Efficient Gain and Loss Amortization and Optimal Funding in Pension Plans

M. Iqbal Owadally and Steven Haberman
Cass Business School, City University London

This paper has been published in *North American Actuarial Journal*, 8, 1, 21--36.

Abstract

Efficient methods of amortizing actuarial gains and losses in defined benefit pension plans are considered. In the context of a simple model where asset gains and losses emerge as a consequence of random (independent and identically distributed) rates of investment return, it has been shown that direct amortization of such gains and losses leads to more variable funding levels and contribution rates compared with an indirect and proportional form of amortization which 'spreads' the gains and losses. Stochastic simulations are performed and they indicate that spreading remains more efficient than amortization with simple AR(1) and MA(1) rates of return. Similar results are obtained when a more comprehensive actuarial stochastic investment model (which includes economic wage inflation) is simulated. Proportional spreading is rationalized as the contribution control that optimizes mean square deviations in the contributions and fund levels when the funding process is Markovian and the fund is invested in two assets (a random risky and a riskfree asset). Efficient spreading and amortization periods are suggested for the U.S., the U.K. and Canada.

JEL classification: D81; G23

1 Introduction

1.1 Aims

Methods of funding defined-benefit pension plans and of amortizing actuarial gains and losses are compared. A simplified model of a pension plan is described in section 1.2 in order to investigate the results of Owadally & Haberman (1999) who suggest that a proportional form of amortization is preferable to the typical practice of direct amortization of gains and losses. These results are reviewed in section 2. In section 3, the results of stochastic simulations, with a more realistic treatment of investment returns and inflation, are presented. Finally, in section 4, the efficiency of one method of amortization over the other is explained by means of the optimization of a quadratic objective criterion.

1.2 Model and Notation

We consider a strictly defined benefit pension plan in which no discretionary benefit improvement is allowed except for benefit indexation. It is also assumed that provision is made only for a retirement benefit at normal retirement age based on final salary and that actuarial valuations are carried out with the following features:

1. Actuarial valuations take place at regular intervals of one time period.
2. The actuarial valuation basis is invariant in time.
3. The market value of pension plan assets is f_t at time t (no smoothing is used).
4. An 'individual' actuarial cost method is used, generating an actuarial liability al_t and a normal cost nc_t at time t.

A simple model for such a pension plan may be projected forward based on the following:

1. The pension plan population is stationary (deterministic) from the start.

2. Mortality and other decrements are assumed to be contingent as per a life table $\{l'_x\}$.

3. A salary scale exactly reflects promotional, merit-based or longevity-based increases in salaries (and may be incorporated in the life-table, $l_x = s_x l'_x$).

Economic wage inflation (the general increase in wages as measured by a national wages index) may be distinguished from the salary scale. The actuarial valuation basis includes $\{l_x\}$ as well as a valuation discount rate and an assumption as to wage inflation (in order to value final-salary benefits). Actual experience is in accordance with the actuarial valuation assumptions except for inflation and returns on plan assets. The model described above bears similarities to the models described by Trowbridge (1952), Bowers *et al.* (1979), Dufresne (1988, 1989) and Owadally & Haberman (1999).

The funding level or funded ratio in the plan is defined as the value of plan assets as a percentage of the actuarial liability. The unfunded liability (ul_t) at the start of year $(t,\, t+1)$ is the excess of the actuarial liability over the value of plan assets (we assume that cash flows occur at the start of the year): $ul_t = al_t \; - \; f_t$.

The outcome of an actuarial valuation at the start of year $(t,\, t+1)$ is to recommend a contribution

$$c_t \;=\; nc_t \;+\; adj_t, \tag{1}$$

where adj_t is a supplementary contribution (or contribution adjustment) paid to amortize past and present experience deviations from actuarial assumptions. These deviations result in actuarial gains or losses.

The actuarial loss l_t is the unanticipated change in the unfunded liability over year $(t-1,\, t)$, that is, it is the excess of the unfunded liability at time t over the unfunded liability anticipated at time t based on information and the valuation basis at time $t-1$. A gain is a negative loss.

1.3 Treatment of Gains and Losses

The calculation of the contribution is crucial to the dynamics of the pension fund when economic experience is volatile. For example, Winklevoss (1982) finds that "the correct treatment of actuarial gains and losses is critical in stochastic simulations because the effect of random fluctuations in salaries and plan assets impact on costs through the funding of such deviations." Typical practice is that the actuarial gain or loss in each year is individually amortized over a fixed term m. The supplementary contribution in year $(t, t+1)$ is (Dufresne, 1989):

$$adj_t = \sum_{j=0}^{m-1} l_{t-j}/\ddot{a}_{\overline{m|}}. \tag{2}$$

Alternatively, gains and losses may be spread by paying a proportion k of the unfunded liability (Dufresne, 1988):

$$adj_t = k\, ul_t = k\,(al_t - f_t), \tag{3}$$

where k is typically defined as

$$k = 1/\ddot{a}_{\overline{m|}}. \tag{4}$$

Equation (3) may be understood as follows. The unfunded liability in the plan is the accumulation with interest of portions of previously incurred losses (as well as of any initial unfunded liability) that have not been fully paid off and have been deferred. When the supplementary contribution in equation (3) is made to the plan, a spread payment equalling a fraction k of the deferred portion of each incurred loss is settled in respect of each loss (as if the loss were taxed at rate k). Therefore, the excess of the unfunded liability over the supplementary contribution in a given year must equal the present value of the unfunded liability, less the newly emergent loss, in the following year: $u(ul_t - k\, ul_t) = ul_{t+1} - l_{t+1}$, where $u = 1+i$ and i is the rate at which cash flows in the plan are discounted. Ignoring any initial unfunded liability, it follows that $ul_t = \sum_{j=0}^{\infty}(1-k)^j u^j l_{t-j}$ and, from equation (3),

$$adj_t = \sum_{j=0}^{\infty} k(1-k)^j u^j l_{t-j}. \tag{5}$$

When $k = m = 1$, gains and losses are not deferred and the unfunded liability consists only of the loss that emerged during the past year, that loss being paid off immediately: $ul_t = adj_t = l_t$. See also Dufresne (1994).

Compare amortization in equation (2) with spreading in equation (5). Under amortization, gains and losses are paid off in level amounts over a finite term m. Under spreading, gains and losses are liquidated in perpetuity by means of exponentially declining payments. A large k (or equivalently a short spreading period m from equation (4)) hastens funding as smaller portions of the losses are deferred. Expressing k as $1/\ddot{a}_{\overline{m|}}$ in equation (4) is a convenient device that enables gains and losses to be removed faster if future cash flows are discounted at a higher rate. A loss is asymptotically liquidated when it is spread since the present value of payments made in respect of a unit loss is $\sum_{j=0}^{\infty} k(1-k)^j = 1$ when $m > 1$ (since $0 < i(1+i)^{-1} < k = 1/\ddot{a}_{\overline{m|}} < 1$). Finally, note that initial unfunded liabilities are disregarded in the following since they can be separately amortized and have no permanent effect (Owadally & Haberman, 1999).

1.4 Volatility of Funding Process

If actuarial assumptions are unbiased and are realized on average, the loss in any year is *expected* to be zero and (once any initial unfunded liability is completely amortized) the unfunded liability (an accumulation of unpaid losses) is also expected to be zero, whether gains and losses are amortized or spread. The uncertain nature of deviations from assumed experience (particularly economic experience) means that gains and losses are volatile, however, and the consequent variability of the funding level and of the required contributions must be examined.

The pension plan is regarded in this paper as an entity, set up under trust, whose financial management is separate from that of the company sponsoring pension benefits for its employees. Trustees (who represent plan members) and sponsor are assumed to have convergent, albeit non-identical, interests in funding the plan on a going-concern basis. One purpose of funding benefits in advance is to maximize the security of benefits (McGill, 1996, p. 592). This is an important objective for trustees but also indirectly for the plan sponsor as it plays a part in the recruitment and remuneration of employees. A reasonable objective of an actuarial funding method is therefore to minimize the variability of the funding level or of the unfunded liability in the plan. An ideal measure of security would be based on second and higher moments and would allow for solvency requirements, but the variance is a reasonable and tractable approximation.

Another motivation for pre-funding pension benefits is to spread the required future pension contributions and hence reduce the strain on the sponsor's cash flows. This is also of interest to the trustees since employees may be contributing to the fund and since successful financial management of the firm benefits not only shareholders but also employees (in terms of continued employment). Trowbridge & Farr (1976, p. 62) thus refer to the "smoothness of contributions" as a desirable objective. This is often expressed relative to the total payroll for plan members. Minimizing the variability of the contribution or contribution rate (that is, total contribution per payroll dollar) is also a reasonable objective. Funding methods therefore seek to reconcile the objectives of both plan trustees and sponsor and achieve stable funding levels as well as stable contribution rates.

In the rest of this paper, the choice between amortizing and spreading volatile gains and losses is discussed in terms of these objectives and under the modeling assumptions of section 1.2. Gains and losses in the pension plan model arise only as a consequence of unforeseen economic variation from actuarial assumptions, that is, in the returns on plan assets and in inflation on plan liabilities. Contributions may be determined as a level percentage of payroll. The annuities in equations (2) and (4) are then calculated at the valuation discount rate net of assumed inflation on wages.

2 Efficient Amortization: Review of Mathematical Results

The treatment of gains and losses and the volatility of funding are studied by Dufresne (1988, 1989) in a simple mathematical model, whose main features are as in the model described in section 1.2. In addition, Dufresne (1988, 1989) assumes that (1) pensions in payment are indexed with wage inflation, so that all monetary quantities can be expressed net of wage inflation (as a consequence of the assumptions made, payroll, actuarial liability etc. are constant when deflated by wage inflation), and (2) the rate of return on the fund (net of wage inflation) is independent and identically distributed from year to year. The second assumption permits mathematical tractability, parsimony and a search for optimal or robust performance.

The market value of plan assets f_t and the contribution rate c_t are random variables and are governed by the recurrence relation $f_{t+1} = (1 + i_{t+1})(f_t + c_t - B)$, where i_t is the rate of return on the fund and B is the benefit outgo. The moments of f_t and c_t are derived by Dufresne (1988, 1989) when gains and losses are spread (equation (3)) as well as amortized (equation (2)). The plan is expected to be fully funded eventually: $\mathrm{E}ul_t \to 0$ and $\mathrm{E}adj_t \to 0$ as $t \to \infty$, provided actuarial assumptions are unbiased. Dufresne (1988, 1989) shows that the stochastic pension funding process is stationary in the limit as $t \to \infty$ provided that gains and losses are not spread or amortized over very long periods. Since the payroll and actuarial liability are constant (net of wage inflation), $\lim \mathrm{Var} f_t$ and $\lim \mathrm{Var} c_t$ represent the long-term variances of the contribution rate and funding level respectively.

The following results pertain to the stationary stochastic pension funding process and are obtained by Dufresne (1988) (in the spreading case) and Owadally and Haberman

(1999) (in the amortization case).

Result 1 (Variance of funding level) *The variance of the funding level increases as gains and losses are spread over a longer period. It also increases as gains and losses are amortized over a longer period. Funding levels are less variable under amortization over a fixed term than under spreading over the same term.*

Gains and losses are volatile. Spreading or amortizing them over shorter periods and recognizing them faster ensures that full funded status is reached faster and is maintained. This should improve the security of benefits as confirmed by Result 1.

Result 2 (Variance of contribution rate) *As the period over which gains and losses are spread increases, the variance of the contribution rate initially decreases, attains a minimum (at m_s^*, say) and then increases. The same occurs as the amortization period increases, with a minimum at $m_a^* > m_s^*$. Contribution rates are less stable under amortization over a fixed term $m < m_a^*$ than under spreading over the same period.*

A graph of the variance of contribution rate against spreading or amortization periods exhibits a minimum. It may be thought that spreading or amortizing, and therefore deferring, gains and losses over longer periods leads to a smoother and more stable contribution rate pattern. Result 2 partly contradicts this: for spreading periods longer than m_s^* (or for amortization periods longer than m_a^*), increasing the spreading period (or amortization period) causes increased volatility in the funding level which feeds back as more volatile gains and losses and, since supplementary contributions are required to liquidate the gains and losses, also as more volatile contribution rates.

Result 3 (Spread and Amortization Periods) *Based on the criterion of minimizing the variances of contribution rate and funding levels, an efficient range of spread periods is $[1, m_s^*]$ and an efficient range of amortization periods is $[1, m_a^*]$.*

For spread periods $m > m_s^*$ (and amortization periods $m > m_a^*$), there will always be a shorter period in the efficient range that yields the same variability in the contribution rate together with less variable funding levels. Dufresne (1988) concludes that, under modern economic conditions, spreading gains and losses over a period $m \in [1, 10]$ is efficient. This result has had some influence on actuarial practice in the United Kingdom. See Wilkie in Dufresne (1994, p. viii). Thornton & Wilson (1992) recommend short spreading periods based partly on Dufresne's (1999) conclusion. Owadally & Haberman (1999) conclude that the common North American practice of amortizing gains and losses over 5 years is also efficient. Faster defrayal of gains and losses arising from amendments to benefits or to valuation bases, rather than experience deviations, have also been promoted (Kryvicky, 1981; Colbran, 1982).

Result 4 (Efficiency) *According to the criterion of minimizing the variances of funding level and contribution rate, it is more efficient to spread gains and losses than to amortize them: for any two spread and amortization periods such that the funding level has the same variance, the variance of the contribution rate is less under spreading than under amortization.*

Graphs of the variance of contribution rate against the variance of funding level for spreading and for amortization both exhibit minimum, but the graph for spreading lies under the graph for amortization (except in the trivial case where $m = 1$ and gains and losses are paid off immediately, in which case spreading and amortization coincide).

Compare Result 1 (amortization leads to less variable funding levels than spreading over the same period), Result 2 (spreading leads to less variable contribution rates than amortization over the same period $m < m_a^*$), Result 3 (there exists an efficient spread period range as well as an efficient amortization period range), and Result 4 (overall, spreading is more efficient than amortization).

3 Efficient Amortization: Simulations

3.1 Description of Stochastic Simulations

The results of section 2 are limited in a number of ways. The first limitation concerns the assumption of independent real rates of return on the pension fund from year to year. First, there is considerable debate in the economics literature about serial correlation in equity returns over long horizons (e.g. Fama & French, 1988). Serial correlation is also demonstrated in the actuarial literature by Panjer & Bellhouse (1980) and Wilkie (1995) among others. Second, debt securities are often held to match certain liability cash flows and dependence in the returns from such securities will occur (Vanderhoof, 1973).

Another limitation of the results in the previous section concerns the fact that inflation and different asset classes were not explicitly modeled. Maynard (1992, p. 245) shows that nominal rates of return have been more volatile than rates of return net of price inflation and that this adds to the volatility of funding for benefits that are not indexed with inflation. Modeling separate asset classes is also more realistic than modeling the return on the whole fund and allows for differences in asset allocation.

It is of interest to consider whether Results 1–4 hold when these limitations are relaxed. Stochastic simulations under two different models were carried out for this purpose. Details of these simulations are given in Appendix A.

Simulation Model 1. The model of Dufresne (1988) (see section 2) was simulated, the only difference being that the logarithmic rate of return (net of wage inflation), denoted by δ_t, is projected as a stationary Gaussian autoregressive process of order 1 or AR(1):

$$\delta_{t+1} - \delta = \varphi(\delta_t - \delta) + e_{t+1}, \tag{6}$$

where $\{e_t\}$ is a sequence of zero-mean independent and identically normally distributed variables. The process is stationary from the start and $|\varphi| < 1$. The AR model is used by Panjer & Bellhouse (1980) and Dhaene (1989) among others.

Simulation Model 2. An asset-liability projection basis described by Wilkie (1995) was used to carry out projections for three countries: the United States, the United Kingdom and Canada. The statistical methodology, historical data and economic theory employed by Wilkie (1995) are exhaustively discussed in the actuarial literature (see Wilkie (1995) for relevant references). This projection model was used because it gives a fairly realistic indication of the relationship between various economic variables that are relevant to pension funding.

There are two noteworthy features of the asset-liability projections. First, stochastic inflation on prices and wages is assumed and pensions are a fraction of final salary and are not indexed with inflation. This is more realistic than considering *real* rates of return, as was done in section 2. The distinction between inflation on wages and on prices, which is important for final-salary pensions, is also made in the projections.

Second, two asset classes (equity and long-term Government debt) are considered. A proportional rebalancing strategy between the two asset types is assumed, with income from an asset being reinvested in that asset. McGill *et al.* (1996, p. 665) state that "a 60-40 split between equities and fixed [income securities] is common" in North America and this was assumed for all the projections. 40:60 and 80:20 equity:bond portfolios were also investigated. These portfolios are typical in practice and no suggestion as to their being optimal or otherwise is being made.

3.2 Simulation Results

More details about the simulations are given in Appendix A. The results are discussed here. For model 1 (AR(1) and equation (6)), the scaled standard deviations of the funding level and contribution rate are given in Table C.1 ($\varphi = +0.3$), Table C.2 ($\varphi = +0.5$), and Table C.3 ($\varphi = -0.1$). For model 2 (asset-liability projections), the standard deviations of the funding level and contribution rate at the time horizon of the simulation study are shown in Tables C.4 and C.6 for the U.S. and Canada projections respectively (assuming a 60:40 equity:bond portfolio). The same results for the U.K. projection with an 80:20 equity:bond portfolio are shown in Table C.5.

Variance of the Funding Level (Result 1). The variance of the funding level in Tables C.1–C.6 increases as spreading and amortization periods increase. This is true both for the AR(1) model (including values of φ not tabulated here for the sake of brevity) and for all three countries in the asset-liability simulations. Result 1 appears to be borne out.

Variance of the Contribution Rate (Result 2). Numerical results, not tabulated here for the sake of brevity, indicate that when the rate of return is highly positively (resp. negatively) autocorrelated at lag 1, contribution rate variability increases (resp. decreases) monotonically as both spreading and amortization periods increase. See for example the graphs for $\varphi = +0.8$ and -0.3 in Figure 1. This feature is also reported by Haberman (1994) in the case of spreading.

For AR(1) rates of return that are moderately autocorrelated at lag 1 (Tables C.1–C.3), the variance of contribution rates decreases and then increases as spreading and amortization periods increase, as stated in Result 2. This is also true for all three countries in the asset-liability projections (Tables C.4–C.6). For the U.S. projections in Table C.4, the variance of the contribution rate is minimized at a spreading period of $m_s^* \approx 7$ years and an amortization period of $m_a^* \approx 13$ years. For $m < 13$, contribution rates are higher under amortization than under spreading. The same feature is reproduced in the other Tables. Result 2 thus appears to hold under practical conditions.

Spread and Amortization Periods (Result 3). The simulations therefore indicate that there exists an efficient range of spreading periods and amortization periods under practical conditions (for moderately autocorrelated rates of return). Table C.4 suggests that gains and losses in the U.S. should be spread over periods no longer than 7 years and amortized over no more than 13 years. (The corresponding numbers for the U.K. are 15 and 35, and for Canada they are 13 and 20.) The gain/loss amortization period range of up to 5 years prescribed by the Employee Retirement Income Security Act, 1974, for single-employer pension plans is well within our suggested range, but the range of gain/loss amortization periods for multi-employer plans under ERISA is between 1 and 15 years (source: McGill *et al.*, 1996, p. 597).

Efficiency (Result 4). The variance of the contribution rate is plotted against the variance of the funding level in Figure 1 for the AR(1) model and in Figure 2 for the asset-liability projections. The 'spreading' curve lies below the corresponding 'amortization' curve in all the cases considered. Ignoring country-specific regulatory requirements, we conclude that, for the most stable funding levels and contribution rates, gains and losses should be paid off by being spread rather than amortized. Result 4 appears to hold.

Further Comments. Simulations were also carried out with the logarithmic rate of return process (net of wage inflation) being projected as a stationary Gaussian moving average process of order 1 or MA(1): $\delta_t - \delta = e_t - \phi e_{t-1}$, where $\{e_t\}$ is as in the AR(1) model and $|\phi| < 1$ (for invertibility). The results in the MA(1) case were very similar to those in the AR(1) case. (The detailed results of these simulations are not shown here for the sake of brevity.) For ϕ = -0.5, -0.3, 0, +0.1, +0.3, Results 1–4 all appear to hold. Finally, Wilkie's (1995) economic time series are essentially linear and autoregressive. It is not entirely surprising that similar results are obtained from the AR(1) and the asset-liability simulations. The feature of an efficient range of spreading periods is indeed reproduced by Haberman & Smith (1997) who perform simulations of Wilkie's (1995) time series based on U.K. economic data.

4 Optimal Funding when Returns are Independent

4.1 An Optimization Problem

The efficiency of spreading as compared to direct gain/loss amortization may be explained by the fact that, in the former, the supplementary contribution is proportional to the contemporaneous unfunded liability in the plan (equation (3)). Proportional adjustment represents an optimal form of contribution control when a quadratic measure of the variability of contribution and market value of plan assets is to be minimized. This result is obtained by O'Brien (1987) and Haberman & Sung (1994) assuming random rates of investment return on the plan assets. Boulier *et al.* (1995) obtain a similar result when asset allocation between two assets is also a decision variable and when the pension plan is being valued continuously and indefinitely.

A linear contribution adjustment is also obtained when regular valuations and cash flows occur at discrete intervals and a finite time horizon is assumed. Consider the pension plan model of section 1.2. Assume that the fund is invested in two assets: a risk-less asset earning risk-free rate r and a risky asset earning $r + \alpha_{t+1}$ in year $(t, t+1)$, where α_{t+1} is a random risk premium. Let y_t be the proportion of the fund invested in the risky asset in year $(t,\, t+1)$, and $1 - y_t$ be the proportion invested in the risk-less asset. The arithmetic rate of return on the fund in year $(t,\, t+1)$ is $r + y_t \alpha_{t+1}$. It is further assumed that $\{\alpha_t\}$ is a sequence of independent and identically distributed random variables over time, with mean $\alpha > 0$ and variance σ^2. It is also simpler to disregard inflation at this stage. Since the plan population is assumed to be stationary, the payroll, actuarial liability and benefit payout are constant. The variability of contributions corresponds to the contribution rate variability while the variability of market values of plan assets corresponds to funding level variability.

The pension fund can be considered as a random system,

$$f_{t+1} \;=\; (1 + r + y_t \alpha_{t+1})\,(f_t + c_t - B), \tag{7}$$

where the market value of plan assets f_t is a state variable and c_t and y_t are contribution and asset allocation control variables respectively. B represents the retirement benefits paid out yearly. By virtue of the independence over time of $\{\alpha_t\}$, f_t exhibits the Markov property: $\mathrm{E}[f_{t+1}|W_t] = \mathrm{E}[f_{t+1}|f_t, y_t, c_t]$, where W_t represents information available up to time t.

The objectives of the funding process are to stabilize contributions, defray any unfunded liabilities and pay off actuarial losses and gains as they emerge. The performance of the pension fund may be judged in terms of the deviations in the values of plan assets and contributions from their desired levels (say FT_t and CT_t respectively) relating to the actuarial liability and normal cost. The 'cost' incurred for any such deviation at time $0 \le t \le N-1$ may be defined as

$$C(f_t, c_t, t) \;=\; \theta_1 (f_t - FT_t)^2 \;+\; \theta_2 (c_t - CT_t)^2 \tag{8}$$

Different weights ($\theta_1 > 0$ and $\theta_2 > 0$) are placed on the twin long-term objectives of fund security and contribution stability. The cost in equation (8) reflects a quadratic utility function. Minimizing the cost also minimizes the risks of contribution instability and of fund inadequacy.

The performance of the fund may be given different importance over time. At the end of the given control period N, a closing cost is incurred if an unfunded liability still exists: $\mathbb{C}_N = \theta_0(f_N - FT_N)^2$. The *discounted* cost of deviation occurring t years ahead is $\beta^t C(f_t, c_t, t)$, where β is a time preference rate and $0 < \beta < 1$. For $0 \le t \le N-1$, the discounted cost-to-go or discounted cost incurred from time t to N is

$$\mathbb{C}_t \;=\; \sum_{s=t}^{N-1} \beta^{s-t} C(f_s, c_s, s) \;+\; \beta^{N-t}\mathbb{C}_N. \tag{9}$$

An objective criterion for the performance of the pension funding system over period N may therefore be defined to be

$$\mathrm{E}[\mathbb{C}_0 | W_0] \;=\; \mathrm{E}\left[\beta^N \mathbb{C}_N \;+\; \sum_{s=0}^{N-1} \beta^s C(f_s, c_s, s) \,\middle|\, W_0\right]. \tag{10}$$

Define the value function $J(f_t, t)$ as the minimum, over the remaining asset allocation and contribution decisions, of the expected discounted cost-to-go from time t given information at time t: $J(f_t, t) = \min_\pi \mathrm{E}[\mathbb{C}_t | W_t]$, where $\pi = \{c_t, y_t, c_{t+1}, y_{t+1}, \ldots, c_{N-1}, y_{N-1}\}$. Objective criterion (10) may be minimized using the Bellman optimality principle (see e.g. Bertsekas, 1976): the minimizing values of c_t and y_t (say, c_t^* and y_t^* respectively) in the optimality equation,

$$J(f_t, t) \;=\; \min_{c_t, y_t} \Big\{ C(f_t, c_t, t) \;+\; \beta\, \mathrm{E}[J(f_{t+1}, t+1) \,|\, f_t, c_t, y_t] \Big\}, \tag{11}$$

with boundary condition $J(f_N, N) = \mathbb{C}_N = \theta_0(f_N - FT_N)^2$, are the optimal contribution and asset allocation controls.

The pension planning objectives above were set over a finite period N. The plan is assumed to remain solvent and not discontinued during these N years, so that the funding process does not terminate unexpectedly. When the pension plan is regarded as a going concern, an infinite planning horizon may be usefully envisaged as a reasonable approximation to long-term funding.

No closing cost is incurred in the infinite horizon case. Suppose that the funding process is time-homogeneous, in that the fund and contribution targets are constant. The instantaneous cost at time t is as in equation (8) with $FT_t = FT$, $CT_t = CT$ $\forall t$. The discounted cost-to-go and objective criterion are as in equations (9) and (10) respectively, except that $\mathbb{C}_N = 0$ and an infinite summation of discounted costs is taken. One naturally expects that the control rules that are optimal at time t based on the infinite discounted cost-to-go are also the optimal control rules at times $t+1$, $t+2$,... since the same infinite discounted cost-to-go exists at all times. The dynamic programming algorithm in equation (11) can in fact be shown to converge as $N \rightarrow \infty$ (see Bertsekas, 1976, p. 251) as it involves a contraction mapping and the instantaneous costs in equation (8) are non-negative and are discounted. The optimal contribution and asset allocation over an infinite horizon (say, c_t^∞ and y_t^∞ respectively) are the minimizing values of c_t and y_t in equation (11) when the value function and the control variables are fixed or time-invariant functions of the state variable f_t.

4.2 Solution of the Optimization Problem

It is shown in Appendix B that the solution to the Bellman equation (11) in the finite-horizon case is

$$J(f_t, t) \ = \ P_t f_t^2 - 2Q_t f_t + R_t, \tag{12}$$

where

$$P_t \;=\; \theta_1 + \theta_2\beta\sigma^2(1+r)^2\tilde{P}_{t+1}P_{t+1}, \tag{13}$$

$$\tilde{P}_{t+1} \;=\; [\theta_2(\alpha^2+\sigma^2) + \beta\sigma^2(1+r)^2P_{t+1}]^{-1}, \tag{14}$$

$$Q_t \;=\; \theta_1 FT_t \;+\; \theta_2\beta\sigma^2(1+r)\tilde{P}_{t+1}[Q_{t+1} - P_{t+1}(1+r)(CT_t - B)], \tag{15}$$

with boundary conditions $P_N = \theta_0$ and $Q_N = \theta_0 FT_N$ (with R_t representing some additional terms independent of f_t).

It is also shown in Appendix B that the equivalent solution in the infinite-horizon case is

$$J(f_t) \;=\; Pf_t^2 - 2Qf_t + R, \tag{16}$$

where $P > \theta_1$ is the positive root of the quadratic equation

$$P^2[\beta\sigma^2(1+r)^2] + P[\theta_2(\alpha^2+\sigma^2) - (\theta_1+\theta_2)\beta\sigma^2(1+r)^2] - \theta_1\theta_2(\alpha^2+\sigma^2) \;=\; 0 \tag{17}$$

and $P = \lim_{t\to\infty} P_t$, and where

$$Q = [\theta_1 FT + (P-\theta_1)(B - CT)]P(1+r)/(Pr+\theta_1) \tag{18}$$

and $Q = \lim_{t\to\infty} Q_t$, and R contains terms independent of f_t.

Define $\Theta_t = \theta_2(\alpha^2+\sigma^2)\tilde{P}_t = \theta_2(\alpha^2+\sigma^2)/[\theta_2(\alpha^2+\sigma^2) + \beta\sigma^2(1+r)^2P_t]$ in the finite-horizon setting and correspondingly $\Theta = \theta_2(\alpha^2+\sigma^2)/[\theta_2(\alpha^2+\sigma^2) + \beta\sigma^2(1+r)^2P]$ in the infinite-horizon setting. The following proposition is proven in Appendix B.

Proposition 1 *For $0 \le t \le N-1$ over a finite horizon, the optimal contribution is*

$$c_t^* = \Theta_{t+1}CT_t + (1-\Theta_{t+1})[B - f_t + Q_{t+1}P_{t+1}^{-1}(1+r)^{-1}] \tag{19}$$

and the optimal amount invested in the risky asset is

$$y_t^*[f_t + c_t^* - B] = [Q_{t+1}P_{t+1}^{-1}(1+r)^{-1} - (f_t + c_t^* - B)]\alpha(1+r)(\alpha^2+\sigma^2)^{-1}. \tag{20}$$

The corresponding optimal decisions over an infinite horizon are:

$$c_t^\infty = \Theta CT + (1-\Theta)[B - f_t + QP^{-1}(1+r)^{-1}], \tag{21}$$

$$y_t^\infty[f_t + c_t^\infty - B] = [QP^{-1}(1+r)^{-1} - (f_t + c_t^\infty - B)]\alpha(1+r)(\alpha^2+\sigma^2)^{-1}. \tag{22}$$

4.3 Optimal Funding Decisions

The behavior of the optimal funding decisions as the fund level varies is described in the following proposition, which is proven in Appendix B.

PROPOSITION 2 *The optimal contribution and asset allocation decisions are decreasing functions of the fund level: in the finite-horizon case,* $\partial c_t^*/\partial f_t < 0$ *and* $\partial y_t^*/\partial f_t < 0$*; in the infinite-horizon case,* $\partial c_t^\infty/\partial f_t < 0$ *and* $\partial y_t^\infty/\partial f_t < 0$*.*

Proposition 2 states that the optimal proportion invested in the risky asset decreases as f_t increases, whatever the planning horizon. A similar result is obtained by Boulier *et al.* (1995) and Cairns (1997) for an infinite horizon and for continuous pension plan valuations. The better the investment performance of the risky asset, the better funded the plan is, the more the pension fund should be invested in the risk-less asset. This is a contrarian strategy that entails buying as the market falls and selling as the market rises. It is reasonable in the sense that, first, liabilities need to be hedged so as to minimize the volatility of both surpluses and contributions and, second, any available surpluses should be 'locked in' by being invested in less risky assets (Exley *et al.*, 1997). Conversely, the optimal strategy requires that an underfunded plan takes a riskier investment position than an overfunded plan, all other things equal. For instance, the assets of a poorly funded immature pension plan (with a young membership) could arguably be invested more aggressively in the early years than the assets of a comparable plan with a healthy surplus. The optimal strategy here may be contrasted with portfolio insurance strategies (Black & Jones, 1988) that require riskier investment as the value of plan assets less some minimum value (possibly determined by a solvency requirement) increases. The contrarian strategy is evidently a consequence of the quadratic utility function implied in criterion (8), which is simplistic as it is symmetric and continuous. and does not admit solvency and full funding constraints. The risk that the plan sponsor winds up the plan or defaults on pension obligations was also disregarded.

The optimal contribution is linear in f_t, whatever the planning horizon: from equation (19), c_t^* may be written as $\Lambda_t - (1 - \Theta_{t+1})f_t$, where $1 - \Theta_{t+1} > 0$ while, from equation (21), c_t^∞ may be written as $\Lambda - (1 - \Theta)f_t$, where $1 - \Theta > 0$. The optimal contribution at the start of year $(t,\, t+1)$ is therefore similar to the contribution calculated when gains and losses are spread (equations (1) and (3)) in that they both depend in a decreasing linear way on the *current* market value of assets. This result is based on the assumption of serially independent rates of return and a Markovian funding process. The market value of plan assets represents the *state* of the funding process and, conditional on knowing the current state of funding and current funding decisions, the future evolution of the fund is statistically independent of its past. The optimal contribution is therefore a function of the current state only. This contrasts markedly with the amortization of gains and losses where the contribution (equations (1) and (2)) is a function of unanticipated changes in the state of the funding process over the past m years. This analysis helps to justify the efficiency of spreading over amortization as stated in Result 4 (which was also based on quadratic or second-moment criteria), despite the simplifying assumptions of a quadratic utility function and of zero inflation.

5 Conclusion

The funding of final-salary defined benefit pension plans was considered and methods of amortizing actuarial gains and losses arising from unforeseen economic experience were investigated within simple models. Dufresne (1988) and Owadally & Haberman (1999) show that there exist efficient periods over which to spread or amortize gains and losses Spreading the gains and losses by paying a proportion of the current unfunded liability appears to yield less volatile funding levels and more stable contribution rates than amortizing past and present gains/losses. These results depend on the assumption that rates of return on the fund are independent from year to year. Stochastic projections simulating Gaussian AR(1) and MA(1) logarithmic rates of return indicated that these results are robust when rates of return are moderately autocorrelated. Further simulations of pension plan assets and liabilities based on published actuarial stochastic time series models of equities, long-term Government debt and wage inflation in three separate jurisdictions and based on typical asset portfolios supported the results. Gains and losses in U.S. pension plans should be amortized over no more than 13 years, but more stable funding levels and contribution rates emerge if gains and losses are spread over suggested periods of no more than 7 years. Finally, the efficiency of spreading over amortization was explained by the fact that the optimal contribution, based on a quadratic utility function and irrespective of the optimization period, for a pension fund invested in one risk-free and one random risky asset resembles the contribution calculated when gains and losses are spread. Both are a function of the current level of funding rather than the past and present gains or losses.

Further work should address the fairly constraining modeling assumptions that were made. In particular, gains and losses arising from uncertain mortality, withdrawal and early retirement experience were ignored. Statutory and regulatory requirements were also ignored. Other areas requiring investigation include: considering a more realistic utility function incorporating solvency and full funding constraints; modeling other asset classes, particularly shorter-term and inflation-indexed bonds, and using asset allocation strategies other than proportional rebalancing (e.g. constant proportion portfolio insurance); introducing a dynamic valuation basis possibly based on corporate bond yields; and investigating efficient amortization mechanisms for expensing purposes.

References

Bertsekas, Dimitri P. 1976. *Dynamic Programming and Stochastic Control.* New York: Academic Press.

Black, Fischer and Robert Jones. 1988. "Simplifying portfolio insurance for corporate pension plans," *Journal of Portfolio Management* 14(4): 33–37.

Boulier, Jean-François, Etienne Trussant and Danièle Florens. 1995. "A Dynamic Model for Pension Funds Management," *Proceedings of the 5th AFIR International Colloquium*, Leuven, Belgium, 1: 361–384.

Bowers, Newton L., James C. Hickman and Cecil J. Nesbitt. 1979. "The Dynamics of Pension Funding: Contribution Theory," *Transactions of the Society of Actuaries* 31: 93–122.

Cairns, Andrew J.G. 1997. "A Comparison of Optimal and Dynamic Control Strategies for Continuous-time Pension Fund Models," *Proceedings of the 7th AFIR International Colloquium*, Cairns, Australia, 1: 309–326.

Colbran, Roy B. 1982. "Valuation of Final-salary Pension Schemes," *Journal of the Institute of Actuaries* 109: 359–385.

Dhaene, Jan 1989. "Stochastic Interest Rates and Autoregressive Integrated Moving Average Processes," *ASTIN Bulletin* 19: 131–138.

Dufresne, Daniel 1988. "Moments of Pension Contributions and Fund Levels when Rates of Return are Random," *Journal of the Institute of Actuaries* 115: 535–544.

Dufresne, Daniel 1989. "Stability of Pension Systems when Rates of Return are Random," *Insurance: Mathematics and Economics* 8: 71–76.

Dufresne, Daniel 1994. *Mathématiques des Caisses de Retraite.* Montréal, Canada: Éditions Supremum.

Exley, C. Jonathan, Shyam J.B. Mehta and Andrew D. Smith 1997. "The Financial Theory of Defined Benefit Pension Schemes," *British Actuarial Journal* 3: 835–966.

Fama, Eugene F. and Kenneth R. French 1988. "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics* 22: 3–25.

Haberman, Steven 1994. "Autoregressive Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme," *Insurance: Mathematics and Economics* 14: 219–240.

Haberman, Steven and David Smith. 1997. "Stochastic Investment Modelling and Pension Funding: a Simulation Based Analysis," Actuarial Research Paper No. 102. Department of Actuarial Science and Statistics, The City University, London, England.

Haberman, Steven and Joo-Ho Sung 1994. "Dynamic Approaches to Pension Funding," *Insurance: Mathematics and Economics* 15: 151–162.

Kryvicky, Robert C. 1981. "The Funding of Negotiated Pension Plans," *Transactions of the Society of Actuaries* 33: 405–472.

Maynard, John C. 1992. "Financing Defined-Benefit Pension Plans with Indexed Benefits," *Transactions of the Society of Actuaries* 44: 193–246.

McGill, Dan M., Kyle N. Brown, John J. Haley and Sylvester J. Schieber. 1996. *Fundamentals of Private Pensions,* 7th ed. Philadelphia, Pennsylvania: University of Pennsylvania Press.

O'Brien, Thomas 1987. "A Two-Parameter Family of Pension Contribution Functions and Stochastic Optimization," *Insurance: Mathematics and Economics* 6: 129–134.

Owadally, M. Iqbal and Steven Haberman. 1999. "Pension Fund Dynamics and Gains/Losses Due to Random Rates of Investment Return," *North American Actuarial Journal* 3 (3): 105–117.

Panjer, Harry H. and David R. Bellhouse. 1980. "Stochastic Modelling of Interest Rates with Applications to Life Contingencies," *Journal of Risk and Insurance* 47: 91–110.

Sharp, Keith P. 1993. "Modeling Canadian Price and Wage Inflation," *Proceedings of the Canadian Institute of Actuaries* 24: 565–587.

Thornton, Paul N. and Andrew F. Wilson. 1992. "A Realistic Approach to Pension Funding," *Journal of the Institute of Actuaries* 119: 229–312.

Trowbridge, Charles L. 1952. "Fundamentals of Pension Funding," *Transactions of the Society of Actuaries* 4: 17–43.

Trowbridge, Charles L. and Charles E. Farr. 1976. *The Theory and Practice of Pension Funding.* Homewood, Illinois: Richard D. Irwin.

Vanderhoof, Irwin T. 1973. "Choice and Justification of an Interest Rate," *Transactions of the Society of Actuaries* 25: 417–458.

Wilkie, Andrew D. 1995. "More on a Stochastic Asset Model for Actuarial Use," *British Actuarial Journal* 1: 777–964.

Winklevoss, Howard E. 1982. "Plasm: Pension Liability and Asset Simulation Model," *Journal of Finance* 37: 585–594.

Appendix A

Details of Stochastic Simulations

2000 scenarios are simulated with a time horizon of 300 years each. The randomization routine generates the same set of 300×2000 random numbers so that sampling error does not occur when results are compared. The standard deviations of the funding level and contribution rate at the time horizon are calculated.

Simulation Model 1. For the AR(1) rate of return model, a simple final-salary pension plan as in section 1.2 is assumed. Pensions in payment are assumed to be indexed with economic wage inflation. All quantities may therefore be considered net of wage inflation, and since the pension plan population is stationary, the liability structure of the pension plan is stable in time. The payroll in real dollars (net of wage inflation) is constant. The valuation discount rate (net of wage inflation) is assumed to be 5%. The actuarial liability (AL), normal cost (NC) and yearly benefit outgo (B) (all deflated by wage inflation) are also constant and hold in equilibrium ($B = NC + AL \times 4.76\%$). The standard deviation of the funding level is calculated ($\sqrt{\text{Var}F/AL}$) while the standard deviation of the contribution rate (contribution per payroll dollar) is proportional to $\sqrt{\text{Var}C/NC}$. The arithmetic rate of return (net of wage inflation) in year $(t-1,\, t)$ is $\exp(\delta_t) - 1$ and its mean is 5% (that is, it is equal to the valuation discount rate net of salary inflation) and its standard deviation is 20%. Similar modelling assumptions were made with the MA(1) model to which allusion is made at the end of section 3.2.

Simulation Model 2. In the asset-liability projections, time series models fitted by Wilkie (1995) for equity returns, long bond yields and price inflation in Canada, the United Kingdom and the United States are used. Time series models, fitted by Wilkie (1995) and

Sharp (1993) for economic wage inflation in the UK and Canada respectively, are also used. Assets are valued at market and the Projected Unit Credit method with a time-invariant valuation basis is used to value the liabilities. Contributions are calculated so as to be a level percentage of payroll. The pension fund is invested in two asset classes, equities and long bonds. For the US, payroll increases at a constant 4.5% every year. For the UK, payroll increases in line with Wilkie's (1995) economic wage inflation time series. For Canada, payroll increases in line with Sharp's (1993) model for Canadian wage inflation. Economic valuation assumptions are as follows. For the US and Canada, wage inflation is assumed at 4.5% and a real (net of wage inflation) discount rate of 4.5% is assumed. For the UK, wage inflation is assumed at 6.5% and a real (net of wage inflation) discount rate of 4.5% is assumed. No early retirement and no salary scale was assumed. For the US and Canada, mortality follows the 1983 Group Annuitant Mortality table for males. For the UK, mortality follows English Life Table No. 14 for males. Demographic valuation and projection assumptions are identical. Demographic experience does not deviate from the actuarial valuation basis and gains and losses emerge only as a result of unforeseen economic experience.

Appendix B

Solution of Bellman Equation and Proof of Proposition 1

Proof of Proposition 1: Finite Horizon. Consider first a finite horizon. Let

$$\Phi_t = f_t + c_t - B, \tag{23}$$

$$\Psi_t = 1 + r + y_t \alpha, \tag{24}$$

and note that

$$y_t^2 = \alpha^{-2}\Psi_t^2 - 2\alpha^{-2}(1+r)\Psi_t + \alpha^{-2}(1+r)^2, \tag{25}$$

$$(c_t - CT_t)^2 = \Phi_t^2 - 2(f_t - B + CT_t)\Phi_t + (f_t - B + CT_t)^2. \tag{26}$$

Since α_t is independent and identically distributed over time, it follows from equation (7) that

$$\mathrm{E}[f_{t+1}|f_t] = \Psi_t \Phi_t, \tag{27}$$

$$\mathrm{Var}[f_{t+1}|f_t] = \sigma^2 \Phi_t^2 y_t^2, \tag{28}$$

and using equation (25)

$$\mathrm{E}[f_{t+1}^2|f_t] = (1 + \sigma^2\alpha^{-2})\Phi_t^2\Psi_t^2 \; - \; 2\sigma^2\alpha^{-2}(1+r)\Phi_t^2\Psi_t \; + \; \sigma^2\alpha^{-2}(1+r)^2\Phi_t^2. \tag{29}$$

The Bellman optimality equation (equation (11)) is

$$J(f_t, t) = \min_{c_t, y_t} J, \tag{30}$$

where

$$J = \theta_1(f_t - FT_t)^2 \; + \; \theta_2(c_t - CT_t)^2 \; + \; \beta\, \mathrm{E}[J(f_{t+1}, t+1)|f_t], \tag{31}$$

with boundary condition, at time $t = N$,

$$J(f_N, N) = \theta_0(f_N - FT_N)^2. \tag{32}$$

A trial solution for equation (30) is

$$J(f_t, t) \;=\; P_t f_t^2 - 2Q_t f_t + R_t. \tag{33}$$

The boundary condition in equation (32) certainly satisfies the trial solution, with $P_N = \theta_0$ and $Q_N = \theta_0 F T_N$.

Proceeding by induction, suppose that the right hand side of equation (33) is a solution of the Bellman equation (30) at $t+1$. Then,

$$\begin{aligned}
&\mathrm{E}[J(f_{t+1}, t+1)|f_t] \\
&\qquad = P_{t+1}\mathrm{E}[f_{t+1}^2|f_t] - 2Q_{t+1}\mathrm{E}[f_{t+1}|f_t] + R_{t+1} \\
&\qquad = (1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t^2\Psi_t^2 - 2[Q_{t+1}\Phi_t + \sigma^2\alpha^{-2}(1+r)P_{t+1}\Phi_t^2]\Psi_t \\
&\qquad\qquad + \sigma^2\alpha^{-2}(1+r)^2 P_{t+1}\Phi_t^2 + R_{t+1}.
\end{aligned} \tag{34}$$

where use is made of equations (27) and (29).

J may be written as a quadratic expression in Ψ_t, Φ_t and f_t, by substituting equations (26) and (34) into equation (31):

$$\begin{aligned}
J \;=\; & [\beta(1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t^2]\Psi_t^2 - 2\beta[Q_{t+1}\Phi_t + \sigma^2\alpha^{-2}(1+r)P_{t+1}\Phi_t^2]\Psi_t \\
& + [\theta_2 + \beta\sigma^2\alpha^{-2}(1+r)^2 P_{t+1}]\Phi_t^2 - 2\theta_2(f_t - B + CT_t)\Phi_t \\
& + \beta R_{t+1} + \theta_2(f_t - B + CT_t)^2 + \theta_1(f_t - FT_t)^2.
\end{aligned} \tag{35}$$

Upon completing the squares in Ψ_t and Φ_t,

$$J \;=\; \Psi_t^A(\Psi_t - \Psi_t^B)^2 + \Phi_t^A(\Phi_t - \Phi_t^B)^2 + P_t f_t^2 - 2Q_t f_t + R_t, \tag{36}$$

where P_t and Q_t are as in section 4, R_t represents additional terms independent of f_t, and

$$\Psi_t^A = \beta(1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t^2, \tag{37}$$

$$\Psi_t^B = \frac{\sigma^2\alpha^{-2}(1+r)P_{t+1}\Phi_t + Q_{t+1}}{(1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t}, \tag{38}$$

$$\Phi_t^A = \tilde{P}_{t+1}^{-1}(\alpha^2+\sigma^2)^{-1}, \tag{39}$$

$$\Phi_t^B = \theta_2(\alpha^2+\sigma^2)\tilde{P}_{t+1}(f_t - B + CT_t) + \beta\sigma^2(1+r)\tilde{P}_{t+1}Q_{t+1}], \tag{40}$$

$$\tilde{P}_{t+1} = [\theta_2(\alpha^2+\sigma^2) + \beta\sigma^2(1+r)^2P_{t+1}]^{-1}. \tag{41}$$

J has a unique minimum in Ψ_t and Φ_t provided $\Psi_t^A > 0$ and $\Phi_t^A > 0$. It is sufficient that $P_t > 0$ for $t \in [1, N]$ for both these conditions to be satisfied (since $P_t > 0 \Rightarrow \tilde{P}_t > 0$). (It is assumed that the plan is partially funded and invests in the two assets at all times and $\Phi_t = f_t + c_t - B > 0$.) The minimum occurs when $\Psi_t = \Psi_t^B$ and $\Phi_t = \Phi_t^B$ simultaneously. There is a direct linear relationship between c_t and Φ_t (equation (23)) and between y_t and Ψ_t ($\alpha > 0$ in equation (24)). Therefore, $\min_{c_t, y_t} J = P_t f_t^2 - 2Q_t f_t + R_t$ which is in the form postulated in the trial solution (33). Since the solution holds for $t = N$, it holds for $t \in [1, N]$. Since $P_N = \theta_0 > 0$, then $P_t > 0$ for $t \in [1, N]$ and the sufficient condition for the existence of a single minimum is satisfied.

Let $\Theta_t = \theta_2(\alpha^2+\sigma^2)\tilde{P}_t$. Then, $1 - \Theta_{t+1} = \beta\sigma^2(1+r)^2\tilde{P}_{t+1}P_{t+1}$, from equation (41).

$$\Phi_t^* = \Phi_t^B = \Theta_{t+1}(f_t + CT_t - B) + (1-\Theta_{t+1})Q_{t+1}P_{t+1}^{-1}(1+r)^{-1}. \tag{42}$$

Using equation (23), $c_t^* = \Phi_t^* + B - f_t$ is readily obtained. Replacing $\Phi_t = \Phi_t^*$ in the right hand side of equation $\Psi_t = \Psi_t^B$ gives

$$\Psi_t^* = \frac{Q_{t+1} + \sigma^2\alpha^{-2}(1+r)P_{t+1}\Phi_t^*}{(1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t^*}. \tag{43}$$

Now, $[1 + r + \alpha y_t^*][f_t + c_t^* - B] = \Psi_t^*\Phi_t^* = [Q_{t+1} + \sigma^2\alpha^{-2}(1+r)P_{t+1}(f_t + c_t^* - B)](1 + \sigma^2\alpha^{-2})^{-1}P_{t+1}^{-1}$, from which equation (20) follows.

It is also possible to express y_t^* independently of c_t^* by using equations (24) and (43), giving

$$\alpha y_t^* = \Psi_t^* - (1+r) = \frac{Q_{t+1} - (1+r)P_{t+1}\Phi_t^*}{(1+\sigma^2\alpha^{-2})P_{t+1}\Phi_t^*}, \tag{44}$$

which, upon substitution of Φ_t^* from equation (42), readily yields

$$y_t^* = \frac{\alpha\Theta_{t+1}(1+r)[Q_{t+1} - P_{t+1}(1+r)(f_t + CT_t - B)]}{(\alpha^2+\sigma^2)[(1-\Theta_{t+1})Q_{t+1} + \Theta_{t+1}P_{t+1}(1+r)(f_t + CT_t - B)]}. \tag{45}$$

Proof of Proposition 1: Infinite Horizon. The proof for the infinite-horizon case follows closely the preceding proof for the finite-horizon case and is only sketched.

CT_t and FT_t are constant and may be replaced by CT and FT. The value function does not depend directly on time and $J(f_t, t)$ and $J(f_{t+1}, t+1)$ may be written as $J(f_t)$ and $J(f_{t+1})$. The optimality equation has no boundary condition and backwards induction is not necessary. Since the optimality equation does converge, its solution in the infinite-horizon case is the steady-state or equilibrium solution of the finite-horizon case. A suitable trial solution is $J(f_t) = Pf_t^2 - 2Qf_t + R$, where $P_t \to P$, $Q_t \to Q$ and $R_t \to R$ as $t \to \infty$. All P_t, Q_t and R_t in the preceding proof may be replaced by constant P, Q and R. $\mathrm{E}[J(f_{t+1})|f_t, c_t, y_t]$ and J are as in equations (34) and (35), with the appropriate constant terms described above. J may be simplified to a quadratic in Ψ_t and Φ_t. It turns out that J may be minimized, leaving a quadratic in f_t, thereby confirming the trial solution.

$$\begin{aligned} J \;=\; & \Psi^A(\Psi_t - \Psi^B)^2 + \Phi^A(\Phi_t - \Phi^B)^2 + f_t^2[\theta_1 + \theta_2\beta\sigma^2(1+r)^2\tilde{P}P] \\ & - 2f_t[\theta_1 FT + \theta_2\beta\sigma^2(1+r)P(Q - (1+r)P(CT - B))] + R, \end{aligned} \tag{46}$$

where

$$\tilde{P} = [\theta_2(\alpha^2 + \sigma^2) + \beta\sigma^2(1+r)^2 P]^{-1}, \tag{47}$$

$$\Psi^A = \beta(1 + \sigma^2\alpha^{-2})P\Phi_t^2, \tag{48}$$

$$\Psi^B = \frac{\sigma^2\alpha^{-2}(1+r)P\Phi_t + Q}{(1+\sigma^2\alpha^{-2})P\Phi_t}, \tag{49}$$

$$\Phi^A = \tilde{P}^{-1}(\alpha^2 + \sigma^2)^{-1}, \tag{50}$$

$$\Phi^B = \theta_2(\alpha^2 + \sigma^2)\tilde{P}(f_t - B + CT) + \beta\sigma^2(1+r)\tilde{P}Q], \tag{51}$$

and R includes additional terms independent of f_t and not required here.

By comparing the coefficients of f_t^2 and f_t in equation (46) with those in the trial solution, it is clear that P satisfies $P = \theta_1 + \theta_2\beta\sigma^2(1+r)^2\tilde{P}P$, which may be rewritten as equation (17), while Q satisfies $Q = \theta_1 FT + \theta_2\beta\sigma^2(1+r)\tilde{P}[Q - (1+r)P(CT - B)]$. This may be simplified into the form given in equation (18) by noting that $\theta_2\beta\sigma^2(1+r)\tilde{P} = (P - \theta_1)/(P + Pr)$ given that $P = \theta_1 + \theta_2\beta\sigma^2(1+r)^2\tilde{P}P$.

J in equation (46) has a unique minimum in Ψ_t and Φ_t provided that $\Psi^A > 0$ and $\Phi^A > 0$. It is sufficient that $P > 0$ for both these conditions to hold. (The plan is assumed to be partially funded at all times and $\Phi_t > 0$.) Since the coefficient of P^2 and the constant term in the quadratic equation (17) are positive and negative respectively, P must have one negative real root and one positive real root, the latter being the admissible solution. Then, $\tilde{P} > 0$ in equation (47), and $P > \theta_1$ since $P = \theta_1 + \theta_2\beta\sigma^2(1+r)^2\tilde{P}P$. J is minimized when $\Psi_t = \Psi^A$ and $\Phi_t = \Phi^A$ simultaneously and, exploiting again the direct linear relationship between Φ_t and c_t and between Ψ_t and y_t, the minimizing values of c_t and y_t, denoted respectively by c_t^∞ in equation (21) and by y_t^∞ in equation (22), may be found as in the finite-horizon case. Again, y_t^∞ may also be written as

$$y_t^\infty = \frac{\alpha\Theta(1+r)[Q - P(1+r)(f_t + CT - B)]}{(\alpha^2 + \sigma^2)[(1-\Theta)Q + \Theta P(1+r)(f_t + CT - B)]}. \tag{52}$$

Proof of Proposition 2. Since $P_N = \theta_0 > 0$, backward recursion in the Riccati difference equation for P_t formed by equations (13) and (14) shows that $P_t > 0$, $\tilde{P}_t > 0$ for $t \in [1, N]$. Clearly, $0 < \Theta_t < 1$ for $t \in [1, N]$. It is reasonable to assume that the fund and contribution targets in any year are such that $FT_t > 0$ and $CT_t < B$ (as otherwise there is no sense to funding in advance for retirement benefits). Then, $Q_t > 0$ for $t \in [1, N]$ from equation (15) and the boundary condition $Q_N = \theta_0 FT_N > 0$. In the infinite-horizon case, $P > \theta_1 > 0$ and clearly $0 < \Theta < 1$. With $FT > 0$ and $B > CT$, it follows that $Q > 0$ (equation (18)).

It is immediately observed that $\partial c_t^* / \partial f_t < 0$ from equation (19). From equation (20), it is easy to show that $\partial y_t^* / \partial [f_t + c_t^* - B]$ is directly proportional to $-Q_{t+1} P_{t+1}^{-1} (1+r)^{-1}$ which is negative. Since $\partial [f_t + c_t^* - B] / \partial f_t = \Theta_{t+1} > 0$, it follows that $\partial y_t^* / \partial f_t < 0$. Likewise, $\partial c_t^\infty / \partial f_t < 0$ and $\partial y_t^\infty / \partial f_t < 0$.

Appendix C

Tables

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	19.1%	19.1%	95.26%	95.26%
3	34.6%	30.5%	61.24%	75.83%
$m_s^* \approx 5$	51.0%	41.2%	54.77%	67.08%
$m_a^* \approx 7$	69.3%	52.0%	56.57%	61.24%
10	109.5%	64.8%	67.08%	63.25%
15	273.9%	96.4%	122.47%	70.71%
20	‡	148.3%	‡	89.44%
25	‡	214.5%	‡	111.80%

Table C.1: AR(1) logarithmic rates of return, $\varphi = +0.3$. ‡ indicates nonstationarity.

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	19.1%	19.1%	95.26%	95.26%
2	31.3%	27.4%	80.62%	90.83%
$m_s^* \approx 3$	43.6%	34.6%	77.46%	88.03%
4	57.4%	44.7%	79.06%	86.60%
$m_a^* \approx 5$	74.2%	52.9%	82.16%	85.15%
6	97.5%	61.6%	90.83%	86.60%
7	122.5%	70.7%	104.88%	88.03%
8	167.3%	81.9%	122.47%	94.87%

Table C.2: AR(1) logarithmic rates of return, $\varphi = +0.5$.

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	19.1%	19.1%	95.26%	95.26%
3	24.5%	23.5%	43.01%	54.77%
5	30.7%	27.6%	33.91%	43.87%
$m_s^* \approx 10$	44.7%	37.4%	27.84%	34.28%
15	54.8%	44.7%	28.28%	31.62%
$m_a^* \approx 20$	80.6%	53.9%	30.82%	31.22%
25	104.9%	63.2%	34.64%	32.02%
30	130.4%	72.1%	40.62%	33.17%

Table C.3: AR(1) logarithmic rates of return, $\varphi = -0.1$.

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	15.7%	15.7%	42.46%	42.46%
3	23.2%	22.1%	21.79%	28.87%
5	29.8%	27.1%	17.45%	23.54%
$m_s^* \approx 7$	37.0%	30.8%	16.12%	21.04%
10	52.0%	35.9%	16.84%	17.84%
$m_a^* \approx 13$	75.9%	42.0%	20.04%	17.40%
15	99.9%	46.4%	23.74%	17.41%
17	135.7%	51.4%	29.57%	17.56%
20	235.0%	60.8%	45.97%	19.00%

Table C.4: U.S. projections with a 60:40 equity:bond portfolio.

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	16.7%	16.7%	29.67%	29.67%
3	20.3%	20.1%	12.57%	16.82%
5	23.9%	22.9%	9.31%	12.83%
10	32.4%	28.0%	7.00%	8.83%
$m_s^* \approx 15$	41.1%	33.0%	6.52%	7.74%
20	51.3%	37.7%	6.68%	7.16%
25	64.3%	42.3%	7.28%	7.01%
30	81.2%	47.0%	8.29%	6.97%
$m_a^* \approx 35$	102.2%	51.8%	9.65%	7.11%
40	127.4%	57.0%	11.33%	7.41%
45	156.5%	62.3%	13.29%	7.46%
50	188.9%	67.6%	15.47%	8.09%
60	258.2%	83.8%	20.09%	10.25%
70	323.9%	111.1%	24.40%	11.84%

Table C.5: U.K. projections with an 80:20 equity:bond portfolio.

m	Standard Deviation			
	Funding Level		Contribution Rate	
	Spreading	Amortization	Spreading	Amortization
1	14.0%	14.0%	37.55%	37.55%
3	18.0%	17.6%	16.81%	22.50%
5	21.7%	20.5%	12.69%	17.47%
10	30.5%	25.6%	9.87%	12.57%
$m_s^* \approx 13$	36.3%	28.5%	9.62%	11.46%
15	40.7%	30.4%	9.71%	11.02%
17	45.4%	32.3%	9.94%	10.66%
$m_a^* \approx 20$	53.3%	35.0%	10.46%	10.29%
23	62.0%	38.1%	11.16%	10.64%
25	68.3%	40.3%	11.71%	10.69%
30	85.7%	46.9%	13.30%	11.21%

Table C.6: Canada projections with a 60:40 equity:bond portfolio.

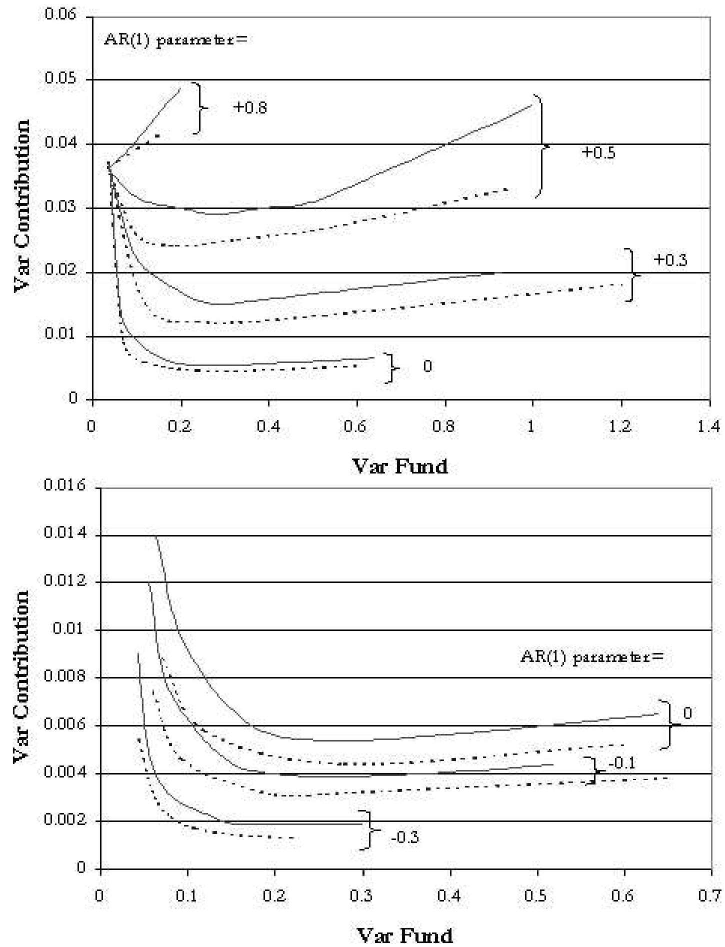
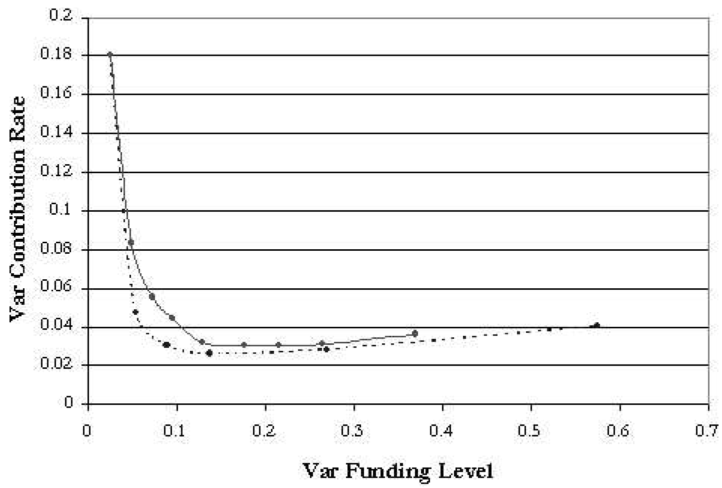
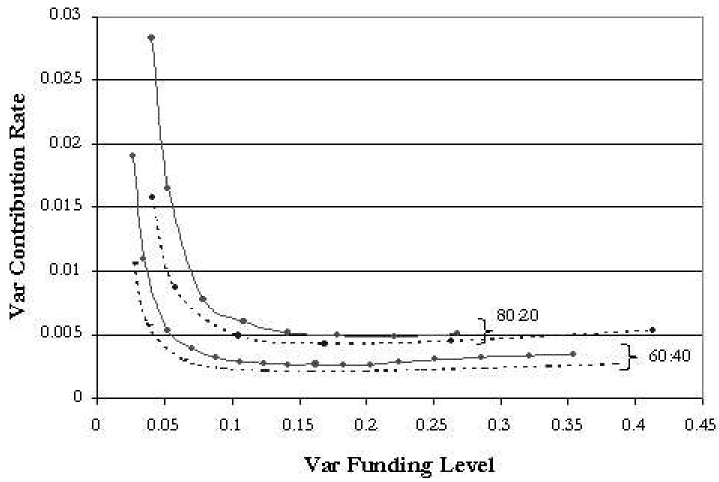


Figure 1: Projections with AR(1) rates of return for various values of φ. Spreading (dashed lines) is more efficient than amortization (solid lines).

US:



UK:



Canada:

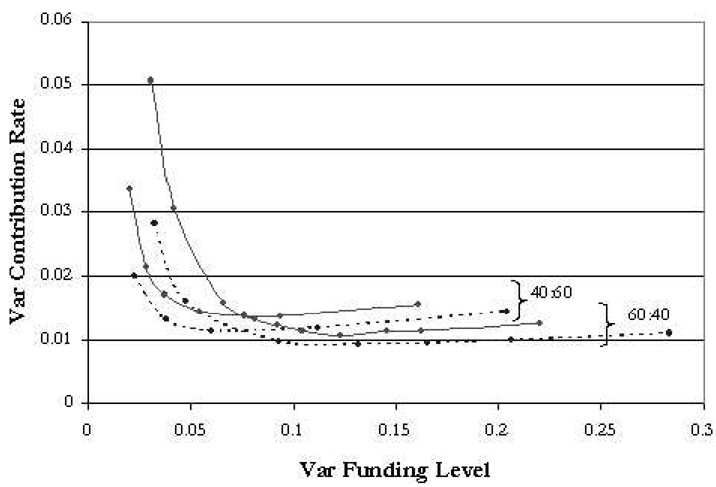


Figure 2: Projections for the US (60:40 equity:bond portfolio), the UK (60:40 and 80:20) and Canada (60:40 and 40:60). Spreading (dashed lines) is more efficient than amortization (solid lines).