make the following assumptions. Some are relaxed in the later discussion (see section 4.7).

1. All actuarial assumptions are consistently borne out by experience, except for investment returns.

2. The population is stationary from the start. (We could alternatively assume that the population is growing at a fixed, deterministic rate i.e. that the population is stable in the sense of Keyfitz (1985)).

3. Inflation on salaries at a deterministic rate is incorporated by considering interest rates that are "real" relative to salaries. In parallel we assume that benefits in payment increase at the same rate as salaries. We therefore consider variables to be in real terms. For simplicity, each active member's annual salary is set at 1 unit at entry. (It is straightforward to incorporate a fixed promotional salary scale simply through a change of notation).

4. The interest rate assumption for valuation purposes is fixed, i.e.
5. The "real" interest rate earned on the fund during the period, \( t \), is \( i(t+1) \). The corresponding "real" force of interest is assumed here to be constant over the interval \( [t, t+1] \) and is written as \( \delta(t+1) \). Thus, \( 1+i(t+1) = \exp{\delta(t+1)} \). \( i(t) \) is defined in a manner consistent with the definition of \( F(t) \).

6. We define \( E[1+i(t)] = E[\exp{\delta(t)}] = 1+i \). We assume that \( i = i(t) \), where \( i(t) \) is the valuation rate of interest. This means that the valuation rate is correct "on average". This assumption is not essential mathematically but it is in agreement with classical ideas on pension fund valuation. We note that the implication is that present value calculations will use an annual discounting factor \( v = [1+F(i(t))]^{-1} \) rather than the theoretically more sound \( v = E[1+i(t)]^{-1} \): Buhlmann (1992). However, we seek to model the approach adopted in current actuarial practice. The effect of \( i = i(t) \), has been explored by Dufresne (1986), Cairns (1994) and Cairns and Parker (1997).

7. It is assumed that the contribution income and benefit outgo occur at the start of each period (or scheme year).

8. The initial value of the fund (at time zero) is known, i.e. Prob \( F(0)=F_0 \) for some \( F_0 \).

9. Valuations are carried out at annual intervals.

Assumptions 1, 2, 3 and 4, imply that the following parameters are constant with respect to time, \( t \) (after rescaling to allow for growth in line with salary inflation):

- NC the total normal contribution
- AL the total actuarial liability
- \( B \) the overall benefit outgo (per unit of time)
- \( \tilde{S} \) the total pensionable payroll
- PVB the present value of future benefits (for active members and pensioners)
- PVS the present value of future pensionable earnings.

Further, assumptions 1, 2, 4, 7 and 9 imply that the following equation of equilibrium holds:

\[
AL = (1+i)(AL + NC - B) \text{ or equivalently } B = dAL + NC
\tag{6}
\]

where \( d = [1+i]^{-1} \), the compound interest discount rate.

This equation of equilibrium can be found in Trowbridge (1952) and Bowers et al (1976).
4.2 INDEPENDENT AND IDENTICALLY DISTRIBUTED i(t)
INVESTMENT RETURNS

As a first model, we assume that the earned real rates of investment return, \( i(t) \) for \( t \geq 1 \), are independent and identically distributed random variables, with \( i(t) > 1 \) with probability 1, and with \( E(i(t)) = i \), and \( \text{Var}(i(t)) = \sigma^2 < \infty \).

Dufresne (1988, 1989a) has described in detail the properties of
- individual funding methods: spread method for ADJ(t)
- aggregate funding methods
- individual funding methods: amortization of losses method for ADJ(t).

For the spread method,
\[
C(t) = NC + k(\text{AL} - F(t))
\]
and
\[
F(t+1) = [1+i(t+1)][F(t)+C(t)-B]
\]

Equation (7) includes a negative feedback component, whereby the current status, \( F(t) \), is compared with a target \( \text{AL} \) and corrective action is taken to deal with any discrepancy.

Then, Dufresne [1988] shows that
\[
E F(t) = q^t F_0 + r(1-q^t)/(1-q)
\]
where \( q = (1+i)(1-k) \) and \( r = (1+i)(\text{NC}+k\cdot \text{AL}-B)/(1-q)\cdot\text{AL} \) after some algebraic simplification using (6),

and
\[
\text{Var } F(t) = b \sum_{i=1}^{\infty} a^i (E F(i))^2
\]

where \( a = q^2 (1+b) \) and \( b = \sigma^2 (1+i)^2 \).

Then, \( q = \frac{\bar{a}_{M-1}}{\bar{a}_M} \) so if \( M > 1, 0 < q < 1 \) and the following limits exist

\[
\lim_{t \to \infty} E F(t) = \frac{r}{1-q} = \text{AL, using (6)}
\]
and
\[
\lim_{t \to \infty} E C(t) = NC, \text{ using (7)}
\]
If \( a < 1 \), then Dufresne (1988) shows that

\[
\lim_{t \to \infty} \text{Var} \, F(t) = \frac{b AL^2}{(1 - a)}
\]

and

\[
\lim_{t \to \infty} \text{Var} \, C(t) = \frac{bk^2 AL^2}{(1 - a)}
\]

If \( a \geq 1 \), then both of these limiting variances would be infinite. The restriction that \( a < 1 \) implicitly places a restriction on the choice of \( M \) viz

\[
a < 1 \text{ is equivalent to } \bar{a}_n < \frac{1}{1 - f} \text{ where } f = \sqrt[1 + b]{1 + b}.
\]

This is equivalent to \( M < M_\infty = \frac{1}{\delta} \ln \left( \frac{(1 + i)(1 + b)^{\delta} - 1}{(1 + b)^{\delta} - 1} \right) \) and provides a restriction on the feasible values of \( M \) for convergence. Analysis shows that \( M_\infty \) decreases as \( i \) and \( \sigma \) each increase.

Dufresne (1988) also considers expressions for the covariances of \( F \) and \( C \) in the limit and deals separately with the special case \( M = 1 \).

For the aggregate funding methods, equation (8) holds with

\[
C(t) = (\text{PVB} - F(t)) S/PVS
\]

so that

\[
E \, F(t) = q' F_0 + r' (1-q')/(1-q')
\]

where

\[
q' = (1+i)(1-S/PVS), \quad r' = (1+i) (S/PVB/PVS-B).
\]

Then \( 0 < q' < 1 \) and we note the similarity between equations (9) and (13) and the definitions of \( q \) and \( q' \). Indeed, by defining \( N \) such that

\[
\bar{a}_N = \frac{PVS}{S},
\]

we can regard \( q \) and \( q' \) as being of the same form. Hence similar results to (11) and (12) apply in this case.

For the amortization of losses method, Dufresne [1989a] shows that

\[
I(t) = (i(t-1)) (\text{UL}(t-1) - \text{ADJ}(t-1) - (1+i) AL)
\]

and, using (5) and (6), we obtain a difference equation for \( I(t) \) viz

\[
I(t) = (i(t-1)) (\sum_{j=1}^{r(t)} l(t-j) - (1+i)^{-1} AL)
\]

(14)
where the coefficients $b_i$ are defined by $\beta_i = \frac{\tilde{a}_{m-i}}{\tilde{a}_{m}}$.

Then $E(l(t)) = 0$ for all $t \geq 1$ and

$\text{Cov}(l(s), l(t+1)) = 0$ for $1 \leq s \leq t$ \hfill (15)

so that the $C(t)$ form an uncorrelated sequence.

If $\sigma^2 \sum_{i=1}^{m-1} \beta_i^2 < 1$, then Dufresne (1989a) shows that

$$
\lim_{t \to \infty} \text{Var}(l(t)) = \frac{\sigma^2 (1 + l)^2 \lambda L^2}{1 - \sigma^2 \sum_{i=1}^{m-1} \beta_i^2} = x, \text{ say}
$$

$$
\lim_{t \to \infty} \text{Var}(C(t)) = \frac{mx}{\tilde{a}_{m}^2}
$$

$$
\lim_{t \to \infty} \text{Var}(F(t)) = \frac{x}{\tilde{a}_{m}^2} \sum_{j=0}^{m-1} (\tilde{a}_{m-j})^2
$$

Further analysis shows that this inequality is equivalent to requiring $M$ to be less than some upper limit, $M^u$, say; see section 4.6 for further discussion.

4.3 **Autoregressive Rates of Investment Return**

In order to investigate the effects of autoregressive models for the earned real rate of return, we follow the suggestion of Panjer and Bellhouse (1980) and consider the corresponding force of interest and assume that it is constant over the interval of time $(t, t+1)$.

4.3.1 **First Order Autoregressive Models**

Now it is assumed that the (earned real) force of interest is given by the following stationary (unconditional) autoregressive process in discrete time of order 1 (AR(1)):

$$
\delta(t) = \theta + \varphi [\delta(t-1) - \theta] + \epsilon(t)
$$

where $\epsilon(t)$ for $t=1, 2, ..., \text{ are independent and identically distributed normal random variables each with mean 0 and variance } \gamma^2$.

This model suggests that interest rates earned in any year depend upon interest rates earned in the previous year and some constant level. Box and Jenkins (1976) have shown that, under the model represented by equation (17),
\[ E[\delta(t)] = 0 \]
\[ \text{Var } [\delta(t)] = \frac{\gamma^2}{1-\phi^2} = \nu^2, \text{ say} \]
\[ \text{Cov } [\delta(t), \delta(s)] = \frac{\gamma^2}{1-\phi^2} \phi^{t-s} = \gamma(t, s), \text{ say}. \]

The condition for this process to be stationary is that \(|\phi| < 1\).

It then follows that \( E(\exp \delta(t)) = \exp (\theta + \frac{1}{2} \nu^2) = 1+i \) and
\[ \text{Var } \{ \exp \delta(t) \} = \exp (2\theta + \nu^2). \{ \exp (\nu^2) - 1 \}. \]

We first consider individual funding methods and the spread method for choosing \( \text{ADJ}(t) \). It is convenient to re-parametrise equation (8) as
\[ F(t+1) = (1+i(t+1)) (QF(t)+R) \]
where \( Q = 1-k \) and \( R = NC-B+k.\text{AL}=AL(k-d) \).

Haberman (1994) then shows that
\[ E F(t) = F_0 Q^t c^t e^{\nu^2 t} + \frac{R}{Q} \sum_{t=2}^{1} Q^t c^t e^{\nu^2 t} e^z \]
\[ \text{where } z = \nu^2 \{ 1-\phi \}^2 \text{ and } c = \exp \left( \theta + \frac{1}{2} \nu^2 \left( \frac{1+\phi}{1-\phi} \right) \right). \]
(18)

If \( Qc < 1 \) then \( \lim_{t \to \infty} E F(t) \) exists and the following approximation to the limit is derived by Haberman (1994):
\[ \lim_{t \to \infty} E F(t) = \frac{R c}{1-Qc} e^{z} \]
(19)

Upper and lower bounds for the limit have been derived by Cairns and Parker (1997).

For convergence of (18), we require that \( Qc < 1 \). This is equivalent to requiring that
\[ M < M_1 = \frac{1}{\delta} \ln \left( \frac{c-1}{\nu c-1} \right) \]
For given \( \phi \), \( \nu \) and \( i \), there is thus a maximum value of \( M \) for which convergence holds. This provides an important restriction on the feasible values of the spread period, \( M \). Analysis shows that \( M_1 \) decreases with increasing \( i, \phi \) and \( \nu \).

Similarly, we can obtain expressions for \( E(F(t)) \) for finite \( t \) and in the limit as \( t \to \infty \).

It is possible also to consider second moments. Thus, Haberman (1994) shows that

\[
E(F(t^2)) = \frac{2R^2}{Q^2} \sum_{r=0}^{i} \sum_{s=0}^{\infty} Q^{r+s} \exp((t-s)\phi + (t-r)\nu + \nu^2) H(t,r,s)
\]

\[
+ \frac{R^2}{Q^2} \sum_{r=0}^{i} \sum_{s=0}^{\infty} Q^{2(r+s)} \exp(2(t-s)\phi + \nu^2) H(t,s,s) \tag{20}
\]

where \( H(t,r,s) = \frac{(1+\phi)}{1-\phi} \left( t - s + 3(t-r) \right) - \frac{\phi(3 - 2\phi^{-r} - 2\phi^{-s} + \phi^{-r+s})}{(1-\phi)^2} \). \tag{21}

Then, \( \lim_{t \to \infty} E(F(t^2)) \) exists if \( Q < 1 \) and \( Q^2 c \nu < 1 \), and we can show that:

\[
\lim_{t \to \infty} E(F(t^2)) = e^{\nu^2} 2R^2 Qc^2 w \left( 1 - Qc(1 - Q^2 c \nu) \right) + e^{\nu^2} \frac{R^2 c \nu}{(1 - Q^2 c \nu)} \tag{22}
\]

where \( w = \exp\left( \phi + \frac{1}{2} \frac{(1+\phi)}{(1-\phi)} \nu^2 \right) \).

Cairns and Parker (1997) have also obtained upper and lower bounds for this limit.

The requirement that \( Q^2 c \nu < 1 \) is more stringent than \( Q < 1 \) since \( w > c \). This requirement is equivalent to

\[
M < M_2 = \frac{1}{\delta} \ln \left( \frac{\sqrt{c} \nu - 1}{\sqrt{c} \nu \sqrt{c} \nu - 1} \right).
\]

Then \( M_2 < M_1 \) and, as before, further analysis shows that \( M_2 \) decreases with increasing \( i, \phi \) and \( \nu \).

For aggregate funding methods similar results can be obtained.
For individual funding methods with the amortization of losses method for choosing $\text{ADJ}(t)$, the discussion is complicated because of the presence of non-linear effects in the resulting equations. It is convenient here to model $i(t)$, rather than $\delta(t)$, as a stationary AR(1) process:

$$i(t) = i + \varphi (i(t-1)-i) + e(t)$$  \hspace{1cm} (23)

where we repeat that $E(i(t)) = i$, $|\varphi|<1$ and $\{e(t)\}$ is a sequence of independent and identically distributed normal random variables with mean 0 and variance $\sigma^2$.

Gerrard and Haberman (1996) demonstrate how equations (14) and (23) can be combined through the use of generating functions to discuss the behaviour of $E(i(t))$ for finite $t$ and, in the limit, as $t\to\infty$. Some progress is also made with $E(i(t)^2)$, and hence with $\text{Var}(i(t))$.

### 4.3.2 Second Order Autoregressive Models

Haberman (1994) discusses briefly the more complicated case of stationary second order autoregressive models for individual funding methods with the spread method. In this case, equation (17) is replaced by

$$\delta(t) = \theta + \varphi_1(\delta(t-1)-\theta) + \varphi_2(\delta(t-2)-\theta) + e(t)$$  \hspace{1cm} (24)

where $e(t)$ for $t=1,2, \ldots$ are independent, identically distributed normal random variables, each with mean 0 and $\gamma^2$. In parallel to the results of section 4.3.1, we quote Box and Jenkins (1976) who have shown that, for the above model,

$$E(\delta(t)) = \theta$$

$$\text{Var}(\delta(t)) = \left(\frac{1 - \varphi_2}{1 + \varphi_2}\right) \left(\frac{\gamma^2}{(1 - \varphi_1)(1 - \varphi_2)}\right) = \nu^2$$, say

$$\text{Cov}(\delta(t), \delta(s)) = \nu^2 \{\lambda \psi_1|t-s| + (1 - \lambda) \psi_2|t-s|\}

where $$ \lambda = \frac{\psi_1(1 - \psi_2)}{(\psi_1 - \psi_2)(1 + \psi_1 \psi_2)} \text{ and } (1 - \lambda) = \frac{\psi_2(\psi_2 - 1)}{(\psi_1 - \psi_2)(1 + \psi_1 \psi_2)}$$

and $\psi_1^{-1}$ and $\psi_2^{-1}$ are the solutions of the characteristic equation: $1 - \varphi_1t - \varphi_2t^2 = 0$.  

14
For stationarity, we now require

\[
\begin{align*}
\phi_1 + \phi_2 &< 1 \\
\phi_2 - \phi_1 &< 1 \\
-1 &< \phi_2 < 1
\end{align*}
\]

It is then possible to construct equations for the moments of \( F(t) \) and \( C(t) \) in finite time and in the limit, as \( t \to \infty \), which correspond in format to those for the AR(1) case. The details are not pursued here.

### 4.3.3 Conditional Autoregressive Models

A disadvantage of the models used in sections 4.3.1 and 4.3.2 is that the uncertainty about \( \delta(t) \) is independent of \( t \), i.e. \( \text{Var} [ \delta(t) ] = \sigma^2 \), a constant. In reality, we would expect this level of uncertainty to depend on \( t \). A model which allows for this feature would be the conditional AR(1) or AR(2) model considered by Bellhouse and Panjer (1981). In this case, it is assumed that the returns of the past years (and the corresponding forces of interest) are known, as initial conditions. We would then expect that the asymptotic results derived in sections 4.3.1 and 4.3.2 would hold also for the conditional processes since, as \( t \to \infty \), the initial values \( \delta_0 \) and \( \delta_1 \) would become increasingly insignificant.

### 4.4 Moving Average Rates of Investment Return

Duforesne (1990) has introduced the use of the moving average process to represent the rate of investment return in pension funding models. This has been explored further by Haberman and Wong (1997).

#### 4.3.1 First Order Moving Average Models

We assume that the (earned real) force of interest is given by the following stationary (unconditional) moving average process in discrete time of order 1

\[
\delta(t) = \theta + e(t) - \varphi e(t-1)
\]

(25)

where \( e(t) \) for \( t = 1, 2 \ldots \) are independent and identically distributed normal random variables each with mean 0 and variance \( \gamma^2 \). \( \varphi \) is the moving average parameter.

From the model, we can see that the interest rates earned in each year depend upon interest rates earned both in same year and previous year. Under the above model (Box and Jenkins 1976), we can show that:
\[ E[\delta(t)] = 0 \]
\[ \text{Var}[\delta(t)] = (1 + \varphi_1^2)\gamma^2 = \sigma^2, \text{ say} \]
\[ \text{Cov}[\delta(t), \delta(s)] = -\varphi_1 \gamma^2, \quad |t-s| = 1 \]
\[ = 0 \quad |t-s| > 1. \]

A moving average process is second-order stationary, regardless of the value of \( \varphi \). The invertibility condition for this model is \(-1 < \varphi < 1\).

It then follows from the properties of the lognormal distribution that:

\[ E[\exp(\delta(t))] = \exp(\sigma + \frac{1}{2}\sigma^2) \]
\[ \text{Var}[\exp(\delta(t))] = \exp(2\sigma + \sigma^2) \{ \exp(\sigma^2) - 1 \}. \]

Haberman and Wong (1997) then obtain explicit expressions for \( E F(t) \), \( \text{Var} F(t) \), \( E C(t) \) and \( \text{Var} C(t) \) for finite \( t \) and in the limit as \( t \to \infty \). For example, it is shown that if \( Qx > 1 \), \( \lim t \to \infty E F(t) \) exists

where \( x = \exp(\theta + \frac{1}{2}\sigma^2 - \varphi \gamma^2) \)

and \( \lim t \to \infty E F(t) = \frac{A L(k-d)xe^{\kappa t}}{1 - Qx} \quad (26) \)

4.4.2 Second Order Moving Average Models.

Now, it is assumed that the (earned real) force of interest is given by the following stationary (unconditional) moving average process in discrete time of order 2:

\[ \delta(t) = \varphi_1 \epsilon(t-1) - \varphi_2 \epsilon(t-2) \quad (27) \]

where \( \epsilon(t) \) for \( t = 1, 2, \ldots \) are independent, identically distributed normal random variables, with mean 0 and variance \( \gamma^2 \).

We obtain the following results (similar to the MA[1] case) from Box and Jenkins (1976).

\[ E(\delta(t)) = 0 \]
\[ \text{Var}(\delta(t)) = (1 + \varphi_1^2 + \varphi_2^2)\gamma^2 = \sigma^2 \]
\[ \text{Cov}(\delta(t), \delta(s)) = (\varphi_1 \varphi_2 - \varphi_1)\gamma^2 \quad |t-s| = 1 \]
\[ = -\varphi_2 \gamma^2 \quad |t-s| = 2 \]
\[ = 0 \quad |t-s| > 2. \]
It is invertible only if the roots of the characteristic equation

\[ 1 - \phi_1 B - \phi_2 B^2 = 0 \]

lies outside the unit circle, that is

\begin{align*}
\phi_1 + \phi_2 &< +1 \\
\phi_2 - \phi_1 &< +1 \\
-1 &< \phi_2 < +1.
\end{align*}

Haberman and Wong (1997) then construct equations for the moments of \( F(t) \) and \( C(t) \) in finite time and in the limit, as \( t \to \infty \), which correspond to those for the MA(1) case. The details are not pursued here.

In the case of moving average models, explicit closed form results are obtained for the limiting values whereas for the corresponding autoregressive models the results are approximate. This arises because the covariance function for the \( s(t) \) process has a simpler form in the moving average case.

4.5. **OPTIMAL SPREAD PERIOD**

In this section we focus on individual funding methods with the spread method, and we shall consider the existence of an "optimal" range of spread periods, \( M \).

In this section, we shall consider the relationship between \( \text{Var} \ F(t) \) and \( \text{Var} \ C(t) \) as \( M \) (or \( k \)) varies, with \( t \) fixed. Rather than take a particular finite \( t \), we shall consider the limiting variances at \( t \to \infty \) and indeed we shall consider these variances relative to the corresponding expectations (i.e. the coefficient of variation). Further, as recognised by Cairns (1994), both \( \text{Var} \ F(t) \) and \( \text{Var} \ C(t) \) are proportional to \( AL^2 \) so that a more secure funding method, with higher \( AL \), would lead to greater variability. The converse suggest that variances could be reduced by choosing a funding method with a lower value of \( AL \). This problem is by-passed by considering the coefficients of variation. Our consideration of the case where \( t \to \infty \) is justified on the grounds that the results are mathematically tractable. We shall now introduce some new notation.

With \( a<1 \) and \( 2 \leq M < \infty \) (so \( d<k<1 \)), we define

\[ \alpha(k) = \frac{\text{Var} F(\omega)}{(\text{EF}(\omega))^2} \quad \text{and} \quad \beta(k) = \frac{\text{Var} C(\omega)}{(\text{EC}(\omega))^2} \]

(28)
and we regard \( \alpha \) and \( \beta \) as functions of \( k \). We could equivalently regard them as functions of \( M \), given the 1-1 correspondence between \( k \) and \( M \). However, it is more convenient to consider \( \alpha(k) \) and \( \beta(k) \).

### 4.5.1 Independent and Identically Distributed \( i(t) \)

For the case of IID \( i(t) \), Dufresne (1988) has considered in detail the trade off between \( \text{Var} F(t) \) and \( \text{Var} C(t) \) in the limit as \( t \rightarrow \infty \), as represented by \( \alpha(k) \) and \( \beta(k) \), and for finite \( t \) under certain conditions. Thus, from (11) and (12), we have that

\[
\alpha(k) = \frac{b}{1 - y(1 - k)^2}
\]

and

\[
\beta(k) = \frac{AL^2}{NC_y}, \quad \frac{bk^2}{1 - y(1 - k)^2}
\]

where \( y = (1 + i)^2 \). Assuming that \( y > 1 \), Dufresne shows that

\[
\frac{d}{dk} \alpha(k) < 0
\]

and

\[
\frac{d}{dk} \beta(k) = 0 \quad \text{where} \quad k' = 1 - \frac{1}{y}.
\]

At \( k = k' \), \( \beta(k) \) takes a minimum value. The value of the spread period corresponding to \( k' \) will be denoted by \( M^* \).

Formally, if \( y > 1 \), then both \( \text{Var} F(\infty) \) and \( \text{Var} C(\infty) \) become infinite for some finite \( M = M^* \) (when \( a \) becomes equal to \( 1 \)) and there exists a value \( M^* \) such that

- for \( M < M^* \), \( \text{Var} F(\infty) \) increases and \( \text{Var} C(\infty) \) decreases with \( M \) increasing
- for \( M > M^* \), both \( \text{Var} F(\infty) \) and \( \text{Var} C(\infty) \) increase with \( M \) increasing.

If \( y = 1 \), \( \text{Var} C(\infty) \rightarrow 0 \) and \( \text{Var} F(\infty) \rightarrow \infty \) as \( M \rightarrow \infty \), although \( \text{Var} F(\infty) \) does stay finite for all \( M \).

If \( y < 1 \), \( \text{Var} C(\infty) \rightarrow 0 \) as \( M \rightarrow \infty \) and \( \text{Var} F(\infty) \) has a finite limit as \( M \rightarrow \infty \).

The particular value of \( M^* \) is determined by

\[
k' = 1 - \frac{1}{y} = \frac{1}{d_M}
\]
\[ i.e. \text{ if } i \neq M' = \frac{1}{\delta} \ln \left( \frac{v_y - 1}{y - 1} \right) \quad (29) \]

and if \( i = 0 \) \( M' = 1 + \frac{1}{\sigma^2} \).

There is thus a trade-off between variability in the fund, represented by \( \sigma \),
and variability in the contribution rate, represented by \( \beta \). This trade-off
takes place but only up to \( M=M' \). Beyond this point, augmenting \( M \)
causes both \( \text{Var F} \) and \( \text{Var C} \) to increase. With the objective of
minimizing variances, any choice of \( M>M' \) should be rejected, for clearly
some \( M<M' \) would reduce both \( \text{Var F} \) and \( \text{Var C} \). If we regard \( M \) as being
a parameter open to the choice of the actuary, then the optimal choices
for \( M \) would lie in the region \( 1 \leq M \leq M' \). Thus, we can describe this region
as an 'optimal' region.

Table 1 provides values of \( M \) as a function of \( i \) and \( \sigma \) (to the nearest
integer). In the UK, it is common to choose \( M \) to correspond to the
average remaining working lifetime of the current membership - with an
average age of membership of 40-45 and a normal retirement age of 65
this would correspond to a choice of \( M \) in the range 20-25. We see from
Table 1, that under particular combinations of \( i \) and \( \sigma \) our model
indicates that this choice is not optimal. If \( i=0.03 \) and \( \sigma=0.20 \) then,
for example, smaller values, namely those in the region \( 1 \leq M \leq 13 \), would be
more satisfactory.

4.5.2 Autoregressive Rates of Return

We shall consider here only the case of stationary AR(1) processes as a
description for \( \delta(t) \). Haberman (1994) has explored the behaviour of the
relative limiting values (as \( t \to \infty \)) of \( \text{Var F(t)} \) and \( \text{Var C(t)} \)
as functions of \( M \) (or equivalently as functions of \( k \)), by analysing the properties of the
results represented by equations (18) and (20) and the corresponding
results for \( C(t) \). The results are reported here, subject to the constraints
that \( Q_c<1 \) and \( Q^c(-c)<1 \) for convergence, and with \( F_i=0 \).

The numerical investigation uses a simple benefit structure and a range of
values for \( i, v, \varphi \) and \( M \).

The detailed calculations indicate the following general features viz

(i) \( \alpha(\varphi, M) \) increases with \( M \) (for fixed \( \varphi \)) and with \( \varphi \) (for fixed \( M \))

(ii) \( \beta(\varphi, M) \) increases with \( \varphi \) (for fixed \( M \)) and decreases with increasing
     \( M \) (for fixed \( \varphi \)) except that for some values of \( \varphi \) (e.g. \( \varphi=0.1 \)) there is
     a minimum at some \( M' \).
When values of $\alpha$ and $\beta$ are plotted for combinations of $i$, $v$ and $\phi$, we find that three distinct patterns emerge, unlike the situation when rates of return are independent, identically distributed random variables (as in section 4.5.1 which corresponds approximately to the case $\phi=0$).

The three patterns in terms of profiles of $\beta(M) \nu \alpha(M)$ are:

**TYPE A:** the profile has a minimum at $M'$ so $1 \leq M \leq M'$ is "optimal".

**TYPE B:** the profile is monotonically decreasing so there is no "optimal" region. The choice of $M$ will depend on the characteristics of the scheme sponsor and their particular attitude to the trade-off between variability in $F$ and in $C$.

**TYPE C:** the profile is monotonically increasing so $M=1$ is "optimal".

Illustrative of the results in Table 2 which corresponds to $\phi=-0.3$, and shows the classification of $\alpha-\beta$ profiles and, where appropriate, the optimal regions for $M$. Tables 3-5 similarly refer to $\phi=-0.1$, 0.1, and 0.3. Because we are interested only in general features, no attempts have been made at this stage to estimate more precisely the turning points in the $\alpha-\beta$ graphs (using, for example, numerical interpolation methods).

From Tables 3 and 4, corresponding to $\phi=\pm0.1$, we note that the implied optimal values of $M$ are consistent with those shown in Table 1 (from Dufresne (1988)) which would correspond approximately to the case $\phi=0$.

The pattern of optimal $M$ values across Tables 2-5 mirrors that for the IID case. In general, the optimal region decreases as $i$ increases (for fixed $v$ and $\phi$) and as $v$ increases (for fixed $i$ and $\phi$). Also, the optimal region decreases as $\phi$ increases from -1 to +1 (for fixed $i$ and $\phi$); thus, an increase in the autoregressive parameter $\phi$ appears to have a similar effect on the optimal region as an increase in the variance parameter $v$.

### 4.5.3 Moving Average Rates of Return

In a similar vein, Haberman and Wong (1997) have investigated numerically the behaviour of $\alpha(k)$ and $\beta(k)$ when the earned real rate of investment return has followed a first order moving average process.

A selection of the results are presented in Tables 6 - 8.

The pattern of optimal $M$ values mirrors the patterns for both the IID and AR(1) cases. The optimal spread period decreases with an increase in the "real" rate of interest (fixed $i$ and $\phi$) and an increase in $v$ (fixed $i$ and $\phi$). These trends are the same in all three models. For the MA(1) case, the optimal spread period decreases as $\phi$ becomes more negative, which is in
contrast to the AR(1) case where the spread period increases as $\phi$ becomes more negative, we should note that the parameter $\phi$ plays a different role in these two types of model (Box and Jenkins 1976).

The invertibility property of MA processes means that the MA(1) process can be represented by an infinite series of AR processes (and vice versa for the AR(1) process). Hence, it is reasonable that similar results should have been obtained in these two cases and that the relationship of the optimal spread period to $\phi$ should be different in the manner described above.

4.5.4 Optimal Spread Periods and Finite Time

A related question would be to investigate whether an optimal spread period range exists at finite values of $t$, rather than only in the limit, for the IID model. For the set of $k$ values in the range $[k^*1]$ to be truly "optimal", it would be necessary to show that the trade-off is maintained over this range of $k$ for all $t>0$. Let $k^*(t)$ be the value of $k$ at which the trade-off is discontinued at time $t$. Then, Owadally and Haberman (1995) prove that, for the case $F_0 = AL$, the trade-off is maintained in finite time between $\text{Var } C(t)$ and $\text{Var } F(t)$ in the range described as optimal by Dufresne (1988) and provide the conditions for the existence of $k^*(t)$. It is then straightforward to transform these values of $k$ into corresponding values of $M$.

4.6 COMPARISON OF SPREAD AND AMORTIZATION OF LOSSES METHODS

In section 2, we described the two commonly used methods for eliminating unfunded liabilities: the spread method (see equation (3)) and the amortization of losses method (see equations (4) and (5)).

In the case of IID investment returns, it is possible to compare the properties of these two methods and obtain definitive results: for fuller details see Haberman and Owadally (1997). For example, it can be shown that:

a) For equal amortization and spread periods

$$\text{Var } F(x)_a < \text{Var } F(x)_a \quad M>1$$
$$\text{Var } F(x)_a = \text{Var } F(x)_a \quad M=1$$

The $\text{Var } F(x)_a$ v. $M$ curve lies under the corresponding curve for the spread method, except that they coincide at $M=1$.

b) There is no turning point in the $\text{Var } F(x)_a$ v $M$ curve. It increases monotonically.
c) For the same $i$ and $\sigma$, the maximum spread period allowable for stability (i.e. finite variances) is less than the maximum allowable amortization period:

$$M^*_s < M^*_a$$, say.

d) The Var $C(\sigma)_a$ v. $M$ curve has only one turning point, which is a minimum point, at which the Var $C(\sigma)_a$ v. $M$ curve intersects it.

e) The Var $C(\sigma)_a$ v Var $F(\sigma)_a$ curve has a minimum point. There exists a non-optimal or inadmissible range of amortization periods $[M^*_s, \infty)$. The existence of an optimal range was not uncovered by the numerical tests of Dufresne (1989a).

f) The amortization of losses method curve Var $C(\sigma)_a$ v Var $F(\sigma)_a$ curve lies above the Spread method curve Var $C(\sigma)_a$ v Var $F(\sigma)_a$ except at $M=1$ where they coincide.

These results are summarized in Figures 1 and 2. The implications are that, for equal spread and amortization periods, the Amortization of Losses method achieves greater fund security than the Spread Method. However, according to the objective of minimizing both the ultimate variances of fund and contribution rate levels, the above results indicate that the Spread method is "more optimal" than the Amortization of Losses method.

This is not a surprising result, since the latter uses information delayed by up to $M$ years.

Cairns (1994) has provided numerical illustrations of c) and f).

4.7 MODEL ENHANCEMENTS

4.7.1 Other Investment Return Models

More sophisticated investment return models could be introduced, along the lines of the autoregressive integrated moving average (ARIMA) process family of Box and Jenkins. These ideas have not been pursued analytically in the pension funding case but have been explored in the field of life insurance mathematics by Giaccotto (1986) and Dhaene (1989, 1992) who have devised recursive methods for the computation of the movements of $\delta(t)$ and applied them to the calculation of the moments of present values of life insurance annuity and insurance contracts.

Simulation based studies do exist in the pension fund literature. Much attention has focused on the stochastic investment model devised by Wilkie for representing different financial time series. This uses ARIMA
methodologies and is of a cascade type: see Wilkie (1995) for the latest version. For example, Loades (1988) and Wright (1997) have investigated assessing the security level of a pension scheme, Bilodeau (1995) has investigated in some detail the variability of assets and contributions albeit in a Canadian context and, in a parallel piece of research, Haberman and Smith (1997) have used the Wilkie investment model to explore the existence of optimal ranges for the spread period (and conclude that the features discussed in section 4.5 persist when these more complex models are introduced for representing the underlying investment returns). Haberman and Smith also explore the suggestion of Winklevoss (1993) and investigate the effect of a shorter spread period for deficiencies than for surpluses.

4.7.2 Random New Entrants

One way of relaxing the assumptions of section 4.1 would be to replace the stationary population assumption by the more realistic one that the number of new entrants into the scheme is subjected to random fluctuations. Mandl and Mazurova (1996) model the number of entrants and the rate of investment return by independent random sequences which are stationary (in a wide sense). They then use harmonic analysis of the random sequences and the introduction of frequency transfer functions to derive formulae for the variances of the fund levels, contributions and discounted cash flows. Numerical illustrations are provided which confirm the results of Haberman (1994).

4.7.3. Frequency of Valuations

A further control variable available to the actuary is the frequency with which valuations are performed. In earlier sections, we have assumed that valuations are annual. Here, we shall briefly consider the case of valuations every 3 years: as noted in section 1, triennial valuations are common in the UK because of legislative and cost considerations. If we consider the case of individual funding methods with the spread method choice for ADJ[t] and independent, identically distributed i(t), it is then straightforward to modify equations (3) and (7) to fit in with this new environment. Equivalent results to (11) and (12) can be obtained and the existence of an optimal range of spread periods can be demonstrated, subject to M≥3 and y^2 = (1+i)^2 + σ^2 > 1. The results are easily generalised to valuations every n years. For further details readers are referred to Haberman (1993a).
4.7.4 Delay in Fixing Contributions

We now introduce a new parameter into the formula for fixing the contribution rate, \( C(t) \). We consider only individual funding methods, with the spread method choice for \( \text{ADJ}(t) \), and aggregate funding methods. We allow for a time delay in the pension scheme’s funding process and use the fund level at time \( t-p \) in order to calculate \( C(t) \). So we would use

\[
\begin{align*}
C(t) &= NC + k(AL - F(t-p)) \\
C(t) &= (PVB - F(t-p))S/PVS
\end{align*}
\] (30) (31)

to replace equations (7) and (7a) respectively where \( p \) is a non negative integer. The delay \( p \) may arise because of the time taken to prepare the financial accounts or to assemble the valuation data and to complete the actuarial valuation exercise. Alternatively, we can think of the parameter \( p \) (like the spread period \( M \)) as being a control variable at the disposal of the actuary and which can be used to control the behaviour over time of \( C(t) \) or \( F(t) \).

In the case of independent and identically distributed \( i(t) \), the recurrence relation for \( E F(t) \) then becomes

\[
E F(t+1) = (1+i)(E F(t) - (1+i)k E F(t-p)) + r.
\] (32)

The solutions for \( EF(t), EC(t), \text{Var} F(t) \) and \( \text{Var} C(t) \) and their limiting values are derived by Haberman (1992) for the case \( p=1 \) and Zimbidis and Haberman (1993) for the case \( p \geq 2 \). They demonstrate that, if \( M>1 \), the limiting variance of \( F(t) \) increases as \( p \) increases. These results are intuitively reasonable, given our understanding of the entropy of systems. When we introduce a time delay, which means that we have lost (or do not have available) some information for the fund between times \( t-p \) and \( t \), we should expect the variance (or, in other words, the entropy) of the fund level and contributions to be greater. These results confirm the findings of Balzer and Benjamin (1980) who report that the longer are the delays in information in a system, the more persistent are the resulting oscillations in that system. Zimbidis and Haberman (1993) also report on the conditions for oscillations to exist in the first two moments of \( F(t) \) and \( C(t) \) as \( p \) varies.

In the case of \( \text{Ar}(1) \) \( i(t) \), Haberman (1993b) investigates the effect of a one year time delay, \( p=1 \).
4.7.5 Distribution of Fund Levels

Dufresne (1990) considers the general problem of determining the distributions of discounted values and specifically applies this to pension funding. He considers the distribution of

\[ Z = \sum_{k=1}^{n} \left( \prod_{j=1}^{k} V_j \right) C_k \]  

(33)

where the sequences \( \{C_k\} \) represents cash flows, the sequence \( \{V_k\} \) represents 1 year discount factors, the distributions of the sequences we known with each being IID and the \( C_k \) and \( V_k \) being mutually independent. \( Z \) then represents the present value of a perpetuity comprising the cash flows, \( C_k \). Dufresne (1990) proves that if \( E \log V_1 < 0 \) and \( E \log |C| < \infty \), then \( Z \) converges absolutely with probability 1, using a result of Vervaat (1979). Following Brandt (1986), Dufresne also shows that the same result holds when the independence assumptions are replaced with ergodicity (so that the law of large numbers applies to each of the sequences \( \{V_k\} \) and \( \{C_k\} \)).

Expressing \( Z \) in continuous time with \( C_k=1 \) and the log \( V_k \) forming a Brownian motion process, Dufresne (1990) proves that \( Z^\dagger \) is distributed as an inverse gamma random variable.

Cairns and Parker (1997) use a different approach. They follow the approach of Parker (1994) and obtain a recursive method for obtaining the joint density of \( F(t) \) and \( \delta(t) \) in the IID and AR(1) cases. The approach is based on the equations:

\[ f_{H|t}(x) = \int_{-\infty}^{\infty} f_{H_{t+1}|t}(x, y) \, dy \]  

(34)

\[ f_{H_{t+1}|t}(x, y) = \int_{-\infty}^{\infty} f_{H_{t+1}|\delta(t-1) = z} \cdot f_{H_{t-1}|\delta(t-1)} \left( \frac{xc - \theta}{1 - k}, z \right) \, dz \]  

(35)

where \( \theta = (k-d)AL = r (1+i)^{t+1} \)

and the starting value \( f_{H_{T+1}|0}(x, y) = f_{H_{T}|\delta(0)} \) if \( x = e^{\nu/[1 - k]F_0 + \theta] \)

is used for the iteration. This approach can be used for processes for which \( f_{H_{T}|\delta(t-1) = z} \) is known e.g. Gaussian processes.
4.7.6 Continuous Time Models

The model can also be expressed in continuous time terms. Thus, Dufresne (1986, 1989b) considers the situation where contributions and benefits are paid N times per year and considers the limit as N→∞ under certain conditions. Dufresne (1986, 1989b) proves that the fund level then converges in the weak sense so that equation (8)

\[ F(t+1) = (1+i(t+1)) \left( F(t) + C(t) - B \right) \]  \hspace{1cm} (8)

becomes

\[ dF(t) = \left( \gamma F(t) + C(t) - B \right) dt + \sigma F(t) dW(t) \]  \hspace{1cm} (36)

where

\[ C(t) = NC + \bar{k} (AL - F(t)) \]  \hspace{1cm} (37)

and \( \bar{k} \) depends on a continuous time annuity value: \( \bar{k}^{-1} = \int_0^X e^{-s} ds \),

\[ \gamma = \log(1 + Ei(t)), \sigma^2 = \log \left( \frac{E[(1 + i(t))^2]}{E(1 + i(t))^2} \right), \] and \( W(t) \) is a standard Wiener process.

Dufresne (1986) is then able to demonstrate that many of the earlier results carry forward into this new framework, including the existence of an optimal range of values for the spread period, M. Dufresne (1990) and Cairns (1996) extends this approach by replacing the geometric Brownian motion component

\[ d\tilde{S}(t) = \gamma dt + \sigma dW(t) \]

by models that allow for non static investment strategies through the consideration of \( d\tilde{S}(t), F(t) \): for example simple rebalancing strategies and the “continuous proportion portfolio insurance strategies” of Black and Jones (1988) and Black and Perold (1992).

5. OPTIMAL CONTROL APPROACH TO PENSION FUNDING

5.1 INTRODUCTION

In this section, we consider a different approach proposed for managing the funding of a pension scheme, which involves the application of control theory. Initially we consider the different risks which confront a defined benefit pension scheme.

Firstly, there is the “contribution rate risk”. Here the sponsor of the scheme, the employer, will be concerned that future investment performance is not such as to expose the pension fund to the risk of significant, unanticipated rises in contribution rate. Traditionally, this risk has been controlled by concentrating on real assets (e.g. equities, property, indeed linked bonds). However, the concern remains about the variability of the levels of the contribution rate. Stability will also be a
feature attractive to the finance manager and the shareholders of the employing/sponsoring company.

Secondly, the trustees, sponsor, members and advising actuary will be concerned that the pension fund can meet its liabilities. This is the "solvency risk".

The choice of future contribution rates can then be considered as a problem involving the minimization simultaneously of these two types of risk: we take a quadratic performance criterion over the chosen time period for optimization. The problem then becomes:

Find the contribution rates \( C_n, C_{n+1}, \ldots, C_{T+1} \) for a finite time span, which minimize the quadratic performance criterion

\[
J_\tau = E \left\{ \sum_{t=1}^{T+1} v^t \left[ \theta(C(t) - CT_t)^2 + (1 - \theta)(F(t) - FT_t)^2 \right] \right\}
\]

(38)

in discrete time.

Alternatively, with an abuse of notation,

\[
J_\tau = E \int_0^\tau e^{-\delta t} \left[ \theta(C(t) - CT_t)^2 + (1 - \theta)(F(t) - FT_t)^2 \right] dt \quad \text{in continuous time.} \quad (50)
\]

(For the deterministic case, we would delete the expectation operator). Here,

\[
\begin{align*}
C(t) & = \text{contribution rate for period \([t, t+1]\) in discrete case or at time } t \text{ in the continuous case} \\
F(t) & = \text{fund level at time } t, \text{ measured in terms of the market value of the underlying assets} \\
CT_t & = \text{contribution target for period at \((t, t+1)\) in the discrete case or at time } t \text{ in the continuous case} \\
FT_t & = \text{Fund } \{F\} \text{ target for period \((t, t+1)\) in the discrete case or at time } t \text{ in the continuous case} \\
v & = (1+i)^{-1}, \text{ } i \text{ is the valuation real rate of interest during the period, and } \delta \text{ is the corresponding force of interest} \\
\theta & = \text{a weighting factor to reflect the relative importance of the contribution rate risk against the solvency risk, } 0 < \theta < 1.
\end{align*}
\]
In this presentation, the first term represents the contribution rate risk and the second term the solvency risk. We next consider the choice of the target values.

One choice appropriate for certain funding methods, would be \( C_t = NC(t) \) and \( F_t = AL(t) \) the normal cost and actuarial liability at time \( t \), which would be appropriate if all actuarial assumptions were realized exactly during the control period.

A second choice would be \( C_t = E(C(t)) \) and \( F_t = E(F(t)) \), so that the stochastic performance criterion would become, in discrete time,

\[
\min \sum_t v^t \left[ \theta \text{Var}(C(t)) + (1 - \theta) \text{Var}(F(t)) \right].
\]

A third choice, adopted by O'Brien (1987), would be to use \( F_t = \eta A(t) \) where \( \eta \) is the funding level (or ratio) and \( A(t) \) is the present value of the future benefits of the active members of the pension scheme at time \( t \).

A number of recent papers have applied the ideas of control theory to pension funding.

Benjamin (1984, 1989) considers the deterministic optimal control of pension funds in discrete time. Benjamin (1984) uses the control approach to analyse the effect of changes in the earned real rate of investment return (the input signal) on the recommended contribution rate (the output signal). Benjamin (1989) defines the optimal set of future contribution rates as being the set which minimizes \( \sum \Delta C(t) \) so that \( C_t \) is effectively \( C(t+1) \) and \( \theta = 1 \), and investigates the effects of a time-varying valuation rate of interest (specified as the arithmetic mean of real investment rates of return earned in a recent period).

O'Brien (1987) considers the stochastic optimal control problem in continuous time, deriving the optimal pension funding controllers in feedback form for active plan members, which are defined as the contribution rates which minimize both the deviations from a target fund (using \( \Delta A(t) \)) and the fluctuations of contribution rates using Bellman's optimality principle. This model does not include a target contribution rate. The benefit outgo is assumed to follow a linear growth function of time. Further, the growth rate in population and salary, and the fund earning rate are considered as stochastic variables which are assumed to be mutually independent.

Vanderbroek (1990) introduces both types of target: the contribution rate target (using a fixed but unknown proportion of the total payroll, \( \alpha S(t) \)) and the fund target (using \( \eta A(t) \) as for O'Brien (1987)). The model deals with the deterministic optimal control of pension funds in continuous time, assuming that the payroll, benefit outgo and the present value of future benefits are each exponential functions of time. The author is
particularly concerned with the application to the case of national social security plans.

Loades (1992) uses a discrete-time deterministic approach, in which he investigates how the contribution rate responds to a periodic oscillation in \( i(t) \), subject to a separate (but not concurrent) change in the valuation methods and a time-varying valuation rate specified by an exponential smoothing mechanism:

\[
i_0(t) = SF. i_0(t-1) + (1 - SF) i(t-1) \text{ for } 0 \leq SF \leq 1.
\]

Fujiki (1994) similarly investigates how to modify effectively the actuarial valuation assumptions to improve the long term stability of contributions under separate (but not concurrent) changes in real investment rates of return, equity dividend growth rates and withdrawal rates in the pension scheme population.

Haberman and Sung (1994) and Sung (1997) draw on and extend this earlier work, using both a deterministic and a stochastic model in discrete time and thus obtain the mathematical form of the optimal contribution rates.

In a related piece of work, Haberman (1997) has investigated the issue of contribution rate risk from the viewpoint of minimizing, as performance criterion,

\[
G(t) = \text{Var} \left[ \sum_{s=t}^{M} C(s) \right]
\]

and considered the existence of optimal choices for the spread period, \( M \), for different initial values of \( F_0 \).

A recent area of research has been the application of stochastic control methods to obtain an optimal asset allocation between a risky asset and a non-risky asset and an optimal contribution policy. A higher proportion of equities in the asset portfolio has the benefit of a lower expected contribution rate but at the price of a higher variability of contribution flows over time. Thus, Boulier et al (1995), 1996) consider as performance criterion

\[
J = \int_0^M e^{-\delta s} \mu(C(s)) \, ds \quad (40)
\]

where \( \mu(C(s)) = \theta C(s)^2 \)

or \( \mu(C(s)) = C(s) + \theta C(s)^2 \)
respectively while Cairns (1997) follows Haberman and Sung (1994) and uses

\[ \mu(C(s)) = \theta (C(s) - CT)^2 + (1-\theta) (F(s) - FT)^2 \]  

(41)

The results so far produced from this line of research unexpectedly show that as the level of surplus in the pension fund increases, the proportion of the fund invested in high-return, high-risk assets (i.e. equities) decreases. It is possible that this unexpected result comes from the formulation of the problem: asset allocation being made to depend on surplus rather than to depend directly on the underlying liabilities and the trade-off between risk and return. Clearly this is an area for future research.

6. COMMENTS AND FURTHER DEVELOPMENTS

Varying levels of inflation and fluctuations in investment returns are problems with which the actuary must contend on an almost daily basis. Unlike mortality and other decrements or movements, for which deterministic and stochastic models are readily available, the movements of these economic factors are more difficult to model. Representation by identically distributed random variables or by simple stationary autoregressive models appear to be very appropriate for this purpose. An objective of this paper has been to show that explicit formulae are available for studying mathematically the variability of contributions and fund levels for a pension scheme. Practical implications for the choice of funding method are then considered as a consequence, and the effect of the choice of control parameter including spread period, valuation frequency and delay in fixing the contribution rate are discussed, as well as a comparison of the two main methods for adjusting contribution rates to eliminate unfunded liabilities.

A number of interesting and potentially useful directions for future research that come from the foregoing review are the following:

- consideration of more realistic rate of investment return processes and models of demographic change, in relation to new entrants or rate of growth.
- consideration of other choices for ADJ(t)
- consideration of the viewpoint of the scheme's sponsoring employer and their aversion to risk
- analysing further the effect of varying the control variables identified and the interactions between them
- further consideration of the introduction of dynamic valuation assumptions using the optimal control approach of section 5.
REFERENCES


Dufresne, D 1986. The Dynamics of Pension Funding, PhD thesis, City University, London, UK.


TABLE 1
M* as a Function of i and σ (IID case)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>-.01</th>
<th>0</th>
<th>.01</th>
<th>.03</th>
<th>.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>-</td>
<td>401</td>
<td>60</td>
<td>23</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>-</td>
<td>101</td>
<td>42</td>
<td>20</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>.15</td>
<td>158</td>
<td>45</td>
<td>28</td>
<td>16</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>.20</td>
<td>41</td>
<td>26</td>
<td>19</td>
<td>13</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>22</td>
<td>17</td>
<td>14</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2
CATEGORY OF α-β PROFILE AND WHERE APPROPRIATE OPTIMAL REGION FOR M, SPREAD PERIOD, φ=0.3 (AR(1) case)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>-.01</th>
<th>.005</th>
<th>.01</th>
<th>.03</th>
<th>.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A(1,30)</td>
<td>A(1,20)</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>B</td>
<td>A(1,400)</td>
<td>A(1,250)</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>.15</td>
<td>B</td>
<td>A(1,250)</td>
<td>A(1,200)</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>.20</td>
<td>B</td>
<td>A(1,200)</td>
<td>A(1,140)</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>B</td>
<td>A(1,150)</td>
<td>A(1,120)</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 3
**CATEGORY OF $\alpha$-$\beta$ PROFILE AND WHERE APPROPRIATE OPTIMAL REGION FOR M, SPREAD PERIOD, $\varphi=0.1$ (AR(1) case)**

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\nu$</th>
<th>.01</th>
<th>.005</th>
<th>.01</th>
<th>.03</th>
<th>.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>B</td>
<td>A(1,130)</td>
<td>A(1,70)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>B</td>
<td>A(1,90)</td>
<td>A(1,60)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td></td>
</tr>
<tr>
<td>.15</td>
<td>B</td>
<td>A(1,55)</td>
<td>A(1,40)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td></td>
</tr>
<tr>
<td>.20</td>
<td>A(1,110)</td>
<td>A(1,35)</td>
<td>A(1,30)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>A(1,45)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 4
**CATEGORY OF $\alpha$-$\beta$ PROFILE AND WHERE APPROPRIATE OPTIMAL REGION FOR M, SPREAD PERIOD, $\varphi=0.1$ (AR(1) case)**

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\nu$</th>
<th>.01</th>
<th>.005</th>
<th>.01</th>
<th>.03</th>
<th>.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>B</td>
<td>A(1,90)</td>
<td>A(1,50)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>B</td>
<td>A(1,40)</td>
<td>A(1,30)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
<td></td>
</tr>
<tr>
<td>.15</td>
<td>A(1,55)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td>A(1,10)</td>
<td>A(1,8)</td>
<td></td>
</tr>
<tr>
<td>.20</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
<td>A(1,8)</td>
<td>A(1,5)</td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>A(1,15)</td>
<td>A(1,9)</td>
<td>A(1,8)</td>
<td>A(1,6)</td>
<td>A(1,3)</td>
<td></td>
</tr>
</tbody>
</table>
## Table 5

<table>
<thead>
<tr>
<th>i</th>
<th>.01</th>
<th>.005</th>
<th>.01</th>
<th>.03</th>
<th>.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>B</td>
<td>A(1,60)</td>
<td>A(1,40)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
</tr>
<tr>
<td>.10</td>
<td>A(1,80)</td>
<td>A(1,20)</td>
<td>A(1,20)</td>
<td>A(1,9)</td>
<td>A(1,6)</td>
</tr>
<tr>
<td>.15</td>
<td>A(1,20)</td>
<td>A(1,9)</td>
<td>A(1,9)</td>
<td>A(1,5)</td>
<td>C</td>
</tr>
<tr>
<td>.20</td>
<td>A(1,7)</td>
<td>A(1,4)</td>
<td>A(1,3)</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>.25</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

### Table 6: Category of $\alpha$-$\beta$ profile and where appropriate optimal region for M*, spread period, $\phi = 0.1$ in MA(1) process.

<table>
<thead>
<tr>
<th>interest rate(i)</th>
<th>-0.01</th>
<th>0.005</th>
<th>0.01</th>
<th>0.03</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu = 0.05$</td>
<td>B</td>
<td>A(1,130)</td>
<td>A(1,70)</td>
<td>A(1,25)</td>
<td>A(1,15)</td>
</tr>
<tr>
<td>0.10</td>
<td>B</td>
<td>A(1,90)</td>
<td>A(1,60)</td>
<td>A(1,25)</td>
<td>A(1,15)</td>
</tr>
<tr>
<td>0.15</td>
<td>B</td>
<td>A(1,55)</td>
<td>A(1,45)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
</tr>
<tr>
<td>0.20</td>
<td>A(1,130)</td>
<td>A(1,35)</td>
<td>A(1,20)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
</tr>
<tr>
<td>0.25</td>
<td>A(1,50)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
</tr>
<tr>
<td>0.30</td>
<td>A(1,30)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
</tr>
<tr>
<td>0.35</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
<td>A(1,10)</td>
</tr>
</tbody>
</table>

### Table 7: Category of $\alpha$-$\beta$ profile and where appropriate optimal region for M*, spread period, $\phi = -0.1$ in MA(1) process.

<table>
<thead>
<tr>
<th>Interest rate(i)</th>
<th>-0.01</th>
<th>0.005</th>
<th>0.01</th>
<th>0.03</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu = 0.05$</td>
<td>B</td>
<td>A(1,85)</td>
<td>A(1,55)</td>
<td>A(1,20)</td>
<td>A(1,15)</td>
</tr>
<tr>
<td>0.10</td>
<td>B</td>
<td>A(1,45)</td>
<td>A(1,30)</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
</tr>
<tr>
<td>0.15</td>
<td>A(1,60)</td>
<td>A(1,25)</td>
<td>A(1,20)</td>
<td>A(1,10)</td>
<td>A(1,8)</td>
</tr>
<tr>
<td>0.20</td>
<td>A(1,25)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
<td>A(1,8)</td>
<td>A(1,7)</td>
</tr>
<tr>
<td>0.25</td>
<td>A(1,15)</td>
<td>A(1,10)</td>
<td>A(1,9)</td>
<td>A(1,6)</td>
<td>A(1,4)</td>
</tr>
<tr>
<td>0.30</td>
<td>A(1,9)</td>
<td>A(1,7)</td>
<td>A(1,6)</td>
<td>A(1,4)</td>
<td>A(1,2)</td>
</tr>
<tr>
<td>0.35</td>
<td>A(1,6)</td>
<td>A(1,5)</td>
<td>A(1,4)</td>
<td>A(1,2)</td>
<td>C</td>
</tr>
</tbody>
</table>

38
Table 8: Category of α-β profile and where appropriate optimal region for M*, spread period, φ = -0.3 in MA(1) process.

<table>
<thead>
<tr>
<th>Interest rate(%)</th>
<th>-0.01</th>
<th>0.005</th>
<th>0.01</th>
<th>0.03</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>ψ = 0.05</td>
<td>B</td>
<td>A(1,70)</td>
<td>A(1,45)</td>
<td>A(1,20)</td>
<td>A(1,10)</td>
</tr>
<tr>
<td>0.10</td>
<td>A(1,250)</td>
<td>A(1,30)</td>
<td>A(1,25)</td>
<td>A(1,10)</td>
<td>A(1,8)</td>
</tr>
<tr>
<td>0.15</td>
<td>A(1,35)</td>
<td>A(1,15)</td>
<td>A(1,15)</td>
<td>A(1,7)</td>
<td>A(1,4)</td>
</tr>
<tr>
<td>0.20</td>
<td>A(1,10)</td>
<td>A(1,8)</td>
<td>A(1,7)</td>
<td>A(1,3)</td>
<td>C</td>
</tr>
<tr>
<td>0.25</td>
<td>A(1,6)</td>
<td>A(1,4)</td>
<td>A(1,2)</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>0.30</td>
<td>A(1,2)</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>0.35</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>
Figure 1:

'Sketches' of the ultimate variances plotted against m. 's' denotes the Spread Method, whereas 'a' denotes the Amortization of Losses Method.
Figure 2:

A 'Sketch' of the ultimate variances. 's' denotes the Spread method, whereas 'a' denotes the Amortization of Losses Method.
DEPARTMENT OF ACTUARIAL SCIENCE AND STATISTICS

Actuarial Research Papers


30. Pusey, A.S. An Investigation into the Effect of Averaging Age Nearest Birthday at the Beginning of a Rate Year to Exact Age. October 1991. 33 pages. ISBN 1 874 770 30 1


   ISBN 1 874 770 84 0

   ISBN 1 874 770 85 9

   ISBN 1 874 770 88 3

   ISBN 1 874 770 89 1

   ISBN 1 874 770 90 5

   ISBN 1 874 770 91 3

   ISBN 1 874 770 93 X

   ISBN 1 874 770 94 8

   ISBN 1 874 770 99 9

   ISBN 1 874 770 00 X

   ISBN 1 901615 03 0

   ISBN 1 901615 04 9

   ISBN 1 901615 05 7

   ISBN 1 901615 06 5
ISBN 1 901615 07 3

ISBN 1 901615 08 1

ISBN 1 901615 09 X

55 Pages.
ISBN 1 901615 12 X

51 Pages.
ISBN 1 901615 16 2

ISBN 1 901615 19 7

103. Rickayzen B.D. A Sensitivity Analysis of the Parameters used in a PHI Multiple State Model. December 1997. 18 Pages.
ISBN 1 901615 20 0

ISBN 1 901615 22 7

ISBN 1 901615 23 5

ISBN 1 901615 24 3

Statistical Research Papers

17 Pages.
ISBN 1 874 770 83 2

ISBN 1 874 770 86 7

ISBN 1 874 770 87 5


ISBN 1 901615 24 3
Department of Actuarial Science and Statistics

Actuarial Research Club

The support of the corporate members

Commercial Union
Coopers & Lybrand
Government Actuary’s Department
Guardian Insurance
Hymans Robertson
KPMG
Munich Reinsurance
Swiss Reinsurance

is gratefully acknowledged.