ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 2

Some Aspects of Statement of Financial Accounting Standards No. 87

by Daniel Dufresne

Research paper presented to the

Society of Actuaries

for Fellowship credit

February 1993

ABSTRACT

This paper focuses on two aspects of Statement of Financial Accounting Standards 87:

(1) the variability of the discount rate, and its consequences; and

(2) the "corridor" approach to gains and losses amortization.

The analysis centers on the variability of pension expense over time. A simplified model is used; its main features are a stationary population and random fluctuations of discount rates and returns on assets. The work is carried out mathematically and with the help of computer simulations. Sensitivity analyses are performed with respect to the parameters of the model (variance of discount rates, variance of returns on assets, length of amortization period and width of corridor). It turns out (under the particular scenario chosen) that the variance of discount rates and the length of the amortization period are the most important determinants of the variability of expense. It is also shown that when the plan population is mature the variability of pension expense is mostly due to the amortization of (accounting) gains and losses, and not to the sensitivity of the service cost to the discount rate.

PREFACE

This paper focuses on two aspects of Statement of Financial Accounting Standards 87 (SFAS 87):

(1) the variability of the discount rate, and its consequences; and

(2) the "corridor" approach to gains and losses amortization.

The analysis centers on the variability of pension expense over time. The work is carried out mathematically and with the help of computer simulations. The ultimate goal of the study is to give a better understanding of (1) and (2) above. It is the author's belief that deterministic case studies (assuming parameters to be constant over time) are not sufficient. The consequences of (1) and (2) appear more clearly when fluctuations of some of the parameters are taken into account. This is why randomness was introduced.

The degree of complexity of a mathematical model is a function of the number of variables involved. Of particular importance is the number of random factors taken into account: if this number is too large, the results may well become impossible to interpret. In the case at hand, it was decided that only the discount rate and the rate of return on the fund's assets would be random. Other factors (e.g. mortality) are supposed static.

Chapter 1 describes the model chosen. Concepts from control theory are also introduced. Chapter 2 deals with the existence of stationary (or steady-state) limits for the stochastic processes considered. Chapter 3 presents the results of the computer simulations performed.

It is hoped that this study will be of interest to those involved in pension accounting and funding. The framework described herein may help in making accounting or funding decisions, for instance in choosing an amortization period or method. Another important use would be to assess the effects on pension plans at large of changing funding or accounting rules. An example of the latter would be to try to answer the following question: "Suppose FASB were to allow a 15% corridor for gains and losses amortization, instead of the current 10% corridor; would this have a significant effect on pension expense (for plans switching from the old to the new minimum amortization requirement)?" Based on the numerical results shown in Section 3.5, it appears that pension expense would be unchanged on average (which makes sense intuitively). Fluctuations over time would be somewhat affected: under the "base scenario" employed (see Section 1.4), the standard deviation would move from 10.90 to 10.13 (a 7% decrease). The change is small, and suggests that allowing a 15% corridor may not by itself bring a significant decrease in the variability of pension expense. A more refined analysis would be required before a definite conclusion can be reached, since the numbers quoted are the results obtained under just *one* scenario. Nevertheless, in the absence of exact mathematical formulas for the variance of pension expense, the methodology suggested should be helpful in studying this type of problem. How the model should be used is discussed further in the Conclusion.

N.B. Unless specifically referenced, all mathematical and numerical results are original and have not, to the author's knowledge, appeared previously.

ACKNOWLEDGMENTS

The author would like to thank the following people for their help: Mr. Lucien Pouliot, FSA, who supervised the research; the Research Paper Committee; and the referees, Mr. Eric Lofgren, FSA, and Professor Arnold Shapiro, FSA.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 DESCRIPTION OF MODEL

1.1. General considerations

Two distinctive features of SFAS 87 are at the center of this study: (1) the variability of the discount rate, and (2) the "corridor" approach to gains and losses amortization. They are described below.

(1) The discount rate is the accounting counterpart of the valuation rate of interest used in pension funding. It directly influences the values of the projected benefit obligation and of the service cost. The Board's requirements are set out in paragraph 44 of the Statement:

> 44. Assumed discount rates shall reflect the rates at which the pension benefits could be effectively settled. It is appropriate in estimating those rates to look to available information about rates implicit in current prices of annuity contracts that could be used to effect settlement of the obligation (including information about available annuity rates currently published by the Pension Benefit Guaranty Corporation). In making those estimates, employers may also look to rates of return on high-quality fixed-income investments currently available and expected to be available during the period to maturity of the pension benefits. Assumed discount rates are used in measurements of the projected, accumulated, and vested benefit oblitations and the service and interest cost components of net periodic pension cost.

The discount rate will therefore vary from year to year, causing fluctuations in the pension benefit obligation and in the service cost.

(2) The minimum requirement for amortization of gains and losses is described in paragraph 32 of the Statement:

32. As a minimum, amortization of an unrecognized net gain or loss (excluding asset gains and losses not yet reflected in market-related value) shall be included as a component of net pension cost for a year if, as of the beginning of the year, that unrecognized net gain or loss exceeds 10 percent of the greater of the projected benefit obligation or the market-related value of plan assets. If amortization is required, the minimum amortization shall be that excess divided by the average remaining service period of active employees expected to receive benefits under the plan. If all or almost all of a plan's participants are inactive, the average remaining life expectancy of the inactive participants shall be used instead of average remaining service.

No amortization is required as long as unrecognized gains and losses (URL) do not exceed 10% of the maximum of the projected benefit obligation (PBO) and the value of assets (F). In other words, the interval ± 10% max (PBO, F) acts as a "corridor" inside which gains and losses need not be recognized. When URL drifts out of the corridor, the amount of gains or losses to be recognized is the excess

$$|URL| - .10 \max(PBO, F)$$

divided by the average remaining service period of active employees. It is important to observe that no schedule of payments is set up, as would be the case with the transition obligation or prior service costs. The whole exercise is done anew every year, no reference being made to amounts previously recognized (other than the fact that they reduced URL).

Two other features of SFAS 87 deserve mention. First, gains and losses include the effect of changes in assumptions (paragraph 29). Thus an important source of gains and losses is the instability of the PBO caused by the discount rate changing from year to year. Second, pension expense being an end-of-year amount, it must include interest to the end of the year. This is done as follows. The service cost is first increased with interest on the PBO, calculated using the discount rate; it is then decreased with the return on plan assets, which is obtained using the "expected long-term rate of return" assumption (not the actual rate of return on assets). The difference between actual and expected return on assets becomes part of gains and losses, and is not eligible for amortization until the following year.

All in all, the components of pension expense (E) are:

(a) Service cost

(b) Interest cost

(c) Return on plan assets

(d) Amortization of gains or losses

(e) Amortization of unrecognized prior service cost

(f) Amortization of transition obligation (or asset).

In order to simplify the model, it is assumed that there is no transition asset or obligation, and that the plan remains the same over time. Items (e) and (f) above are therefore nil. Furthermore, the population of members and retirees is supposed not to change over time.

The evolution of discount rates has to be considered very carefully. Of course a constant discount rate is out of question. Another possibility would be to use past financial data (e.g. rates of return on fixed-income securities) to generate a sequence of discount rates to be used in the model. This approach has a number of advantages, but lacks the flexibility required to perform sensitivity analyses (e.g. how does one "increase the volatility" of discount rates?). It was decided to let the discount rate vary randomly over time. More precisely, discount rates will be modelled using an autoregressive process. This use of stochastic processes to represent financial parameters deserves a few words of explanation.

The author does not believe that discount rates, or rates of return, are drawn out of a hat every month or quarter. On the contrary, most people will agree that those rates are the outcome of decisions made by a large number of economic agents and also, no doubt, of technological breakthroughs, the weather, epidemics, earthquakes, and so on. The part played by "randomness" is difficult to ascertain; one of the main difficulties lies in defining randomness itself. (In relation to this question, a common objection to the quantitative analysis of financial data is "arbitrary political decisions", e.g. a sudden increase in tax rates. Insofar as those decisions are indeed arbitrary, either in their amplitude or their timing, they could in fact be viewed as *supporting* the random hypothesis, rather than contradicting it.)

In the case at hand, the most compelling arguments in favour of stochastic modelling of interest rates are (i) that they have fluctuated in the past, and (ii) that future interest rates can only be partially predicted (in a statistical sense) from previous ones. Whether this uncertainty arises out of "pure randomness", or whether it is ultimately caused by "chaotic" behaviour of a completely deterministic (but yet unknown) system, the conclusion is the same: some way of generating unpredictable changes has to be devised. Stochastic processes are a convenient (and defensible) solution.

Each possible set of rules for calculating pension amounts (pension expense, funded status, ...) transforms the stochastic processes representing discount rates and rates of return into new stochastic processes. This yields a very flexible way of studying these sets of rules.

The choice of the stochastic processes representing the inputs (discount rates and rates of return) is not an easy one. This study uses autoregressive processes, which allow dependence between time periods. These processes have been widely used to describe interest rates, but it is not clear that they best fit historical data. This should not be a major problem, since the research did not aim at representing interest rates as closely as possible, but rather at quantifying how fluctuations in interest rates translate into fluctuations in pension amounts.

Thus, pension expense and unrecognized losses become the outputs of the "pension system". What transforms the inputs (discount rates and rates of return on assets) into the outputs are the methods used to fund and account for the pension plan. The results of the research are the effect on the outputs of (1) varying the inputs (e.g. changing the variance of the discount rates) and (2) modifying the accounting rules (e.g. widening the corridor, or spreading gains/losses over a shorter period).

Sections 1.2 to 1.4 explain the framework of analysis in greater detail.

1.2. The meaning of stationary distributions

The approach taken in studying SFAS 87 is the following: generate discount rates and rates of return possessing specified distributions (described in Section 1.4) to simulate the evolution of the pension "system" over a very long period. The fluctuations of E and URL over time can then be measured and compared for different sets of hypotheses. This section tries to justify the approach chosen, and explains the concept of ergodicity, which is of great theoretical importance in the present context.

"A very long period" refers here to a period so long that initial conditions (i.e. funded status, unrecognized gains or losses, ... at the *start* of the simulation) have no influence on the statistics obtained. For the sake of brevity, let us call the period "infinite"; this is correct mathematically but, naturally, computer simulations could only be performed over a finite period. One reason, perhaps the most important, for choosing an infinite time horizon is convenience: fluctuations over shorter periods would depend on the particular set of initial conditions chosen; using an infinite period removes this dependence.

Of course in concrete cases initial conditions *are* an essential part of the problem, and short time horizons (e.g. 10 or 20 years) are a natural choice. It would be interesting to complete the data given in Chapter 3 by showing the progression of distributions over the first, say, 20 years, given some set of initial conditions.

For the subsequent discussion, it is useful to define the concepts of time averages and ensemble averages. Suppose pension expense is simulated over a 20-year period, yielding value $E_{t,i}$ at time t for the ith simulation run. *Time averages* are computed by letting the time parameter vary, e.g.

$$E_{\cdot,i} = \frac{1}{20}\sum_{t=1}^{20} E_{t,i}.$$

Ensemble averages are computed at a specified point in time and with respect to the distribution at that time; for example

$$E_{t,\cdot} = \frac{1}{M}\sum_{i=1}^{M} E_{t,i}$$

approaches the mean of E_t, if the number of runs M is large enough. In these expressions, the word "average" refers not just to mean values, but more generally to averages of any function of the variable considered.

It is easy to see that, in general, time and ensemble averages do not convey the same information. A simple example will nevertheless be given, in order to shed some light on an important point. Suppose a certain sum is to be invested in a special fund with a fixed rate of return. The rate is decided on once and for all at the time of the investment, but is unknown at present. Assume that, from whatever previous experience, a probability distribution has been obtained for R, the said rate of return, and that the distribution of annual returns for the next 20 years is printed out (really 20 copies of the same distribution). Examining that sequence of distributions, a person unaware of the way they were obtained might conclude that, over time, the returns will be moving up and down between the limits of the distribution, with a greater frequency around its modes, etc. But this would be incorrect: once $R = r$ is decided upon, it's fixed and there are no fluctuations at all. The annual distributions *per se* would not show the total dependence between returns in different years. Although this example is a little artificial, the same reasoning applies to E and URL: ensemble distributions at specified points in time (e.g. $t = 5$ or 10) would *not* adequately show fluctuations over time.

An alternative, not adopted here, would be to simulate the distribution of E and URL for next year (time 1), given particular values of E and URL at time 0. Performing this for a wide enough class of initial states would show how likely fluctuations are over time. This approach appears feasible, but is definitely more complicated than the one chosen here. Numerical results would certainly be more difficult to obtain and harder to interpret.

Take another example. Suppose some phenomenon $\{X_0, X_1, \ldots\}$ is modelled by a uniform distribution on the integers $0, \ldots, 9$,

$$P(X_t = k) = \frac{1}{10}, \qquad 0 \le k \le 9.$$

It is also assumed that the X's are independent. One possible sequence of X's is

$$\begin{array}{cccccccccc} 2 & 4 & 1 & 4 & 8 & 9 & 1 & 1 & 9 & 2 \\ 4 & 0 & 8 & 9 & 6 & 2 & 5 & 0 & 4 & 0 \end{array}$$

(these random numbers were generated by my calculator). The characteristics of this realization of $\{X_t, 1 \leq t \leq 20\}$ may be expressed with the usual statistics: mean, variance, etc. We find

$$\bar{x}_{20} = \frac{1}{20}\sum_{i=1}^{20} x_i = 3.95$$

$$s_{20}^2 = \frac{1}{20}\sum_{i=1}^{20}(x_i - \bar{x}_{20})^2 = 9.9475.$$

These time averages are also called the empirical mean and variance, because they are based on observations only. The ensemble (or theoretical) mean and variance are based on the probability distribution:

$$EX = \sum_{k=0}^{9} k \cdot 1/10 = 4.5$$

$$\text{Var } X = \sum_{k=0}^{9}(k - EX)^2 \cdot 1/10 = 8.25.$$

The Law of Large Numbers states that the empirical mean will approach the theoretical mean as the number of observations increases; in symbols, $\bar{x}_n \to EX$ as $n \to \infty$. In fact, the Law of Large Numbers says a lot more. It also implies that $s_n^2 \to \text{Var } X$; furthermore, let

$$\begin{aligned} Y_t &= 1 \quad \text{if} \quad X_t = 9 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

We have $\bar{y}_{20} = 3/20 = .15$. Since $EY_t = P(Y_t = 9)$, we know that $\bar{y}_n \to P(Y_t = 9) = .10$ as $n \to \infty$. More generally, it can be said that *the empirical distribution approaches the theoretical distribution as the number of observations increases.* In other words, time and ensemble averages are the same provided the time period is infinite.

This property of a sequence of variables is called ***ergodicity***. For any ergodic sequence, the proportion of the time the observations are in a certain region A always approaches, in the long run, the theoretical probability that one observation lies in region A.

For an ergodic process, therefore, the distribution at one point in time already tells a lot concerning the behaviour over time of the process. A sequence of independent and identically distributed random variables is always ergodic. Sequences of dependent variables may or may not be ergodic. The autoregressive sequence

$$X_t = aX_{t-1} + e_t, \qquad |a| < 1, \tag{1.1}$$

where $\{e_t\}$ are independent and possess the same normal distribution, is ergodic. But the sequence

$$X_t = R, \qquad t = 1, 2, \ldots,$$

is not ergodic. Indeed this implies $\bar{x}_n = r$ for all n, and this does not converge to ER, except by chance. This will be recognized as the investment example given at the beginning of this section.

It will be shown in Chapter 2 that the stochastic processes representing pension amounts (E and URL) are ergodic provided (i) some stationarity conditions are satisfied and (ii) realistic assumptions are made regarding the treatment of gains and losses. Conditions (i) are simply that the underlying population, rates of return on assets, discount rates, etc., be ergodic themselves. Requirement (ii) ensures that unrecognized gains or losses do not become "too large" over time (if there is no amortization of gains or losses then $|URL_t|$ drifts to $+\infty$ as $t \to \infty$).

The meaning of the stationary distribution is somewhat similar to that of the probability of ruin in risk theory. The latter is the probability of insolvency if the current portfolio of insurance policies is replicated *ad infinitum*. No one really believes insurance portfolio will last forever, but the probability of ruin nevertheless measures the risk associated with a particular premium/liability structure. In pension accounting, the stationary distribution of (say) pension expense reflects its potential variability, given a particular set of economic/actuarial assumptions and accounting rules.

[*N.B.* Notwithstanding what precedes, the parallel between the model presented here and classical risk theory is subject to serious limitations. For one thing, in the classical risk theoretic model (e.g. Bowers *et al*, 1986) the surplus is the excess of premiums over claims in the period $[0, t]$. Premiums are never adjusted, and are greater than average claims. Hence (1) the amount of surplus, as well as claims experience, have no effect on premiums, and (2) the surplus becomes infinite with probability one as time goes to infinity. In the present model pension expense (corresponding to premiums) *is* adjusted through gains/losses amortization. This implies that unrecognized gains/losses (corresponding to surplus) remain bounded over time. All the amounts calculated reach a steady-state distribution as time passes; this is not the case in the classical risk theoretic model. The model presented here is closer in spirit to the ones described in Beard *et al* (1984, chapters 8 and 10) and Martin-Löf (1983).]

The ergodicity of the processes (E_t) and(URL_t) is studied in depth in Chapter 2. The results show that under the assumptions made the processes are ergodic. One consequence is that initial conditions have no effect on time averages over an infinite period. We are therefore justified in believing that the numerical results shown in Chapter 3 are independent of the values of E, URL, etc. at the start of the simulations.

1.3. Interpretation according to control theory

Jacobs (1974, p. v) describes control theory in the following terms:

> Control theory is a branch of applied mathematics devoted to analysis and design of control systems. Control systems are systems in which a controller interacts with a real process in order to influence its behaviour. A primary objective for most control systems is to make some real variable take a desired value, for example to regulate the temperature of an oven or to make the direction of a receiving aerial track a moving target. The objective is usually to be achieved by adjusting some other variable, such as heat input to the oven or force applied to the aerial, although the response to such adjustments in most real controlled processes is neither instantaneous nor certain. The non-instantaneous response is accounted for by regarding the controlling and controlled variables as input and output of a *dynamic* system described by differential or difference equations. The effect of uncertainties is reduced by using *feedback* to provide the controller with continuous indication of what adjustment is needed; for example, if the oven is too cold more heat must be supplied and if the aerial points to the left of its target it must be forced to turn to the right.

Control theory was initially developed in relation to engineering problems. Over the past thirty years or so it has also been applied to biology and economics. There have been a few attempts to apply the theory to insurance (Balzer and Benjamin, 1980; Balzer, 1982; Martin-Löf, 1983; Smith, 1984) and pension funding (Benjamin, 1989; Dufresne, 1986a, 1991).

Control theory should be distinguished from a more recent development known as optimal control theory. The latter aims at finding the best control of a system given a performance criterion or "cost function". This theory, which will not be discussed here, was applied to pension funding by O'Brien (1986, 1987) and to insurance by Vandebroek and Dhaene (1991).

This section will formulate some pension funding and accounting problems in the language of control theory. A simple example will be used to illustrate the concepts.

In pension funding, the variable to be controlled is the fund level or, to be more precise, the unfunded liability (UL). The desired value of UL would usually be zero. The control used is contributions; once again it might be more correct to say that UL is controlled by the amount by which total contributions exceed the normal cost.

An important feature of pension funding, from the point of view of control theory, is that the behaviour of the control is itself of importance. For instance, amortizing gains and losses over one year will certainly keep UL on target, but then the control applied may become unacceptable (e.g. it may fluctuate too much, or it may be too large in some years). Intuitively, one imagines that if gains/losses are amortized over a longer period, then the contributions fluctuate less, although the unfunded liability may fluctuate more. This "trade-off" has been examined in Dufresne (1986a, 1986b, 1988, 1989, 1991); a similar phenomenon occurs in pension accounting (see Section 3.6).

Variable rates of return on plan assets are in effect "multiplicative" disturbances (UL grows at a variable rate); this complicates the equations describing pension funding systems. Pension accounting also has its own complexities. For those reasons, a simpler system will be used to illustrate the concepts of control theory. Let

$$X_{t+1} = X_t + C_t + D_{t+1}, \tag{1.2}$$

$$C_t = \text{control at time } t.$$

Here X is the variable to be controlled, and $(D_t, t = 1, 2, \dots)$ are the disturbances affecting the system. Eq. (1.2) could represent the evolution of the unfunded liability (X) if the following assumptions were made: the valuation rate of interest is zero, $-C_t$ is the payment made towards amortization of the unfunded liability (that is to say in excess of the normal cost), and D_{t+1} is the actuarial loss experienced during the year $(t, t+1)$. In pension accounting, X could stand for unrecognized losses.

Suppose X has target value zero, and that the control applied is

$$C_t = -kX_t, \qquad 0 < k < 1. \tag{1.3}$$

This is called "proportional" control, since C is a fraction of the difference between the target and the actual value of X. The minus sign on the right hand side of Eq. (1.3) is not surprising: C and X must have opposite signs if X is to be steered towards zero (hence the name "negative feedback" applied to such controls).

The Preliminary Views on pension accounting issued by FASB in 1982 (i.e. prior to SFAS 87) included a "spreading" of gains and losses over the average future working lifetime ($AFWL$) of members. This is an example of proportional control, with $k = 1/AFWL$. As was explained in Section 1.1, SFAS 87 transformed this into a minimum requirement, to be applied only when URL exceeds 10% of the greater of PBO and F. The current minimum requirement is thus a modified form of proportional control.

Let us examine the response of system (1.2) when control (1.3) is applied, for various inputs D_t. First let us suppose that $X_0 = 0$, and that there is only one disturbance, occurring at time 1 : $D_1 = L$, $D_t = 0$ for $t \neq 1$. We find

$$\begin{aligned} X_1 &= L \\ X_{t+1} &= (1-k)X_t, \quad t \geq 1 \\ \Rightarrow \quad X_t &= (1-k)^{t-1}L. \end{aligned}$$

The state of the system decreases geometrically to 0. The system never completely gets rid of the disturbance.

Next, consider disturbances which are constant at every step: $D_t = D$ for all $t \geq 1$. Then

$$\begin{aligned} X_t &= D[1 + \cdots + (1-k)^{t-1}] \\ &= \frac{D}{k}[1 - (1-k)^t]. \end{aligned}$$

A first observation is that X_t remains bounded (i.e. does not become arbitrarily large) even though disturbances of equal magnitude and direction are experienced at each step. Another observation is that the steady-state value of X_t is not zero,

$$\lim_{t \to \infty} X_t = \frac{D}{k}.$$

The fraction k has a magnifying effect on this value. If the control amounts to $k = 10\%$ of the current state of the system, then the latter will eventually settle at $1/.10 = 10$ times the value of the disturbance. With $k = .5$, the ultimate level of X is $D/.5 = 2D$ only.

Now suppose disturbances are random. In actual applications the stochastic process representing the disturbances may be quite complex. For illustrative purposes, however, it is simplest to assume that the variables (D_t) have a common distribution with mean 0 and variance $\sigma^2 > 0$, and are independent, i.e. for any $s \neq t$

$$\text{Prob}(D_s \in A,\ D_t \in B) = \text{Prob}(D_s \in A) \cdot \text{Prob}(D_t \in B)$$

for all sets A and B.

It is essential, in the first place, to observe that although it is assumed that they are zero on average, the disturbances do not "cancel out" over time. It is wrong to think that "the law of averages implies the sum of the disturbances will approach zero as time increases". Any one who has simulated a pension fund with no amortization of gains/losses will have noticed this. What the Law of Large Numbers says is that the *average* disturbance approaches zero:

$$\lim_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} D_s = ED_1 = 0$$

(refer to the discussion of ergodicity in the previous section). In fact, what really happens to the sums of the disturbances is this:

$$\lim_{t\to\infty} \max(D_1, \sum_{1}^{2} D_s, \dots, \sum_{1}^{t} D_s) = +\infty$$

$$\lim_{t\to\infty} \min(D_1, \sum_{1}^{2} D_s, \dots, \sum_{1}^{t} D_s) = -\infty$$

The proof of these statements is beyond the scope of this paper, but a rough intuitive justification can be given. Imagine that $D_t \sim \mathrm{N}(0, \sigma^2)$ ("D_t is normally distributed with mean zero and variance σ^2"). Then

$$\sum_{s=1}^{t} D_s \sim \mathrm{N}(0, t\sigma^2)$$

and so

$$\mathrm{Prob}\Big(\sum_{s=1}^{t} D_s > 10^{10}\Big) = \mathrm{Prob}\Big(Z > \frac{10^{10}}{\sigma\sqrt{t}}\Big).$$

where $Z \sim \mathrm{N}(0,1)$. As t increases, this probability approaches $1/2$. This makes it plausible that, given enough time, the sum of the disturbances will become as large as we choose (the same argument works for negative values).

In order to "stabilize" X some form of control has to be applied. The proportional control (1.3) implies

$$X_{t+1} = (1-k)X_t + D_{t+1}.$$

X_{t+1} is a fraction $(1-k)$ of X_t, plus the new disturbance D_{t+1}. Let us assume that at time 0 the system is on target ($X_0 = 0$). Then

$$
\begin{aligned}
EX &= (1-k)\cdot 0 + ED_1 = 0 \\
EX_2 &= (1-k)EX_1 + ED_2 = 0 \\
&\vdots \\
EX_t &= 0 \quad , t \geq 0.
\end{aligned}
$$

The disturbances being zero on average implies the same for X_t. Now turn to variances:

$$X_1 = D_1, \quad X_2 = D_2 + (1-k)D_1, \ldots,$$

$$X_t = D_t + (1-k)D_{t-1} + \cdots + (1-k)^{t-1}D_1$$

$$
\begin{aligned}
\Rightarrow \operatorname{Var} X_t &= \sigma^2[1 + (1-k)^2 + \cdots + (1-k)^{2t-2}] \\
&= \sigma^2 \frac{1-(1-k)^{2t}}{1-(1-k)^2}.
\end{aligned}
$$

Since $0 < k < 1$, $\operatorname{Var} X_t$ increases from 0 to

$$\lim_{t\to\infty} \operatorname{Var} X_t = \frac{\sigma^2}{1-(1-k)^2} = \frac{\sigma^2}{k(2-k)}. \tag{1.4}$$

Compare this with the uncontrolled system:

$$
\begin{aligned}
\operatorname{Var} X_t &= \operatorname{Var}(D_1 + \cdots + D_t) \\
&= t\sigma^2 \\
&\to \infty \quad \text{as } t \to \infty.
\end{aligned}
$$

A proportional control therefore brings a stable steady-state response, with mean 0 and variance given by (1.4). Once again it is seen that the fraction

k acts as a magnifier, but this time of the variance of X. If $k = 10\%$, the variance of X is $1/(.10 \cdot 1.9) = 5.3$ times that of the disturbances. If $k = 50\%$, the multiplier is only 1.3.

In many situations, including pension funding and accounting, it is also important to look at the behaviour of the control C_t. We get:

$$EC_t = E(-kX_t) = 0$$

$$\begin{aligned} \text{Var } C_t &= k^2 \text{Var } X_t \\ &= \frac{\sigma^2 k^2 [1-(1-k)^{2t}]}{1-(1-k)^2} \end{aligned}$$

$$\Rightarrow \lim_{t \to \infty} \text{Var } C_t = \sigma^2 \frac{k}{2-k}.$$

As a function of k, Var C_∞ behaves otherwise than Var X_∞. The variance of the control is directly proportional to k; it can be made artibrarily small by choosing k small enough.

Table 1.1 shows the limit standard deviations of X_t and C_t as t tends to infinity, for selected values of k in the interval $[0, 1]$. [*N.B.* The standard deviation is the square root of the variance. The variance is the average *squared* deviation from the mean, and its units are those of X squared — if X is in dollars, Var X is in dollars squared. Here it is perhaps better to deal with standard deviations, which are in the same units as X and C.]

Table 1.1 shows that there is a trade-off between the variance of the state of the system and the variance of the control applied: to increase k makes Var C_∞ larger, and Var X_∞ smaller. When both variances are considered, no single value of k appears better than the others.

In pension funding, X represents the unfunded liability and $-C$ the payments made to liquidate it; the disturbance D_t is the total actuarial loss arising in year $(t-1, t)$. Intuitively it makes sense that to liquidate the unfunded liability more rapidly (i.e. to increase k) has a destabilizing effect

k	$\sqrt{\text{Var } X_\infty}/\sigma$	$\sqrt{\text{Var } C_\infty}/\sigma$
0	∞	–
.1	2.294	.229
.2	1.667	.333
.3	1.400	.420
.4	1.250	.500
.5	1.155	.577
.6	1.091	.655
.7	1.048	.734
.8	1.021	.816
.9	1.005	.905
1.0	1.000	1.000

Table 1.1. Limits of standard deviations of X_t and C_t when t tends to ∞, as multiples of σ. (X_t is the state of the system, C_t the control applied, σ^2 the variance of the disturbances (D_t), k is the fraction of X fed back into the system.)

on contributions and the reverse effect on the unfunded liability. This is another way of interpreting the numbers in Table 1.1. Nevertheless there is a significant difference in pension funding, caused by *interest*. Consider the system described above but with interest at constant rate $i > 0$:

$$X_{t+1} = (1+i)(X_t + C_t) + D_{t+1}, \tag{1.5}$$

$$C_t = -k\, X_t.$$

k	$\sqrt{\text{Var } V_\infty}/\sigma$	$\sqrt{\text{Var } C_\infty}/\sigma$
0	∞	–
.0196 (= d)	∞	∞
.025	9.548	.239
.030	6.886	.207
.035	5.666	.198
.03883(= k^*)	5.075	.19706
.04	4.929	.19714
.06	3.520	.211
.08	2.894	.232
.10	2.522	.252
.20	1.730	.346
.40	1.264	.506

Table 1.2. Limits of standard deviations of X_t and C_t when t tends to ∞, as multiples of σ. The interest rate is $i = .02$.

Table 1.2 shows the standard deviations of X and C as functions of k when $i = .02$. [*N.B.*. Since we are dealing with limits as $t \to \infty$, it is better to use deflated (constant) monetary values; i is therefore net of wage increases.] For values of k greater than $k^* = .03883$, there is the same trade-off as before, that is to say to increase k decreases Var X but increases Var C. But for $k < k^*$ things are radically different: to decrease k increases *both* Var X and Var C. This means that any k smaller than k^* is to be rejected, since it implies values of Var X and Var C which are both larger than they would be for $k = k^*$. It is shown in Appendix 1.1 that

$$\begin{aligned} k^* &= d(1+v) \\ &\approx 2d \end{aligned}$$

where $v = 1/(1+i)$, $d = i/(1+i)$. In other words, in order to reduce variances it is necessary to pay at least twice the interest on X. Observe that when $i = .02$

$$k^* = .03883 = 1/\ddot{a}_{\overline{35.5}|}.$$

Variances are reduced if the unfunded liability is "spread" over a period shorter than 36 years. This maximum period is a function of i, but not of σ (see Appendix 1.1). For example, it is 18.2 years when $i = .04$, and 12.4 years when $i = .06$. It can be shown that

$$m^* \approx \log 2/\log(1+i) + .5,$$

which indicates that the maximum period is a decreasing function of i.

In the above example the disturbances were *additive* (see Eq. (1.5)). A more realistic model would include random rates of return on pension assets, which are *multiplicative* disturbances. The situation then becomes a little different, e.g. k^* depends on the distribution of the rates of return. The interested reader is referred to Dufresne (1988, 1989, 1991) for more details on this topic.

This section illustrated the type of analysis which will be performed in Chapter 3. The problems studied are not as tractable as the simple one described above; this is why computer simulations have to be relied on.

1.4. The model

The description of the model has been devided into three parts:

1. Population and pension plan
2. Economic and actuarial assumptions
3. Equations describing the system

1.4.1. Population and pension plan

In order to focus more accurately on fluctuations caused by rates of return and discount factors, a stationary (i.e. static) population has been chosen. Of course real pension plan populations are at best only approximately stationary, but including population fluctuations would only serve to cloud up the analysis.

The two factors influencing the choice of the population are (1) simplicity and (2) sensitivity to changes in the discount rate (which is the valuation rate of interest for accounting purposes). The first requirement is that numerical results be easy to check or reproduce. Since this study is not aimed at a particular real pension plan, the population chosen does not have to be "realistic", as long as the numbers obtained conform to what would be expected in real situations. The only way the population comes into play is in the computation of the actuarial liability and normal actuarial cost (funding), and projected benefit obligation and service cost (accounting). Of crucial importance is the ***sensitivity*** of these values to changes in valuation rates of interest (or discount rates).

The same considerations apply to the choice of the pension plan terms. Accordingly the simplest plan was chosen: one retirement age, benefits equal to a fixed fraction of final salary, payable for life.

The assumptions concerning the plan and population are shown below

Benefits	$b\%$ of salary per year of service payable annually for life
Entry age	30
Retirement age	65
Pre-retirement mortality	none
Post-retirement mortality	$q_x = 0,\ x < 79,\ q_{79} = 1$

(*N.B.* The constant b will be set equal to a computationally convenient value whenever needed; of course this does not affect the conclusions reached.)

Appendix 1.2 shows that the sensitivities of the actuarial values calculated on the above basis are comparable to those obtained if the population is stationary and in accordance with the 1983 Group Annuity Mortality Table.

The assumption of no mortality before retirement would be appropriate if each member were to withdraw his entire actuarial liability on termination before retirement. This is approximately true when the plan offers full portability. In any case, Appendix 1.2 proves a rather surprising fact: if (i) the plan has no pre-retirement benefits; (ii) the population is in accordance with the life table at all ages before retirement (this is the case with a stationary population); and (iii) the valuation method is projected unit credit, then *pre-retirement decrements have no influence on the sensitivities* of the service cost and benefit obligation to changes in the discount rate.

The post-retirement mortality table chosen greatly simplifies calculations and programming. With a life expectation around 15 or 20 years at age 65, the sensitivity of $\ddot{a}_{65}$ should compare with that of $\ddot{a}_{\overline{15}|}$. This is examined further in Appendix 1.2.

1.4.2. Economic and actuarial assumptions

The following have to be specified:

- funding:
  - actuarial cost method and assumptions
  - treatment of gains and losses
  - rates of return on plan assets
- expensing:
  - expected long-term rate of return on plan assets
  - discount rates.

The plan will be funded according to the projected unit credit method (another method could have been chosen). The mortality table used for valuation purposes (funding and expensing) is the one described in Subsection 1.4.1.

Now turn to the valuation interest rate for funding purposes. Rates of return on plan assets (described below) fluctuate around a fixed mean value. In practice the valuation rates of interest follow earned rates of return to some extent, though a long-term approach is taken, meaning that valuation rates usually vary little from one year to the next. It was decided to let the valuation rate of interest be constant and equal to the long-term average earned rate. This may seem questionable, but is in complete agreement with the ideas expressed in Section 1.2. The goal of the work is to study the variability of pension amounts caused by some of the requirements of SFAS 87; it is essential, as a first step, to remove other sources of fluctuations which could make results harder to interpret. A subsequent step (not taken here) would be to introduce valuation rates of interest which are based on, say, the last 10 years' experience.

No explicit inflation assumption is required, since rates of return on assets and discount rates are net of wage increases and benefits are supposed fully indexed. This means, for instance, that in a final salary plan retirement benefits are indexed in line with wage increases.

The treatment of (funding) gains and losses has an effect on the value of the fund and therefore influences pension expense. There are many ways of amortizing gains and losses; none of them seemed more appropriate than the others. The simulations were done on the assumption that gains and losses are recognized immediately (recall that fund returns are the only source of gains and losses on the funding side). Intuitively, one expects that slower recognition of gains/losses would bring a small increase in the variability of funded status and pension expense. It would be interesting to look into this more closely.

Given that the valuation rate of interest (for funding purposes) is a constant, the expected long term rate of return on plan assets (for expensing purposes) was also held constant.

Before turning to discount rates and rates of return on plan assets, let me first remark on vocabulary. If the value of some asset grows by a factor

$$U = 1 + R = e^G$$

during a certain period, R will be called the *arithmetic* rate of return, and G the *geometric* rate of return. These terms are more descriptive than the usual "rate" and "force of interest". They also agree with the vocabulary used in other disciplines. (The expression "force of interest" probably comes from "force of mortality"; both expressions refer to continuous models, as these rates are instantaneous. Furthermore, when speaking of rates of return, and not rates of interest, it feels awkward to use the expression "force of return".)

Paragraph 44 of SFAS 87, quoted in Section 1.1, gives the requirements concerning the discount rate. Since rates implicit in annuity contracts and

PBGC rates are not easy to model, the author chose to focus on high-quality bond yields. There is no single accepted model describing bond yields (or bond rates of return) over time. Very sophisticated models have been devised, for example including transition probabilities between different yield curves, or other "functional processes" specifying yield curves over time. For the purpose of this research it was thought sufficient to model the discount rate directly, i.e. without using a model describing the evolution of the whole yield curve over time. A sequence of independent random variables would have been convenient, but historical bond yields do not appear to follow this kind of process. The approach taken in Panjer and Bellhouse (1980) was retained. Those authors fitted autoregressive (AR) processes to a number of financial series. The results were AR processes of order one or two. I chose to use an AR(1) process

$$X_t = M + A(X_t - M) + V_t$$

with a parameter $A = .9$. In the cases of Standard and Poor's Composite yield on high grade corporate bonds, Panjer and Bellhouse had found $A = .957$. A mean (arithmetic) discount rate of 1% was chosen after examining long-term US bond yields, deflated by wage increases, over the period 1926-1988 (Tables 9A and 11A of *Economic Statistics for Pension Actuaries*, 1990). The geometric net rates have an average of 0.71% and a standard deviation of 4.96%, which mean an arithmetic average of about

$$\exp(.0071 + .5 \cdot .0496^2) - 1 = 0.84\%.$$

The situation is less clear with respect to rates of return on assets. Typically, pension fund assets are not invested for the most part in high quality bonds, so that the above approach is not appropriate. Pension funds often invest significant amounts in stocks. A number of authors have contended that the rates at which individual stock prices grow (over time) are independent random variabales (e.g. Fontaine, 1990). True or false, this does not really help us, as pension funds are managed and, therefore, stocks are not kept indefinitely, but bought and eventually sold.

The approach adopted here is based on two considerations:

(1) that there should be a dependence between discount rates and rates of return on the fund's assets; and

(2) that, besides the dependence stated in (1), the additional randomness present in the sequence of rates of return on assets should result from random variables which are independent over time.

The idea is that part of the fund's assets have returns similar to those of long-term bonds, while the rest of the fund has rates of return which are completely unpredictable. The rates of return on plan assets are thus less predictable than are discount rates.

The processes representing discount rates and rates of return on assets will now be described. The required inputs are:

EDS	mean arithmetic discount rate
ER	mean arithmetic rate of return on assets
$VARDS$	variance of geometric discount rates
$VAROR$	variance of geometric rates of return on assets
COR	correlation coefficient of geometric discount rates and rates of return on assets.

The above parameters are sufficient to completely specify the two processes. Define:

$DSCR$	arithmetic discount rate
X	geometric discount rate = $\log(1 + DSCR)$
R	arithmetic rate of return on assets
Y	geometric rate of return on assets = $\log(1 + R)$
HX	mean geometric discount rate
HY	mean geometric rate of return on assets
$(V_t), (W_t)$	two independent sequences of independent $\mathbf{N}(0, 1)$ random variables.

Geometric discount rates and rates of return on assets satisfy the following equations:

$$X_t = HX + A(X_{t-1} - HX) + B \cdot V_t$$
$$Y_t = HY + D(X_t - HX) + G \cdot W_t$$

for $t \geq 1$. In Chapter 3 the parameter A will be set equal to .9. Whatever the value of X_0, as t increases the average value of X_t approaches HX, since $EV_t = 0$. (X_t) is an AR(1) process, but (Y_t) is not. Y_t is a combination of X_t and W_t. The mean of Y_t approaches HY as t increases.

The variances of geometric discount rates satisfy

$$\text{Var } X_t = A^2 \text{ Var } X_{t-1} + B^2$$

and thus

$$\lim_{t \to \infty} \text{Var } X_t = B^2/(1 - A^2)$$
$$\lim_{t \to \infty} \text{Var } Y_t = D^2 B^2/(1 - A^2) + G^2.$$

From now on, suppose the process (X_t) is stationary. From what has just been said, this implies that $X_t \sim \mathrm{N}(HX, B^2/(1-A^2))$, and $Y_t \sim \mathrm{N}(HY, D^2B^2/(1-A^2)+G^2)$ for all t. Then

$$\begin{aligned}
\mathrm{Cov}(X_t, Y_t) &= E(X_t - HX)(Y_t - HY) \\
&= E[D(X_t - HX)^2 + G \cdot W_t(X_t - HX)] \\
&= D \ \mathrm{Var}\ X_t
\end{aligned}$$

and thus

$$\begin{aligned}
\mathrm{Corr}(X_t, Y_t) &= \mathrm{Cov}(X_t, Y_t)/(\mathrm{Var}\ X_t\ \ \mathrm{Var}\ Y_t)^{1/2} \\
&= D(\mathrm{Var}\ X_t/\mathrm{Var}\ Y_t)^{1/2} \\
&= D \cdot B/[D^2B^2 + G^2(1-A^2)]^{1/2}
\end{aligned}$$

(B is chosen positive). The inputs EDS, ER, $VARDS$, $VAROR$ and COR completely specify the parameters B, D, G, HX and HY. The correlation COR can take any value between -1 and $+1$. It is shown in Appendix 1.3 that

$$\begin{aligned}
B &= [(1-A^2)VARDS]^{1/2} \\
D &= (VAROR/VARDS)^{1/2}COR \\
G &= [VAROR(1-COR^2)]^{1/2} \\
HX &= \log(1+EDS) - (1/2)VARDS \\
HY &= \log(1+ER) - (1/2)VAROR.
\end{aligned}$$

	Mean	Standard Deviation
1. Long-term US Government Bond Yields (nominal)	4.95%	2.96%
2. Wage Index (annual rate of increase)	4.24%	4.31%
3. Series 1 deflated by Series 2	0.71%	4.96%
4. Standard and Poor's Composite Composite Value Index (nominal)	9.23%	20.10%
5. Series 4 deflated by Series 2	5.00%	20.20%

Table 1.3. Means and standard deviations of five economic series. All rates are geometric. Each series covers the period 1926-1988. (Computed by author based on Tables 9A, 11A and 15A of *Economic Statistics for Pension Actuaries*, 1990.)

The "base" scenario is shown below.

Valuation rate of interest (arithmetic)	.02
Expected long-term rate of return on plan assets (arithmetic)	.02
Mean arithmetic discount rate (EDS)	.01
Variance of geometric discount rate ($VARDS$)	.0009
Mean arithmetic rate of return on assets (ER)	.02
Variance of geometric rate of return on assets ($VAROR$)	.0025
Correlation between geometric discount rate and rate of return (COR)	.60

Although some guidance was sought from published statistics (see Table 1.3), these assumptions are not meant to portray any historical period accurately.

As was explained previously, the mean arithmetic discount rate of 1% is approximately equivalent to the values on the third line of Table 1.3. The variance of discount rates was set equal to .0009 (i.e. a standard deviation of 3%), which is smaller than the observed value of $.0496^2$. As will be seen in Chapter 3, the volatibility of pension amounts is very sensitive to this parameter, and it is plausible that plan sponsors will try to limit the fluctuations of the discount rates used.

Now turn to rates of return on assets. Many pension funds are heavily invested in stocks and bonds, with other assets playing a relatively minor role. Corporate bond yields are higher than those of US Government bonds, while stocks have returns which, in the long run, clearly surpass those of bonds. This would suggest a mean rate of return on pension fund assets of around $3 - 4\%$ (net of wage increases). However, to take into account the "prudence" exercised in pension fund investments, a mean rate of 2% only was chosen. For the same reason the standard deviation of rates of return was set at 5%, which is low considering that stock returns historically show a 20% standard deviation.

The valuation rate of interest and the long-term expected rate of return were set equal to the mean rate of return on assets, although those three are probably seldom equal in practice.

The model requires the correlation between discount rates and returns on assets. It would have been possible to use US statistics to try to estimate this parameter; it was quicker to use the correlation coefficient of bond returns and total pension fund returns, shown on p. 25 of the *Report on Canadian Economic Statistics* (1990). The value given is .61, which was rounded to .60.

As a final remark on the base scenario, observe that one of its distinctive features is probably the fact that the average discount rate is lower than the valuation rate of interest (1% versus 2%). The reverse situation is probably often seen in practice, especially when the valuation rate of interest is chosen conservatively. The base scenario assumes that the latter is a best estimate of the long-term returns of the fund (since it equals ER). (Of course in practice the actuary only has a small number of observations to work with, and could not estimate the long-term average rate of return with perfect accuracy, as is supposed here.)

1.4.3. Notation and mathematical analysis

The following notation will be used throughout.

$AFWL$	average future working lifetime of active employees
AL	actuarial liability (funding)
AM	amortization of gains and losses (expensing)
B	annual (aggregate) benefit payments (constant)
C	contributions (funding)
COR	correlation coefficient of geometric discount rates and rates of return on assets (see Subsection 1.4.2)
$DSCR$	arithmetic discount rate (expensing)
E	pension expense
EDS	mean arithmetic discount rate (see Subsection 1.4.2)
$ELTR$	expected long-term rate of return on plan assets
ER	mean arithmetic rate of return on assets (see Subsection 1.4.2)
F	fund value
HX	mean geometric discount rate (see Subsection 1.4.2)

HY	mean geometric rate of return on assets (see Subsection 1.4.2)
k	$1/AFWL$
L	loss (gain if negative) during one fiscal year (expensing)
NC	normal actuarial cost (funding)
PBO	projected benefit obligation (expensing)
R	arithmetic rate of return on assets
SC	service cost (expensing)
URL	unrecognized losses (gains)
V	sequence of independent variables used in defining $DSCR$ and R (see Subsection 1.4.2)
$VARDS$	variance of geometric discount rate (see Subsection 1.4.2)
$VAROR$	variance of geometric rate of return on assets (see Subsection 1.4.2)
VI	valuation rate of interest (funding)
X	geometric discount rate (see Subsection 1.4.2)
Y	geometric rate of return on assets (see Subsection 1.4.2)
W	sequence of independent variables used in defining R (see Subsection 1.4.2)

Some of the equations below are given in Berin and Lofgren (1987), though with a different notation.

As explained before, the population is static, and all amounts are deflated by increases in benefits. Annual benefit outgo is therefore constant. Assets are valued at year-end. Contributions and benefits are paid in totality at the beginning of the year.

Fund values fluctuate solely because rates of return on assets vary over time. We have

$$F_t = (1 + R_t)(F_{t-1} + C_{t-1} - B).$$

Under the base scenario, gains/losses are amortized over one year, i.e.

$$C_t = NC + AL - F_t.$$

and

$$\begin{aligned} F_t &= (1 + R_t)(AL + NC - B) \\ &= AL(1 + R_t)/(1 + VI) \end{aligned}$$

since, under static conditions

$$AL = (1 + VI)(AL + NC - B).$$

The pension benefit obligation (PBO) shown in the financial statement at time t is the projected unit credit liability valued at $DSCR_t$. By contrast, the service cost (SC) in the financial statement at time t is computed at the beginning of the year, i.e. at time $t-1$.

Unrecognized losses (URL) are set equal to 0 at time 0. Afterwards, URL is increased with emerging losses (or decreased with emerging gains) and decreased by the amount recognized in expense:

$$URL_t = URL_{t-1} + L_t - AM_t.$$

Here

$$\begin{aligned} L_t &= PBO_t - PBO_{t-1} + (ELTR - R_t)(F_{t-1} + C_{t-1} - B) \\ &= PBO_t - PBO_{t-1} + F_t(ELTR - R_t)/(1 + R_t) \end{aligned}$$

The first part of this expression is the unexpected increase of the PBO during the year. In the present model this loss (gain) is entirely due to the fact that $DSCR_t \neq DSCR_{t-1}$. The second part is the interest loss on assets during year $t-1$ to t. It is the difference between the projected and the actual fund at time t. One notable feature of SFAS 87 is that pension expense (a year-end amount) includes interest on PBO_{t-1} at rate $DSCR_{t-1}$, but is reduced by return on assets at rate $ELTR$. The difference between projected and actual return on assets in pushed into the loss for that year, and is therefore not eligible for amortization until the following year. Observe that in the model chosen contributions and benefit payments are paid at the beginning of the year, and thus get a full year's interest.

Let

$$M = \max(PBO_{t-1}, F_{t-1}).$$

M is the greater of the PBO and the fund at the beginning of year $(t-1, t)$. It is this amount which is compared to unrecognized losses (also at the beginning of the year) in order to determine the minimum amount to be recognized at time t:

$$\begin{aligned} AM_t &= 0, && \text{if} \quad |URL_{t-1}| \leq 10\%M \\ &= k(URL_{t-1} - 10\%M), && \text{if} \quad URL_{t-1} > 10\%M \\ &= k(URL_{t-1} + 10\%M), && \text{if} \quad URL_{t-1} < -10\%M \end{aligned}$$

(see Paragraph 32 of the Statement, reproduced at the beginning of Section 1.1). Here k is the reciprocal of the average future working lifetime of active employees.

The model excludes prior service cost and transition obligation. Pension expense therefore consists in

– the service cost,

– plus interest on PBO,

– minus return on assets,

– plus recognition of part of URL, if any.

In symbols

$$E_t = SC_{t-1}(1 + DSCR_{t-1}) + DSCR_{t-1}(PBO_{t-1} - B) \\ - ELTR(F_{t-1} - B) + AM_t. \tag{1.6}$$

($N.B.$ Benefits are paid at the beginning of the year, and thus get a full year's interest. Including interest on benefits at two different rates may seem a little strange; see Section 3.2 for an explanation.)

The service cost (with interest to the end of the year) is typically very sensitive to changes in $DSCR$. The same thing can be said about the PBO. This source of volatility of pension expense has been widely recognized (see for example paragraph 182 of the Statement). The two remaining terms in Eq. (1.6) also cause volatility in pension expense. In particular, AM may bring about important fluctuations (whether or not the minimum amortization described above is applied); this is because losses include the inexpected increase/decrease of the PBO, which can be quite large by comparison to other losses (mortality, terminations, etc.).

The part of the right hand side of Eq. (1.6) which directly depends on the discount rate is

$$Q = SC(1 + DSCR) + DSCR(PBO - B).$$

It is instructive to try to analyse mathematically the sensitivity of Q to changes in $DSCR$. Of course it is difficult to do this in general, as the results obviously depend on plan benefit and demographic profile. The remainder of this subsection examines three specific cases in some detail.

First, suppose the population is stationary (as will be done in Chapters 2 and 3). The equation of equilibrium

$$PBO_{t-1} = (1 + DSCR_{t-1})(PBO_{t-1} + SC_{t-1} - B)$$

implies

$$0 = (1 + DSCR_{t-1})SC_{t-1} + DSCR_{t-1}(PBO_{t-1} - B) - B.$$

Thus $Q = B$ and

$$E_t = B - ELTR(F_{t-1} - B) + AM_t.$$

In a stationary population, service cost plus interest on PBO (minus benefits) is always precisely equal to benefits. The latter quantity is not sensitive at all to changes in the discount rate. The sensitivities of the service cost, on one hand, and interest on the PBO (minus benefits), on the other, cancel out completely. In this case the volatility of pension expense is entirely due to AM_t.

REMARK. That Q does not depend on $DSCR$ when the population is stationary can be given another (more intuitive) explanation. Consider an unfunded plan with a stationary population. The cost of the plan can be calculated in either of two ways. First, it is simply the cost of benefits paid; in our model, this is $(1 + DSCR)B$, at the end of the year. Second, suppose any valuation method is used. The cost (at the end of the year) is the service cost (with interest) plus interest on the pension benefit obligation. These two costs have to be equal, i.e.

$$(1 + DSCR)B = (1 + DSCR)SC + DSCR \cdot PBO$$

$$\Longleftrightarrow B = (1 + DSCR)SC + DSCR(PBO - B). \qquad \square$$

As a second example, suppose that the active population conforms exactly to the life table, but that retirees are separated from the active lives. Use primes to denote this new situation, and no primes for the stationary case above. Define

$$\begin{aligned} RPBO &= \text{retirees' pension benefit obligation} \\ &= \text{actuarial present value of retirees' benefits.} \end{aligned}$$

Then $SC' = SC$, $PBO' = PBO - RPBO$ and

$$\begin{aligned} &SC'(1 + DSCR) + DSCR \cdot PBO' \\ &\qquad = SC(1 + DSCR) + DSCR \cdot PBO - DSCR \cdot RPBO \\ &\qquad = B(1 + DSCR) - DSCR \cdot RPBO. \end{aligned}$$

Now suppose r is the retirement age, and p the annual benefit paid to one retiree. Then

$$(1 + DSCR)(RPBO - B) = RPBO - p\ell_r\ddot{a}_r.$$

(this is an equilibrium equation for the retired population only; see Dufresne (1986a), p. 84). This implies

$$DSCR \cdot RPBO = (1 + DSCR)B - p\ell r\ddot{a}_r$$

and

$$SC'(1 + DSCR) + DSCR \cdot PBO' = p\ell_r\ddot{a}_r.$$

The annuity $\ddot{a}_r$ is valued at rate $DSCR$. Therefore, when there are no retirees, SC plus interest on PBO is very sensitive to variations in $DSCR$.

As a third and final example, suppose there are no active members, only retirees with a population in accordance with the life table. Then there is no service cost, and interest on the PBO (minus benefits) becomes

$$DSCR \cdot (RPBO - B) = B - p\ell_r\ddot{a}_r.$$

This is very sensitive to changes in $DSCR$, but this time it is an *increasing* function of $DSCR$.

Another way of analysing the problem is to consider individual service costs and benefit obligations. Suppose e is the entry age into the plan. Then for one member age x

$$
\begin{aligned}
Q_x &= (1 + DSCR)SC_x + DSCR \cdot (PBO_x - B_x) \\
&= \frac{1}{r-e}[1 + DSCR + (x-e)DSCR](1 + DSCR)^{-(r-x)}(\ell_r/\ell_r)p\ddot{a}_r, \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad e \leq x < r \\
&= DSCR \cdot pa_x, \qquad\qquad \text{if} \quad x \geq r.
\end{aligned}
$$

Since

$$
\begin{aligned}
\frac{\partial}{\partial i} i\ a_x &= \frac{\partial}{\partial i} i\ E\ a_{\overline{K}|} \qquad (1.7) \\
&= \frac{\partial}{\partial i}\ i\ E\ \left(\frac{1-(1+i)^{-K}}{i} \right) \\
&= \frac{\partial}{\partial i} \left[1 - E(1+i)^{-K} \right] \\
&= EK(1+i)^{-K-1} > 0
\end{aligned}
$$

we see that Q_x is a decreasing function of $DSCR$ for young members and an increasing one for older members. (For an explanation of Eq. (1.7) the reader is refered to Section 5.4 of Bowers et al, 1986).

We therefore conclude: *the service cost plus interest on the pension benefit obligation is a decreasing function of $DSCR$ for a relatively young ("undermature") population, and an increasing function of $DSCR$ for a relatively old ("overmature") population; when the population is approximately stationary ("mature"), service cost plus interest on PBO shows little sensitivity to changes in $DSCR$.*

What precedes somewhat weakens (at least for mature populations) the claim that fluctuating discount rates produce large fluctuations in pension expense. Nevertheless it should be noted that fluctuations do arise in all cases, since unexpected variations in the PBO have to be amortized. In the case of a mature population, these variations can be quite large. The final outcome is that emphasis is shifted from the service cost to the amortization of gains and losses. This is explored in Chapter 3.

Appendix 1.1. Formula for k^*

From (1.5),

$$\begin{aligned} X_{t+1} &= (1+i)(1-k)X_t + D_{t+1} \\ &= q\, X_t + D_{t+1} \end{aligned}$$

where $q = (1+i)(1-k)$. Then

$$\begin{aligned} X_1 &= D_1 + qx_0 \\ X_2 &= D_2 + qD_1 + q^2 x_0 \\ &\vdots \\ X_t &= D_t + qD_{t-1} + q^2 D_{t-2} + \cdots + q^t x_0 . \end{aligned} \tag{1.8}$$

It can be shown that X_t will have a limit as $t \to \infty$ (in the sense of convergence in distribution) as soon as $|q| < 1$, i.e.

$$\begin{aligned} |(1+i)(1-k)| &< 1 \\ |1-k| &< v \\ -v < k-1 &< v \\ d < k &< 1+v \end{aligned}$$

$(v = 1/(1+i), d = 1-v)$. It is therefore necessary to pay at least interest on X_t in order to obtain a meaningfull steady-state response. From (1.8) we get

$$\begin{aligned} \text{Var } X_t &= \sigma^2(1 + q^2 + q^4 + \cdots + q^{2t-2}) \\ &= \sigma^2(1-q^{2t})/(1-q^2). \end{aligned}$$

As $t \to \infty$, this approaches

$$\text{Var } X_\infty = \sigma^2/(1-q^2).$$

Since $C_t = -kX_t$, we conclude that

$$\text{Var } X_\infty = \sigma^2/[1-(1+i)^2(1-k)^2]$$

$$\text{Var } C_\infty = \sigma^2 k^2/[1-(1+i)^2(1-k)^2].$$

It is possible to specify the value k^* below which variances are too high. It is the point at which Var C_∞ is a minimum, and is thus the solution of

$$0 = \frac{d}{dk}\{k^2/[1-(1+i)^2(1-k)^2]\}$$

i.e.

$$\begin{aligned} k^* &= 1 - v^2 = (1-v)(1+v) \\ &= d(1+v). \end{aligned}$$

The corresponding "maximum period" m^* can also be calculated:

$$\begin{aligned} k^* &= 1/\ddot{a}_{\overline{m^*}|} \\ \Leftrightarrow d(1+v) &= d/(1-v^{m^*}) \\ \Leftrightarrow 1 - v^{m^*} &= 1/(1+v) \\ \Leftrightarrow v^{m^*} &= v/(1+v) \\ \Leftrightarrow m^* &= 1 + \log(1+v)/\log(1+i) \\ &= \log(2+i)/\log(1+i). \end{aligned}$$

m^* does not depend on the distribution of the disturbances (D_t), and is a decreasing function of i. A very good approximation of m^* can be found as follows:

$$\begin{aligned} m^* &= [\log 2 + \log(1+i/2)]/\log(1+i) \\ &= (\log 2)/\delta + \log(1+i/2)/\log(1+i), \quad \delta = \log(1+i) \\ &\approx (\log 2)/\delta + .5 \end{aligned}$$

since the function $x \mapsto \log(1+x)$ is nearly linear for small values of x.

Appendix 1.2. Sensitivity of actuarial values to changes in valuation interest rate

The criteria which were applied in deriving the demographic assumptions given in subsection 1.4.1 are (1) simplicity and (2) that the sensitivities of the service cost (SC) and projected benefit obligation (PBO) be comparable to those which would be observed in real-life situations. The latter depend on the demographic profile of the plan population; the yardstick which will be used here is a stationary population following the 1983 Group Annuity Mortality (GAM) Table. The results would be different if the population were younger (undermature) or older (overmature).

The discount rate will be represented by i. The notations $d = i/(1+i)$ and $v = 1/(1+i)$ will also be used. The rate of interest is net of wage increases, and SC and PBO are correspondingly valued in dollars deflated by the wage index. By suitably choosing the constant b defining the benefit level (for example if there are ℓ_x member age x, each receiving a pension of 1 unit from age 65), we have

$$\begin{aligned} SC &= \frac{1}{65-30} \sum_{30}^{64} \ell_x \; {}_{65-x|}\ddot{a}_x \\ &= \frac{1}{35} \sum_{30}^{64} \ell_x \; v^{65-x} (\ell_{65}/\ell_x) \; \ddot{a}_{65} \\ &= \ell_{65} \; \ddot{a}_{65} \; a_{\overline{35}|}/35 \end{aligned} \tag{1.9}$$

To calculate PBO, use the equation of equilibrium

$$\begin{aligned} & PBO = (1+i)(PBO + SC - B) \\ \Rightarrow\ & PBO = (B - SC)/d \end{aligned}$$

where

$$B = \sum_{65}^{\omega-1} \ell_x = \ell_{65} \, \ddot{a}_{65}^{(0)}$$

is the annual benefit outgo. Thus

$$PBO = \ell_{65} \left(\ddot{a}_{65}^{(0)} - \ddot{a}_{65}^{(i)} a_{\overline{35}|}/35 \right) /d \tag{1.10}$$

(These relationships are simple consequences of the stationarity of the population; the reader is referred to Trowbridge (1952) or Dufresne (1988) for more details.)

Observe that pre-retirement decrements do not appear in Eq. (1.9) nor in Eq. (1.10); hence pre-retirement decrements have no influence on the sensitivities of SC or PBO to changes in the discount rate. This is essentially a consequence of three assumptions: (i) there are no pre-retirement benefits; (ii) the population conforms to the life table at all ages before retirement; (iii) the projected unit credit method is used. The conclusion would in general be incorrect if any of these assumptions did not hold.

Table 1.4 shows the sensitivities of SC and PBO for the model population and the 1983 GAM Table. The base rate is 2% (recall that this rate is net of wage increases). The model population produces slightly smaller variations then 1983 GAM. For comparison purposes, two other sets of figures are shown. The first part of Table 1.5 uses the 1983 GAM Table with ℓ_x replaced with $\ell_x(1.02)^{-(x-65)}$ at ages $x \geq 65$. The point in doing this is that using the same net valuation rate of interest before and after retirement means that benefits in payment increase at the same rate as salaries. This is rarely the case in practice: at best benefits in payment would get full CPI indexation, which is less than indexation in accordance with wage increases. The modified ℓ_x function corrects this situation by supposing that benefit payments are increased by the wage index minus 2%. The result is slightly smaller sensitivities, since benefits paid have a shorter duration.

The second set of figures in Table 1.5 is based on the assumption that survival is certain up to age 81, instead of age 79 under the model population. Thus $\ddot{a}_{65}^{(0)} = 17$ (instead of 15), which is closer to $17.19 = \ddot{a}_{65}^{(0)}$ under 1983 GAM.

In conclusion, the model population produces sensitivities which are comparable to (though slightly smaller than) those obtained using the 1983 GAM Table.

REMARK. Some plan features have an effect on the duration of benefits paid, for instance: pre-retirement benefits, early retirement, guaranteed periods for benefits in payment, joint and survivor pensions. Intuitively the first two should decrease duration, and the last two increase it. These plan features should therefore have similar effects on the sensitivities of SC and PBO. It would be interesting to try to quantify these effects. $\square$

	Model population		1983 GAM	
i	*SC*	*PBO*	*SC*	*PBO*
.01	11.767	326.58	13.093	414.05
	(25.7%)	(13.6%)	(29.0%)	(15.3%)
.02	9.361	287.58	10.151	359.14
	(-)	(-)	(-)	(-)
.03	7.549	255.82	8.004	315.50
	(-19.4%)	(-11.0%)	(-21.2%)	(-12.2%)

Table 1.4. Service cost (SC) and projected benefit obligation (PBO). Relative variations from values at $i = 2\%$ are shown in brackets.

	Modified 1983 GAM		Modified model population	
i	*SC*	*PBO*	*SC*	*PBO*
.01	10.936	330.88	13.207	383.11
	(27.7%)	(14.8%)	(26.8%)	(14.0%)
.02	8.561	288.22	10.412	335.98
	(-)	(-)	(-)	(-)
.03	6.811	254.10	8.325	297.83
	(-20.4%)	(-11.8%)	(-20.0%)	(-11.4%)

Table 1.5. Service cost (SC) and projected benefit obligation (PBO). On the left the ℓ_x function from the 1983 GAM Table has been replaced with $\ell_x(1.02)^{65-x}$ for $x \geq 65$. On the right $q_x = 0$, $x < 81$, $q_{81} = 1$.

Appendix 1.3. Formulas for B, D, G, HX, HY

The problem is to obtain the values of B, D, G, HX and HY, from those of the inputs $EDS, ER, VARDS, VAROR$ and COR. The parameter A has been set equal to .9. Recall that the process X is supposed stationary.

From

$$VARDS = \operatorname{Var} X_t = B^2/(1 - A^2)$$

we get

$$B = [(1 - A^2)VARDS]^{1/2}.$$

Next,

$$\begin{aligned} VAROR &= \operatorname{Var} Y_t = D^2 B^2/(1 - A^2) + G^2 \\ COR &= D \cdot B/[D^2 B^2 + G^2(1 - A^2)]^{1/2} \end{aligned}$$

imply

$$\begin{aligned} VAROR/VARDS &= [D^2 B^2 + G^2(1 - A^2)]/B^2 \\ D &= [VAROR/VARDS]^{1/2} \cdot COR. \end{aligned}$$

The parameter G can be found from

$$\begin{aligned} 1 - COR^2 &= G^2(1 - A^2)/[D^2 B^2 + G^2(1 - A^2)] \\ &\Rightarrow VAROR(1 - COR^2) = G^2 \\ &\Rightarrow G = [VAROR(1 - COR^2)]^{1/2}. \end{aligned}$$

Now turn to HX. For any variable $Z \sim \mathrm{N}(\mu, \sigma^2)$,

$$Ee^Z = e^{\mu + (1/2)\sigma^2}.$$

Thus

$$1 + EDS = \exp(HX + (1/2)VARDS)$$
$$\Rightarrow HX = \log(1 + EDS) - (1/2)VARDS.$$

Similarly

$$1 + ER = \exp(HY + (1/2)VAROR)$$
$$\Rightarrow HY = \log(1 + ER) - (1/2)VAROR.$$

CHAPTER 2 EXISTENCE OF LIMIT DISTRIBUTIONS

2.1. Introduction

This chapter demonstrates the existence and uniqueness of the limit distributions of unrecognized losses (URL) and pension expense (E) under the model described in Chapter 1. As explained below, these questions are of importance when building a model which will be simulated over long periods of time, as will be the case in Chapter 3.

Section 2.2 first defines the concepts used: stochastic processes, Markov processes, stationary distributions, and embedding. The general theorem which will be used to prove the ergodicity of URL and E is then stated; finally, several relatively simple examples show the application and meaning of the theorem. The last two sections (2.3 and 2.4) deal with the ergodicity of unrecognized losses and pension expense, respectively. The rest of the present section is a discussion of two aspects of the problems studied later in this chapter: (1) the *existence* of a limit distribution; (2) the *dependence* of the limit distribution *on initial conditions.*

Not all stochastic processes have "meaningful" limit distributions. For example, consider the so-called "random walk" model for rates of return. Under this model the deviation of the rate of return in year t from some base rate r is S_t, with

$$S_t = \varepsilon_1 + \cdots + \varepsilon_t$$

where $(\varepsilon_t, t \geq 1)$ are independent and identically distributed (i.i.d.) and have mean zero. This model has the interesting property that

$$E(S_{t+1}|S_t) = S_t$$

i.e. given what is known at time t, next year's rate $(S_{t+1} + r)$ is "expected" to be this year's rate $(S_t + r)$. This model has been used in finance and actuarial science (e.g. de Jong, 1984). The less interesting property of the random walk model is that S_t cannot have a limit distribution. In fact the probability that S_t lies in the range $[-100\%, +100\%]$ approches 0 when t becomes very large. This kind of process could not have been used for the purposes of the present research.

The same thing occurs in a pension model if gains/losses are not amortized. For instance, consider

$$URL_t = URL_{t-1} + L_t - AM_t$$

If we suppose $AM_t = 0$ for all t, then URL_t is the cumulative sum of past losses (gains). In this respect paragraph 184 of SFAS 87 contains the following sentences:

> 184. The Board noted that, if assumptions prove to be accurate estimates of experience over a number of years, gains or losses in one year will be offset by losses or gains in subsequent periods. In that situation, all gains and losses would be offset over time, and amortization of unrecognized gains and losses would be unnecessary.

The view expressed by FASB appears to be that

$$\sum_{s=1}^{t} L_s \approx 0,$$

if assumptions are correct on average and t is large enough. This is mathematically incorrect: $\sum_{s=1}^{t} L_s$ behaves very much like the random walk S_t, i.e. its probable values are further and further away from 0 as t increases. (The main difference between the two processes is that the losses (L_t) are not independent; this does not change the conclusion). Thus, no amortization of gains and losses implies that there is no limit distribution for URL.

Now turn to the dependence of limit distributions on initial conditions. The following example will show what can happen when a "corridor" approach is adopted. Consider a pension funding model, with a stationary population, no inflation and constant valuation rate of interest $i = .06$. Earned rates of return are independent and have distribution

$$\begin{aligned} R_t &= .065 \quad \text{with probability} \quad 1/2 \\ &= .055 \quad \text{with probability} \quad 1/2. \end{aligned}$$

Since the population is stationary and there is no inflation,

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

Fund values evolve according to

$$F_{t+1} = (1 + R_{t+1})(F_t + AM_t + NC - B).$$

Subtracting this equation from the previous one, we obtain a recurrence relation for the unfunded liability:

$$\begin{aligned} UL_{t+1} &= AL - F_{t+1} \\ &= (1 + R_{t+1})(AL + NC - B) - (1 + R_{t+1})(F_t + NC - B) \\ &\qquad - (1 + R_{t+1})AM_t - (R_{t+1} - i)(AL + NC - B) \\ &= (1 + R_{t+1})(UL_t - AM_t) - (R_{t+1} - i)vAL \\ &= (1 + R_{t+1})(UL_t - AM_t) - vAL\Delta R_{t+1} \end{aligned}$$

where $\Delta R_{t+1} = R_{t+1} - i$. Assume that the gains/losses amortization payment AM_t is the excess of the unfunded liability over 10% of AL, divided by 5:

$$\begin{aligned} AM_t &= .20(UL_t - .10AL), \quad UL_t > .10AL \\ &= 0, \qquad\qquad\qquad\quad |UL_t| \leq .10AL \\ &= .20(UL_t + .10AL), \quad UL_t < -.10AL. \end{aligned}$$

First, suppose $|UL_0| < .10AL$, i.e. that the system is initially inside the corridor. Then UL_t will sooner or later drift out of the corridor. One possibility is that $UL_t \geq .10AL$ for some $t \geq 1$. Let us see what happens at time $t+1$: with probability 1/2, $R_{t+1} = .065$ and

$$\begin{aligned} UL_{t+1} &= 1.065[UL_t - .20(UL_t - .10AL)] - vAL(.005) \\ &= 1.065 \cdot .80UL_t + 1.065 \cdot .02AL - .005vAL \\ &\geq 1.065 \cdot .80 \cdot .10AL + 1.065 \cdot .02AL - .005AL/1.06 \\ &= .1018AL. \end{aligned}$$

With probability 1/2, $R_{t+1} = .055$ and

$$\begin{aligned} UL_{t+1} &= 1.055[UL_t - .20(UL_t - .10AL)] + vAL(.005) \\ &\geq 1.055 \cdot .80 \cdot .10AL + 1.055 \cdot .02AL + .005AL/1.06 \\ &= .1102AL. \end{aligned}$$

Thus, once UL is above $10\%AL$, it can never become smaller than that amount again.

The other possibility is that for some $t \geq 1$, $UL_t \leq -.10AL$. Then with probability 1/2

$$\begin{aligned} UL_{t+1} &= 1.065[UL_t - .20(UL_t + .10AL)] - vAL(.005) \\ &= 1.065 \cdot .80UL_t - 1.065 \cdot .02AL - .005AL/1.06 \\ &\leq -.1112AL \end{aligned}$$

or

$$\begin{aligned} UL_{t+1} &= 1.055[UL_t - .20(UL_t + .10AL)] + vAL(.005) \\ &= 1.055 \cdot .80UL_t - 1.055 \cdot .02AL + .005AL/1.06 \\ &\leq -.1008AL. \end{aligned}$$

The same phenomenon is seen to take place: if the surplus ever exceeds $10\% AL$, it can never get smaller than that amount again.

It can be shown that there is a limit distribution for UL_t (and therefore AM_t), whatever the initial state of the system. If the system starts in the region $\{UL \geq .10AL\}$, then the limit distribution is concentrated on that region, since entry into the corridor is impossible. The existence of a limit distribution is a consequence of

$$E(1-.20)(1+R_{t+1}) < 1$$

(loosely speaking, this means that more than interest on $UL - .10AL$ is fed back into the system). The situation is identical if $UL_0 \leq -.10AL$, with a limit distribution concentrated on the region $\{UL \leq -.10AL\}$. If UL_0 is inside the corridor, then the limit distribution depends on which threshold ($+.10AL$ or $-.10AL$) is eventually reached. There is no ergodicity; for example, the empirical mean

$$\overline{UL}_t = \frac{1}{T}\sum_{t=1}^{T} UL_t$$

will converge to the theoretical average of one or the other limit distribution. These values are

$$\overline{UL}^{\,+}_{\infty} = 1.06\,[\overline{UL}^{\,+}_{\infty} - .20(\overline{UL}^{\,+}_{\infty} - .10AL)]$$

$$\Rightarrow \overline{UL}^{\,+}_{\infty} = 1.06 \cdot .02AL/(1-1.06\cdot .80)$$

$$= .1395AL$$

(for the distribution on $(.10AL, \infty)$) and

$$\overline{UL}^{\,-}_{\infty} = 1.06\,[\overline{UL}^{\,-}_{\infty} - .20(\overline{UL}^{\,-}_{\infty} + .10AL)]$$

$$\Rightarrow \overline{UL}^{\,-}_{\infty} = -1.06 \cdot .02AL/(1-1.06\cdot .80)$$

$$= -.1395AL$$

(for the distribution on $(-\infty, -.10AL)$).

Could the same thing happen with the model described in Section 1.4? The problem has to be studied, if the results of Chapter 3 are to be believed.

2.2. Some definitions, and a theorem

A *stochastic process* is a collection of random variables. In the present context, stochastic processes are indexed by the non-negative integers, that is to say there is one random variable for each integer $0, 1, 2, \ldots$. Some of the stochastic processes encountered in the following sections take values in $\mathbf{R}^2$ or $\mathbf{R}^3$, that is to say to each integer there corresponds a random vector in two or three dimensions. The distribution of a random variable X is that function

$$m_X(A) = P(X \in A)$$

which associates to a set A the probability that X lies in A. In some cases a stochastic process $(X_1, X_2, \ldots)$ has a *limit distribution*, i.e. a distribution m such that

$$m(A) = \lim_{t \to \infty} P(X_t \in A)$$

for all relevant sets A. This does *not* mean that $\lim_{t\to\infty} X_t$ exists. Take a simple example: call X_t the result of throwing a die for the t^{th} time. Then X_t has no limit: each X_t is completely unpredictable from past values. But the (X_t) do have a limit distribution, simply the distribution common to all of them. Something similar happens when we simulate pension funds. The distribution of, say, total contributions (in real terms) may approach a particular distribution as time increases, though in any one simulation total contributions continue to fluctuate indefinitely.

A stochastic process is said to possess the *Markov property*, or to be *markovian*, if its movement from time t to time $t+1$ depends only on its position at time t, and not on its previous positions at times $t-1, t-2, \ldots$. An example is the random walk

$$X_t = \varepsilon_1 + \cdots + \varepsilon_t$$

where $(\varepsilon_1, \varepsilon_2, \ldots)$ are independent. Here $P(X_{t+1} \in A \mid X_t \in B_t, \cdots, X_1 \in B_1) = P(X_{t+1} \in A \mid X_t \in B_t)$ for all sets $A, B_1, \ldots, B_t$. What happens before time t has no effect on the transition from time t to time $t+1$.

A stochastic processes $(X_1, X_2, \ldots)$ is stationary if the distribution of any vector

$$(X_{t_1}, \ldots, X_{t_n})$$

is invariant under translation, that is to say if it is the same as that of

$$(X_{t_1+h}, \ldots, X_{t_n+h})$$

for any h.

The process (X_t) representing geometric discount rates (Section 1.4) satisfies

$$X_{t+1} - HX = A(X_t - HX) + B \cdot V_{t+1}$$

where $A = .9$ and (V_t) is an i.i.d. sequence of $\mathrm{N}(0,1)$ variables. This process is markovian, since the transition from $X_t = x$ to X_{t+1} depends only on x and V_{t+1}, which are independent of $((X_s, V_s), s \leq t-1)$. If we set $X_0 = x_0$ (fixed), then the process X is not stationary. Nevertheless, it is known that if $X_0 \sim \mathrm{N}(HX, B^2/(1-A^2))$ and X_0 is independent of $V_1, V_2, \ldots$, then X is stationary. This may intuitively be explained as follows: the distribution $\mathrm{N}(HX, B^2/(1-A^2))$ is the limit distribution obtained for X_t as $t \to \infty$, whatever $X_0 = x_0$. If we give the initial condition X_0 the limit distribution, then X_1 can only reproduce that distribution, and so on for $X_2, X_3, \dots$.

On the contrary, the process Y representing geometric rates of return is *not* markovian. Recall that

$$Y_{t+1} = HY + D(X_{t+1} - HX) + G \cdot W_{t+1}.$$

Thus Y_{t+1} depends on X_{t+1} and W_{t+1}. The variable W_{t+1} is independent "the past" (i.e. of $Y_{t-1}, Y_{t-2}, \dots$) but X_{t+1} is not (since it depends on X_t, and X_t and Y_t are dependent).

(X_t) is markovian, (Y_t) is not (because of its dependence on (X_t)), *but the two processes considered jointly form a two-dimensional markov process.* Consider $Z_t = (X_t, Y_t)^T$. (*N.B.* "$(a, b)^T$" denotes the transpose of the vector (a, b), that is to say the column vector with elements a and b.) Then

$$\begin{aligned} Z_{t+1} = \begin{pmatrix} X_{t+1} \\ Y_{t+1} \end{pmatrix} &= \begin{pmatrix} HX + A(X_t - HX) + B \cdot V_{t+1} \\ HY + D(X_{t+1} - HX) + G \cdot W_{t+1} \end{pmatrix} \\ &= \begin{pmatrix} HX + A(X_t - HX) + B \cdot V_{t+1} \\ HY + D[(X_t - HX) + B \cdot V_{t+1}] + G \cdot W_{t+1} \end{pmatrix} \\ &= g(Z_t, V_{t+1}, W_{t+1}). \end{aligned}$$

The variables V_{t+1} and W_{t+1} are independent of the past (i.e. of $Z_{t-1}, Z_{t-2}, \dots$). Thus the transition from Z_t to Z_{t+1} is independent of $Z_{t-1}, Z_{t-2}, \dots$, and (Z_t) is markovian.

The technique used above (looking at one variable as a component of a higher dimensional variable with "better" properties) is called *embedding*. It will be used to prove the ergodicity of (URL_t) and (E_t).

Markov processes indexed by the integers are called *Markov chains*. The theorem given below concerns the ergodicity of certain types of Markov chains. A few more definitions are required first. The *transition law* (at time t) of a Markov chain (M_t) is the function

$$\begin{aligned} Q_t(m, A) &= \text{Prob}(M_{t+1} \in A \ \text{ given } \ M_t = m) \\ &= P(M_{t+1} \in A \mid M_t = m) \end{aligned}$$

where A is a subset of the space in which (M_t) takes values (this space will be denoted by $\mathcal{M}$). We will only consider Markov chains with the same transition law at each step; these chains are called *time-homogeneous*. Let φ be a σ-finite measure on $\mathcal{M}$. (M_t) is said to be *φ-irreducible* if, whenever $\varphi(A) > 0$, for every $m \in \mathcal{M}$ there is $t \geq 0$ such that

$$P(M_t \in A \mid M_0 = m) > 0.$$

The only case considered here is when φ is Lebesgue measure on the whole d-dimensional space $\mathbf{R}^d$; irreducibility then means that any state m' can be reached from any other state m in a finite number of steps. The last technical definition is the *strong continuity* of a transition law. Q is said to be strongly continuous if the function

$$x \longmapsto Q(x, A)$$

is continuous for each subset A of $\mathcal{M}$.

The next theorem is taken from Tweedie (1975), Section 5.

THEOREM. *Suppose (M_t) is a time-homogeneous, φ-irreducible Markov chain taking values in a finite-dimensional Banach space $\mathcal{M}$, with strongly continuous transition law $Q(m, A)$. Let $\|\cdot\|$ denote the norm on $\mathcal{M}$ and*

$$\begin{aligned} h(m) &= E(\|M_{t+1}\| - \|M_t\| \mid M_t = m) \\ &= \int Q(m, dm')\|m'\| - \|m\|. \end{aligned}$$

Then (M_t) is ergodic if there exist strictly positive constants a and b such that

(i) *$h(m)$ is uniformly bounded if $\|m\| \le a$, and*

(ii) *$h(m) \le -b$ for all m such that $\|m\| > a$.*

This theorem can be interpreted as follows: a Markov chain is ergodic if it "drifts" consistently back to the "center" when it takes values far from it. (*N.B.* The norm $\|m\|$ represents the distance from the origin to m.) The following examples illustrate the meaning and application of the theorem. In all examples (ε_t) is a sequence of i.i.d. normal variables with mean μ and variance σ^2.

Example 1. The random walk

$$S_t = \varepsilon_1 + \cdots + \varepsilon_t, \qquad t \ge 1$$

is a Markov chain. Given $S_t = s$, S_{t+1} may take any value in $\mathbf{R}$, since the variable $\varepsilon_{t+1} = S_{t+1} - S_t$ has a continuous distribution with a positive density on $(-\infty, \infty)$; (S_t) is thus irreducible with respect to Lebesgue measure. The transition law is normal and thus strongly continuous. Let us try to apply the theorem to (S_t). We obtain

$$h(s) = E(|S_{t+1}| - |S_t| \;\;|\;\; S_t = s) = E|s + \varepsilon_{t+1}| - |s|.$$

There is no $a > 0$, $b > 0$ such that $h(s) \le -b$ if $|s| > a$. In fact it is easy to see that

$$\lim_{|s| \to \infty} E|s + \varepsilon_{t+1}| - |s| = 0.$$

This is consistent with the fact that (S_t) does *not* attain an "equilibrium" distribution as t increases, as was pointed out earlier.

Example 2. Consider the autoregressive process

$$X_{t+1} = \alpha X_t + \varepsilon_{t+1}, \qquad t \geq 1,$$

with $|\alpha| < 1$. Irreducibility and strong continuity are obtained as in the previous example. Here

$$\begin{aligned} h(x) &= E(|X_{t+1}| - |X_t| \mid X_t = x) \\ &= E|\alpha x + \varepsilon_{t+1}| - |x| \\ &\leq E(|\alpha||x| + |\varepsilon_{t+1}|) - |x| \\ &= (|\alpha| - 1)|x| + c, \qquad c = E|\varepsilon_{t+1}|. \end{aligned}$$

The conditions in the theorem are satisfied, since $(|\alpha| - 1)|x| \to -\infty$ as $|x| \to \infty$. We conclude that (X_t) is ergodic. Intuitively, the fact that $|\alpha| < 1$ means that X_t is pulled back towards the origin, unlike the process S_t in Example 1. (*N.B.* Example 1 is obtained by setting $\alpha = 1$.)

Example 3. Consider the "autoregressive process with autoregressive noise" (Y_t):

$$\begin{aligned} Y_{t+1} &= \beta Y_t + X_{t+1} \quad , \quad |\beta| < 1 \\ X_{t+1} &= \alpha X_t + \varepsilon_{t+1} \quad , \quad |\alpha| < 1 \end{aligned}$$

Y_t is not markovian, but the vector $M_t = (Y_t, X_t)^T$ is. The space $\mathcal{M}$ is now $\mathbf{R}^2$. Irreducibility and strong continuity are easily verified. Given $M_t = (y, x)^T$, the transition from M_t to M_{t+1} can be represented as

$$\begin{pmatrix} y \\ x \end{pmatrix} \to \begin{pmatrix} \beta y + \alpha x + \varepsilon_{t+1} \\ \alpha x + \varepsilon_{t+1} \end{pmatrix}.$$

It is better not to use the ordinary euclidean norm $\|(y, x)\|_e = (y^2 + x^2)^{1/2}$, but rather

$$\|(y, x)\| = |y| + d|x|$$

where d is a strictly positive constant such that $(1 + d)|\alpha| < d$ (take any $d > |\alpha|/(1 - |\alpha|)$). We then get

$$h(y, x) \leq (|\beta| - 1)|y| + [(1 + d)|\alpha| - d]|x| + (1 + d)E \mid \varepsilon_{t+1}|.$$

The conditions of the theorem are satisfied, and (M_t) is ergodic. This also proves the ergodicity of (Y_t) since Y_t is a function of M_t (namely the projection $(y, x) \longrightarrow y$).

Example 4. Consider a modification of Example 3:

$$Y_{t+1} = \beta Y_t + e^{X_{t+1}}, \quad |\beta| < 1$$

$$X_{t+1} = \alpha Y_t + \varepsilon_{t+1}, \quad |\alpha| < 1.$$

The noise is now the exponential of an autoregressive process. Intuitively (Y_t) should once again be ergodic, since basically the same situation prevails: Y_t is brought back towards the origin by the constant β, and the noise added reaches an equilibrium distribution itself (since $|\alpha| < 1$). However, using the norm $\|\cdot\|$ of the previous example, we get

$$h(y, x) \leq (|\beta| - 1)|y| + d(|\alpha| - 1)|x| + dE|\varepsilon_{t+1}| + e^{\alpha x} E e^{\varepsilon_{t+1}}.$$

It cannot be said that $h(y, x) \leq -b$ for all (y, x) such that $\|(y, x)\|$ is greater than some $a > 0$. This is because $e^{\alpha x}$ is unbounded as $|x|$ increases (unless α equals 0). The theorem given can apparently not be used to prove the ergodicity of (Y_t) (some clever trick might do it, but the author was unable to find such a trick). (Y_t) *is* nevertheless ergodic. This is proved by noting that

$$Y_t = e^{X_t} + \beta e^{X_{t-1}} + \cdots + \beta^{t-1} e^{X_1}$$

has a limit distribution, since $|\beta| < 1$ and (X_t) itself has a limit distribution; see Brandt (1986) or Dufresne (1991, Section 3) for more details.

The above difficulty will surface once more in Section 2.3 and 2.4. Essentially the difficulty here is that $Ee^{\alpha x + \varepsilon_{t+1}}$ is unbounded as a function of "the present"; in the pension model the same will apply to the average gain or loss arising in one year.

Example 5. Now consider an autoregressive process with "thresholds":

$$X_{t+1} = \alpha(X_t)X_t + \varepsilon_{t+1}$$

where

$$\begin{aligned} \alpha(x) &= \alpha \quad \text{if} \quad |x| > 1, \quad |\alpha| < 1 \\ &= \beta \quad \text{if} \quad |x| \leq 1. \end{aligned}$$

The constant β does not have to be smaller than 1 in absolute value. Here X_t drifts back towards the origin, but only when it is outside the "corridor" $[-1, 1]$. We find

$$h(x) \leq (|\alpha| - 1)|x| + E|\varepsilon_{t+1}|, \quad \text{if} \quad |x| > 1,$$

and conclude that (X_t) is ergodic.

Example 6. Consider a pension funding model with a stationary population, no inflation and constant valuation rate of interest i. The earned rates of return on assets (R_t) are i.i.d. with a lognormal distribution (i.e. $1 + R_t = \exp\, \varepsilon_t$). Then (see Section 2.1)

$$UL_{t+1} = (1 + R_{t+1})(UL_t - AM_t) - vAL\ \Delta R_{t+1}$$

where UL is the unfunded liability, AM is the special payment towards amortization of UL and $\Delta R_{t+1} = R_{t+1} - i$. Suppose

$$AM_t = k\ UL_t$$

(this was called "proportional control" in Chapter 1). Then, given $UL_t = u$,

$$\begin{aligned} UL_{t+1} &= (1 + R_{t+1})(1 - k)u - vAL\ \Delta R_{t+1} \\ \Rightarrow h(u) &= E|(1 + R_{t+1})(1 - k)u - vAL\Delta R_{t+1}| - |u| \\ &\leq [E(1 + R_{t+1})(1 - k) - 1]|u| + vAL \cdot E|\Delta R_{t+1}|. \end{aligned}$$

A sufficient condition for the ergodicity of (UL_t) is therefore

$$E(1 + R_t)(1 - k) < 1$$

or

$$\begin{aligned} k &> 1 - 1/E(1 + R_t) \\ &= 1 - \exp(-\mu - \sigma^2/2) \end{aligned}$$

if $1 + R_t \sim \log\mathrm{N}(\mu, \sigma^2)$. It can be shown that if $k \leq 1 - 1/E(1 + R_t)$, then (UL_t) does not have a limit distribution.

The theorem requires the distribution of R_t to be continuous and positive over a sufficiently wide interval (this is to satisfy irreducibility and strong continuity). In this example, however, other more direct methods can be used to show that (UL_t) is ergodic as soon as $k > 1 - 1/E(1 + R_t)$, whatever the distribution of R_t. The next example is different in this respect.

Example 7. Consider the same model, but let

$$\begin{aligned} AM_t &= k(UL_t - C), &&\text{if } UL_t > C \\ &= 0, &&\text{if } |UL_t| \leq C \\ &= k(UL_t + C), &&\text{if } UL_t < -C \end{aligned}$$

for some $C > 0$. This generalizes the example given in the latter part of Section 2.1. Now assume $UL_t = u$; then

(i) if $u > C$,

$$\begin{aligned} UL_{t+1} &= (1 + R_{t+1})[u - k(u - C)] - vAL\ \Delta R_{t+1} \\ &= (1 + R_{t+1})(1 - k)u + (1 + R_{t+1})kC - vAL\ \Delta R_{t+1}; \end{aligned}$$

(ii) if $|u| \leq C$,

$$UL_{t+1} = (1 + R_{t+1})u - vAL\ \Delta R_{t+1};$$

(iii) if $u < -C$,

$$\begin{aligned} UL_{t+1} &= (1 + R_{t+1})[u - k(u + C)] - vAL\ \Delta R_{t+1} \\ &= (1 + R_{t+1})(1 - k)u - (1 + R_{t+1})kC - vAL\ \Delta R_{t+1}. \end{aligned}$$

Then

$$\begin{aligned} h(u) &= E(|UL_{t+1}| \mid UL_t = u) - |u| \\ &\leq [E(1 + R_{t+1})(1 - k) - 1]|u| + kCE(1 + R_{t+1}) + vAL \cdot E|\Delta R_{t+1}| \\ &\leq -b \end{aligned}$$

for all u such that $|u|$ is large enough, if $E(1 + R_t)(1 - k) < 1$. Observe that this is the same condition as in the previous example, i.e. the presence of the corridor does not change anything this far. But (U_t) may yet fail to be ergodic. In Section 2.1, the rates of return on assets take the two values .0065 and .0055 with equal probability, and $k = .2$; thus

$$E(1 + R_t)(1 - k) = 1.06(1 - .20) < 1.$$

Nevertheless (UL_t) did not have a unique limit distribution. This is because the condition of irreducibility was not satisfied. As was explicitly shown in Section 2.1, it is impossible for UL to move from the region $(-\infty, -.10AL]$ to the region $[+.10AL, +\infty)$, or vice-versa. Thus the range space can be "reduced" to three sets (namely $(-\infty, -.10AL]$, $(-.10AL, +.10AL)$, $[+.10AL, +\infty)$), two of which are "closed" (i.e. cannot be escaped from). This particular distribution for the rates of return therefore violates one of the conditions of the theorem. The theorem will apply if the rates of return have a continuous distribution with a range sufficiently wide to allow UL to move (with positive probability) from any one point to any other one.

2.3. Ergodicity of unrecognized losses (URL)

The ergodicity of (URL_t) will now be proved, assuming the model described in Section 1.4. The population may be any stationary population, not necessarily the one adopted for the simulations; all the parameters are left unspecified, except that $|A| < 1$ (to ensure ergodicity of discount rates and rates of return on assets) and $0 < k \leq 1$; it will be assumed that funding gains and losses are liquidated over one year, although the result would also hold for less rapid amortization.

From Section 1.4,

$$URL_{t+1} = URL_t + L_{t+1} - g(URL_t, F_t \vee PBO_t)$$

where

$$\begin{aligned} L_{t+1} &= \text{accounting loss during } (t, t+1) \\ &= PBO_{t+1} - PBO_t + (ELTR - R_{t+1})AL/(1 + VI) \end{aligned}$$

$$g(URL_t, F_t \vee PBO_t) = \text{losses recognized in expense at time } t+1$$
$$= AM_{t+1}.$$

Here $F_t \vee PBO_t = \max(F_t, PBO_t)$ and

$$\begin{aligned} g(u,q) &= 0, & \text{if } |u| \leq cq \\ &= k(u - cq), & \text{if } u > cq \\ &= k(u + cq), & \text{if } u < -cq. \end{aligned}$$

The constant $c > 0$ determines the width of the corridor (for example $c =$ 10% in SFAS 87), and the function g takes as inputs u (= unrecognized losses) and q (= maximum of fund and projected benefit obligation) and outputs the minimum amount to be recognized at time $t+1$.

As pointed out before, (URL_t) is not markovian; however, the vector (URL_t, X_t, F_t) *is* a Markov chain; thus it is sufficient to prove that this vector is ergodic in order to show that URL_t is ergodic. It is technically easier to deal with $H_t = \log F_t$, instead of F_t itself. We have

$$H_{t+1} = HY + D[A(X_t - HX) + B \cdot V_{t+1}] + G \cdot W_{t+1} + \log[AL/(1+VI)].$$

Consider the vector $(M_t) = (URL_t, X_t, H_t)^T$. Given $M_t = m = (u, x, h)^T$, we have

$$M_{t+1} = \begin{pmatrix} u + L_{t+1} - g(u, e^h \vee p(x)) \\ HX + A(x - HX) + B \cdot V_{t+1} \\ HY + D[A(x - HX) + B \cdot V_{t+1}] + G \cdot W_{t+1} + \log[AL/(1+VI)] \end{pmatrix}$$

Here the function $p(x)$ represents the projected benefit obligation valued at geometric rate x (so that $PBO_t = p(X_t)$).

In Section 1.4 the variables (V_t) and (W_t) had $\mathbf{N}(0,1)$ distributions. This is not required here. The calculations below suppose that each sequence is i.i.d. with $E|V_t| < \infty$, $E|W_t| < \infty$. The transition law of (M_t) will be strongly continuous if the distributions of V_t and W_t are continuous. Just as in Example 7 of Section 2.2, irreducibility requires that the ranges of these variables be wide enough to make it possible for URL to move (with positive probability) from any one point u to any other point u' in a finite number of steps.

The following norm will be used:

$$\|m\| = \|(u, x, h)\| = |u| + |x| + d|h|$$

where $d > 0$ is such that $|A| + d|D \cdot A| < 1$. Let

$$h(m) = E(\|M_{t+1}\| - \|M_t\| \;\; \big| \;\; M_t = m).$$

We have to show that there exist strictly positive constants a and b such that $h(m)$ is bounded for $\|m\| \leq a$, and $h(m) \leq -b$ for $\|m\| > a$. The first condition clearly holds. In order to check the second one, observe that

$$\begin{aligned} u - g(u,q) &= u, & &\text{if} \quad |u| \leq cq \\ &= (1-k)u + kcq, & &\text{if} \quad u > cq \\ &= (1-k)u - kcq, & &\text{if} \quad u < -cq. \end{aligned}$$

Since $0 < k \leq 1$ and $c > 0$, $q > 0$, the foregoing imply

$$\begin{aligned} |u - g(u,q)| - |u| &= 0, & &\text{if} \quad |u| \leq cq \\ &= -k(|u| - cq), & &\text{if} \quad |u| > cq. \end{aligned}$$

A more concise way of writing this is

$$|u - g(u,q)| - |u| = -k(|u| - cq)_+$$

where $(z)_+$ is the "positive part of z" (i.e. z if positive, zero otherwise).

Using the triangle inequality ($|\alpha + \beta| \le |\alpha| + |\beta|$) repeatedly, we find

$$h(m) \le -k(|u| - cq)_+ + (|A| + d|D \cdot A| - 1)|x| - d|h| + C(x)$$

where $q = e^h \vee p(x)$ and

$$\begin{aligned} C(x) &= E(|L_{t+1}| \mid M_t = m) + (1 - A + d|D \cdot A|)|HX| + |HY| \\ &\quad + (1 + d|D|)|B| \cdot E|V_{t+1}| + d|G| \cdot E|W_{t+1}| + d \left| \log[AL/(1 + VI)] \right| . \end{aligned}$$

$C(x)$ depends on x only through the first term $E(|L_{t+1}| \mid M_t = m)$. This term depends on x because

$$\begin{aligned} E(|L_{t+1}| \mid M_t = m) &= E[|PBO_{t+1} - PBO_t + (ELTR - R_{t+1}) \\ &\qquad \times AL/(1 + VI)| \mid M_t = m] \\ &= E|p(HX + A(x - HX) + B \cdot V_{t+1}) - p(x) \\ &\qquad + (1 + ELTR - \exp\{HY + D[A(x - HX) \\ &\qquad + B \cdot V_{t+1}] + G \cdot W_{t+1}\})AL/(1 + VI)|. \end{aligned}$$

When the possible range of V_t is $(-\infty, +\infty)$, x also has range $(-\infty, \infty)$, and the above expression is unbounded as a function of x. We are then faced with the same problem as in Example 4 (Section 2.2). In order to be able to use Tweedie's theorem to prove the ergodicity of (M_t), the following assumption is made: *there exists* $C < \infty$ *such that*

$$E(|L_{t+1}| \mid M_t = m) \le C \qquad (*)$$

for all $m \in \mathbf{R}^3$. This assumption will be discussed in the remarks below.

The proof may now be completed. There is a constant $C_1 < \infty$ such that

$$h(m) \le -k(|u| - cq)_+ + (|A| + d|D \cdot A| - 1)|x| - d|h| + C_1.$$

Let $b > 0$. There exist $C_2 < \infty$ and $C_3 < \infty$ such that

$$|u| - c[e^h \vee p(x)] \geq C_2 \Rightarrow h(m) \leq -b,$$

$$|x| + d|h| \geq C_3 \Rightarrow h(m) \leq -b.$$

Let

$$C_4 = c \cdot \sup_{|x|+d|h| \leq C_3} [e^h \vee p(x)]$$

and suppose $\|m\| > C_2 + C_3 + C_4$. Then there are two possibilities:

(1) $|x| + d|h| < C_3$ which implies (a) $c[e^h \vee p(x)] \leq C_4$ and (b) $|u| \geq C_2 + C_4$ which in turn imply $|u| - c[e^h \vee p(x)] \geq C_2$ and $h(m) \leq -b$.

(2) $|x| + d|h| \geq C_3$ which implies $h(m) \leq -b$.

The proof that (URL_t) is ergodic is complete.

REMARKS 1. The result appears plausible even without assumption $(*)$. In fact, the result can be so proved when there is no corridor (i.e. $c = 0$), see Brandt (1986). Unfinished calculations also show that the result might be proved for $c > 0$, without $(*)$, but only when $0 \leq A < 1$.

2. From a modelling point of view, there is no difficulty in accepting assumption $(*)$. The loss during one year depends on the increase in the PBO and on the value of the fund at the end of the year. Most of us would accept that PBO and F have natural limits e.g. the earth's total wealth, or the largest number the computer can handle. These limits ensure that $(*)$ is satisfied.

3. The reason why PBO and F may take unlimited values under the model described in Chapter 1 is that X_t (the geometric discount rate) and Y_t (the geometric rate of return on assets) may take values in the whole interval $(-\infty, +\infty)$. For most purposes this is not a problem, as the probability of

"very large" values of either X or Y is very small as long as their variances agree with historical observations. For instance, under the base scenario X_t has a normal distribution with a mean value around $\mu = .01$ and a standard deviation of $\sigma = .03$; the probability that a normal variable takes a value outside the interval $(\mu - 5\sigma, \mu + 5\sigma)$ is less than 10^{-6}. In some cases, however, it might be appropriate to limit the possible ranges of X or Y. One simple way of achieving this would be to limit the possible values of the disturbances (V_t) and (W_t). This would be another way of making sure that $(*)$ holds.

4. Other ways of amortizing unfunded liabilities could be considered, for example (1) amortization payments equal to a fraction k' of UL, or (2) separate amortization of each annual gain/loss over a number of years. In those cases F_t would not be a function of R_t alone, but would also depend on its own past. Proving ergodicity in these cases is possible by using a higher dimension vector. $\square$

2.4. Ergodicity of pension expense (E)

We have (see Section 1.4)

$$E_t = B - ELTR(F_{t-1} - B) + AM_t.$$

Thus E_t is a function of F_{t-1}, PBO_{t-1} and URL_{t-1}, which means that E_t may also be seen as a function of the vector $(URL_{t-1}, X_{t-1}, H_{t-1})$ considered in Section 2.3. This automatically proves the ergodicity of (E_t) under the assumptions made in that section.

CHAPTER 3 RESULTS OF SENSITIVITY ANALYSIS

3.1. Methodology

This chapter quantifies the variability of pension expense and unrecognized losses under the model described in Chapter 1. Chapter 2 has shown that the stochastic processes representing these amounts are ergodic, implying that their stationary distributions can be obtained from one "long" realization of the processes. The stationary distributions obtained under different sets of assumptions can be compared, showing the relative importance of each assumption. This section describes how the computer simulations were performed. The next section analyses in some detail the numbers obtained under the base scenario. The rest of the chapter presents the results of the sensitivity analysis conducted with respect to the following parameters: the variance of discount rates (Section 3.3); the variance of the rates of return on assets (Section 3.4); the width of the corridor (Section 3.5); and, finally, the fraction of losses recognized in each year (Section 3.6).

A Fortran program was written to simulate the model described in Chapter 1. The program takes the parameters of the model as inputs and outputs the mean, variance as well as the frequency distributions of pension expense (E) and unrecognized losses (URL). Each simulation ran for one million periods.

Pseudo-random numbers (U_t) possessing a uniform distribution on interval $(0, 1)$ were generated using the combined congruential method

$$X_t = 40014X_{t-1} \pmod{2147483563}$$
$$Y_t = 40692Y_{t-1} \pmod{2147483399}$$
$$Z_t = (X_t + Y_t) \pmod{2147483563}$$

with $X_0 = 33$ and $Y_0 = 99$. U_t is then equal to Z_t rescaled to $(0, 1)$. This generator has period $2.3 \cdot 10^{18}$ (Bratley et al, 1987, p. 204). The normal variables (V_t) and (W_t) were then obtained from the so-called Box-Muller transformation

$$V_t = \cos(2\pi U_{2t})\sqrt{-2\log U_{2t-1}}$$
$$W_t = \sin(2\pi U_{2t})\sqrt{-2\log U_{2t-1}}.$$

REMARK. It is well known that using a linear congruential generator

$$Z'_{t+1} = (aZ'_t + b) \pmod{m}$$

in conjunction with the Box-Muller method produces very poor normal variables (the pair of variables obtained are certainly not independent, for one thing). The combined congruential method used here avoids this problem. For more details, the reader is referred to Bratley et al (1987), pp. 204, 223-224. □

3.2. Analysis of results under the base scenario

The parameters chosen for the base scenario are as follows:

Arithmetic valuation rate of interest (VI)	.02
Expected long-term arithmetic rate of return on plan assets ($ELTR$)	.02
Mean arithmetic discount rate (EDS)	.01
Standard deviation of geometric discount rate ($VARDS^{1/2}$)	.03
Mean arithmetic rate of return on assets (ER)	.02
Standard deviation of geometric rate of return on assets ($VAROR^{1/2}$)	.05
Correlation between geometric discount rate and rate of return on assets (COR)	.60
Fraction of $\max(PBO, F)$ used for corridor (C)	.10
Fraction of excess of $\lvert URL \rvert$ over $C \cdot \max(PBO, F)$ recognized in expense	1/15

Economic and actuarial assumptions were analysed in Subsection 1.4.2. In all simulations the geometric discount rates (X) and rates of return on assets are generated from

$$X_{t+1} = HX + .9(X_t - HX) + B \cdot V_{t+1}$$
$$Y_{t+1} = HY + D(X_{t+1} - HX) + G \cdot W_{t+1}$$

where (V_t) and (W_t) are two independent i.i.d. $N(0, 1)$ sequences, and HX, HY, B, D and G are such that EDS, $VARDS$, ER, $VAROR$ and COR take the desired values. Under the base scenario the corridor used is the one prescribed by SFAS 87, i.e. $\pm 10\%$ of $\max(PBO, F)$. When unrecognized losses fall outside the corridor, the excess, multiplied by 1/15, is included in expense for that year. This corresponds to an average future working lifetime of active employees ($AFWL$) equal to 15 years.

The pension plan (see Subsection 1.4.1) provides $b\%$ of salary per year of service. For convenience, b is chosen so that the total annual benefit outgo is equal to 15 units (in constant currency). This corresponds to setting $\ell_{65} = 1$ in the formulas shown in Appendix 1.2. On the funding side, the actuarial liability and normal cost are valued using the projected unit credit method:

$$AL(@\ VI = 2\%) = 287.58$$
$$NC(@\ VI = 2\%) = 9.361.$$

If the discount rate is equal to its mean value, then

$$PBO(@\ DSCR = 1\%) = 326.58$$
$$SC(@\ DSCR = 1\%) = 11.767.$$

Table 3.1 shows the frequency distributions of pension expense and unrecognized losses. Means and standard deviations are summarized in Table 3.2. Figures 3.1 and 3.2 are graphic representations of the distribuitons of E and URL. Observe that pension expense has a very large frequency around 9.56; this is a consequence of the use of the corridor: if $|URL_{t-1}| \leq 10\%\max(F_{t-1}, PBO_{t-1})$, then $AM_t = 0$ and

$$E_t = B - ELTR(F_{t-1} - B)$$

which has a relatively small variance. Otherwise the data has three striking features: (1) the distributions have very wide ranges; (2) their variances are large, and (3) the distributions are skewed (i.e. not symmetrical). These points will be discussed in turn.

FREQUENCY DISTRIBUTIONS

Expense (E)		Unrecognized losses (URL)	
Interval	*Frequency*	*Interval*	*Frequency*
$-\infty$		$-\infty$	
	1,154		33
-13.13		-683.0	
	1,930		13
-10.74		-611.1	
	5,621		38
-8.35		-539.2	
	13,931		89
- 5.97		-467.3	
	29,245		438
-3.58		-395.4	
	50,398		2,965
-1.19		-323.5	
	72,860		22,912
1.19		-251.6	
	90,715		90,111
3.58		-179.7	
	97,087		176,204
5.97		-107.8	
	94,365		201,577
8.35		-36.0	
	246,193		168,761
10.74		36.0	
	63,486		120,298
13.13		107.8	
	50,831		77,964
15.52		179.7	
	39,467		48,978
17.90		251.6	
	30,602		30,572
20.29		323.5	
	23,912		19,040
22.68		395.4	
	18,332		12,340
25.06		467.3	
	14,143		8,116
27.45		539.2	
	10,809		5,492
29.84		611.1	
	8,544		3,816
32.23		683.0	
	36,375		10,243
∞		∞	

Table 3.1. Frequency distributions of pension expense (E) and unrecognized losses (URL) under base scenario (one million iterations). Second column shows the number of times the variable took a value in the interval given in first column.

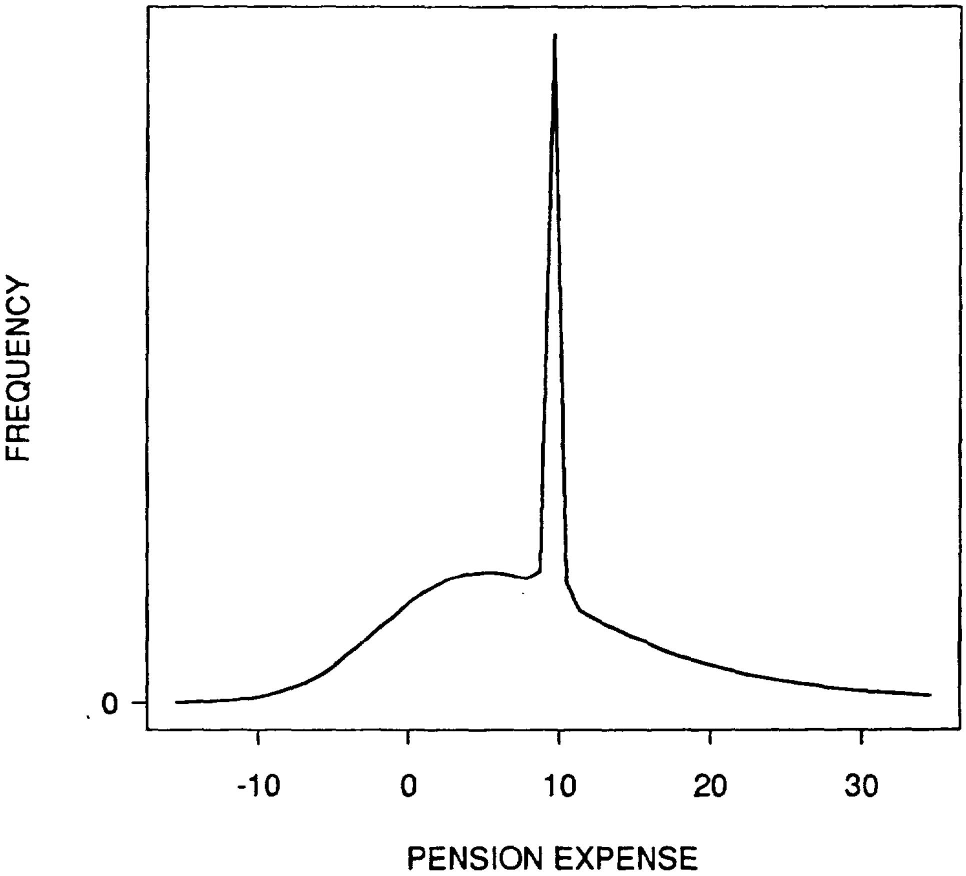

FIGURE 3.1. Frequency distribution of pension expense

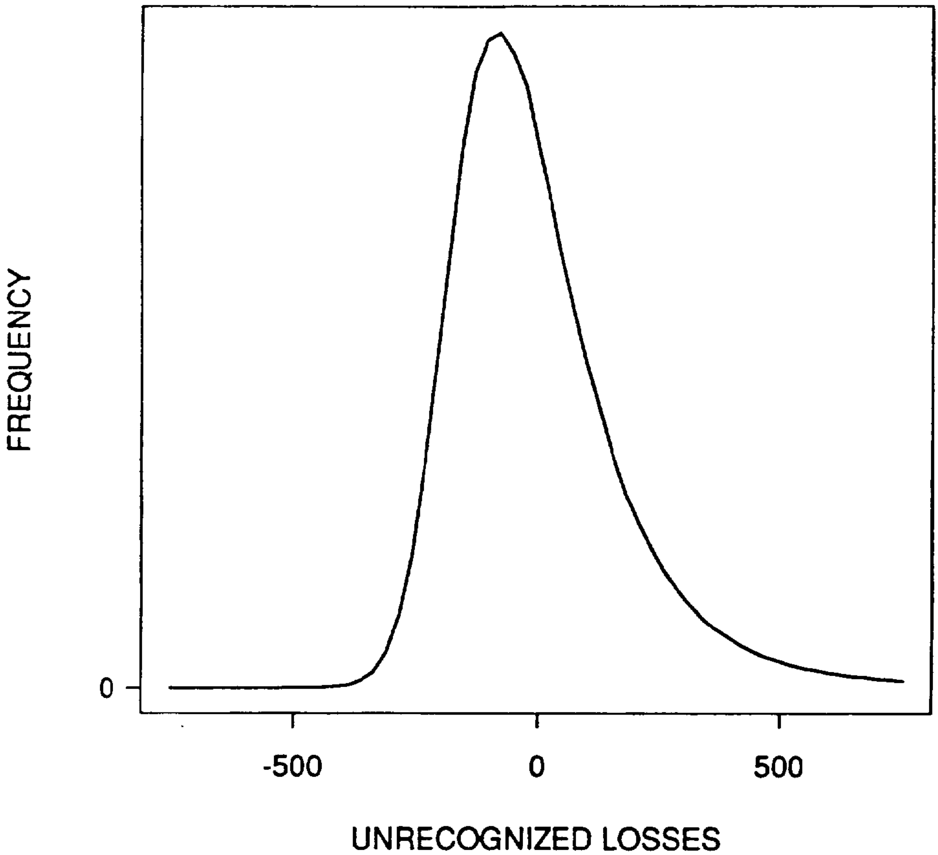

FIGURE 3.2. Frequency distribution of unrecognized losses

Variable	Mean	Standard deviation
Pension expense (E)	9.558	10.90
Unrecognized losses (URL)	3.83	197.94
Amortization payment (AM)	0.01	10.74
Losses (L)	0.01	100.60
Losses due to increase (decrease) in PBO ($LPBO$)	0.00	98.09
Losses due to return on fund (LF)	0.01	14.38
Pension benefit obligation (PBO)	377.97	195.64
Geometric discount rate (X)	0.009504126	0.029976002
Geometric rate of return (Y)	0.018520661	0.049982874
Arithmetic discount rate ($DSCR$)	0.010003097	0.030281681
Arithmetic rate of return (R)	0.019966476	0.051009802

Table 3.2. Observed means and standard deviations of some of the variables, under base scenario (one million iterations).

That the distributions of E and URL have wide ranges is easy to explain. The loss in year $(t-1, t)$ was defined in Chapter 1 as

$$\begin{aligned} L_t &= PBO_t - PBO_{t-1} + (ELTR - R_t)AL/(1+VI) \\ &= (\text{unexpected increase (decrease) in } PBO) \\ &\qquad + (\text{loss or return on assets}). \end{aligned} \tag{3.1}$$

The PBO is valued at (geometric) rate $X = \log(1+DSCR)$, X having a normal distribution. When X becomes very large, PBO approaches zero. When X becomes large in magnitude but negative, PBO increases without bounds. (Consider one unit discounted s years at geometric rate G, G a normal random variable; the discounted value is then $\exp(-sG)$. Since the possible values of G are $(-\infty, +\infty)$, the possible values of $\exp(-sG)$ are $(0, +\infty)$.

The same thing applies to PBO.) Thus the range of $PBO_t - PBO_{t-1}$ (and of L_t by way of consequence) is the whole real line. Unrecognized losses and pension expense therefore have values ranging from $-\infty$ to $+\infty$.

Of course, just as very large values (positive or negative) of X have very low probabilities, the frequencies of very large values of E or URL are quite small (see Figures 3.1 and 3.2). Nevertheless the standard deviations of these amounts remain quite high (Table 3.1). The variances of E and URL are determined by the distribution of the process (L_t). The standard deviation of the second component of L_t (see Eq. (3.1) above) is $14.38 = \text{Stdev}(R) \cdot 287.58/1.02$; that of the first component is 98.10 (see Table 3.2). Thus, the large variances of E and URL result mostly from the great variability of $PBO_t - PBO_{t-1}$. Rates of return on assets vary more than discount rates, but the fluctuations of the latter have far greater consequences than those of the former.

The distributions of PBO_t and $PBO_t - PBO_{t-1}$ are themselves of some importance. Figure 3.3 shows PBO as a function of $DSCR$. PBO is seen to be a convex function of $DSCR$ (i.e. its second derivative is always positive). It follows that for any (non-degenerate) distribution for $DSCR_t$,

Expected value of $PBO_t > PBO$ valued at expected value of $DSCR_t$

(from Jensen's inequality). This is clearly seen here, as $E(PBO_t) = 377.97 > 326.58 = PBO(@1\%)$. The PBO is very sensitive to $DSCR$, as evidenced by its large standard deviation. The distribution of PBO_t is shown in Figure 3.4; it is not symmetric. Nevertheless the distribution of $PBO_t - PBO_{t-1}$ (not shown) *is* perfectly symmetric, as explained in Appendix 3.1.

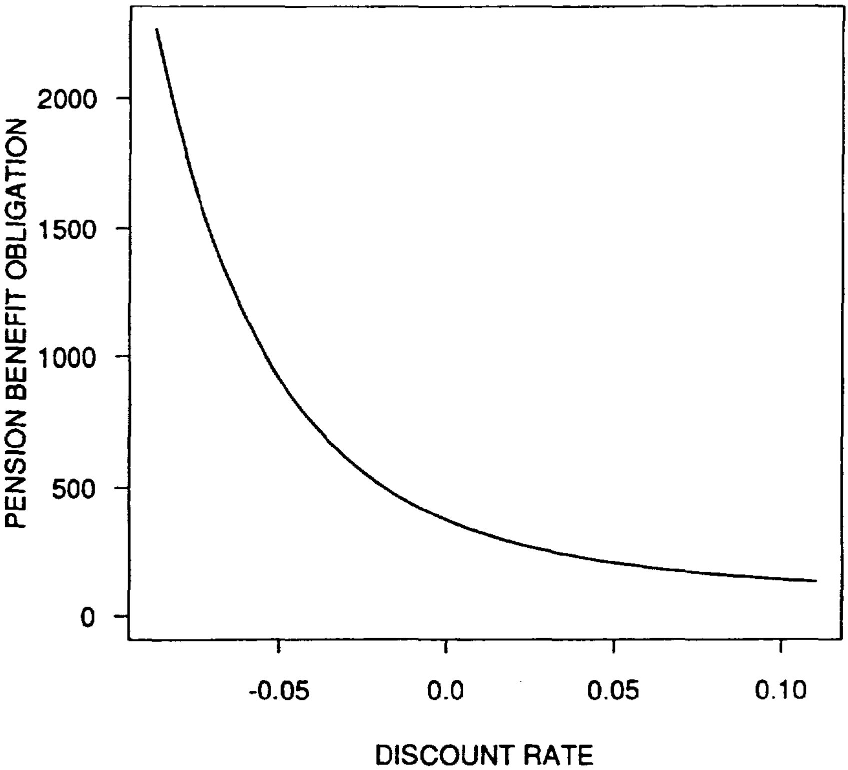

FIGURE 3.3. Pension benefit obligation as a function of the discount rate

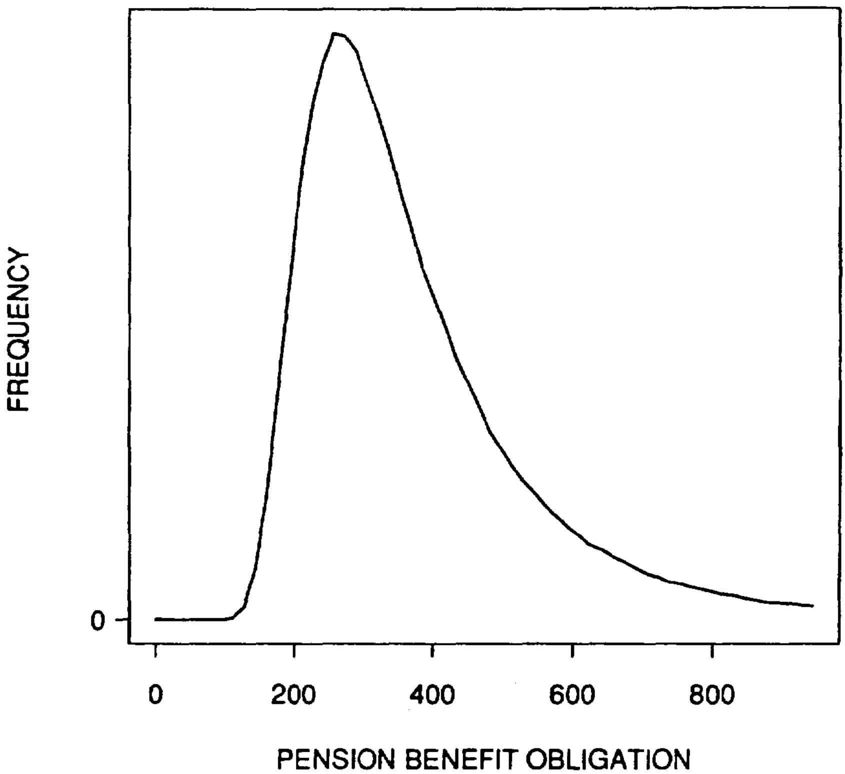

FIGURE 3.4. Frequency distribution of pension benefit obligation

The author was puzzled for some time by the skewness of the distributions of E and URL. Intuitively, there are two reasons why one would expect symmetric distributions: (1) the distribution of L_t is dominated by that of $PBO_t - PBO_{t-1}$, which is symmetric about 0, and (2) negative losses are treated the same way as positive ones: AM_t is a symmetric function of URL_{t-1}. What was more intriguing is that E and URL do not have symmetric distributions even if the second term in Eq. (3.1) is removed. In this case the distribution of losses and the "system" itself (i.e. the way AM_t is obtained from URL_{t-1}) are perfectly symmetric about the origin, but the outputs E and URL still have significantly skewed distributions. The explanation was finally found: the variable $PBO_t - PBO_{t-1}$ may have a symmetric distribution, but this cannot be said of the *process* $(PBO_t - PBO_{t-1}, t \geq 1)$. Details are given in Appendix 3.1. The skewness of the distributions of E and URL ultimately results from the skewness of that of PBO_t.

The skewness of some of the distributions may explain why URL is not precisely zero on average. On average pension expense is very close to the normal actuarial cost plus interest (9.558 versus $9.361 \cdot 1.02 = 9.548$). This is not a coincidence, as will now be explained. Define

$$
\begin{aligned}
LPBO_t &= \text{accounting loss on } PBO \\
&= PBO_t - PBO_{t-1}; \\
LF_t &= \text{accounting loss on } F \\
&= (ELTR - R_t)(F_{t-1} + C_{t-1} - B)
\end{aligned}
$$

so that

$$L_t = LPBO_t + LF_t.$$

We find

$$
\begin{aligned}
LPBO_t &= PBO_t - (1 + DSCR_{t-1})(PBO_{t-1} + SC_{t-1} - B) \\
&= (PBO_t - PBO_{t-1}) - [SC_{t-1}(1 + DSCR_{t-1}) \\
&\qquad + DSCR_{t-1}(PBO_{t-1} - B) - B]
\end{aligned}
$$

and

$$
\begin{aligned}
LF_t &= [(1 + ELTR) - (1 + R_t)](F_{t-1} + C_{t-1} - B) \\
&= F_{t-1} + C_{t-1} - B + ELTR(F_{t-1} + C_{t-1} - B) - F_t \\
&= ELTR(F_{t-1} + C_{t-1} - B) + (F_{t-1} - F_t) + C_{t-1} - B.
\end{aligned}
$$

Thus

$$
\begin{aligned}
E_t &= SC_{t-1}(1 + DSCR_{t-1}) + DSCR_{t-1}(PBO_{t-1} - B) \\
&\qquad - ELTR(F_{t-1} - B) + AM_t \qquad (3.1) \\
&= (PBO_t - PBO_{t-1}) + (F_{t-1} - F_t) - L_t + AM_t \\
&\qquad + (1 + ELTR)C_{t-1}.
\end{aligned}
$$

Recalling that

$$URL_t = URL_{t-1} + L_t - AM_t$$

we finally obtain

$$
\begin{aligned}
E_t &= (PBO_t - PBO_{t-1}) + (F_{t-1} - F_t) \\
&\qquad + (URL_{t-1} - URL_t) + (1 + ELTR)C_{t-1}. \qquad (3.2)
\end{aligned}
$$

On average the three terms in brackets should equal zero. Given that the valuation interest rate is equal to the earned rate of return on assets, on average $(1 + ELTR)C_{t-1}$ is equal to $1.02 \cdot 9.361 = 9.548$.

REMARKS 1. We can now justify our including interest on benefits at two different rates in Eq (3.1). Failure to do so would have resulted in an extra term $(DSCR_{t-1} - ELTR)B$ in Eq. (3.2). In our model this term is not zero on average, and E_t would have been equal to $9.558+(.01-.02)15 = 9.408$ on average. In practice the treatment of interest on benefits is probably not a major concern, especially since benefits only get half a year's interest in most cases.

2. Eq. (3.2) is the counterpart of Eq. (9) in Berin and Lofgren (1987). Those authors avoid the problem noted above by using the same symbol $({}_0B_1)$ to represent benefits with interest at rate $DSCR$ (in their Eq. (1)) and also at rate $ELTR$ (in their Eq. (2)), even though these two rates are in general different. □

3.3. Sensitivity to the variance of discount rates

It is expected that the variability of pension expense would be significantly affected by changing the variance of discount rates (everything else remaining the same as in the base scenario). This is substantiated by Table 3.3 and Figure 3.5. When the standard deviation of $X(= \log(1 + DSCR))$ is close to zero, the standard deviation of E is close to 5.48. This is the variability attributable to the other source of gains and losses, namely returns on assets. As the standard deviation of X is increased, that of E increases more and more rapidly; the relationship appears more exponential than linear. This can be explained by the fact that PBO is extremely sensitive to $DSCR$. Table 3.4 shows the mean and variance of PBO for Stdev(X) ranging from 0 to .05. The very large figures obtained for the higher values of Stdev(X) are caused by the higher probability of low values of $DSCR$, which

Stdev(X)	**Stdev(E)**
.0000	5.480
.0025	5.531
.0050	5.625
.0075	5.761
.0100	5.944
.0125	6.181
.0150	6.483
.0175	6.866
.0200	7.348
.0225	7.953
.0250	8.716
.0275	9.679
.0300	10.902
.0325	12.466
.0350	14.487
.0375	17.125
.0400	20.618
.0425	25.308
.0450	31.707
.0475	40.575
.0500	53.054

Table 3.3. Standard deviation of expense (Stdev(E)) as a function of the standard deviation of geometric discount rates (Stdev(X)).

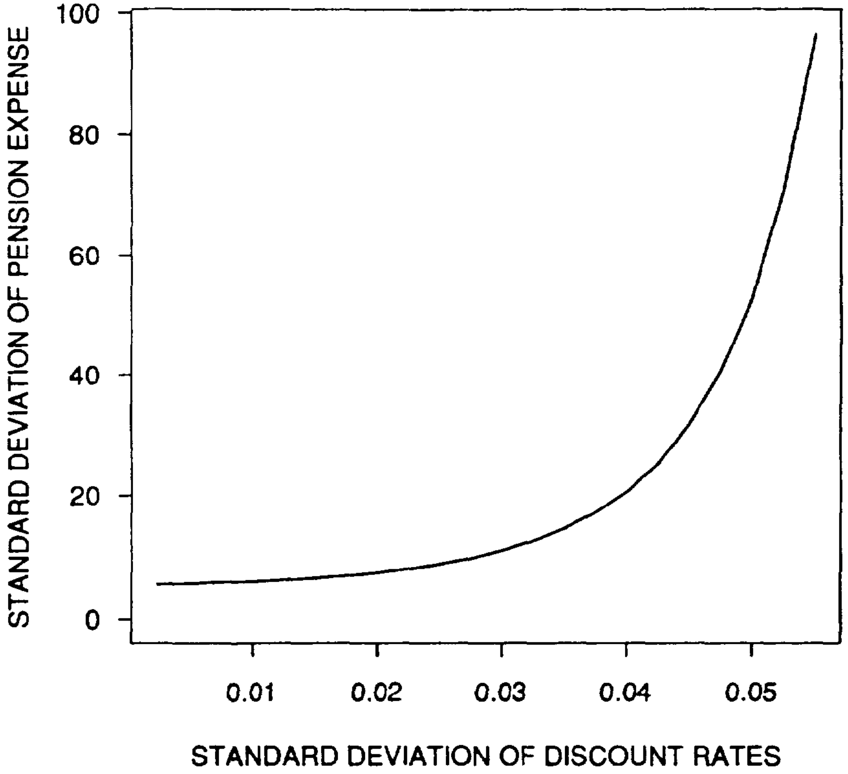

FIGURE 3.5. Standard deviation of pension expense as a function of the standard deviation of discount rates

result in a very large PBO (see Figure 3.3). The mean value is also affected; this is a consequence of the convexity of PBO as a function of $DSCR$.

Stdev(X)	Mean	Standard deviation
.00	326.58	0.00
.01	331.43	45.42
.02	347.16	102.87
.03	377.97	195.64
.04	433.96	394.09
.05	539.89	986.06

Table 3.4. Mean and standard deviation of pension benefit obligation (PBO) as functions of the standard deviation of geometric discount rates (Stdev(X)). The mean discount rate (EDS) remains equal to .02.

3.4. Sensitivity to the variance of rates of return on assets

Table 3.5 and Figure 3.6 show the standard deviations of pension expense for standard deviations of rates of return on assets ranging from 0 to .09. When Stdev(Y) = 0, Stdev(E) = 8.82, which thus represents the variability attributable to discount rate fluctuations only. (*N.B.* This value, together with the variability attributable to rates of return only, 5.48, add up to more than Stdev(E) under the base scenario, i.e. 10.902. This is because the variables $LPBO_t$ and LF_t are dependent.)

It is plain that Stdev(Y) has a much smaller influence on Stdev(E) than Stdev(X) has. This is what should be expected, since $LPBO_t$ has a much greater variability than LF_t.

Stdev(Y)	Stdev(E)
.000	8.82
.005	8.87
.010	8.96
.015	9.09
.020	9.25
.025	9.45
.030	9.69
.035	9.95
.040	10.25
.045	10.56
.050	10.90
.055	11.26
.060	11.64
.065	12.03
.070	12.44
.075	12.87
.080	13.30
.085	13.75
.090	14.21

Table 3.5. Standard deviation of expense (Stdev(E)) as a function of the standard deviation of geometric rates of return on assets (Stdev(Y)).

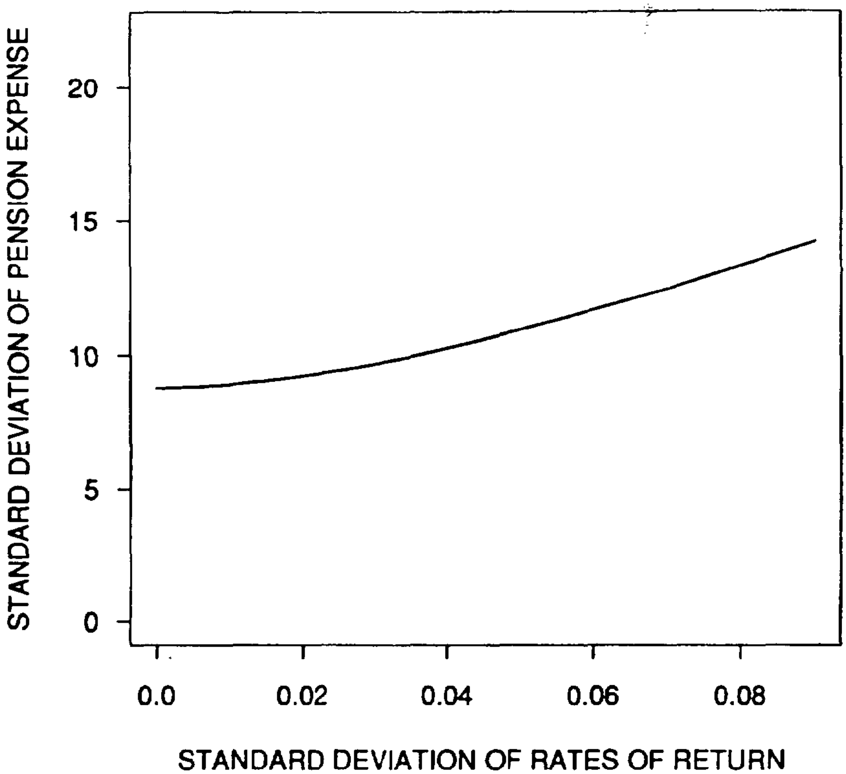

FIGURE 3.6. Standard deviation of pension expense as a function of the standard deviation of rates of return on assets

3.5. Sensitivity to the width of the corridor

Table 3.6 and Figure 3.7 show the dependence of Stdev(E) on the width of the corridor, other things equal. As expected, pension expense fluctuates less when the corridor is wider. The dependence is perhaps not as dramatic as one might have thought. For instance, if there is no corridor Stdev (E) is equal to 12.61; when a 10% corridor is allowed Stdev (E) becomes 10.30, a decrease of less than 15%.

Percentage	Stdev(E)
0.00	12.61
0.05	11.73
0.10	10.90
0.15	10.13
0.20	9.42
0.25	8.77
0.30	8.16
0.35	7.60
0.40	7.09
0.45	6.62
0.50	6.20

Table 3.6. Standard deviation of expense (Stdev(E)) as a function of percentage used for corridor

3.6. Sensitivity to the fraction of losses recognized when *URL* is outside corridor

The fraction k, representing the reciprocal of the average future working lifetime of active members, has a significant effect on the variability of

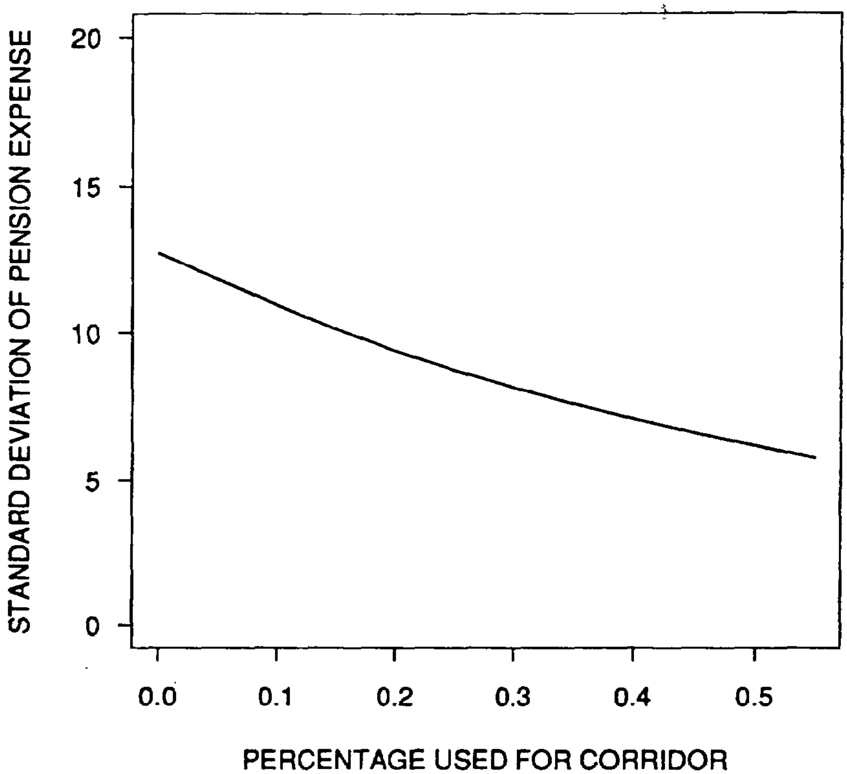

FIGURE 3.7. Standard deviation of pension expense as a function of the percentage used for the corridor

pension expense. This is shown in Table 3.7 and Figure 3.8.

When $k = 0$, there is no amortization of losses ($AM_t = 0$) and

$$\begin{aligned} E_t &= SC_{t-1}(1 + DSCR_{t-1}) + DSCR_{t-1}(PBO_{t-1} - B) \\ &\qquad - ELTR(F_{t-1} - B) \\ &= B - ELTR(F_{t-1} - B) \end{aligned}$$

so that

$$\begin{aligned} \text{Stdev}(E_t) &= ELTR \cdot \text{Stdev}(F_{t-1}) \\ &= .02 \cdot \text{Stdev}(R_{t-1}) \cdot AL/(1 + VI) \\ &= .02 \cdot .051 \cdot 287.58/1.02 \\ &= .288 \end{aligned}$$

(Notice that URL has no limit distribution in this case.)

Stdev(E) is an increasing function of k. The dependence is nearly linear (Figure 3.8). This is not surprising. In Section 1.3, for the system

$$\begin{aligned} X_{t+1} &= X_t + C_t + D_{t+1} \\ C_t &= -kX_t \end{aligned}$$

where (D_t) are independent adn identically (i.i.d.) disturbances, we had found

$$\text{Var } C = \frac{k}{1-k}\sigma^2, \qquad \sigma^2 = \text{Var } D_t$$

$$\Rightarrow \text{Stdev}(C) = \sigma\left(\frac{k}{2-k}\right)^{1/2}.$$

The graph of this function is very similar to Figure 3.8. In fact the approximation

$$\text{Stdev}(E) \doteq \left(A + B\frac{k}{2-k}\right)^{1/2}$$

k=1/AFWL	AFWL	Stdev(E)
0.000	∞	0.288
0.001	1000.00	0.956
0.005	200.00	2.18
0.010	100.00	3.22
0.020	50.00	4.89
0.025	40.00	5.64
0.033	30.00	6.81
0.040	25.00	7.69
0.050	20.00	8.94
0.100	10.00	14.44
0.150	6.67	19.13
0.200	5.00	23.32
0.250	4.00	27.17
0.300	3.33	30.76
0.333	3.00	33.05
0.350	2.86	34.17
0.400	2.50	37.43
0.450	2.22	40.59
0.500	2.00	43.67
0.550	1.82	46.69
0.600	1.67	49.68
0.650	1.54	52.64
0.700	1.43	55.58
0.750	1.33	58.52
0.800	1.25	61.48
0.850	1.18	64.46
0.900	1.11	67.46
0.950	1.05	70.51
1.000	1.00	73.62

Table 3.7. Standard deviation of expense (Stdev(E)), as a function of amortization period allowed ($AWFL$), or its reciprocal (k).

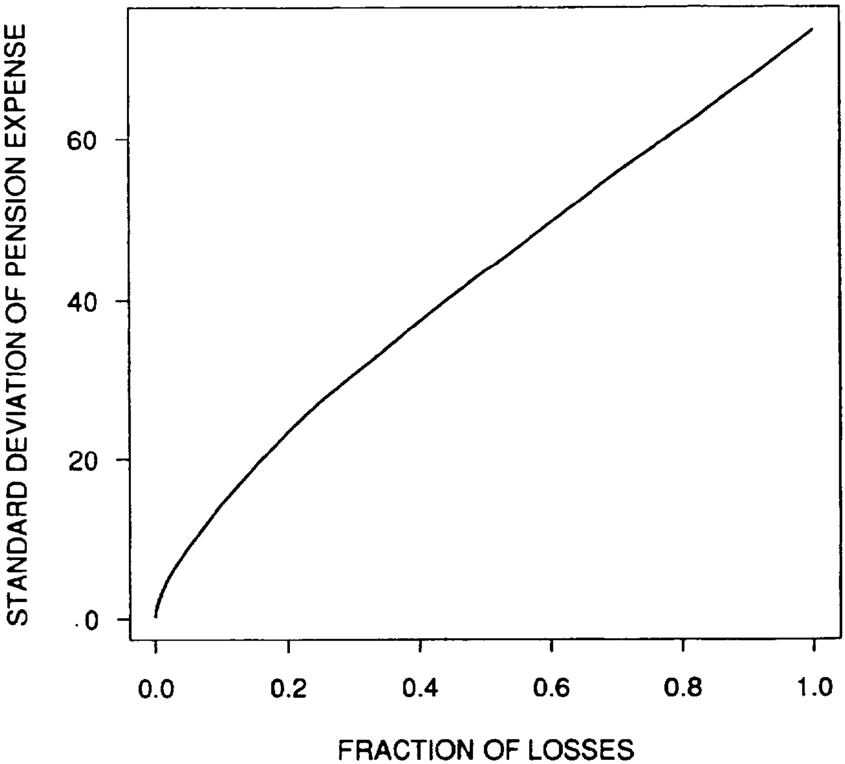

FIGURE 3.8. Standard deviation of pension expense as a function of the fraction of losses recognized when URL is outside corridor

is relatively good for well chosen constants A and B. This shows that the qualitative response of the system is not totally different when disturbances are dependent and a corridor is used, compared to the case of i.i.d. disturbances and no corridor.

Regarding the corridor approach to gains/losses amortization, one interesting question is the following: given that Stdev(E) $= 10.90$ when $k = 1/15$ and there is a 10% corridor, for what k' do we get the same value for Stdev(E), but when there is *no* corridor? In this case we find $k' = .0545 = 1/18.35$. In other words, as far as variability of pension expense is concerned, extending the amortization period ($AFWL$) from 15 to 18.35 years has the same effect as allowing a 10% corridor.

Appendix 3.1. The skewness of the distributions of pension expense and unrecognized losses

The corridor approach to gains/losses amortization operates in a symmetric fashion, and the gains or losses themselves have a distribution which is roughly symmetric about 0. So, why do E and URL have significantly skewed distributions? The answer to this question follows.

Under the base scenario, the major component of the loss L_t is

$$LPBO_t = PBO_t - PBO_{t-1}.$$

When the discount rate is a stationary normal process, as is the case in the present model, $LPBO_t$ will always have a symmetric distribution. This can be proved as follows. Let

$$\begin{aligned} p(x) &= PBO \quad \text{valued at geometric rate } x \\ f(x_1, x_2) &= p(x_1) - p(x_2) \\ X_t &= \text{geometric discount rate at time } t \end{aligned}$$

so that

$$\begin{aligned} PBO_t &= p(X_t) \\ LPBO_t &= p(X_t) - p(X_{t-1}) \\ &= f(X_t, X_{t-1}). \end{aligned}$$

The notation $U \stackrel{d}{=} V$ ("U equals V in distribution") will mean that the variable U and V have the same distribution. If (X_t) is a stationary normal process, then

$$(X_t, X_{t-1}) \stackrel{d}{=} (X_{t-1}, X_t).$$

This is because means and covariances uniquely determine normal distributions. Thus

$$\begin{aligned} -LPBO_t &= p(X_{t-1}) - p(X_t) \\ &= f(X_{t-1}, X_t) \\ &\overset{d}{=} f(X_t, X_{t-1}) \\ &= LPBO_t \end{aligned}$$

i.e. $PBO_t - PBO_{t-1}$ and $PBO_{t-1} - PBO_t$ have the same distribution, which therefore has to be symmetric about 0.

The equations describing the evolution of (E_t) and (URL_t) are

$$E_t = B - ELTR(F_{t-1} - B) + AM_t \tag{3.3}$$

$$URL_t = URL_{t-1} + L_t - AM_t \tag{3.4}$$

$$\begin{aligned} AM_t = [&\text{excess of } |URL_{t-1}| \text{ over } 10\% \\ &\times \max(F_{t-1}, PBO_{t-1}] \cdot \text{sign}(URL_{t-1}) \end{aligned} \tag{3.5}$$

$$L_t = PBO_t - PBO_{t-1} + (ELTR - R_t)AL/(1 + VI) \tag{3.6}$$

This system lacks symmetry in the following respects: (1) the term $ELTR \cdot F_{t-1}$ in (3.3); (2) $\max(F_{t-1}, PBO_{t-1})$ and URL_{t-1} are dependent; (3) the distribution of $ELTR - R_t$ in (3.6) is not symmetric about 0. These facts appear relatively unimportant; indeed, E and URL are still skewed even if these aspects are changed to make the system completely symmetric (e.g. remove term $-ELTR \cdot F_{t-1}$ in (3.3), etc). The same holds even when the corridor is removed. The only way the author found (by trial and error) of making E and URL symmetric was to replace PBO_t by a *normal* variable in (3.6). This is what led to the following interpretation of the problem.

Consider a system

$$U_t = aU_{t-1} + V_t, \qquad |a| < 1$$
$$V_t = e_t - e_{t-1}$$

where (e_t) is a sequence of i.i.d. random variables with mean 0 and variance 1. This is a simpler system than the one described above, but both have one essential feature in common: the disturbances (V_t) are dependent and have a symmetric distribution.

Suppose (U_t) is stationary (this is possible because $|a| < 1$). Let us calculate the first and third moments of U_t. Clearly $EU_t = 0$. Thus (U_t) cannot be symmetric if we find that $EU_t^3 \neq 0$. We have

$$EU_t^3 = a^3 EU_{t-1}^3 + 3a^2 \underset{\text{(I)}}{EU_{t-1}^2 V_t} + 3a \underset{\text{(II)}}{EU_{t-1} V_t^2} + \underset{\text{(III)}}{EV_t^3}$$

(I) $$EU_{t-1}^2 V_t = E(a^2 U_{t-2}^2 + 2aU_{t-2}V_{t-1} + V_{t-1}^2)(e_t - e_{t-1})$$
$$= -2a \underset{\text{(Ia)}}{EU_{t-2}V_{t-1}e_{t-1}} - \underset{\text{(Ib)}}{EV_{t-1}^2 e_{t-1}}$$

(Ia) $$EU_{t-2}V_{t-1}e_{t-1} = EU_{t-2}(e_{t-1} - e_{t-2})e_{t-1}$$
$$= 0$$

(Ib) $$EV_{t-1}^2 e_{t-1} = E(e_{t-1}^2 - 2e_{t-1}e_{t-2} + e_{t-2}^2)e_{t-1}$$
$$= Ee_{t-1}^3 = Ee_t^3$$

$(\therefore)$ $$EU_{t-1}^2 V_t = -Ee_t^3$$

(II)
$$EU_{t-1}V_t^2 = EU_{t-1}(e_t^2 - 2e_te_{t-1} + e_{t-1}^2)$$
$$= EU_{t-1}e_{t-1}^2$$
$$= E(aU_{t-2} + V_{t-1})e_{t-1}^2$$
$$= EV_{t-1}e_{t-1}^2$$
$$= E(e_{t-1} - e_{t-2})e_{t-1}^2$$
$$= Ee_{t-1}^3 = Ee_t^3$$

(III)
$$EV_t^3 = E(e_t - e_{t-1})^3 = 0$$

$(\therefore)$
$$EU_t^3 = a^3EU_{t-1}^3 - 3a^2Ee_t^3 + 3aEe_t^3$$

In these calculation we have repeatedly used the independence assumption regarding (e_t), and assumed that $E|e_t|^3 < \infty$. If (U_t) is stationary, we therefore obtain
$$EU_t^3 = 3a(1-a)Ee_t^3/(1-a^3).$$
Thus U_t cannot be symmetric if e_t is not itself so. (However it can be shown that the skewness coefficient of U_t will always be less than that of e_t.)

Another way of viewing the problem is as follows. We have
$$U_t = V_t + aV_{t-1} + a^2V_{t-2} + \cdots$$
$$= f(\tilde{V}_t)$$

where $\tilde{V}_t = (V_t, V_{t-1}, \dots)$ and f is the function from $\mathbf{R}^\infty$ to $\mathbf{R}$ defined as
$$(x_1, x_2, \dots) \longmapsto f(x_1, x_2, \dots) = \sum_{j=1}^{\infty} a^{j-1}x_j.$$

U_t is symmetric about 0 if, and only if, $-U_t \stackrel{d}{=} U_t$. Here

$$-U_t = -f(\tilde{V}_t) = f(-\tilde{V}_t)$$

Thus, a sufficient condition for U_t to have a symmetric distribution is that $-\tilde{V}_t \stackrel{d}{=} \tilde{V}_t$. (Open question: is the condition also necessary?) In general this is not the case here, since

$$\tilde{V}_t = (e_t - e_{t-1}, e_{t-1} - e_{t-2}, \dots)$$
$$-\tilde{V}_t = (e_{t-1} - e_t, e_{t-2} - e_{t-2}, \dots)$$

do not have the same distribution, except in some special cases (e.g. if e_t has a symmetric distribution to start with).

The above arguments break down when there is a corridor, since (1) we can no longer calculate EU_t^3 explicitly, and (2) we can not express U_t as $f(\tilde{V}_t)$ for some function $f : \mathbf{R}^\infty \to \mathbf{R}$. But it intuitively makes sense to believe that the same situation prevails.

CHAPTER 4 CONCLUSIONS

This chapter summarizes and complements the paper. The first section states the main conclusions; Section 4.2 discusses the practical utility of the methodology presented; Section 4.3 describes how the model is to be used; Section 4.4 studies an alternative scenario, in an attempt to see whether the conclusions reached in Chapter 3 hold more generally; Section 4.5 discusses two points which were raised in relation to the model used in the paper: (1) why negative discount rates arise, and (2) the exact methodology used to perform the simulations; finally, Section 4.6 provides some ideas for future research.

4.1. Main conclusions

This paper focused primarily on two aspects of SFAS 87:

(1) the consequences of the variability of the discount rate, and

(2) the minimum requirement for amortization of gains and losses.

The model adopted includes a stationary population and stochastic processes representing discount rates and rates of return on assets. Pension expense is therefore also a stochastic process, which can be studied mathematically or with the help of computer simulations.

The four main conclusions of the paper are listed below. The first and second are mathematical results which hold in all cases. The third conclusion is based on computer simulations and, therefore, may not hold with the same generality as the previous two. At the time of writing the last one is still a conjecture.

1. When the plan population is mature, the quantity "service cost plus interest on PBO" is not sensitive to variations in the discount rate. This is proved at the end of Section 1.4.

2. Given the stochastic processes assumed for discount rates and rates of return on assets, pension expense and unrecognized losses have a limit (or "steady-state") distribution. This is proved mathematically in Chapter 2. Observe that it is essential that appropriate amortization rules be applied, for both funding and accounting purposes. One case where this requirement is not met is described in Section 4.4.

3. It is not possible to calculate explicitly the limit distribution (or even the moments) of pension expense. But this can be done using simulations. A "base scenario" was chosen, specifying the behaviour of discount rates, rates of return on assets, etc. (see Section 1.4). A sensitivity analysis was then conducted with respect to four factors:

(a) variance of discount rates;

(b) variance of rates of return on assets;

(c) width of corridor;

(d) fraction of unrecognized losses included in expense when outside the corridor.

(For example, in case (a), the limit distribution of pension expense was computed assuming that discount rates have variance 0, then .0001, and so on, yielding the variance of pension expense for variances of discount rates in the interval [0, .0025]. The results show how important the first factor is in determining the volatility of pension expense. The same was done for the other factors, changing only one of them at a time.)

The first factor was by far the most important determinant of the variability of pension expense. Even moderate fluctuations in the discount rate produce sizeable fluctuations in expense. The sensitivity of the variance of expense with respect to the variance of discount rates is very high. The second and third factors were not very important by comparison. In the case of rates of return on assets this is easily understood, since gains/losses on return on assets are of smaller magnitude than those caused by variations of the discount rate. As to the width of the corridor, it was observed that the 10% corridor allowed under SFAS 87 decreased the standard deviation of expense by 14%, when compared with the case where no corridor is allowed. The same thing could be achieved by slightly increasing the amortization period permitted (the "average future working lifetime" of active employees under SFAS 87). The last factor turned out to be relatively important in influencing the variability of expense. (This makes sense intuitively when the pension accounting "system" is interpreted from the point of view of control theory, as was done in Chapter 1.)

The results of the sensitivity analysis relate only to the base scenario chosen, and, strictly speaking, it is impossible to predict what the results would be if a different base scenario were used. Other simulation results are presented in Section 4.4

4. As was pointed out above, allowing a corridor based on 10% of the maximum of the pension benefit obligation and fund value does not drastically reduce the variability of pension expense. One possible explanation is that gains and losses do not "cancel over time", as some apparently believe; on the contrary, it appears that their cumulative sum eventually becomes arbitrarily large, even when actuarial assumptions are "correct on average".

This claim has been shown to be correct when successive gains and losses are not correlated. (This was discussed in two talks recently given by the author, one at the International Congress of Actuaries (June 1992) and the other at the most recent Actuarial Research Conference (August 1992). A written account of these talks will appear in ARCH.) Under the model described in this paper gains and losses are correlated, and a mathematical proof has yet to be found. The author is currently studying this problem.

4.2. Practical utility of the methodology presented

The paper describes how the limit distribution of pension expense can be computed and then used to study the effects of some of the accounting rules contained in SFAS 87. It is stated in the Preface that the methodology may be useful in two situations:

(a) when making accounting or funding decisions concerning a specific pension plan, and

(b) when assessing the effects of new funding or accounting rules on pension plans at large.

In case (a), a shorter horizon would usually be appropriate, say 10 or 20 years. The variability of pension expense may be obtained for each future year within that period, for any given funding and/or accounting strategy. This "methodology" is not new to actuaries, since they have been performing pension plan simulations for a long time. The author has done a number of short-term simulations using the model population and base scenario. It appeared that the distributions of pension expense, unrecognized losses, etc., at duration 20 were not very different from the limit distributions, though initial conditions still had some importance. It is thus plausible that in some specific cases an infinite horizon may be relevant, for instance if the plan population is initially mature and is supposed to remain so for some time.

The methodology should be more useful in situation (b). The effects of alternative accounting (or funding) rules can be assessed by comparing the limit distributions obtained. The fact that these limit distributions are independent of initial conditions now becomes an advantage. The procedure could be applied to the rules concerning discount rates, amortization periods, width of corridor allowed, etc. One restriction is that the limit distribution may not exist, in cases where the rules for amortizing gains/losses do not constitute a proper "control" of the system. This problem arises with the Alternative Scenario as it is initially described in Section 4.4.

4.3. How the model is to be used

One should distinguish between cases (a) and (b) discussed above.

In case (a), the following have to be determined:

– time horizon (for example $n = 20$ years);

– evolution of population and benefits;

– economic scenario (rates of return on the various asset classes, indicators used to set the discount rate, inflation);

– actuarial assumptions (may be "path-dependent", that is to say dependent on the evolution of the economic/financial scenario);

– how accounting parameters (discount rate, expected long-term rate of return on assets, amortization period) are determined;

– funding method (actuarial cost method plus amortization of gains and losses and unfunded liabilities);

– accounting methods (e. g. faster/slower recognition of gains/losses and liabilities).

The simulations then yield frequency distributions for the variables of interest (e. g. pension expense) for each of the n future years under consideration. There remains the problem of making a decision based on these n distributions. There is no clear-cut answer here. Many "decision functions" are possible, among others:

- considering only the results of the last year;

- averaging some the results (e. g. variances);

- "discounting" results (i. e. giving a relatively smaller weight to more distant years).

Case (b) requires similar assumptions and parameters. Choosing an infinite horizon is not mandatory , but avoids the problem of multiple distributions (since there is only one limit distributions for each variable). In this case all amounts have to be deflated (otherwise they grow without bounds), for example by expressing them as fractions of payroll.

4.4. Study of an alternative scenario

After a first draft of this paper had been written, an alternative scenario was suggested. This scenario differs from the base scenario in the following respects:

1. Standard Deviation of Return on Assets: 10%, rather than 5%.

2. Standard Deviation of Discount Rates: 0.5%, rather than 3%.

3. Relation of Mean Discount Rate to Valuation Interest Rate: the former exceeds the latter by 0.25%.

4. Funding rules: no negative contributions.

This scenario is of great interest. Some comments follow.

First, the higher standard deviation for rates of return appears more realistic, in view of the high variability of returns on stocks (see for instance Table 1.3, p. 29). (I have not found American statistics on the subject, but according to Table 7 of the *Report on Canadian Economic Statistics, 1924-1991*, rates of return on Canadian pension plan assets showed a standard deviation of 8.88% over the period 1967-1991.)

As to the third assumption, I chose to let the mean discount rate remain unchanged at 1%, which implies a valuation rate equal to .75%.

There is a technical problem with the last assumption, because what the simulations determine is a limit distribution which does not always exist, even if the accounting or funding rules are justifiable in the real world. (This is why the paper had to include mathematical proofs for the existence of the limit distributions of pension expense and unrecognized losses.) If negative contributions are not allowed, then fund values may not have a steady-state distribution. I will give two justifications for this claim, one theoretical and the other more intuitive.

First justification. Let us return to the theorem given on pp. 56 and 57. The theorem says that a process (if it satisfies the other conditions stated) will have a limit distribution if it has the property of "reverting to the center of the space". If no negative contributions are allowed, the equation describing the evolution of the fund becomes:

$$\begin{aligned} F_{t+1} &= (1 + R_{t+1})AL/(1 + i), \quad \text{if} \quad NC + AL - F_t \geq 0; \\ &= (1 + R_{t+1})(F_t - B), \quad \text{if} \quad NC + AL - F_t < 0. \end{aligned}$$

It can be seen that when F_t is larger than $NC + AL$ no control is applied to keep it from becoming even larger. Since there is a positive probability that

F_t will become larger than $NC + AL$, it is certain that this will eventually cause the fund to grow without bounds.

Second justification. Consider the following example. Suppose a certain amount of money is invested initially, and that the return on the fund is paid out every year (no new money is deposited into the fund after it is established). To keep things simple, just ignore the possibility of negative returns. Suppose that you simulate the operation of this fund over a long period. Then clearly the value of the fund will reach a stationary distribution (if it is assumed that the returns on the fund themselves reach such a distribution). Every year the fund will revert to its initial value, and the only randomness left is the effect of the rates of return over one year. Now suppose that you modify the rules, and say that returns will only be paid out up to a certain fixed level, say 5% of the fund value (this is similar to the interdiction of negative contributions in pension funding). Then every time returns exceed 5% there will be a net addition to the fund, and over time the fund will get larger and larger (without bounds). Consequently, there will not be a limit distribution for the value of the fund.

Disallowing negative contributions may not cause any difficulty in practice because, among other things,

(a) gains and losses are amortized over more than one year, which lowers the variability of contributions (see Dufresne, 1989);

(b) the plan sponsor will take "contribution holidays" long enough to use up the surplus; and

(c) benefits are increased, actuarial assumptions are changed, etc., implying that negative contributions are not very likely to occur.

It is not possible to investigate the Alternative Scenario as it was suggested (running the computer programs produces meaningless overflows); pension expense apparently does not have a limit distribution if negative contributions are not permitted. The author therefore decided to study two modifications of that scenario which do yield limit distributions:

Modified Alternative Scenario I (MAS I) : Alternative scenario as described above, except for the last assumption (i. e. negative contributions *are* allowed).

Modified Alternative Scenario II (MAS II) : Same as Alternative Scenario I, except that

(a) funding gains and losses are amortized over 15 years;

(b) rates of return on assets are independent.

The scenarios are summarized in Tables 4.1 and 4.2. The first modification does lead to a limit distribution for pension expense, but negative contributions occur very often, due to the large standard deviation of returns on assets . The theoretical probability of a negative contribution is computed as follows (using the data in Table 4.1 and on p. 110):

$$R = \text{ rate of return on assets } = e^Y - 1$$

$$\begin{aligned} \mathrm{P}(F > AL + NC) &= \mathrm{P}(e^Y AL/(1 + ER) > AL + NC) \\ &= \mathrm{P}\Big(Y > \log[(1 + ER)(1 + NC/AL)]\Big) \\ &= .3398, \end{aligned}$$

where Y is the geometric rate of return on assets. Out of the one million iterations performed, there were 339,679 for which the contribution was negative.

Arithmetic valuation rate of interest (VI)	.0075
Expected long-term arithmetic rate of return on plan assets ($ELTR$)	.0075
Mean arithmetic discount rates (EDS)	.01
Standard deviation of geometric discount rate ($VARDS^{1/2}$)	.005
Mean arithmetic rate of return on assets (ER)	.0075
Standard deviation of geometric rate of return on assets ($VAROR^{1/2}$)	.10
Correlation between geometric discount rate and rate of return on assets (COR)	.60
Fraction of max (PBO, F) used for corridor (C)	.10
Fraction of excess of $\|URL\|$ over $C \cdot \max(PBO, F)$ recognized in expense	1/15
Amortization period for funding gains and losses (years)	1

Table 4.1. Modified Alternative Scenario I.

Under the second modification the variance of contributions is much lower . Out of the one million iterations performed, there were 96,840 for which the contribution was negative (a frequency of about 10%, which is significantly less than with the first modified scenario).

REMARKS. 1. When returns are lognormal the distribution of gains has range $(-\infty, +\infty)$. Hence, whatever the way gains are amortized there is always a positive probability that contributions will become negative.

2. The present remark explains why assumption (b) was added in MAS II. Rates of return on assets were supposed independent because otherwise the average accounting gain or loss would not be zero any more. To see why

Arithmetic valuation rate of interest (VI)	.0075
Expected long-term arithmetic rate of return on plan assets ($ELTR$)	.0075
Mean arithmetic discount rates (EDS)	.01
Standard deviation of geometric discount rate ($VARDS^{1/2}$)	.005
Mean arithmetic rate of return on assets (ER)	.0075
Standard deviation of geometric rate of return on assets ($VAROR^{1/2}$)	.10
Correlation between geometric discount rate and rate of return on assets (COR)	0
Fraction of max (PBO, F) used for corridor (C)	.10
Fraction of excess of $\|URL\|$ over $C \cdot \max(PBO, F)$ recognized in expense	1/15
Amortization period for funding gains and losses (years)	15

Table 4.2. Modified Alternative Scenario II.

this is so, consider the expression for the accounting loss (p. 33):

$$L_t = PBO_t - PBO_{t-1} + (ELTR - R_t)(F_{t-1} + C_{t-1} - B).$$

The first part of the loss is the increase of the PBO during the year, and has mean zero. When funding gains/losses are amortized over one year the second part (LF_t) boils down to

$$(ELTR - R_t)AL/(1 + VI)$$

(Eq. (3.1), p. 76) which also has mean zero. For longer amortization periods there is no such simplification, and the dependence between rates of return and fund values comes into play. With fifteen year amortization and a correlation of .60 between discount rates and rates of return, there is a correlation of -.063 between $(ELTR - R_t)$ and $(F_{t-1} + C_{t-1} - B)$. Even though the average value of R_t is exactly equal to $ELTR$, on average LF is equal to -7.00, which significantly decreases average pension expense. It was feared that this would distort comparisons between MAS II and the other scenarios. The problem is avoided by making rates of returns on assets independent (which is achieved by setting $COR = 0$ in the model). □

Simulations were performed using the altenative scenarios. The results are shown in Tables 4.3 and 4.4; the corresponding results for the base scenario are shown in Table 3.2 (p. 76). Average pension expense is higher than before; this is because the valuation rate of interest and rates of return on assets are lower, producing larger funding contributions (see Eq. (3.2), p. 81). We now have

$$AL(@\ VI = .75\%) = 337.70$$

$$NC(@\ VI = .75\%) = 12.486.$$

The standard deviations of pension expense (E), unrecognized losses (URL) and amortization payments (AM) are higher under MAS I than under the base scenario. Nevertheless the standard deviation of annual losses is significantly smaller. This deserves a few words of explanation. Under the base scenario, the part of the loss due to the increase or decrease of the pension benefit obligation ($LPBO$) has a very large standard deviation, in fact ten times larger than under MAS I. But the losses on the fund (LF) have a greater variability under MAS I than under the base scenario. All these facts are explained by the lower standard deviation of discount rates and the higher standard deviation of rates of return on assets. The

Variable	Mean	Standard deviation
Pension expense (E)	12.603	13.37
Unrecognized losses (URL)	0.67	226.66
Amortization payment (AM)	0.02	13.24
Losses (L)	0.02	36.50
Losses due to increase (decrease) in PBO ($LPBO$)	0.00	9.89
Losses due to return on fund (LF)	0.02	33.84
Pension benefit obligation (PBO)	327.77	22.05
Geometric discount rate (X)	0.009938463	0.004996001
Geometric rate of return (Y)	0.002408076	0.099965752
Arithmetic discount rate ($DSCR$)	0.010000618	0.005045970
Arithmetic rate of return (R)	0.007431756	0.1009481282

Table 4.3. Observed means and standard deviations of some of the variables, under Modified Alternative Scenario I (one million iterations).

reason why URL has a larger variability under MAS I, while annual losses have a standard deviation three times smaller, is that the losses on the fund are positively correlated. This can be seen from the expression for these losses, which is

$$LF_t = (ELTR - R_t)AL/(1 + VI)$$

when funding gains and losses are amortized over one year (see Eq. (3.1), p. 76). The variance of the sum of the LF's is larger than the sum of the variances, because of the positive correlation between the R's; thus URL tends to take larger values. This compounding effect does not occur for the

Variable	Mean	Standard deviation
Pension expense (E)	12.610	6.18
Unrecognized losses (URL)	0.75	113.94
Amortization payment (AM)	0.03	5.66
Losses (L)	0.03	36.05
Losses due to increase (decrease) in PBO ($LPBO$)	0.00	9.89
Losses due to return on fund (LF)	0.03	34.65
Pension benefit obligation (PBO)	327.77	22.05
Geometric discount rate (X)	0.009938463	0.004996001
Geometric rate of return (Y)	0.002408076	0.099965752
Arithmetic discount rate ($DSCR$)	0.010000618	0.005045970
Arithmetic rate of return (R)	0.007431756	0.1009481282

Table 4.4. Observed means and standard deviations of some of the variables, under Modified Alternative Scenario II (one million iterations).

other part of the losses ($LPBO$), because the latter partly cancel over time:

$$LPBO_1 + \cdots + LPBO_t = PBO_t - PBO_0.$$

It can also be seen that the average value of the PBO has significantly decreased, though it is still larger than when it is valued at $i = .01 =$ average discount rate; this is not surprising (see p. 77).

It is also interesting to compare MAS I and MAS II. The standard deviations of L, LF and $LPBO$ are nearly identical, but those of E and URL are much smaller under MAS II. This is entirely attributable to the

Percentage	Stdev(E)
0.00	13.97
0.05	13.65
0.10	13.37
0.15	13.10
0.20	12.85
0.25	12.61
0.30	12.39
0.35	12.18
0.40	11.99
0.45	11.80
0.50	11.62

Table 4.5. Standard deviation of expense (Stdev(E)) as a function of percentage used for corridor, under Modified Alternative Scenario I.

fact that under MAS II rates of return are independent: the variance of the sum of the LF's is now equal to the sum of the variances. This is a striking example of the fact that in this sort of model it is not sufficient to know the distributions of interest rates and rates of return at each point in time. The correlation structure of these processes is also very important.

Sensitivity analyses were conducted (for both MAS I and MAS II), with respect to the same four factors as in Chapter 3: (1) variance of discount rates, (2) variance of rates of return on assets, (3) width of the corridor and (4) fraction of excess unrecognized losses included in expense. Only the results concerning the width of the corridor are shown (see Tables 4.5 and 4.6, which correspond to Table 3.6, p. 88, for the base scenario).

Percentage	Stdev(E)
0.00	7.11
0.05	6.60
0.10	6.18
0.15	5.82
0.20	5.51
0.25	5.24
0.30	5.01
0.35	4.80
0.40	4.62
0.45	4.45
0.50	4.30

Table 4.6. Standard deviation of expense (Stdev(E)) as a function of percentage used for corridor, under Modified Alternative Scenario II.

Overall the results are very similar to those described in Section 4.1 for the base scenario. The variance of discount rates was found to be the most important factor determining the volatility of pension expense. The last factor was found to be relatively important, while the second and third were relatively less important.

Once again the effect of allowing a 10% corridor is not very great. Under MAS I it decreases the standard deviation of pension expense by 4.5%, while there is a 15% decrease under MAS II.

4.5. Answers to two specific questions

This section clarifies two points which were raised in connection with the model: (1) Why is it that discount rates can take negative values under the base scenario? (2) Describe the exact methodology used to perform the simulations, in enough detail to enable one to reproduce the results.

4.5.1. Negative discount rates

In the model all amounts are deflated by wage increases; correspondingly, discount rates are net of rates of wage increase. In some years the rate of increase of wages is higher than the nominal discount rate, thus producing a negative (*net*) discount rate. (This possibility is mentioned on page 76.) Under the base scenario (Chapter 3), the probability that the discount rate takes a negative value can be calculated as follows:

$$
\begin{aligned}
& DSCR = e^{X} - 1, \qquad X \sim \mathrm{N}(.00950, .0009) \\
\Longrightarrow \quad & \mathrm{P}(DSCR < 0) = \mathrm{P}(X < 0) \\
& \qquad = \mathrm{P}\left(Z < \frac{0 - .00950}{.03}\right), \quad \text{where} \quad Z \sim \mathrm{N}(0,1) \\
& \qquad = .376.
\end{aligned}
$$

This probability is surprisingly high. What is even more surprising is that it *understates* what actually happened in US economic history: during the period 1926-1990, the annual rate of increase of the Wage Index was higher than annual average long-term US bond yields in 27 out of the 65 years, representing a frequency of 27/65 = .415 (*Economic Statistics for Pension Actuaries*, August 1991, Table 11A).

4.5.2. Exact methodology used to perform the simulations

All the variable names and equations, as well as a description of the base scenario, are given in Section 1.4 (see pp. 25-35). A Fortran program simulated the operation of the pension fund and pension accounts, year after year. Initially the fund was set equal the actuarial liability, and unrecognized losses were set equal to zero. One million iterations were done (double-precision) for each combination of the parameters considered. The formulas used to generate the random numbers (including initial values) are given on page 70.

4.6. Future research

The author has found this subject a fascinating area of research. The paper goes some way in formulating a model and answering some basic questions, but a lot more could be done. Here are a few ideas for future work:

(1) Sensitivity of the pension benefit obligation to variations in discount rates, given explicit assumptions as to salary and post-retirement benefit increases.

(2) Speed of convergence of distributions to their limits.

(3) Modelling discount rates and rates of return on assets using other sto- chastic processes.

(4) Suppose $URL_0 = 0$; how long does it take before URL escapes from the corridor?

REFERENCES

Balzer, L. A. (1982). Delayed profit/loss sharing feedback and persisting unpredicted claims. *Journal of the Institute of Actuaries* **109**: 285-311.

Balzer, L. A. and 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.

Beard, R. E., Pentikäinen, T. and Pesonen, E. (1984). *Risk Theory*. Chapman and Hall, London.

Benjamin, S. (1989). Driving the pension fund. Paper presented at the Joint Seminar on Applications of Mathematics in Insurance, Finance and Actuarial Work, London (England), 6-7 July 1989.

Berin, B. N. and Lofgren, E. P. (1987). Financial Accounting Standards No. 87: Recursion formulas and other related matters. *Transactions of the Society of Actuaries* **39**: 9-15.

Bowers, N. L. Jr et al. (1986). *Actuarial Mathematics*. Society of Actuaries, Itasca, Illinois.

Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Prob.* **18**: 211-220.

Bratley, P., Fox, B. and Schrage, L. E. (1987). *A Guide to Simulation*. Springer-Verlag, New York.

Canadian Institute of Actuaries. *Report on Canadian Economic Statistics 1924-1989*. June 1990.

de Jong, P. (1984). The mean square error of a randomly discounted sequence of uncertain payments. *Insurance: Mathematics and Economics* **3**: 173-178.

Dufresne, D. (1986a). The Dynamics of Pension Funding. Ph.D. thesis, School of Actuarial Science and Statistics, The City University, London (England).

Dufresne, D. (1986b). Pension funding and random rates of return. In: *Insurance and Risk Theory*, M. Goovaerts et al. (eds.): 277-291.

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

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

Dufresne, D. (1991). The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuarial J.* **1990**: 39-79.

Financial Accounting Standards Board (1985). *Statement of Financial Accounting Standards No. 87: Employers' Accounting for Pensions*. Financial Accounting Series No. 012, December 1985.

Fontaine, P. (1990). Peut-on prédire l'évolution des marchés d'actions à partir des cours et des dividendes passés? *Proceedings of the 1st AFIR International Colloquium*, Paris, 1990 **1**: 235-257.

Hicks, E. L. and Trowbridge, C. L. (1985). *Employer Accounting for Pensions: An Analysis of the Financial Accounting Standards Board's Preliminary Views and Exposure Draft.* Pension Research Council. Richard D. Irwin, Homewood, Illinois.

Jacobs, O. R. L. (1974). *Introduction to Control Theory.* Oxford University Press.

Martin-Löf, A. (1983). Premium control in an insurance system: An approach using linear control theory. *Scand. Actuarial J.* **1983**: 1-27.

O'Brien, T. (1986). A stochastic dynamic approach to pension funding. *Insurance: Mathematics and Economics* **5**: 141-146.

O'Brien, T. (1987). A two-parameter family of pension contribution functions and stochastic optimization. *Insurance: Mathematics and Economics* **6**: 129-134.

Panjer, H. H. and Bellhouse, D. R. (1980). Stochastic modelling of interest rates with applications to life contingencies. *Journal of Risk and Insurance* **47**: 91-110.

Smith, J. C. McK. (1984). A ballistic approach to actuarial problems. *Transactions of the Society of Actuaries* **36**: 501-522.

Society of Actuaries (1990). *Economic Statistics for Actuaries.* Chicago, Illinois.

Trowbridge, C. L. (1952). Fundamentals of pension funding. *Transactions of the Society of Actuaries* **4**: 17-43.

Tweedie, R. L. (1975). Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **3**: 385-403.

Vandebroek, M. and Dhaene, J. D(1991). Optimal premium control in a non-life insurance business. *Scand. Actuarial J.* **1990**: 3-13.