

Department of Actuarial Science and Statistics

City University
Northampton Square
London EC1V 0HB

Telephone: 44 171 477 8471
Facsimile: 44 171 477 8572
email: S.Haberman@city.ac.uk
<http://www.city.ac.uk/actstat/dept/shwww.html>

**PENSION FUND DYNAMICS AND SURPLUSES/DEFICITS DUE TO
RANDOM RATES OF RETURN****ABSTRACT**

A simple model for defined benefit occupational pension schemes with i.i.d rates of investment return and a stationary membership is considered. Two methods of adjustment to the normal cost, as surpluses/deficits arise, are compared and a suitable choice of spread or amortization period is made. We also investigate the evolution in time of the first and second moments of the pension fund level and contribution rate.

Keywords: Pension fund dynamics

*Steven Haberman and
M Iqbal Owadally*

Pension Fund Dynamics and Surpluses/Deficits due to Random Rates of Investment Return

S. Haberman & M.I. Owadally

Abstract

A simple model for defined benefit occupational pension schemes with i.i.d. rates of investment return and a stationary membership is considered. Two methods of adjustment to the normal cost, as surpluses/deficits arise, are compared and a suitable choice of spread or amortization period is made. We also investigate the evolution in time of the first and second moments of the pension fund level and contribution rate.

Keywords:

Pension fund dynamics

1 Introduction

In a defined benefit pension scheme, the benefits to be paid out to members are defined according to some formulae, and actuarial valuations determine what the annual contributions must be so that the fund assets plus investment return plus future contributions cover liabilities. Demographic and financial assumptions form the valuation basis. Since these assumptions are unlikely to be borne out in reality, a surplus or deficiency will emerge at successive valuations and will need to be adjusted for. In addition, there may be initial unfunded liabilities to be amortized. The actuary seeks to control the pension funding process by determining stable contribution rates and by securing assets to meet all liabilities. In order to have a tractable model to investigate this, we make a number of simplifying assumptions.

Assumption 1 *Actual experience is in accordance with actuarial assumptions except for investment returns.*

Assumption 2 *The population is stationary from the start.*

Assumption 3 *There is no promotional salary scale; interest rates and benefits are in real terms; inflation on salaries is the same as that on benefits and is assumed to be deterministic. (A salary scale could be introduced by appropriately scaling terms. Haberman, 1994.)*

Assumption 4 *The valuation rate of interest is fixed.*

Assumption 5 *Contribution income and benefit outgo occur at the start of each period.*

Assumption 6 *Valuations are effected at regular intervals of one time unit.*

The assumptions above are similar to those used in the Trowbridge (1952) model. We now need to introduce a stochastic rate of investment return $i(t)$. The fund level $f(t)$ and contribution rate $c(t)$ are therefore discrete-time stochastic processes.

Assumption 7 *The real rate of return on the fund earned during the period $(t - 1, t)$ is $i(t)$ (a stochastic process). $Ei(t)$ equals the valuation rate of interest.*

Assumption 8 *The fund levels at $t \leq 0$ are known with certainty, i.e. $f(0) = f_0$ w.p. 1, $f(-1) = f_{-1}$ w.p. 1 etc.*

Finally, an assumption is required about the random rate of investment return. We restrict our analysis to the simple case explored by Dufresne (1988).

Assumption 9 *$i(t), t \geq 1$, is a stochastic process such that it is a sequence of independent and identically distributed random variables. $Ei(t) = i$. $\text{Vari}(t) = \sigma^2 < \infty$.*

The actuarial liability, normal cost (or standard contribution) and benefit outgo, are AL , NC and B respectively and are constant in time as the membership is deterministic and stationary. As in Trowbridge (1952), the following equation of equilibrium (or 'equation of maturity') applies on the 'liabilities' side and holds on the basis used for valuation:

$$AL = (AL + NC - B)(1 + i). \quad (1)$$

With aggregate funding methods (e.g. Attained Age), the normal cost and actuarial liability are not hypothecated to individuals. The contribution rate is calculated as follows:

$$c(t) = [PVB - f(t)] S / PVS, \quad (2)$$

where PVS is the payroll; PVB is present value of future benefits; S is present value of future salaries. The assumptions we have made fix all the above in time.

Individual methods (e.g. Projected Unit) require the normal cost and actuarial liability for each member to be calculated separately and added up to give totals for the employee population. The unfunded liability $ul(t)$ is

$$ul(t) = AL - f(t), \quad (3)$$

and the normal cost is adjusted by an amount $adj(t)$ to give the contribution rate,

$$c(t) = NC + adj(t). \quad (4)$$

The calculation of the adjustment to the normal cost is crucial to the dynamics of the pension fund. According to Dufresne (1986) (and following his terminology), two methods are in use, as described hereunder.

1. The 'Spread' method. This is commonly used in the United Kingdom. It requires the unfunded liability to be spread into the future, over a 'spread period' of m :

$$adj(t)_s = k.ul(t), \quad k = 1/\ddot{a}_{\overline{m}|}. \quad (5)$$

The annuity is calculated at the valuation rate of interest. The subscript s refers to the 'Spread' method. We assume that the same spread period is used whether there is a surplus or deficit. In practice, this may not be so.

Zimbidis & Haberman (1993) introduce a delay q , at the point of feedback into the system, arising because of the accounting and actuarial effort required during a valuation. From equations (3), (4) and (5):

$$ul(t) = AL - f(t - q)_s. \quad (6)$$

$$c(t)_s = NC + k(AL - f(t - q)_s). \quad (7)$$

2. The 'Amortization of Losses' method. In the United States, the practice is to amortize the initial unfunded liability, as well as the intervalation gains/losses. The actuarial loss $l(t)$ experienced during an intervalation period $(t - 1, t)$ is defined as the difference between $ul(t)$ and the value of the unfunded liability had all actuarial assumptions held true. Thus,

$$adj(t)_a = \begin{cases} ul(0)/\ddot{a}_{\overline{m}|} + \sum_{j=0}^{m-1} l(t-j)/\ddot{a}_{\overline{m}|} & 0 \leq t \leq n-1 \\ \sum_{j=0}^{m-1} l(t-j)/\ddot{a}_{\overline{m}|} & t \geq n \end{cases} \quad (8)$$

and $l(t) = 0$ for $t \leq 0$. The subscript a denotes the 'Amortization of Losses' method. We note that both the initial unfunded liability and the set of intervalation losses are amortised entirely, the former within n years and the latter within m years.

Dufresne (1989) considers the simpler case where $m = n$, such that

$$adj(t)_a = \sum_{j=0}^{m-1} l(t-j)/\ddot{a}_{\overline{m}|},$$

$$l(t) = 0, \quad (t < 0),$$

$$l(0) = ul(0).$$

2 The 'Spread' Method

It is obvious that the period chosen to spread payments for the unfunded liabilities into the future will affect the pace of funding. Dufresne (1988) has also shown that the choice of this parameter affects the stationarity in the limit of the fund and contribution rate levels, and that an optimal range of spread periods exists.

2.1 Spread Period and Stochastic Dynamics

The basic recurrence relationship applying is:

$$f(t+1) = [1 + i(t+1)][f(t) + c(t) - B]. \quad (9)$$

Consider the individual funding methods first. Substituting equation (6) into equation (9) and letting $r = NC - B + kAL$, we obtain:

$$f(t+1)_s = [1 + i(t+1)][f(t)_s - kf(t-q)_s + r]. \quad (10)$$

$f(t)_s$ and $c(t)_s$ form non-linear time series. We need to ascertain that there is at least some weak form of convergence of $f(t)_s$ to AL (to achieve security) and of $c(t)_s$ to NC (for stability). It is difficult to establish in any meaningful closed-form expression how the probability distributions of $f(t)_s$ and $c(t)_s$ vary in finite-time (and even in the limit); it is instructive to consider their moments instead.

Two special and simple cases arise.

1. Immediate cash injection to meet unfunded liability ($m = 1$ and $q = 0$). Letting $k = 1$ in equation (7), substituting in (9) and applying the equation of equilibrium (1) yields:

$$f(t)_s = AL[1 + i(t)]/(1 + i), \quad t \geq 1. \quad (11)$$

On average the fund is able to meet its liabilities as from the end of the first time period (see Result 2.1).

2. Spreading in perpetuity ($m = \infty$ and $q = 0$). Letting $k = d$ in equation (7), substituting in (9) and applying the equation of equilibrium (1) yields:

$$f(t+1)_s = f(t)_s [1 + i(t+1)]/(1 + i), \quad (12)$$

$$f(t)_s = f_0 v^t \prod_{\tau=1}^t [1 + i(\tau)], \quad t \geq 1. \quad (13)$$

The fund level is constant on average (see Result 2.1).

Consider aggregate funding methods now. A delay of q implies that equation (2) is transformed to:

$$c(t) = [PVB - f(t-q)]S/PVS, \quad (14)$$

and substitution in equation (9) leads to:

$$f(t+1) = [1 + i(t+1)][f(t) - k'f(t-q) + r'], \quad (15)$$

where $k' = S/PVS$ and $r' = S.PVS/PVB - B$. There is an obvious similarity between equations (10) and (15), i.e. aggregate funding methods are similar to individual funding methods.

We may use equations (1) and (10) to find moments of the fund and contribution levels. Dufresne (1988) gives

Result 2.1 For delay $q = 0$:

$$Ef(t)_s = [f_0 - AL][u(1 - k)]^t + AL,$$

where $u = 1 + i$. If $m = 1$, then $k = 1$ and $Ef(t)_s = AL$ for $t \geq 1$. If $m = \infty$, then $k = d$ and $Ef(t)_s = f_0 \forall t$. Furthermore,

$$\begin{aligned} Ec(t)_s &= NC + kAL - kEf(t)_s \\ &= NC - [f_0 - AL][u(1 - k)]^t. \end{aligned}$$

This is easily obtained by solving the difference equation obtained by taking expectations across recurrence relationship (10) (with $q = 0$). Taking expectations across equation (7) yields the second part. Note that since $i(t)$ is a sequence of i.i.d. random variables, $i(t + 1)$ is independent of $i(t)$ and of $f(t)_s$.

Clearly, convergence will only occur if $|u(1 - k)| < 1$. It can be shown that for $i > -1$, $0 < u(1 - k) < 1$, so that $\lim_{t \rightarrow \infty} Ef(t)_s = AL$ and $\lim_{t \rightarrow \infty} Ec(t)_s = NC$. This implies that the fund will eventually meet all pension liabilities on average and that the average contribution will eventually be equal to the normal cost.

By squaring and taking expectations across equation (10), and solving the resulting difference equation, it is also possible to consider second moments of the fund and contribution levels. For zero delay, $\text{Var}f(t)_s$ and $\text{Var}c(t)_s$ do converge given certain conditions.

Proposition 2.1 For delay $q = 0$, $\text{Var}f(t)_s$ will vary in time according to the following: (Let $k_{\min} = 1 - 1/\sqrt{(u^2 + \sigma^2)}$.)

1. $\{d < k < 1, k \neq 1 - u/(u^2 + \sigma^2), k \neq k_{\min}\}$

$$\text{Var}f(t)_s = \Theta + \Omega[u(1 - k)]^t + \Psi[u(1 - k)]^{2t} - (\Theta + \Omega + \Psi)[(u^2 + \sigma^2)(1 - k)^2]^t$$

where

$$\begin{aligned} \Theta &= \sigma^2 v^2 AL^2 / [1 - (u^2 + \sigma^2)(1 - k)^2], \\ \Omega &= 2\sigma^2 v (f_0 - AL) AL / [u - (u^2 + \sigma^2)(1 - k)], \\ \Psi &= -(f_0 - AL)^2. \end{aligned}$$

If $k > k_{\min}$ then $\lim_{t \rightarrow \infty} \text{Var}f(t)_s = \Theta$.

2. $k = k_{\min}$

$$\text{Var}f(t)_s = \Psi[u(1 - k)]^{2t} + \Phi[u(1 - k)]^t + \Upsilon t - (\Phi + \Psi)$$

where

$$\begin{aligned} \Phi &= -2AL(f_0 - AL) [1 + v\sqrt{(u^2 + \sigma^2)}], \\ \Upsilon &= \sigma^2 v AL^2 / \sqrt{(u^2 + \sigma^2)}. \end{aligned}$$

$\text{Var}f(t)_s$ does not converge as $t \rightarrow \infty$.

$$3. k = 1 - u/(u^2 + \sigma^2)$$

$$\text{Var}f(t)_s = \Psi[u(1-k)]^{2t} - (\Xi + \Psi)[u(1-k)]^t + \Xi + \Lambda t[u(1-k)]^t$$

where

$$\Lambda = 2\sigma^2 v^2 AL(f_0 - AL),$$

$$\Xi = AL^2(1 + \sigma^2 v^2).$$

$$\lim_{t \rightarrow \infty} \text{Var}f(t) = \Xi.$$

$$4. k = d$$

$$\text{Var}f(t)_s = [(1 + \sigma^2 v^2)^t - 1] f_0^2$$

Convergence does not occur.

$$5. k = 1$$

$$\text{Var}f(t)_s = \sigma^2 v^2 AL^2$$

Refer to Owadally & Haberman (1995) for a proof of the above. From equation (7), it follows that

$$\text{Var}c(t)_s = k^2 \text{Var}f(t)_s. \quad (16)$$

$\text{Var}f(t)_s$ and $\text{Var}c(t)_s$ will converge to a finite value as $t \rightarrow \infty$ only if $k_{\min} < k \leq 1$. Hence, k_{\min} is a minimum value for k (to which there is a corresponding maximum spread period) for convergence to occur. Provided the spread period is such that $k_{\min} < k \leq 1$,

$$\lim_{t \rightarrow \infty} \text{Var}f(t)_s = \frac{\sigma^2 v^2 AL^2}{1 - (u^2 + \sigma^2)(1 - k)^2}. \quad (17)$$

Thus, the choice of spread period affects

- the 'pace of funding';
- the stationarity of the pension funding process in the limit;
- all moments and the probability distributions of $f(t)_s$ and $c(t)_s$ over time.

Spreading payments for unfunded liabilities over long periods may make the fund and contributions non-stationary in the limit, so that the pension funding process becomes 'unstable'. A short spread period means that the contribution rate becomes highly variable very quickly, whereas the fund level remains subject to much less stochastic variation. The spread period therefore affects the stochastic dynamics of the fund over time. This is illustrated in terms of the second central moments in Figure 1.

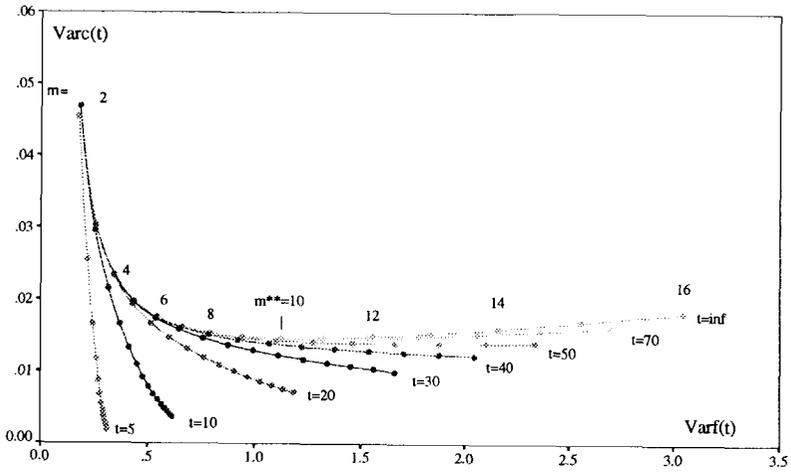


Figure 1: Time profile of a pension fund using the 'Spread' method with $AL = 1.5$, $NC = 0.2$, $f_0 = 1$, $i = 3\%$, $\sigma = 0.25$. Contours of constant k (or m) and of constant t are shown.

2.2 Optimal Spread Periods

A tradeoff exists between fund security and contribution stability. Dufresne (1988) notes that this tradeoff, in terms of ultimate variances of fund level and contribution rate, breaks down for the ‘Spread’ method at $t \rightarrow \infty$, for a certain range of values of the spread period. Dufresne (1988) gives:

Result 2.2 *Let $k^{**} = 1 - 1/(u^2 + \sigma^2)$. For delay $q = 0$ and $(u^2 + \sigma^2) > 1$:*
*if $k_{\min} < k < k^{**}$ then $\text{Var}f(\infty)_s$ and $\text{Var}c(\infty)_s$ decrease with increasing k ;*
*if $k^{**} \leq k < 1$ then $\text{Var}f(\infty)_s$ decreases and $\text{Var}c(\infty)_s$ increases with increasing k .*

A $\text{Var}c(\infty)_s$ v. $\text{Var}f(\infty)_s$ curve therefore exhibits a minimum. For spread periods such that $k < k^{**}$, there will always be a shorter spread period for which both $\text{Var}f(\infty)_s$ and $\text{Var}c(\infty)_s$ are reduced, and therefore they would be unacceptable.

2.3 Delays

The introduction of delays in the valuation leads to a loss of ‘information’ and this should make the system less stable. Zimbidis & Haberman (1993) demonstrate that as $t \rightarrow \infty$, $\text{Var}f(t)_s$ and $\text{Var}c(t)_s$ are larger as the delay q increases. In addition we can show that if there is a delay in the funding process,

- conditions for which the first and second moments of the fund level converge are more constrained (i.e. the spread period range is more restricted);
- the various moments may exhibit oscillatory behaviour (on average, there may be a succession of surpluses and then deficits) for certain spread periods, indicating that $f(t)$ and $c(t)$ do not converge ‘smoothly’ to certain probability distributions in the limit.

Proposition 2.2 *Let delay $q = 1$. There will be no oscillatory behaviour in $\text{E}f(t)_s$,*

- *if $m > -\ln[1 - 4v(1 - v)]/\delta$ (approximately 4) (let α and β be the distinct, real and positive roots of $z^2 - uz + uk = 0$):*

$$\text{E}f(t)_s = AL + \frac{uk[f_{-1} - AL] - \alpha[f_0 - AL]}{\beta - \alpha} \alpha^t + \frac{uk[f_{-1} - AL] - \beta[f_0 - AL]}{\alpha - \beta} \beta^t;$$

- *if $m = -\ln[1 - 4v(1 - v)]/\delta$ (approximately 4) (i.e. if the roots of $z^2 - uz + uk = 0$ are real and coincident):*

$$\text{E}f(t)_s = AL - AL \left(\frac{u}{2}\right)^{t+1} + \left[f_0 + AL \left(\frac{u}{2} - 1\right)\right] (t + 1) \left(\frac{u}{2}\right)^t - ukf_{-1}t.$$

If $m < -\ln[1 - 4v(1 - v)]/\delta$, there will be transient oscillatory behaviour:

$$Ef(t)_s = AL + \frac{(f_0 - AL)}{\sin \theta} a^t \sin(t + 1)\theta - \frac{uk(f_{-1} - AL)}{\sin \theta} a^{t-1} \sin t\theta,$$

where $ae^{\pm i\theta}$ are conjugate complex roots of $z^2 - uz + uk = 0$ and $a = \sqrt{(uk)}$, $\theta = \cos^{-1}(u/(2\sqrt{(uk)}))$.

For a proof, see Owadally & Haberman (1995). Haberman (1992) proves that $Ef(t)_s \rightarrow AL$ as $t \rightarrow \infty$ if the spread period is greater or equal to 2 and for all but very large interest rates (e.g. for $m = 2$, $i < 61.8\%$ is necessary for convergence). In addition, Proposition 2.2 shows that under current economic conditions, no transient oscillatory behaviour is expected in the mean fund and contribution rate level, if the spread period is chosen to be greater or equal to 4 (approximately) time periods. Haberman (1992) also shows that the second moments evolve dynamically in time according to a characteristic equation of the third order as opposed to second order for a zero-delay situation. Conditions for stability of $\text{Var}f(t)_s$ and $\text{Var}c(t)_s$ therefore restrict further the allowable spread period range.

For delay $q \geq 2$, the expected value of the fund level may converge under certain conditions, but will always exhibit some oscillatory behaviour.

Proposition 2.3 For delay $q \geq 2$, $Ef(t)_s$ exhibits transient oscillatory behaviour.

Proof. The recurrence relationship (10) applies. Since $i(t)$ is a sequence of i.i.d. random variables, $i(t + 1)$ is independent of $f(t)$ and $f(t - 2)$ and therefore:

$$\begin{aligned} Ef(t + 1)_s &= (1 + i)[Ef(t)_s - kEf(t - q)_s + r], \\ &= uEf(t)_s - ukEf(t - q)_s + r''. \end{aligned}$$

The characteristic equation here is $z^{q+1} - uz^q + uk = 0$. We know that there will be oscillations if it has imaginary roots or negative real roots. Suppose that the polynomial has at least one pair of complex conjugate roots, then $Ef(t)$ will exhibit oscillations. Suppose that all its roots are real. Since the coefficients of z^{q-2}, \dots, z are zero, application of an elementary theory of algebra (Burnside & Panton, 1935:36) indicates that the roots of the polynomial are not all positive. Since $z = 0$ is not a root, then there must be at least one negative root. Hence, $Ef(t)_s$ will exhibit transient oscillatory behaviour. \square

For example, for $q = 2$, if α, β, γ are distinct roots of $z^3 - uz^2 + uk = 0$, then

$$\begin{aligned} Ef(t)_s &= AL + \frac{\alpha^2[f_0 - AL] - uk\alpha[f_{-2} - AL] - uk[f_{-1} - AL]}{(\alpha - \beta)(\alpha - \gamma)} \alpha^t \\ &+ \frac{\beta^2[f_0 - AL] - uk\beta[f_{-2} - AL] - uk[f_{-1} - AL]}{(\beta - \alpha)(\beta - \gamma)} \beta^t \\ &+ \frac{\gamma^2[f_0 - AL] - uk\gamma[f_{-2} - AL] - uk[f_{-1} - AL]}{(\alpha - \gamma)(\beta - \gamma)} \gamma^t. \end{aligned}$$

If $f_0 = f_{-1} = \dots = f_{-q} = AL$, then it can be proven (Zimbidis & Haberman, 1993) that there are no transients and $Ef(t)_s = AL \forall t$.

Only if the roots of the characteristic equation $z^{q+1} - uz^q + uk = 0$ have magnitude less than unity will $Ef(t)_s$ (and $Ec(t)_s$) converge. (Zimbidis & Haberman (1993) show that convergence in $Ef(t)_s$ will always occur if $i < 50\%$ for a delay $q = 2$ and also that as q tends to infinity, $Ef(t)_s$ diverge.) Thus, as the delay in the system increases, stability conditions for the first moments become more restrictive, and hence stationarity of the fund and contribution levels in the limit becomes more difficult to achieve.

If there are long delays in the pension funding process, oscillations in the first and other moments of the fund level also mean that the pension system is less ‘stable’. The actuary therefore has less control; it becomes more difficult to judge fund solvency and recommend a contribution rate.

(The fact that delays cause instability in actuarial systems is well-known. Balzer & Benjamin (1980) report that time delays of 2 or more time periods lead to oscillations in the simple linear model they used for an insurance system with delayed profit/loss-sharing feedback. Proposition 2.3 for our stochastic model shows that the expected fund level will also oscillate if time delays of 2 or more are introduced in the system. This is not altogether surprising since some kind of negative proportional feedback is used in both models.)

3 The ‘Amortization of Losses’ Method

The first and second moments of $f(t)_a$ and $c(t)_a$ for the ‘Amortization of Losses’ method are derived by Dufresne (1989). He shows that under the assumptions we have made, $Ef(t)_a \rightarrow AL$ and $Ec(t)_a \rightarrow NC$ as $t \rightarrow \infty$. Furthermore, when $m = 1$, the ‘Spread’ and ‘Amortization of Losses’ methods are identical and

$$\text{Var}f(\infty)_s = \text{Var}f(\infty)_a = \text{Var}c(\infty)_s = \text{Var}c(\infty)_a = \sigma^2 AL^2 v^2. \tag{18}$$

In general, Dufresne (1989) proves that, if $\sigma^2 \sum \beta_j^2 < 1$, then

$$\lim_{t \rightarrow \infty} \text{Var}f(t)_a = \frac{\sigma^2 AL^2 v^2 \sum \lambda_j^2}{1 - \sigma^2 \sum \beta_j^2}, \tag{19}$$

$$\lim_{t \rightarrow \infty} \text{Var}c(t)_a = \frac{\sigma^2 AL^2 v^2 m}{(1 - \sigma^2 \sum \beta_j^2) \ddot{a}_{\overline{m}|}^2}, \tag{20}$$

where

$$\beta_j = \frac{a_{\overline{m-j}|}}{\ddot{a}_{\overline{m}|}}, \quad j \in [1, m], \tag{21}$$

$$\lambda_j = \frac{\ddot{a}_{\overline{m-j}|}}{\ddot{a}_{\overline{m}|}}, \quad j \in [0, m]. \tag{22}$$

(For equations (19) and (20) to hold for $m = 1$, we define $\ddot{a}_{\overline{0}|} = a_{\overline{0}|} = 0$.)

Note also that:

$$\sum \beta_j^2 = v^2 \left(\sum \lambda_j^2 - 1 \right). \tag{23}$$

We wish to determine

- whether there exists an optimal range of amortization periods, analogous to the optimal spread period range;
- which of the two methods is more efficient in terms of minimising ultimate variances of fund and contribution rate levels.

3.1 Optimal Amortization Periods

In order to proceed with investigating the tradeoff between variabilities of contribution rate and fund levels (i.e. contribution rate stability and fund security), we need to define the following, from equations (16), (17), (19) and (20):

$$\alpha_s(m) = 1 - (u^2 + \sigma^2)(1 - k)^2 \tag{24}$$

$$\alpha_a(m) = \frac{1 - \sigma^2 \sum \beta_j^2}{\sum \lambda_j^2} \tag{25}$$

$$\beta_s(m) = (1 - (u^2 + \sigma^2)(1 - k)^2)/k^2 \tag{26}$$

$$\beta_a(m) = \left(1 - \sigma^2 \sum \beta_j^2 \right) \ddot{a}_{\overline{m}|}^2 / m \tag{27}$$

where the α s are proportional to the reciprocal of the variance of the fund level in the limit, and the β s are proportional to the reciprocal of the variance of the contribution rate level in the limit. $\beta_a(m)$ is distinct from β_j .

We need the following two propositions, proven in the Appendix.

Proposition 3.1 For $m \geq 2$,

$$\frac{\ddot{a}_{\overline{m-2}|}^2 + \dots + \ddot{a}_{\overline{0}|}^2}{\ddot{a}_{\overline{m-1}|}^2} < \frac{\ddot{a}_{\overline{m-1}|}^2 + \dots + \ddot{a}_{\overline{0}|}^2}{\ddot{a}_{\overline{m}|}^2}.$$

Proposition 3.2 For $m \geq 2$,

$$\left(\ddot{a}_{\overline{m}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2 \right) \left(\ddot{a}_{\overline{m-1}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2 \right) > \left(\ddot{a}_{\overline{m}|} \ddot{a}_{\overline{m-1}|} + \dots + \ddot{a}_{\overline{2}|} \ddot{a}_{\overline{1}|} \right)^2.$$

Dufresne (1986) obtains the sensible conclusion from a numerical investigation that “under the Amortization of Losses method, greater emphasis is laid on security of benefits (i.e. VarF is smaller) than under the Spread method.” This is encapsulated in the following proposition.

Proposition 3.3 *For equal amortization and spread periods,*

$$\begin{aligned} \text{Var}f(\infty)_a &< \text{Var}f(\infty)_s, & m > 1, \\ \text{Var}f(\infty)_a &= \text{Var}f(\infty)_s, & m = 1. \end{aligned}$$

Proof in Appendix. Thus, the $\text{Var}f(\infty)_a$ v. m curve lies under the corresponding curve for the ‘Spread’ method, except that they coincide at $m = 1$.

We also expect that amortizing intervaluation gains/losses over a longer period means that the fund is less secure, i.e. the fund is more variable.

Proposition 3.4 *$\text{Var}f(\infty)_a$ increases monotonically with amortization period m .*

$$\nabla \text{Var}f(\infty)_a > 0.$$

Proof in Appendix.

The stability conditions on equations (17) and (19) provide maximum spread and amortization periods, which we shall denote by m_s^∞ and m_a^∞ respectively, for the pension funding process to be stationary in the limit. k_{\min} in Result 2.2 is clearly related to m_s^∞ .

Proposition 3.5 *For the same i and σ , the maximum spread period allowable for stability is not greater than the maximum allowable amortization period:*

$$m_s^\infty \leq m_a^\infty.$$

Proof in Appendix.

We now prove that, as in the ‘Spread’ method and somewhat counterintuitively, the contribution rate variability does not always decrease as the amortization period increases. There is a minimum point in the $\text{Var}c(\infty)_a$ v. m curve. Let m_s^{**} (resp. m_a^{**}) be the spread (resp. amortization) period for which $\text{Var}c(\infty)_s$ (resp. $\text{Var}c(\infty)_a$) is a minimum. Clearly, m_s^{**} is related to k^{**} in Result 2.2.

Proposition 3.6 *The $\text{Var}c(\infty)_a$ v. m curve has only one turning point, which is a minimum point, at which the $\text{Var}c(\infty)_s$ v. m curve intersects it.*

$$\begin{aligned} m_s^{**} &< m_a^{**} \\ \text{Var}c(\infty)_s|_{m=m_s^{**}} &< \text{Var}c(\infty)_a|_{m=m_s^{**}} \end{aligned}$$

Proof in Appendix.

Result 2.2 and Propositions 3.3, 3.4 and 3.5 enable us to ‘sketch’ the variation of $\text{Var}f(\infty)_a$ and $\text{Var}f(\infty)_s$ v. m . Result 2.2 and Proposition 3.6 can be used to sketch $\text{Var}c(\infty)_a$ and $\text{Var}c(\infty)_s$ v. m . See Figure 2.

Finally, Propositions 3.4 and 3.6 lead us directly to the following, which corresponds to Result 2.2, for the ‘Spread’ method. This shows that the tradeoff between fund security and contribution rate stability breaks down for larger amortization periods.

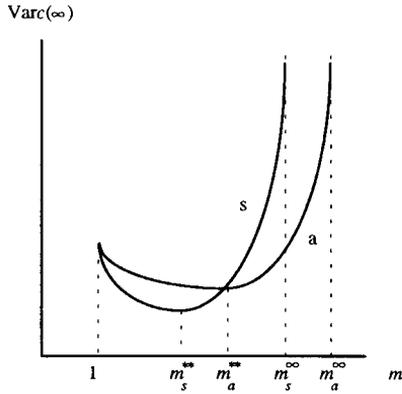
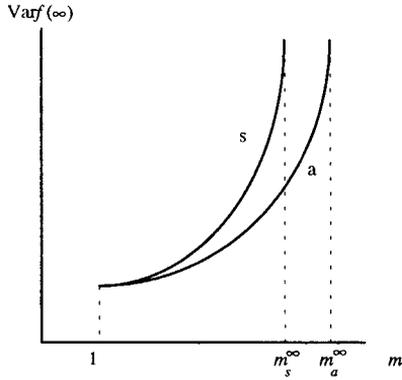


Figure 2: Ultimate variances v. m . 's' denotes the 'Spread' method, whereas 'a' denotes the 'Amortization of Losses' method.

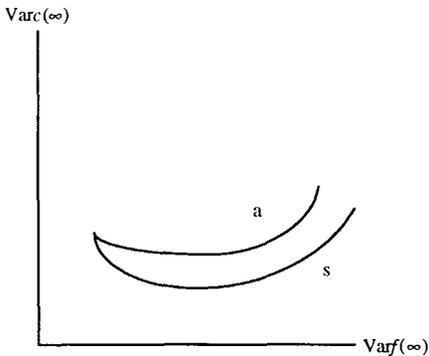


Figure 3: The Spread method ('s') is more efficient than the Amortization of Losses method ('a').

Proposition 3.7 *There exists a non-optimal range of amortization periods (m_a^{**}, ∞) :*
*if $1 < m < m_a^{**}$, $\text{Var}f(\infty)_a$ increases and $\text{Var}c(\infty)_a$ decreases with increasing m ;*
*if $m_a^{**} < m < m_a^\infty$, both $\text{Var}f(\infty)_a$ and $\text{Var}c(\infty)_a$ increase with increasing m ;*
if $m > m_a^\infty$, then $f(t)_a$ and $c(t)_a$ are not stationary in the limit.

The $\text{Var}c(\infty)_a$ v. $\text{Var}f(\infty)_a$ curve has a minimum point at m_a^{**} . For amortization periods in the non-optimal range (m_a^{**}, ∞) , there will always be an amortization period in $[1, m_a^{**}]$ that yields the same $\text{Var}c(\infty)_a$ and a lower $\text{Var}f(\infty)_a$.

We have therefore shown the existence of an optimal amortization period range, $[1, m_a^{**}]$, which is larger than the optimal spread period range. The numerical test, based on small σ and i , performed by Dufresne (1989) fails to show that the tradeoff breaks down for a large enough amortization period and that an optimal range does therefore exist. The purely numerical work of Cairns (1994) illustrates some of the above propositions.

Based on the numerical investigation of these authors, we deduce that a real rate of return (i.e. net of salary inflation) of 1%, with standard deviations 0.025, 0.05 or 0.1, yields a value for m_a^{**} of over 40 years; whereas a real rate of return of 5%, with standard deviation 0.2, yields m_a^{**} of about 16 years. We conclude that, under current economic conditions, the common practice of amortizing gains/losses over periods of 5 years is optimal.

3.2 Comparison of the Two Methods

We now show that the 'Spread' method should be preferred to the 'Amortization of Losses' method, on the grounds of minimising variabilities of fund and contribution rate levels,

Proposition 3.8 *According to the objective of minimising ultimate variances of fund and contribution rate levels, the ‘Spread’ method is more optimal than the ‘Amortization of Losses’ method since for $\{m_a, m_s \neq 1\}$ such that $\text{Var}f(\infty)_a = \text{Var}f(\infty)_s$, then $\text{Var}c(\infty)_a > \text{Var}c(\infty)_s$.*

Proof in Appendix. The $\text{Var}c(\infty)_a$ v. $\text{Var}f(\infty)_a$ curve lies above the $\text{Var}c(\infty)_s$ v. $\text{Var}f(\infty)_s$ curve (except at $m = 1$ where they coincide): see Figure 3. We have shown that, for a given ultimate variance of the fund level, the ‘Spread’ method will always yield a lower ultimate variance of the contribution rate level than the ‘Amortization of Losses’ method (except at $m = 1$). Hence, for equivalent ‘fund security’, the ‘Spread’ method achieves better ‘contribution rate stability’. This is not surprising since the ‘Amortization of Losses’ method uses information delayed by up to m_a years (see section 2.3).

Of course, the ‘Spread’ method is not necessarily the most optimal method. Finite-time horizon optimal methods are discussed by O’Brien (1987) and Haberman & Sung (1994), using stochastic control theory. These methods are not used in practice and can be heavily computational.

4 Conclusion

We have set up a stochastic model for defined benefit pension schemes, using i.i.d. rates of investment return, similar to that described by Dufresne (1988, 1989). Two different methods of adjusting the normal cost are considered. For the ‘Spread’ method, delay in the funding process was introduced, as per Zimbidis & Haberman (1993). By investigating the behaviour of the pension fund in finite time, we have shown that $Ef(t)$ and $Ec(t)$ will converge under given conditions and for certain spread periods. These conditions become more restrictive for increasing delays, and the first moments may exhibit oscillations. We have also shown that the exact evolution of $\text{Var}f(t)$ and $\text{Var}c(t)$ is critically dependent on the spread period chosen. Dufresne (1988) has also shown that an optimal range of spread periods, $[1, m_s^{**}]$, exists; the tradeoff between fund security and contribution stability is broken outside this range.

For the ‘Amortization of Losses’ method, we have similarly established that there exists a range of amortization periods, $[1, m_a^{**}]$, that is optimal and for which there is a tradeoff between ultimate variabilities of the fund and contribution rate levels. This range is wider than the corresponding range in the ‘Spread’ method. For equal spread and amortization periods, we have shown that the ‘Amortization of Losses’ method achieves greater fund security than the ‘Spread’ method. However, we have also shown that the ‘Spread’ method may be regarded as *more* optimal or efficient than the ‘Amortization of Losses’ method, based on the criterion of minimising both variabilities of the fund and contribution rate levels ultimately.

Potential areas of further development include the use of more realistic rate of return processes (AR, MA) and of dynamic rather than fixed valuation assumptions. Other methods of adjusting contributions may also be considered.

References

- [1] Balzer, L.A. & Benjamin, S. (1980). Dynamic response of insurance systems with delayed profit/loss-sharing feedback to isolated unpredicted claims. *Journal of the Institute of Actuaries*, 107, 513-528.
- [2] Burnside, W.S. & Panton, A.W. (1935). *The Theory of Equations*, Vol. 1, 7th ed. Dublin University Press, Dublin.
- [3] Cairns, A.J.G. (1994). An Introduction to Stochastic Pension Plan Modelling. Paper presented to the workshop on Interest Rate Risk, Vancouver.
- [4] Dufresne, D. (1986). The Dynamics of Pension Funding. Ph.D. thesis. School of Mathematics, The City University, London.
- [5] 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.
- [6] Dufresne, D. (1989). Stability of pension systems when rates of return are random. *Insurance: Mathematics and Economics*, 8, 71-76.
- [7] Haberman, S. (1992). Pension funding with time delays. *Insurance: Mathematics and Economics*, 11, 179-189.
- [8] Haberman, S. (1994). Pension Funding Modelling and Stochastic Investment Returns. Paper presented at Joint Meeting of the Staple Inn Actuarial Society and the Royal Statistical Society, General Applications Section, London.
- [9] Haberman, S. & Sung, J.-H. (1994). Dynamic approaches to pension funding. *Insurance: Mathematics and Economics*, 15, 151-162.
- [10] O'Brien, T. (1987). A two-parameter family of pension contribution functions and stochastic optimization. *Insurance: Mathematics and Economics*, 6, 129-134.
- [11] Owadally, M.I. & Haberman, S. (1995). Finite-time Pension Fund Dynamics with Random Rates of Return. Actuarial Research Paper No. 75, School of Mathematics, Actuarial Science and Statistics, The City University, London.
- [12] Trowbridge, C.L. (1952). Fundamentals of pension funding. *Transactions of the Society of Actuaries*, 4, 17-43.
- [13] Zimbidis, A. & Haberman, S. (1993). Delay, feedback and variability of pension contributions and fund levels. *Insurance: Mathematics and Economics*, 13, 271-285.

Appendix

Proof of Proposition 3.1

We first need to show that:

$$\ddot{a}_{t-1} \ddot{a}_{t+1} < \ddot{a}_t^2. \tag{28}$$

$$\begin{aligned} \frac{\ddot{a}_{t-1} \ddot{a}_{t+1}}{\ddot{a}_t^2} - 1 &= \frac{(1 - v^{t-1})(1 - v^{t+1})}{(1 - v^t)^2} - 1 \\ &= \frac{1 - v^{t-1} - v^{t+1} + v^{2t} - 1 + 2v^t - v^{2t}}{(1 - v^t)^2} \\ &= -v^{t-1} \frac{(v - 1)^2}{(1 - v^t)^2} \\ &< 0 \end{aligned}$$

Proof by induction. Let $m = 2$; r.h.s. of the equation in Proposition 3.1 is zero, and the l.h.s. is positive; hence the proposition is true for $m = 2$. (It can also easily be proven for $m = 3$.)

Assume Proposition 3.1 is true for $m = t$, $t \in \mathcal{Z}^+$. Then, replace the denominator on the l.h.s. by \ddot{a}_t^2 and the one on the r.h.s. by \ddot{a}_{t+1}^2 . The inequality is maintained as a result of equation 28 and we have:

$$\frac{\ddot{a}_{t-2}^2 + \dots + \ddot{a}_0^2}{\ddot{a}_t^2} < \frac{\ddot{a}_{t-1}^2 + \dots + \ddot{a}_0^2}{\ddot{a}_{t+1}^2}.$$

We can then subtract the following from the l.h.s. to yield the required result (the inequality is maintained again as a result of equation 28):

$$\frac{\ddot{a}_t^2}{\ddot{a}_{t+1}^2} - \frac{\ddot{a}_{t-1}^2}{\ddot{a}_t^2} \quad (> 0).$$

Since the result holds for $m = 2$, it must hold for $\{ m \in \mathcal{Z}^+ : m \geq 2 \}$. \square

Proof of Proposition 3.2

Proof by induction. Let $m = 2$ and $(\ddot{a}_2^2 + \ddot{a}_1^2)\ddot{a}_1^2 > \ddot{a}_2^2 \ddot{a}_1^2$ and therefore Proposition 3.2 holds for $m = 2$. Suppose it also holds for $m = t$, $t \in \mathcal{Z}^+$. Then, we may expand

$$\left[\ddot{a}_{t+1}^2 + (\ddot{a}_t^2 + \dots + \ddot{a}_1^2) \right] \left[\ddot{a}_t^2 + (\ddot{a}_{t-1}^2 + \dots + \ddot{a}_1^2) \right] - \left[\ddot{a}_{t+1} \ddot{a}_t + (\ddot{a}_t \ddot{a}_{t-1} + \dots + \ddot{a}_2 \ddot{a}_1) \right]^2,$$

cancel out $\ddot{a}_{t+1}^2 \ddot{a}_t^2$ and obtain $T_1 + T_2$, where

$$T_1 = (\ddot{a}_t^2 + \dots + \ddot{a}_1^2)(\ddot{a}_{t-1}^2 + \dots + \ddot{a}_1^2) - (\ddot{a}_t \ddot{a}_{t-1} + \dots + \ddot{a}_2 \ddot{a}_1)^2 > 0,$$

and

$$\begin{aligned}
 T_2 &= \ddot{a}_{t+1}^2(\ddot{a}_{t-1}^2 + \dots + \ddot{a}_1^2) + \ddot{a}_1^2(\ddot{a}_t^2 + \dots + \ddot{a}_1^2) - 2\ddot{a}_{t+1}\ddot{a}_t(\ddot{a}_t\ddot{a}_{t-1} + \dots + \ddot{a}_1\ddot{a}_1), \\
 &> \left[\ddot{a}_{t+1}(\ddot{a}_t^2 + \dots + \ddot{a}_1^2)^{1/2} - \ddot{a}_t(\ddot{a}_{t+1}^2 + \dots + \ddot{a}_1^2)^{1/2} \right]^2 > 0.
 \end{aligned}$$

Therefore $T_1 + T_2 > 0$ and the result holds for $m = t + 1$. Hence, the proposition is true for $\{m \in \mathcal{Z}^+ : m \geq 2\}$. \square

Proof of Proposition 3.3

For $m > 1$, we need to show that $\alpha_a(m) > \alpha_s(m)$. Starting from Proposition 3.1 and noting that $u(1 - k) = \ddot{a}_{m-1}/\ddot{a}_m$, it is easily shown that

$$\begin{aligned}
 [1 - u^2(1 - k)^2] \sum \lambda_j^2 &< 1 \\
 u^2 \sum \lambda_j^2 \cdot \{1 - (u^2 + \sigma^2)(1 - k)^2\} &< u^2 + \sigma^2 - \sigma^2 \sum \lambda_j^2
 \end{aligned}$$

where we have multiplied across by $(u^2 + \sigma^2)$. Rearranging and using equation (23), the inequality in Proposition 3.3 is proven.

For $m = 1$, it is easy to show identity of the ‘Amortization of Losses’ and ‘Spread’ methods (equation (18)). \square

Proof of Proposition 3.4

We may equivalently prove:

$$\nabla \alpha_a(m) < 0.$$

Using the customary backward difference operator rules,

$$\nabla \alpha_a(m) = \frac{(\ddot{a}_{m-1}^2 + \dots + \ddot{a}_0^2)[\ddot{a}_m^2 - (1 + \sigma^2 v^2)\ddot{a}_{m-1}^2] - [\ddot{a}_m^2 - \sigma^2 v^2(\ddot{a}_{m-1}^2 + \dots + \ddot{a}_0^2)]\ddot{a}_m^2}{(\ddot{a}_{m-1}^2 + \dots + \ddot{a}_0^2)(\ddot{a}_{m-1}^2 + \dots + \ddot{a}_0^2)}.$$

The denominator is positive, and the numerator is proved to be negative by rearranging it into

$$(1 + \sigma^2 v^2) \{(\ddot{a}_{m-2}^2 + \dots + \ddot{a}_0^2)\ddot{a}_m^2 - (\ddot{a}_{m-1}^2 + \dots + \ddot{a}_0^2)\ddot{a}_{m-1}^2\}$$

and using Proposition 3.1. \square

Proof of Proposition 3.5

The $\text{Var}f(\infty)_s$ v. m curve has a vertical asymptote at $m = m_s^\infty$ (equation (17)), whereas the $\text{Var}f(\infty)_a$ v. m curve has a vertical asymptote at $m = m_a^\infty$ (equation (19)). Propositions 3.3 and 3.4 mean that the latter asymptote must be at an amortization period not less than the former asymptote. \square

Proof of Proposition 3.6

$$\begin{aligned}
 \sum_{i=1}^m \beta_s(i) &= \sum_{i=1}^m (\ddot{a}_{i|}^2 - \ddot{a}_{i-1|}^2)(1 + \sigma^2 v^2) \\
 &= \ddot{a}_{m|}^2 - \sigma^2 v^2 (\ddot{a}_{m-1|}^2 + \cdots + \ddot{a}_{0|}^2) \\
 &= m \cdot \beta_a(m) \\
 \beta_a(m) &= \frac{\sum_{i=1}^m \beta_s(i)}{m} \\
 \nabla(m\beta_a(m)) &= \beta_s(m).
 \end{aligned}$$

A number of alternative methods can be followed.

1. $\beta_a(m)$ is an average of $\beta_s(m)$ over all spread periods up to m . By Result 2.2 for the ‘Spread’ method, $\beta_s(m)$ has a maximum; $\beta_s(m_s^\infty) = \beta_a(m_a^\infty) = 0$ with $m_s^\infty < m_a^\infty$ (Proposition 4); $\beta_s(m=1) = \beta_a(m=1)$. As $\beta_s(m)$ increases with m , $\beta_a(m)$ increases, but $\beta_a(m) < \beta_s(m)$. When $\beta_s(m)$ decreases and intersects $\beta_a(m)$, $\beta_a(m)$ then starts decreasing.

There will only be one maximum in $\beta_a(m)$. It occurs where the two curves intersect and the maximum in $\beta_s(m)$ will occur before the maximum in $\beta_a(m)$ as m increases, i.e. $m_s^{**} < m_a^{**}$. $\text{Varc}(\infty)_s$ and $\text{Varc}(\infty)_a$ are ‘reciprocals’ of $\beta_s(m)$ and $\beta_a(m)$ respectively.

2. Differencing equation (11),

$$\begin{aligned}
 \nabla\beta_a(m) &= \frac{m\beta_s(m) - \sum_{i=1}^m \beta_s(i)}{m(m-1)} \\
 &= \frac{\beta_s(m) - \beta_a(m)}{m-1} \\
 \nabla^2\beta_a(m) &= \frac{\nabla\beta_s(m) + 2\nabla\beta_a(m)}{m-2}
 \end{aligned}$$

For a turning point, $\nabla\beta_a(m) \approx 0$, or $\beta_s(m) \approx \beta_a(m)$. Hence, turning points in $\beta_a(m)$ will only occur where it crosses $\beta_s(m)$. Consider the shape of $\beta_s(m)$; $\beta_s(m_s^\infty) = \beta_a(m_a^\infty) = 0$ with $m_s^\infty < m_a^\infty$ (Proposition 3.5); $\beta_s(m=1) = \beta_a(m=1)$. Hence, $\beta_a(m)$ can only intersect $\beta_s(m)$ once when it is decreasing, i.e.

$$\nabla^2\beta_a(m) < 0,$$

giving rise to a single maximum in $\beta_a(m)$, i.e. a single minimum point in $\text{Varc}(\infty)_a$.
□

Proof of Proposition 3.8

Consider a spread period m_s (and $k_s = 1/\ddot{a}_{\overline{m_s}|}$) and an amortization period m_a (with corresponding $\sum \lambda_j^2$ and $\sum \beta_j^2$) such that

$$\text{Var}f(\infty)_s = \text{Var}f(\infty)_a. \tag{29}$$

Using equations (24) to (27), we find that

$$\begin{aligned} \alpha_s(m_s) &= \alpha_a(m_a), & (30) \\ \beta_s(m_s)k_s^2\ddot{a}_{\overline{m_s}|}^2 \sum \lambda_j^2 &= \beta_a(m_a) m_a. & (31) \end{aligned}$$

If $m_a = m_s = 1$, then clearly $\text{Var}c(\infty)_a = \text{Var}c(\infty)_s$. For other $\{m_a, m_s\}$, if we show that $k_s^2\ddot{a}_{\overline{m_s}|}^2 \sum \lambda_j^2 < m_a$, then we will show that $\text{Var}c(\infty)_a > \text{Var}c(\infty)_s$.

Substituting equations (24) and (25) into equation (30), we have

$$[1 - (u^2 + \sigma^2)(1 - k_s)^2] \sum \lambda_j^2 = 1 - \sigma^2 \sum \beta_j^2.$$

Replacing $\sum \beta_j^2$ by equation (23), multiplying across by u^2 and rearranging, we obtain

$$[1 - u^2(1 - k_s)^2] \sum \lambda_j^2 = 1,$$

from which we may solve for k_s^2 ,

$$k_s^2 = \frac{[(\sum \lambda_j^2)^{1/2} - v(\sum \lambda_j^2 - 1)^{1/2}]^2}{\sum \lambda_j^2}.$$

Hence,

$$\begin{aligned} &k_s^2\ddot{a}_{\overline{m_s}|}^2 \sum \lambda_j^2 \\ &= \left[(\sum \lambda_j^2)^{1/2} - v(\sum \lambda_j^2 - 1)^{1/2} \right]^2 \ddot{a}_{\overline{m_s}|}^2 \\ &= \left[(\ddot{a}_{\overline{m_s}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2)^{1/2} - v(\ddot{a}_{\overline{m_s-1}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2)^{1/2} \right]^2 \\ &= (\ddot{a}_{\overline{m_s}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2) + v^2(\ddot{a}_{\overline{m_s-1}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2) - 2v[(\ddot{a}_{\overline{m_s}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2)(\ddot{a}_{\overline{m_s-1}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2)]^{1/2} \\ &< (\ddot{a}_{\overline{m_s}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2) + v^2(\ddot{a}_{\overline{m_s-1}|}^2 + \dots + \ddot{a}_{\overline{1}|}^2) - 2v(\ddot{a}_{\overline{m_s}|}\ddot{a}_{\overline{m_s-1}|} + \dots + \ddot{a}_{\overline{2}|}\ddot{a}_{\overline{1}|}) \end{aligned}$$

where we use Proposition 3.2 in the last step. Since $(\ddot{a}_{\overline{t}|} - v\ddot{a}_{\overline{t-1}|})^2 = 1$, $t \in \mathcal{Z}^+$, we find that

$$k_s^2\ddot{a}_{\overline{m_s}|}^2 \sum \lambda_j^2 < m_a.$$

From equation 31, we have therefore proven that $\text{Var}c(\infty)_a > \text{Var}c(\infty)_s$. \square

