Pension Funding in a Stochastic Environment: The Role of Objectives in Selecting an Asset Allocation Strategy

Andrew J.G. Cairns
Department of Actuarial Mathematics and Statistics,
Heriot-Watt University, Edinburgh, EH14 4AS
United Kingdom

Tel: 0131-451 3245
Fax: 0131-451 3249
e-mail: andrewc@ma.hw.ac.uk

Summary

For larger defined benefit pension plans variability in funding levels and contribution rates arises primarily from variability in real investment returns relative to salary growth. Recent work by Dufresne (1988, 1989, 1990) and Haberman (1994) has shown how this uncertainty can be reduced by chosing an appropriate amortization strategy. In the present paper we first consider to what extent the effectiveness of the decision making strategy is compromized by uncertainty in the model parameters.

We then extend previous work by considering, in a simple fashion, how the asset allocation strategy can also be used to control variability. The obvious approach is to make use of less volatile assets. This reduces uncertainty in funding levels while the lower expected investment returns raise the mean contribution rate. However, lower risk assets have a tendency to produce returns which are positively correlated through time. This has the effect of increasing the variability in the funding level over that which might be expected, reducing the benefits of, for example, a switch fom equities into bonds.

It is argued that clearly defined, mathematical objectives must exist for a fund to settle on an appropriate asset allocation strategy. Such objectives should define what levels of uncertainty are tolerable and which events are to be avoided (for example, the event that the solvency level falls below 90%).

It is described how the Inverse-Gamma distribution provides a good approximation to the stationary distribution of the fund size. Using this approximation, a simple argument shows that an objective which provides an upper bound on the probability of insolvency favours a strategy which holds a fixed *amount* in a low risk asset and any surplus in a higher risk asset over the rebalancing strategy which maintains a fixed *proportion* in each asset class independent of the current funding level.

Keywords: stochastic pension fund model; stochastic interest rate models; stationary distribution; Inverse-Gamma distribution; objectives; asset allocation; rebalancing; constant proportion portfolio insurance.

1 Introduction

In this paper we will consider the uncertainty which arises in defined benefit pension plans as a result of the inherent randomness in investment returns. In this context a stochastic framework is the only sensible one to use. Within a deterministic framework there is no concept of uncertainty: the very thing we are attempting to quantify and control.

In this paper we will review and develop the results for simple stochastic models derived by Dufresne (1988, 1989, 1990) and Haberman (1992, 1993 a,b, 1994). This will involve providing further insight into the problems investigated by these authors; development of some distributional theory behind funding levels; and consideration of how to treat more than one asset class.

In all cases there is more than one obviously optimal strategy: that is, there will not be a single strategy which minimizes variances and maximizes returns. This introduces the need for clearly defined objectives, and these will be discussed towards the end of the paper.

In this, introductory section we review the existing results, and in subsequent sections the results are generalized and extended to produce new insight into the problem of uncertainty in pension funding.

1.1 Defined benefit pension plans

Defined benefit pension plans provide benefits to members which are defined in terms of a member's final salary (according to some definition), and the length of membership in the plan. For example,

$$\begin{aligned} \text{Annual pension} &= \frac{N}{60} \times FPS \\ \text{where } N &= \text{number of years of plan membership} \\ FPS &= \text{final pensionable salary} \end{aligned}$$

In defined benefit pension plans pension and other benefits do not depend on past investment performance. Instead the risk associated with future returns on a fund's assets is borne by the employer. This manifests itself through the contribution rate which must vary through time as the level of the fund fluctuates above and below its target level. If these fluctuations are not dealt with (that is, if the contribution rate remains fixed) then the fund will ultimately either run out of assets from which to pay the benefits or grow exponentially out of control.

1.2 A simple model

A number of the factors which we will look at can be first investigated by looking at a very simple stochastic model. By doing so we are able to focus quite quickly on the problem and to give ourselves a good feel for what might happen when we look at more realistic and complex models. This approach follows that of Dufresne (1988, 1989, 1990), Haberman (1992, 1993 a,b, 1994), Zimbidis and Haberman (1993), Cairns (1995) and Cairns and Parker (1995).

Suppose, then, that we have a fund which has a stable membership and a stable level of benefit outgo. Assuming that all benefits and contributions are paid at the start of each year we have the following relationship:

$$AL(t+1) = (1+i'_v)\,(AL(t)+NC(t)-B(t))$$

where

$$\begin{aligned} AL(t) &= \text{actuarial liability at time } t \\ B(t) &= \text{benefit outgo at time } t \\ NC(t) &= \text{normal contribution rate at time } t \\ \text{and } i'_v &= \text{valuation rate of interest} \end{aligned}$$

(Strictly speaking, $AL(t+1)$ is the liability which would be calculated at time $t+1$ given the current conditions at time t and if individual valuation assumptions were borne out over the next year.)

Suppose that salary inflation is at the rate s per annum and that benefit outgo increases in line with salaries each year. Then

$$\begin{aligned} B(t) &= B.(1+s)^t \\ AL(t) &= AL.(1+s)^t \\ NC(t) &= NC.(1+s)^t \end{aligned}$$

giving

$$\begin{aligned} AL.(1+s) &= (1+i'_v)(AL+NC-B) \\ \text{or } AL &= (1+i_v)(AL+NC-B) \\ \text{where } i_v &= (1+i'_v)/(1+s)-1 = (i'_v-s)/(1+s) \\ &= \text{real valuation rate of interest} \end{aligned}$$

Hence

$$\begin{aligned} NC &= B-(1-v_v)AL \\ \text{where } v_v &= 1/(1+i_v) \end{aligned}$$

In this ideal situation B will be known, while AL will be determined by the valuation method and its associated assumptions, of which the real valuation rate of interest is one.

For convenience we will work in real terms relative to salary growth. In effect this means that we may assume that $s=0$, without losing any level of generality.

Now let $F(t)$ be the actual size of the fund at time t. Then

$$F(t+1)=(1+i(t+1))\,(F(t)+C(t)-B)$$

where $i(t+1)$ is the effective rate of interest earned on the fund during the period t up to $t+1$, and $C(t)$ is the contribution rate at time t.

$C(t)$ can be split into two parts: the normal contribution rate, NC; and an adjustment $ADJ(t)$ to allow for surplus or deficit in the fund relative to the actuarial liability. Thus

$$C(t)=NC+ADJ(t)$$

We will deal with the calculation of this adjustment in the next two sections.

The deficit or unfunded liability at time t is defined as the excess of the actuarial libility over the fund size at time t. Hence we define

$$\begin{aligned} UL(t) &= \text{unfunded liability at time } t \\ &= AL-F(t) \end{aligned}$$

In North America it is common also to look at the loss which arises over each individual year. This is defined as the difference between the expected fund size (based on the valuation assumptions) and the actual fund size at the end of the year given the history of the fund up to the start of the year. This gives us

$$\begin{aligned} L(t) &= \text{loss in year } t \\ &= E[F(t)]-F(t) \;\; \text{given the fund history up to time } t-1 \\ &= UL(t)-E[UL(t)] \;\; \text{given the fund history up to time } t-1 \end{aligned}$$

(for example, see Dufresne, 1989).

We will make use of $UL(t)$ and $L(t)$ in the next section.

No mention has been made so far of the interest rate process $i(t)$. Initially we will assume that $i(1), i(2), \ldots$ form an independent and identically distributed sequence of random variables with

$$\begin{aligned} i(t) &> -1 \text{ with probability } 1 \\ E[i(t)] &= i \\ Var[i(t)] &= Var[1+i(t)] = \sigma^2 \\ \Rightarrow E[(1+i(t))^2] &= (1+i)^2 + \sigma^2 \end{aligned}$$

For notational convenience we will define

$$\begin{aligned} v_1 &= \frac{1}{E[1+i(t)]} = \frac{1}{1+i} \\ v_2 &= \frac{1}{E[(1+i(t))^2]} = \frac{1}{(1+i)^2+\sigma^2} < v_1^2 \end{aligned}$$

These will be made use of in later sections.

1.3 Two methods of amortization

The Spread Method: This is in common use in the UK. The adjustment to the contribution rate is just a fixed proportion of the unfunded liability: that is,

$$\begin{aligned} ADJ(t) &= k.UL(t) \\ \text{where } k &= \frac{1}{\ddot{a}_{\overline{m}|}} \text{ at rate } i_v \\ \text{and } m &= \text{the period of amortization.} \end{aligned}$$

The period of amortization is chosen by the actuary, and commonly ranges from 5 years to over 20 years. For accounting purposes in the UK m must be set equal to the average future working lifetime of the membership.

The Amortization of Losses Method: This is in common use in the USA and Canada. The adjustment is calculated as the sum of the losses in the last m years divided by the present value of an annuity due with a term of m years calculated at the valuation rate of interest: that is,

$$ADJ(t) = \frac{1}{\ddot{a}_{\overline{m}|}} \sum_{j=0}^{m-1} L(t-j)$$

The interpretation of this is that the loss made in year s is recovered by paying m equal instalments of $L(s)/\ddot{a}_{\overline{m}|}$ over the next m years. These m instalments have the same present value as the loss made in year s.

Dufresne (1989) showed that the unfunded liabilities and the losses are linked in the following way:

$$\begin{aligned} UL(t) &= \sum_{j=0}^{m-1} \lambda_j L(t-j) \\ \text{where } \lambda_j &= \frac{\ddot{a}_{\overline{m-j}|}}{\ddot{a}_{\overline{m}|}} \end{aligned}$$

Intuitively this makes sense, since $\lambda_j L(t-j)$ is just the present value of the future amortization instalments in respect of the loss made at time $t-j$. Hence $UL(t)$ is equal to the present value of the outstanding instalments in respect of all losses made up until time t.

The Spread Method can also be defined in terms of the loss function. Whereas the Amortization of Losses Method recovers the loss at time t by taking in m *equal* instalments of $L/\ddot{a}_{\overline{m}|}$, the Spread Method recovers this by making a geometrically decreasing, infinite sequence of instalments which starts at the same level.

We are now in a position to calculate the long term mean and variance of the fund size and of the contribution rate. Details of these are provided in Dufresne (1989) (in the case when the valuation and the true mean rate of interest are equal) and Cairns (1995) (covering the case when $i \neq i_v$). For the Spread method we find that

$$\begin{aligned} E[F(t)] &= \frac{(1-k-v_v)AL}{(1-k-v_1)} \\ E[C(t)] &= B - \frac{(1-k-v_v)(1-v_1)AL}{(1-k-v_1)} \\ Var[F(t)] &= \frac{(1-k-v_v)^2(v_1^2-v_2)}{(1-k-v_1)^2(v_2-(1-k)^2)} AL^2 \end{aligned}$$

$$Var[C(t)] \;=\; k^2 \frac{(1-k-v_v)^2(v_1^2-v_2)}{(1-k-v_1)^2(v_2-(1-k)^2)} AL^2$$

When $i = i_v$ these simplify to

$$\begin{aligned}
E[F(t)] &= AL \\
E[C(t)] &= B-(1-v_1)AL \\
Var[F(t)] &= \frac{(v_1^2-v_2)}{(v_2-(1-k)^2)} AL^2 \\
Var[C(t)] &= k^2 \frac{(v_1^2-v_2)}{(v_2-(1-k)^2)} AL^2
\end{aligned}$$

Now $v_1^2 > v_2$ and we must have $Var[F(t)]$ and $Var[C(t)]$ greater than 0. Hence we must have $(1-k)^2 < v_2 \;\Rightarrow\; k > 1-\sqrt{v_2}$. This then automatically implies that $k > 1-v_1$ and if this is combined with $k > 1-v_v$ it ensures that the mean fund size is also positive.

Looking at the Amortization of Losses Method we have, when the valuation rate of interest is equal to the true long term mean rate of interest,

$$\begin{aligned}
Var[L(t)] &= \frac{\sigma^2(1+i)^{-2}AL^2}{1-\sigma^2(1+i)^{-2}\sum_{j=1}^{m-1}\lambda_j^2} = V_\infty \text{ say} \\
Var[F(t)] &= V_\infty \sum_{j=0}^{m-1} \lambda_j^2 \\
Var[C(t)] &= \frac{m.V_\infty}{(\ddot{a}_{\overline{m}|})^2}
\end{aligned}$$

1.4 The period of amortization

One factor which we have within our control is the period of amortization, m.

For the time being, assume that the valuation and the true long term mean rates of interest are equal: we will look at the more general case in a later section. The following results can be shown to hold for the Spread Method (for example, see Dufresne, 1989)

- $Var[F(t)]$ increases as m increases.

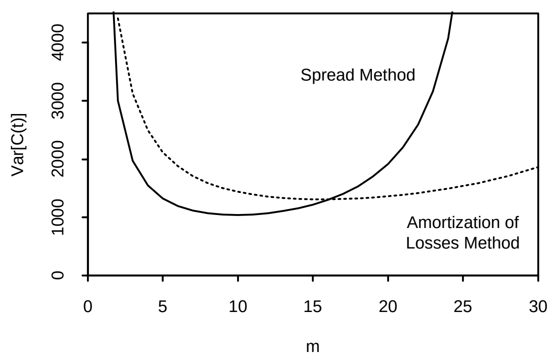

Figure 1: The effect of the period of amortization on the variance of the contribution rate with $E[i(t)] = 0.05$ and $Var[i(t)] = 0.04$.

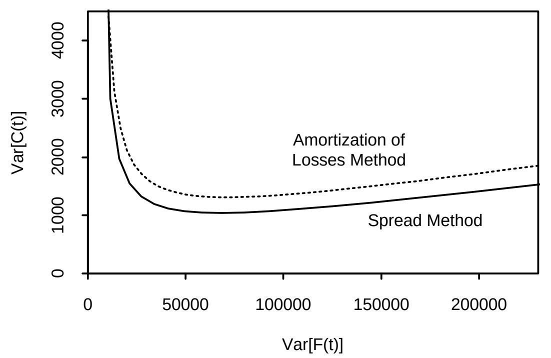

Figure 2: $E[i(t)] = 0.05$ and $Var[i(t)] = 0.04$. Comparison of $Var[F(t)]$ with $Var[C(t)]$. Notes: $Var[F(t)]$ increases as m increases; the efficient frontier for the Spread Method is always more efficient than that for the Amortization of Losses Method.

- $Var[C(t)]$ decreases initially as m increases from 1 up to some value m^* and then increases as m increases beyond m^*. The optimal value, m^*, is such that
  $k^* = 1/\ddot{a}_{\overline{m^*}|} = 1 - v_2$.

Looking at the Amortization of Losses Method no such analytical results have been proved but numerical examples show that the same qualitative behaviour holds, as illustrated in the following example.

Suppose that the mean and the variance of the long term rate of interest are equal to 0.05 and 0.04 respectively. Figure 1 illustrates how the variance of the contribution rate (with $AL = 1$) depends on m. The Spread Method has its minimum at about 10 while the Amortization of Losses Method has its minimum at about 16, and this minimum is higher.

In Figure 2 we compare the variance of the fund size against the variance of the contribution rate. We do this because we may be interested in controlling the variance of the fund size as much as the variance of the contribution rate (since this is linked to the security of members' interests). As m increases each curve moves to the right, first decreasing and then increasing as m passes through m^*. Above m^* both the variance of the fund and the variance of the contribution rate are increasing. It is clear then that no value of m above m^* can be 'optimal' because the use of some lower value of m (say, m^*) can lower the variance of both the fund size and the contribution rate. The range $1 \leq m \leq m^*$ is the so-called *efficient* region: that is, given a value of m in this range there is no other value of m which can lower the variance of both the fund size and the contribution rate. There is therefore a trade-off between variability in the fund size and the contribution rate and settling on what we regard as an optimal spread period can only be done with reference to a more specific objective than 'minimize variance'.

It is significant that the Amortization of Losses Method curve always lies above the Spread Method curve. This means that the Spread Method is certainly more efficient than the Amortization of Losses Method: that is, for any value of m in combination with the Amortization of Losses Method there is a (different) value m' for which the variance of both the fund size and the contribution rate can be reduced by switching to the Spread Method.

2 The Strength of the Valuation Basis

So far we have concentrated on the case where the valuation rate of interest is equal to the mean long term rate of interest. It is common, however, for valuations to be carried out on a strong (occasionally weak) basis: that is, to use a valuation rate of interest which is lower (greater) than the true long term mean rate of interest. This gives rise to a wider variety of results.

Recall that

$$\begin{aligned}
E[F(t)] &= \frac{(1-k-v_v)AL}{(1-k-v_1)} \\
E[C(t)] &= B - \frac{(1-k-v_v)(1-v_1)AL}{(1-k-v_1)} \\
Var[F(t)] &= \frac{(1-k-v_v)^2(v_1^2-v_2)}{(1-k-v_1)^2(v_2-(1-k)^2)}AL^2 \\
Var[C(t)] &= k^2\frac{(1-k-v_v)^2(v_1^2-v_2)}{(1-k-v_1)^2(v_2-(1-k)^2)}AL^2
\end{aligned}$$

We concentrate on the variance of the contribution rate and look for the existence of a minimum with respect to the period of amortization, m. There are a number of cases to consider which are defined by the relationship between the valuation rate of interest and the true mean and variance of the long term rate of interest.

1. **Strong basis:** (valuation rate less than true mean rate)

   (these are currently observations, and not proved)

   (a) $E(C_t)$ is an increasing function of k for $k > 1-\sqrt{v_2}$.

   (b) $Var(C_t)$ has a minimum for some $1-\sqrt{v_2} < k^* < 1$.

   (c) $Var(F_t)$ is a decreasing function of k.

   From this we can see that for $k > k^*$ both the expected value and the variance of the contribution rate are increasing so that increasing k above k^* is not worthwhile. If k is decreased then we trade off a lower contribution rate for a higher variance. The optimal value therefore depends on the pension fund's utility function or objectives.

   For some values of k the mean contribution rate will be negative, indicating that the fund is large enough to pay for itself and at times requiring refunds to the employer. Although this seems an ideal situation, the reality is that the company must first have built up the fund to this level. It would also be likely to violate statutory surplus regulations.

It is possible to have smaller expected fund levels and higher contribution rates, but these do not arise if the projected unit method is used in the calculation of the funding rate and using a conservative valuation rate of interest.

2. **Best estimate:** (valuation rate equal to true mean rate)

   The results of Dufresne (1989) hold.

   (a) $E(C_t)$ is a constant function of k for $k > 1 - \sqrt{v_2}$.

   (b) $Var(C_t)$ has a minimum for some $1 - \sqrt{v_2} < k^* < 1$.

   (c) $Var(F_t)$ is a decreasing function of k.

3. **Weak basis:** (valuation rate slightly greater than true mean rate)

   Defined by $i < i_v < \sqrt{(1+i)^2 + \sigma^2} - 1$.

   (a) $E(C_t)$ is a decreasing function of k for $k > 1 - \sqrt{v_2}$.

   (b) $Var(C_t)$ has a minimum for some $1 - \sqrt{v_2} < k^* < 1$.

   (c) $Var(F_t)$ is a decreasing function of k.

   This time we find that it may be acceptable to increase k above k^*, trading off lower contributions for higher variability.

4. **Very weak basis:** (valuation rate significantly greater than true mean rate)

   Defined by $\sqrt{(1+i)^2 + \sigma^2} - 1 < i_v$.

   (a) $E(C_t)$ is a decreasing function of k for $k > 1 - v_v$ at which point it equals B and the scheme is funded on a pay as you go basis. For $1 - v_v > k > 1 - \sqrt{v_2}$ $E(C_t)$ is still a decreasing function.

   (b) $Var(C_t)$ has a minimum equal to zero at $k = 1 - v_v$. This is because the scheme is now funded on a pay as you go basis and contributions equal the constant B.

   (c) $Var(F_t)$ has a local minimum at $k = 1$, a maximum at some $1 - v_v < k^* < 1$ and a global minimum equal to zero at $k = 1 - v_v$ when the fund stays constant at zero.

The efficient frontier

Pooling these results together we can determine a curve $m(\mu_c)$ where

$$m(\mu_C) = \min\{Var(C_t) : E(C_t) = \mu_C, 1 > k > \max(1 - v_v, 1 - \sqrt{v_2}), v_v < 1\}$$

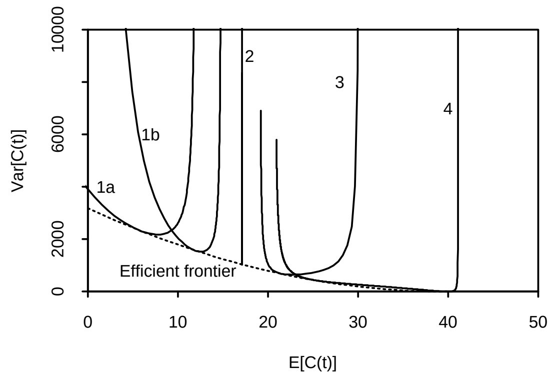

Figure 3: Effect of different valuation rates of interest. Moving from left to right: $i_v = 0.03$ and $i_v = 0.04$ (type 1, strong basis); $i_v = 0.05$ (type 2, best estimate basis); $i_v = 0.06$ (type 3, weak basis); $i_v = 0.07$ (type 4, very weak basis). The dotted line is the efficient frontier.

This curve defines the minimum variance which can be attained for a given mean contribution rate. In fact, it can be shown that this curve is convex (quadratic).

These different types of outcome are illustrated in Figure 3, with $i = 0.05$ and $\sigma^2 = 0.2^2$.

3 Sensitivity Testing

In carrying out such analyses it is important to realize that the model for the rate of return including its parameter values are uncertain. First, the model we use here is only one of a range of possible models of varying complexity which all fit past data reasonably well. All of these models are, however, only an approximation to a much more complex reality. Second, the parameter values which we have used (here $i = 0.05$ and $\sigma^2 = 0.04$) are not known with certainty: for example i could equally well be 0.04 or 0.06.

In fact this can have a very significant effect on level the variability. Figures 4 and 5 illustrate this point. The true mean rate of interest is successively given the values 0.04, 0.05 and 0.06. In Figure 4 the effect on $Var[C(t)]$ is very significant, particularly for larger values of m. However, these results are distorted by the fact that, when the valuation and the true long term mean rates of interest are not equal, the mean fund size depends on m. The normalized variance of $C(t)$ is plotted in

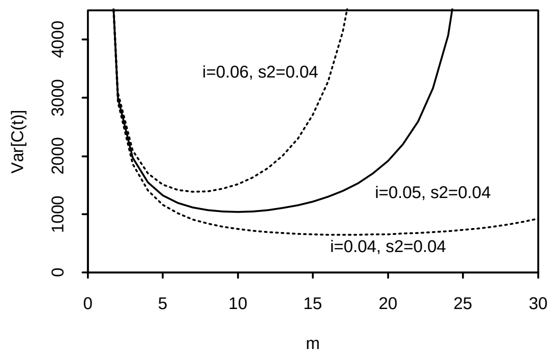

Figure 4: $E[i(t)] = i = 0.04, 0.05, 0.06$ and $Var[i(t)] = s2 = 0.04$. $Var[C(t)]$ plotted against m for different long term rates of return. The valuation rate of interest is fixed.

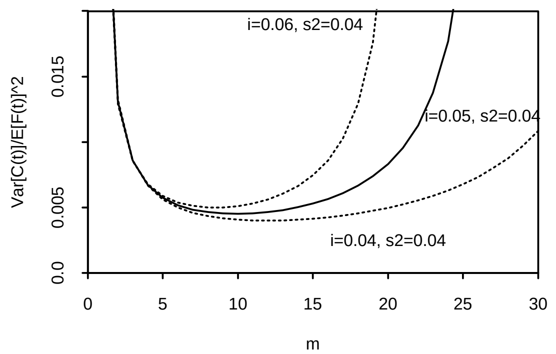

Figure 5: $E[i(t)] = i = 0.04, 0.05, 0.06$ and $Var[i(t)] = s2 = 0.04$. $Var[C(t)]/E[F(t)]^2$ plotted against m for different long term rates of return. The valuation rate of interest is fixed.

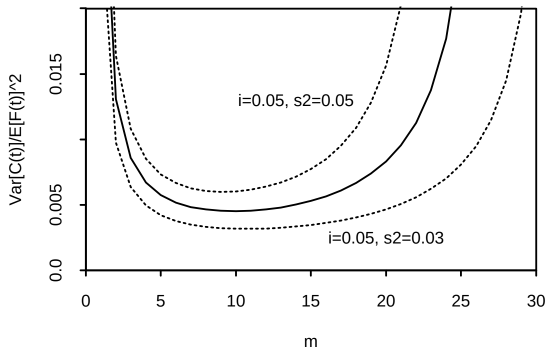

Figure 6: $E[i(t)] = 0.05$ and $Var[i(t)] = s2 = 0.03, 0.04, 0.05$. $Var[C(t)]/E[F(t)]^2$ plotted against m for varying levels of volatility in the rate of return. The valuation rate of interest is fixed.

Figure 5 and the effect can be seen to be reduced but still significant.

A change in the value of i of 1% makes a difference in m^* of about 2 years (for example, moving from $i = 0.05$ to $i = 0.06$ changes m^* from 10 to 8).

The result of these changes is not as significant as might first appear. For example, suppose we settled upon $m^* = 10$ on the basis that $i = 0.05$. If in fact the long term mean turned out to be $i = 0.06$ then the decision to amortize over 10 years would turn out to have been only marginally worse than if the true optimum $m^* = 8$ had been used. The fact that the actual variance of the contribution rate was perhaps 20% higher than that expected is irrelevant since the lower value would never, in fact, have been attainable.

Figure 6 shows the effects of uncertainty in σ^2 (with σ^2 taking the values 0.03, 0.04 and 0.05). The effect is again substantial, but much more uniform over the whole range of values for m. This is because σ^2 has a much more direct effect on the variance of the fund size and the contribution rate. However, as with uncertainty in i, the normalized variance is relatively stable over a range of values about the minimum, so choosing the wrong value of m will only marginally increase the long term variance.

The point to take in from this section is that we need to take care in ensuring that we look at the right quantities. We therefore need to compare the *actual* outcome based on the decision which was based on incorrect assumptions with the outcome which would have *actually* happened had the decision been based on the correct assumptions. Here the differences have been shown to be minimal but if we were to find that they were significant then we may need to look carefully at our estimates to see if they can be refined and improved upon.

With only a limited amount of past data it is extremely plausible that the true long term mean rate of interest may be one, two or even three percent different from our best estimate. Similarly, the long term variance could be quite different from our best estimate.

4 Objectives

We have already discussed that within the efficient region for m ($1 \leq m \leq m^*$) there is a trade off between higher variance of $F(t)$ and higher variance of $C(t)$. To settle on an optimal spread period therefore requires the use of a specific objective or utility function. For example, we may be concerned about containing the fund size within a specified band (bounded below, say, by the minimum solvency level and above by a statutory surplus limit). We could accommodate this by specifying that $E[F(t)]$ lie in the middle of this band and that the standard deviation of $F(t)$ be no more than 10% of this mean fund size. In this case the optimum would be m^{**} which pushes the variance of $F(t)$ up to the maximum level allowable or m^* if this is lower.

If a proper optimum is to be found then the fund must have a well defined objective which will allow optimization to take place. Examples of some objectives are:

1. Minimize $Var[C(t)]$ subject to $Var[F(t)] \leq V_{max}$;

2. Minimize $Var[C(t)]$ subject to $E[F(t)] = \mu_F$;

3. Minimize the variance of the present value of all future contributions (that is, $\sum_{t=0}^{\infty} v^t C(t)$) subject to ......;

4. Maximize $E[u(F(t))]$ where $u(f)$ is utility function which depends on the fund size. For example, if $u(f) = -(f - f_0)^2$ then $E[u(F(t))] = -\{Var[F(t)] + (E[F(t)] - f_0)^2\}$, the second term being a penalty for deviation of the mean from the target of f_0;

5. Minimize $Var[C(t)]$ subject to $Pr\{F(t) < AL_{\min}\} \leq 0.05$, where, for example, $AL_{\min}$ is the minimum solvency liability.

Care should be taken when formulating an objective. For example, the fourth of these makes less sense if $E[F(t)]$ is constant for all values of m (that is if $i_v = i$); and constraints should have reasonable rather than extreme or even impossible values (for example, do not impose a requirement that the mean rate of return on the assets should equal 50%).

We will return to the role of objectives later in this paper after we have introduced the possibility of investment in several asset classes.

5 Asset Allocation Strategies

So far we have considered only a simple stochastic interest model (independent and identically distributed returns) which provides us with some intuitively-appealing, analytical results. From a simple point of view this can be regarded as a single asset model. However, it can be applied equally well to funds with more than one asset class.

5.1 The Rebalancing Strategy

This strategy dictates that the fund maintains a constant proportion of its assets in each asset class and is therefore independent of the current funding level. In practice the fund is rebalanced only periodically (to keep down transaction costs) so that the proportions in each class may drift away temporarily from their target values. For example, the fund may be rebalanced once a month or once a year, or whenever the proportion of the fund in a given asset class deviates from its target value by more than 2%, say.

Suppose we rebalance the fund once a year and there are m assets $1,2,\ldots,m$. Asset k ($k = 1,2,\ldots,m$) will produce returns of $j_k(t)$ in year t ($t = 1,2,\ldots$). For modelling purposes we will assume that the process $j(t)^T = (j_1(t),\ldots,j_m(t))^T$ is stationary and ergodic. Furthermore, let $\pi^T = (\pi_1,\ldots,\pi_m)^T$ represent the target proportions in each asset class. Then $i(t) = \pi^T j(t)$ is the return on the fund in year t. Since $j(t)$ is stationary and ergodic, so is $i(t)$. If $j(1), j(2),\ldots$ are independent then so are $i(1), i(2),\ldots$.

We can therefore model the process $i(t)$ directly and apply this to the simple model for a pension plan described in Section 2.

If the process $j(t)$ has some sort of correlation structure then it is likely that $i(t)$ does also. Again, however, it may be possible to model $i(t) = \pi^T j(t)$ directly allowing us to carry out a relatively simple investigation similar to that reviewed in Section 2.

Haberman (1993a, 1994) has investigated the use of an AR(1) time series model for $i(t)$. It was found that $\rho > 0$ (positively correlated returns) decreases the value of m^* (for example, if we maintain $E[i(t)]$ and $Var[i(t)]$ at 0.05 and 0.04 respectively but increase ρ from 0 (i.i.d. returns) to only 0.1 then m^* falls from 10 to 5). Conversely a negative value of ρ (as could be the case for equities) will increase the value of m^*.

Now it is intuitively clear that an alternative way to the period of amortization of reducing the variance of the funding level is to invest in lower risk assets. However, such assets generally tend to provide more positively correlated returns. Cairns and Parker (1995) have shown that if $E[i(t)]$ and $Var[i(t)]$ remain fixed while ρ increases then both the mean and the variance of the funding level will increase. The first of these effects is beneficial to the fund, whereas the second is not.

When we move to lower risk assets both $E[i(t)]$ and $Var[i(t)]$ will fall but the results of Cairns and Parker (1995) suggest that the likely increase in ρ will mean that the variance of the funding level will not fall by as much as we might expect. In fact this can be shown to be the case in the following simulation study. Returns on equities, irredeemable bonds (consols) and short term bonds, and the growth of salaries were generated by the Wilkie model (1994a, b). Returns for bonds of intermediate duration were generated by using an exponential yield curve to interpolate between the short term and irredeemable bond yields. Figure 7 shows the standard risk-return profile for a selection of portfolios: equities and 15-year bonds; equities and consols; equities and short term bonds; and single bond holdings. What this graph does not show is the autocorrelations which exist in certain of the portfolios. The correlations between real returns over salary growth in successive years ranged from (approximately) 0 (for equities), through 0.3 (consols) to 0.6 (short term bonds).

The situation is quite different when we look at the mean and standard deviation of the fund size (Figure 8). Here we see that, if the Wilkie model gives an accurate representation of the future, then the benefits of a move into bonds which the ordinary risk-return profile would suggest (in terms of risk reduction) are much reduced when they are applied to a pension fund. As discussed above this is a result of the autocorrelations which are present in bond returns. In particular note that:

- consols give rise to a fund size which has a lower mean but has a higher variance than an equity based fund;
- all-bond portfolios are inefficient (that is, there always exists a mixed equity-bond portfolio which raises the mean fund size and reduces the variance of the fund size: for example, equity/short-

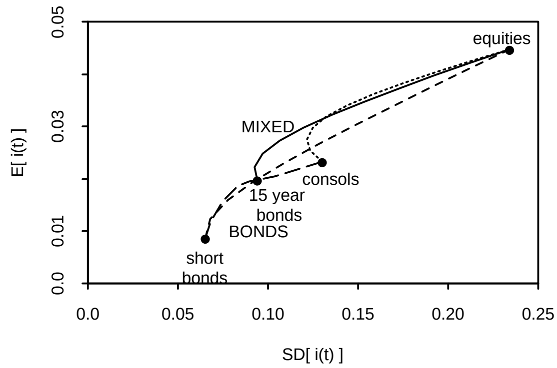

Figure 7: Risk-Return profile for portfolios simulated by the Wilkie model. Solid curve: different mixtures of equities and 15-year bonds. Dotted curve: equities and consols. Short dashed curve: equities and short term bonds. Long dashed curve: single bond holdings.

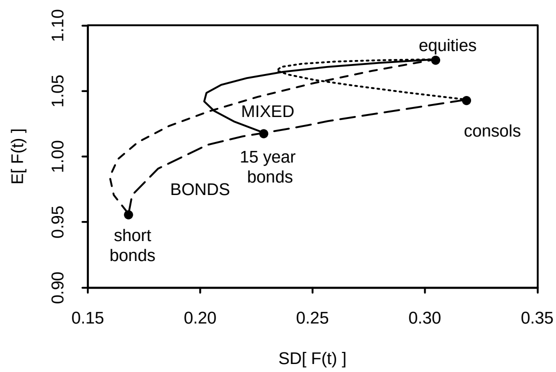

Figure 8: Standard deviation against mean of the fund size (rebalancing strategy) for given portfolios simulated by the Wilkie model. The curves are defined in Figure 7.

bond curve in Figure 8);

- if bond returns in successive years had (like equities) been more-or-less independent of one another then Figure 8 would have looked much more like the much more conventional situation presented in Figure 7;
- for this model, mixed bond portfolios appear to be less efficient than single bond portfolios (but not greatly so);
- portfolios which maintain 50% or 60% in equities and the rest in long-bonds are significantly less risky for a fund than a pure equity fund without reducing the mean fund size by a great amount.

It should be reiterated that these observations have been drawn from a Wilkie model based simulation. It is therefore possible that other asset models will lead to different conclusions – further work needs to be done here.

5.2 Constant proportion portfolio insurance

This is a strategy described by Black and Jones (1987) and Black and Perold (1992). Here, there are two model portfolios, one low risk and one high risk, into which we can invest the assets of the fund. Constant proportion portfolio insurance requires that we invest a certain multiple of the fund's surplus in a risky portfolio and the remainder in the low risk portfolio. Surplus is defined here as being the excess of assets over some 'floor'. For example, the floor may be that defined by minimum solvency regulations, and need not be the actuarial liability which is used in the calculation of the on giong funding level. If the amount of surplus is precisely zero (that is, the value of the assets is equal to the floor) then the amount of the fund invested in the high risk portfolio will be zero.

Mathematically we have

$$
\begin{aligned}
S(t) &= \max\{F(t) - L_{\min}, 0\} = \text{amount of surplus at time } t \\
F_1(t) &= c.S(t) = \text{amount of fund in risky portfolio} \\
F_2(t) &= F(t) - c.S(t) = \text{amount of fund in low risk portfolio} \\
L_{min} &= \text{`floor'} \\
c &= \text{multiple of surplus invested in risky asset}
\end{aligned}
$$

The advantage of using such a strategy is that it provides a simple mechanism for the avoidance of the floor (a mechanism which is missing from the rebalancing strategy) and this will hopefully reduce the probability that the floor is breached.

Suppose that π_1 and π_2 are the vectors representing the proportions in the various assets under the high risk and the low risk strategies respectively. Then a fund which is wholly invested in the high risk fund will obtain a return on its assets in year t of

$$i_1(t) = \pi_1^T j(t)$$

while the return on a fund which is wholly invested in the low risk fund will obtain a return on its assets in year t of

$$i_2(t) = \pi_2^T j(t)$$

If the fund is invested according to the constant proportion portfolio insurance strategy then the return in year t will be

$$\begin{aligned} i(t) &= p(t-1)i_1(t) + (1-p(t-1))i_2(t) \\ \text{where } p(t-1) &= cS(t-1)/F(t-1) \end{aligned}$$

It therefore follows that the m-dimensional asset model can be replaced by a 2-dimensional stochastic process $(i_1(t), i_2(t))$. This mimics the reduction to 1-dimension when the rebalancing strategy is used.

[Note: From a mathematically tractible point of view it helps to measure the amount of surplus immediately *after* the payment of benefits and contributions. This will be discussed elsewhere (Cairns, 1995).]

6 The Stable Distribution of the Fund Size

So far we have looked at the unconditional (or stationary) mean and variance of the fund size. Sometimes (for example, in the setting of objectives) it is of interest to know how often the funding level will fall below or rise above a certain level. In discrete time (as we are considering here) it is only possible to derive the stationary distribution for $F(t)$ when $i(t)$ takes one of a small number of distributions (see Dufresne, 1990). Alternatively, the distribution of $F(t)$ can be found by using the recursive methods described by Parker (1994) (see also Cairns and Parker, 1995).

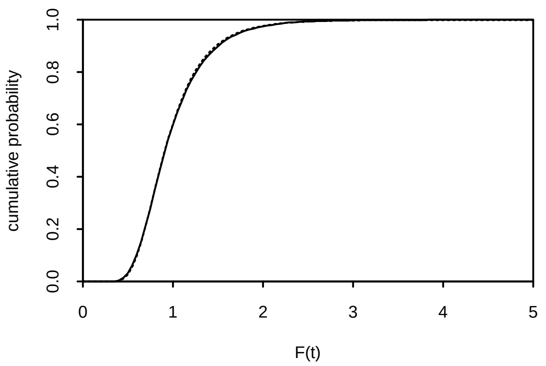

Figure 9: Inverse Gamma approximation to the distribution of $F(t)$. Solid curve: empirical distribution of $F(t)$. Dotted curve: Inverse Gamma approximation.

6.1 The Inverse-Gamma approximation

In the continuous time version of the model it can be shown that if we are following the rebalancing strategy then $F(t)$ has an Inverse-Gamma distribution (for example, see Dufresne, 1990, Cairns, 1995). (If a random variable, X, has an Inverse-Gamma distribution with parameters α and λ then $1/X$ has a Gamma distribution with parameters α and λ, and X has mean $\lambda/(\alpha-1)$ and variance $\lambda^2/[(\alpha-1)^2(\alpha-2)]$.)

Furthermore, if we are following the constant proportion portfolio insurance strategy instead then it can be shown (see Cairns, 1995) that $F(t)-M$ has an Inverse-Gamma distribution (for some constant M).

This exact result has been matched to the distribution for $F(t)$ in the discrete time model, with the conclusion that the Inverse-Gamma distribution provides a very good approximation in a wide variety of cases. This can be seen in Figure 9 where we compare the empirical and approximate distributions for $F(t)$. This example was generated by independent and identically distributed Log-Normal returns:

$$\begin{aligned} \log i(t) &\sim N(0.0286, 0.0399) \\ \Rightarrow E[i(t)] &= 0.05 \\ Var[i(t)] &= 0.0449 \end{aligned}$$

Space prevents further illustration but the approximation is just as good in cases where the process $1+i(t)$ is generated by:

- both the rebalancing and constant proportion portfolio insurance strategies;
- various independent and identically distributed processes including the Log-Normal, Gamma, Log-Normal-with-a-minimum, and Translated-Gamma distributions (the last two mimicking a portfolio of equities and options);
- processes with correlated returns including the AR(1) (for which the application to pension fund modelling is described by Haberman, 1994) and the Wilkie model (see Wilkie, 1987, 1994 a,b).

Having a good idea about what the distribution of $F(t)$ looks like is important when we are considering certain types of objectives. In particular, those which involve probabilities and certain utility functions rather than just means and variances require the use of the full distribution function for $F(t)$. For example, suppose the minimum solvency level will be 60% of the ongoing actuarial liability. A suitable objective for a scheme might then be to minimize the variance of the contribution rate subject to the constraint that the probability of falling below the minimum solvency level in any one year is at most 0.05.

7 Comparison of Strategies

The question arises as to which of the two strategies (rebalancing and constant proportion portfolio insurance) is to be preferred. Clearly this will depend on the objective which the pension plan has set itself. However, a number of facts can be drawn together which will clarify the situation before we consider the objective.

We look at the continuous time model dealt with by Dufresne (1990) and Cairns (1995). Following the analysis of Cairns (1995) we assume the existence of 2 assets: one risk-free and the other risky (but offerring a higher expected return).

First, consider the constant proportion portfolio insurance strategy. At any given time, the amount of surplus dictates how much of the fund should be invested in each asset and consequently we are able to calculate the mean, the variance and the (Inverse-Gamma) distribution of the fund size (Figure 10, dotted curve).

Second, consider the rebalancing strategy. We can choose what proportion of the fund to invest in each asset. Suppose, then, we choose this proportion in such a way as to ensure that the mean fund size under this strategy matches that under the strategy described above. (We

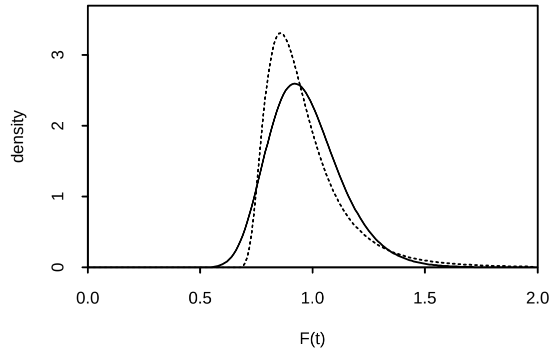

Figure 10: Comparison of Rebalancing strategy (solid curve) and Constant Proportion Portfolio Insurance strategy (dotted curve). Both strategies give the same mean fund size. Rebalancing gives a lower variance of the fund size. Constant Proportion Portfolio Insurance is better at avoiding low funding levels.

do this by using the formula for $E[F(t)]$ given in Section 1.3 to derive the mean rate of return. This mean can then be achieved by choosing an appropriate mix of assets.) Once the portfolio mix has been chosen we can derive the variance and the distribution of the fund size (Figure 10, solid curve). It can be shown (Cairns, 1995) that this variance will always be lower than that under the constant proportion portfolio insurance strategy.

In terms of variances the rebalancing strategy is the more efficient of the two and should therefore be preferred (backing up the conslusions of Lee, 1994). However, there are circumstances under which constant proportion portfolio insurance will be the preferred strategy: in particular, when the objective requires that the probability of falling below a given level (for example, the minimum solvency level) is no more than 0.05, say. It is a strategy which is much better at avoiding specific undesirable levels of funding.

8 References

Black, F. and Jones, R. (1987) Simplifying portfolio insurance for corporate pension plans. *Journal of Portfolio Management* **Summer**, 33-37.

Black, F. and Perold, A. (1992) Theory of constant proportion portfolio insurance. *Journal of Economic Dynamics and Control* **16**, 403-426.

Cairns, A.J.G. (1995) Stochastic methods in pension funding: asset allocation and objectives. *In preparation* , .

Cairns, A.J.G. and Parker, G. (1996) Stochastic pension fund modelling. *In preparation* , .

Dufresne, D. (1988) Moments of pension contributions and fund levels when rates of return are random. *Journal of the Institute of Actuaries* **115**, 535-544.

Dufresne, D. (1989) Stability of pension systems when rates of return are random. *Insurance: Mathematics and Economics* **8**, 71-76.

Dufresne, D. (1990) The distribution of a perpetuity, with applications to risk theory and pension funding. *Scandinavian Actuarial Journal* **1990**, 39-79.

Haberman, S. (1992) Pension funding with time delays: a stochastic approach. *Insurance: Mathematics and Economics* **11**, 179-189.

Haberman, S. (1993a) Pension funding with time delays and autoregressive rates of investment returns. *Insurance: Mathematics and Economics* **13**, 45-56.

Haberman, S. (1993b) Pension funding: the effect of changing the frequency of valuations. *Insurance: Mathematics and Economics* **13**, 263-270.

Haberman, S. (1994) Autoregressive rates of return and the variability of pension fund contributions and fund levels for a defined benefit pension scheme. *Insurance: Mathematics and Economics* **14**, 219-240.

Lee, P.J. (1994) Maximising Long Term Return. *4th AFIR, Orlando* **1**, 311-328.

Parker, G. (1994) Limiting distribution of the present value of a portfolio of policies. *ASTIN Bulletin* **24**, 47-60.

Wilkie, A.D. (1987) Stochastic investment models – theory and practice. *Insurance: Mathematics and Economics* **6**, 65-83.

Wilkie, A.D. (1994a) Stochastic modelling of long-term investment risks. *submitted to the IMA Journal of Mathematics Applied in Business and Finance* , .

Wilkie, A.D. (1994b) The risk premium on ordinary shares. *with discussion. To appear in British Actuarial Journal* **1**, .

Zimbidis, A. and Haberman, S. (1993) Delay, feedback and variability of pension contributions and fund levels. *Insurance: Mathematics and Economics* **13**, 271-285.