Some Notes on the Dynamics and Optimal Control of Stochastic Pension Fund Models in Continuous Time

Andrew J.G. Cairns [1] [2]
Department of Actuarial Mathematics and Statistics,
Heriot-Watt University,
Edinburgh, EH14 4AS,
United Kingdom

Abstract

This paper discusses the modelling and control of pension funds.

A continuous-time stochastic pension fund model is proposed in which there are n risky assets plus the risk-free asset as well as randomness in the level of benefit outgo. We consider Markov control strategies which optimise over the contribution rate and over the range of possible asset-allocation strategies.

For a general (not necessarily quadratic) loss function it is shown that the optimal proportions of the fund invested in each of the risky assets remain constant relative to one another. Furthermore, the asset allocation strategy always lies on the capital market line familiar from modern portfolio theory.

A general quadratic loss function is proposed which provides an explicit solution for the optimal contribution and asset-allocation strategies. It is noted that these solutions are not dependent on the level of uncertainty in the level of benefit outgo, suggesting that small schemes should operate in the same way as large ones. The optimal asset-allocation strategy, however, is found to be counterintuitive leading to some discussion of the form of the loss function. Power and exponential loss functions are then investigated and related problems discussed.

The stationary distribution of the process is considered and optimal strategies compared with dynamic control strategies.

Finally there is some discussion of the effects of constraints on contribution and asset-allocation strategies.

Keywords: continuous time; stochastic differential equation; asset-allocation; contribution strategy; Bellman equation; optimal control; constraints.

[1] E-mail: A.Cairns@ma.hw.ac.uk
[2] WWW: http://www.ma.hw.ac.uk/∼andrewc/

1 Introduction

The analysis and control of pension fund dynamics is becoming increasingly important as members start to pay more attention to the security of promised benefits and as sponsoring employers become more concerned about the timing and stability of cashflows.

This paper discusses some current problems in the analysis and control of defined benefit pension funds. Under a pure defined benefit pension fund the benefits payable to an individual member depend only upon his or her salary and length of past service.

The principal alternative to a defined benefit scheme is a defined contribution occupational pension scheme. Here the benefits are defined by the level of contributions which are paid into an individual member's fund or 'pot' and by the investment returns which are achieved over the period up to retirement. Since the pot is used to purchase an annuity at the time of retirement the level of pension is also determined by the annuity rate which prevails at the date of retirement and, in particular, the term structure of interest rates on that date. Generally the rates of contribution by the sponsor and by the member are fixed. All of the investment risk is borne by the member and there is no opportunity for the member to smooth out the effects of adverse investment returns. Existing literature on defined contribution problems typically deals with the case where the terminal utility is a function of the fund size at retirement (for example, see Merton, 1990, Gerber & Shiu, 2000, and Deelstra *et al.,*). The case where the terminal utility is a function of pension purchased at retirement (that is, fund divided by annuity rate) in a stochastic interest-rate environment has been considered by Cairns *et al.* (1999).

Under a defined benefit scheme the sponsoring employer has no ability to vary the timing or amount of the benefits payable. In contrast to this and to a defined contribution scheme the rate at which contributions are paid into the fund are (within limits) very flexible. Typically this flexibility rests fully with the fund sponsor while individual members contribute a fixed percentage of their salaries.

Increasingly, we also see schemes which provides elements of both defined benefit and defined contribution. Most common are schemes which allow for discretionary increases to pensions in payment with the size of the increase depending upon recent investment returns. Other 'hybrid' schemes provide a pension which is equal to the maximum of a defined benefit pension and a defined contribution pension.

Within the pure defined benefit framework there is considerable scope for freedom:

- in how the variable contribution rate should be varied;
- in the choice of asset allocation strategy.

1.1 Contributions

By-and-large, the fund sponsor has considerable freedom in how the contribution rate can be varied. The basic principle underlying how the contribution rate is set is that it should take account of the amount of surplus or deficit (that is, the excess of assets over liabilities). Thus, in some sense, the contribution rate can be reduced during periods of

surplus and increased above the normal rate when the scheme is in deficit. The role of the actuary is to take account of the needs of the sponsor and of the members before recommending to what extent surplus or deficit should affect the contribution rate.

The overall level of flexibility may be restricted by the presence of certain constraints:

- There may be a legal requirement to keep the funding level (the asset/liability ratio) above a certain minimum level (the method of calculation of which can take a number of forms). If the funding level drops below this minimum the sponsor may be compelled to make up the deficit immediately.
- Similarly there may be a restriction on the maximum size of the fund. This may require refunds to the sponsor or improvements to the benefits (although, in the latter case, the fund would cease then to be a 'pure' defined benefit scheme).
- The fund sponsor may wish to keep the contribution rate below a certain level (for example, twice the normal rate).
- Regulations or plan rules may prevent refunds to the employer, or perhaps refunds are only permitted when the funding level is sufficiently high.

1.2 Assets

A pension fund will normally fall under the responsibility of a group of *trustees* or *managers* who must act in the best interests of the fund members. Within this remit they can choose how to invest the assets of the fund. Appropriate investment strategies will take account of:

- prudence;
- requirements to
  - maximise returns;
  - minimise risk;
  - diversify;
  - avoid self-investment;
- immediate cashflow requirements;
- security;
- the tax status of the fund and of the various potential assets.

Besides taking the advice of their fund managers, trustees may also seek the advice of the fund actuary before deciding upon an appropriate strategy. How the funds available should be allocated presents an interesting problem for the actuary. The solution to such a problem must take account of many things:

- the balance between the conflicting interests of the members and the sponsor;
- the expected returns on the various assets and the associated risks and dependencies (both between individual assets and through time);
- the current level of funding;
- constraints on short selling of assets.

1.3 Objectives

For an actuary to set an optimal contribution rate and asset allocation strategy it is necessary to use a well defined objective function with appropriate constraints. Objective functions must be sufficiently precise to avoid ambiguous or non-sensical solutions. For example, the imprecise objective *minimise variance* leads to various outcomes which minimise the variance of the funding level and/or the contribution rate.

Other apparently precise objectives lead to optimal solutions which do not entirely make sense. In such circumstances it may be necessary to revise the objective function.

1.4 Types of model

A basic question which must be answered first is should we use a deterministic or a stochastic model. Deterministic models are adequate for cashflow projections and valuations but little else. Stochastic models, on-the-other-hand, allow us to investigate fully the dynamics of the fund through time and, for example, devise suitable control strategies. Here we consider stochastic models only.

A separate question is whether models should be kept simple or be made very realistic. The answer here depends on the reasons for modelling. In a more academic study we are looking for the major drivers of pension fund dynamics. Simple models allow detailed study of these factors. Often it is possible to derive analytical results which can then be used to provide specific links between causes and effects. A more complex model, on the other hand, may be required if the modeller has in mind a specific pension fund with a very specific benefit structure. As models become more complex we input more and more factors and find that more detail comes in the output from each simulation. It then becomes very difficult to identify why certain effects are evident. However, simple models provide the backup in the analysis of complex models. Such models give pointers to what we should be investigating. Thus we may be able empirically to observe the same links between causes and effects as were found analytically in the simple model. More-often-than-not such comparisons can explain, with ease, the majority of the variation in the dynamics of a complex model.

In some problems the aim may be to devise an optimal control strategy. As we show here it is possible using simple models to derive precisely an optimal control. This then gives us the starting point for further study and optimisation within a more complex model.

This paper has a number of aims. First, it will pull together some recent results in continuous-time pension fund modelling (O'Brien, 1986, 1987, Dufresne, 1990, Boulier

et al., 1995, 1996, and Cairns, 1996, 1997). Fresh proofs of these results will be presented as appropriate along with further discussion of their implications. Second, some new avenues will be developed to show how this earlier work can be modified to consider some generalisations and to pull the results closer to current practice. Third, the paper will discuss some open problems.

Within this framework the paper will proceed as follows. Section 2 introduces the continuous-time stochastic model for the dynamics of a pension fund in its most general form which will be used in the majority of the paper.

Section 3 considers dynamic stochastic control of the model by making reference to a value function which discounts exponentially future random values of a quadratic loss function. The section proceeds by looking at various cases both constrained and unconstrained. The advantages and disadvantages of the quadratic loss function are discussed in detail here. Finally, power and exponential loss functions are considered with problems similar to those under the quadratic loss function identified.

In Section 4 we take the longer-term view and consider the stationary distribution of the process (although the distribution of the model nears its stationary form within 10 to 15 years usually). This includes a look at the continuous proportion portfolio insurance approach to asset allocation introduced by Black and Jones (1988) and compares this with a static investment strategy. Section 5 compares the results of dynamic versus stationary optimisation derived in Sections 3 and 4 and shows how sensitive these results are to changes in the control parameters.

Finally Section 6 discusses how the model and value function might be developed in the future to come closer to reality.

2 A general model

In this paper we consider continuous-time stochastic models for pension fund dynamics which allow for n risky assets and for noise in the level of benefit outgo. The general form of this simple model is:

$$\begin{aligned} dX(t) &= X(t).d\delta_X(t,X(t)) + c(t).dt - B.dt - \sigma_b.dZ_b(t) \qquad (1)\\ \text{where } X(t) &= \text{fund size at } t\\ d\delta_X(t,X(t)) &= \text{instantaneous return on assets between } t \text{ and } t+dt\\ c(t) &= c(t,X(t))\\ &= \text{contribution rate}\\ B &= \text{expected rate of benefit outgo}\\ \text{and } \sigma_b &= \text{volatility in benefit outgo} \end{aligned}$$

Discrete-time models have been considered by Cairns (1995), Cairns & Parker (1997), Dufresne (1988, 1989, 1990) and Haberman & Sung (1994). Such models have yielded a number of useful analytical results with wider applications. Continuous-time models,

which are, in some ways, more idealised, yield further analytical results (for example, see Dufresne, 1990, Boulier *et al.*, 1995, and Cairns, 1996). Similar results can then be sought empirically in discrete-time models.

The contribution rate, $c(t)$, is a predictable process and provides us with one of the means of controlling the dynamics of the pension fund. Dufresne (1990) and Cairns (1996) considered continuous-time models in which the contribution rate was a linear function of the current fund size, $X(t)$. Boulier *et al.* (1995) considered more general forms for $c(t)$ but found that the optimal solution to a simple control problem was that the contribution rate should indeed be linear in $X(t)$. These results are discussed in detail in Sections 3 and 4 of this paper. O'Brien (1987) considered a similar objective function where the contribution rate only was controllable and where there was a stochastic reserve (in contrast to the constant target x_p relative to salary roll used in Section 3 of this paper). He found that the optimal contribution rate was linear in the amount of surplus. However, other aspects of the model used by O'Brien (1987) were unrealistic even for a simple pension scheme, making a fresh start here appropriate.

The other means of control is through the asset-allocation strategy. First we may allow for the possibility of a risk-free asset (or *cash*) which has a value at time t of $R_0(t) = R_0(0)\exp(\delta_0 t)$. There are, in addition, n risky assets, the prices of which (including reinvestment of dividend income) we assume follow correlated geometric brownian motion: that is,

$$\begin{aligned}
\frac{dR_i(t)}{R_i} &= d\delta_i(t) = \delta_i.dt + \sum_{j=1}^{n} \sigma_{ij}.dZ_j(t) && (2)\\
\text{or } d\delta(t) &= \delta.dt + S.dZ && (3)\\
\text{where } d\delta(t) &= (d\delta_1(t),\ldots,d\delta_n(t))^T\\
\delta &= (\delta_1,\ldots,\delta_n)^T\\
S &= (\sigma_{ij})_{i,j=1}^{n}\\
dZ &= (dZ_1,\ldots,dZ_n)^T
\end{aligned}$$

and $Z(t)$ is standard n-dimensional Brownian motion. We assume that $Z(t)$ and $Z_b(t)$ are independent.

For convenience later on, we define $D = SS^T$ (the instantaneous covariance matrix) and $\lambda = (\lambda_1,\ldots,\lambda_n)^T$ where $\lambda_i = \delta_i - \delta_0$ is the risk premium attached to asset i.

Let us assume that $\delta_i > \delta_0$ for all $i \geq 1$ (that is, investors are rewarded with higher expected returns for taking on some risk). No assumption is made about the level of correlation between the returns on the various stocks including, for example, the benefits (or otherwise) of diversification. The proportion of the assets invested in asset i ($i = 0,1,\ldots,n$) is denoted by $p_i(t,X(t))$. It follows that $\sum_{i=0}^{n} p_i(t,X(t)) = 1$. In the development below we write $p = p(t,X(t)) = (p_1(t,X(t)),\ldots,p_n(t,X(t)))^T$. The instantaneous rate of return on the fund is then:

$$(1-\sum_{i=1}^{n} p_i)\delta_0.dt + \sum_{i=1}^{n} p_i d\delta_i(t) = \delta_0.dt + p^T\lambda.dt + p^T SdZ \tag{4}$$

In this paper we will consider a range of constraints on the proportions invested in each asset. These include the possibility that we hold no cash (or a fixed percentage of the fund in cash) and that there shall be no short-selling of assets.

We allow for more than one risky asset for two reasons. First, it allows for a degree of realism without complicating substantially the analysis. Second, the experience of the UK pension funding scene is that pension funds only use cash for short-term liquidity rather than as a serious asset. Instead funds use government bonds (fixed interest and index linked) as low-risk (but non-zero-risk) assets. This situation is modelled in Section 3.3.

3 Optimal dynamic stochastic control

3.1 The general quadratic case

We consider first the case where there is no constraint on the amount invested in cash. Following Boulier *et al.* (1995) we define the value function for a general controlled pension fund process

$$W(t,x)(c,p) = E\left[\int_t^\infty \exp(-\beta s)L(s,c(s,X(s)),X(s))ds \mid X(t)=x\right] \tag{5}$$

Here $\exp(-\beta s)$ is a discount function and $L(s,c,x)$ is a loss function given that at time s, $X(s)=x$. This value function is also a function of the chosen, Markov contribution strategy $c(s,X(s))$ and investment strategy $p(s,X(s))$ which we abbreviate, where appropriate to c and p respectively.

Let $V(t,x) = \inf_{(c,p)} W(t,x)(c,p) = W(t,x)(c^*,p^*)$ assuming that such optimal control strategies c^* and p^* exist. Then $V(t,x)$ satisfies the Hamilton-Jacobi-Bellman equation (for example, see Merton, 1990, Øksendal, 1998, or Fleming & Rishel, 1975):

$$\begin{aligned} 0 &= \inf_{c,p}\left(e^{-\beta t}L(t,c,x) + V_t + \left[(\delta_0 + p^T\lambda)x + c - B\right]V_x + \frac{1}{2}V_{xx}\left(x^2 p^T Dp + \sigma_b^2\right)\right) \\ \text{where } V_t &\equiv \partial V/\partial t \\ V_x &\equiv \partial V/\partial x \\ V_{xx} &\equiv \partial^2 V/\partial x^2 \end{aligned} \tag{6}$$

We differentiate the expression in brackets with respect to c and p to find that:

$$\frac{\partial}{\partial c}\Big(\cdot\Big) = e^{-\beta t}L_c + V_x = 0, \qquad \text{where } L_c = \partial L/\partial c \tag{7}$$

$$\Rightarrow c^*(t,x) = L_c^{-1}(-e^{\beta t}V_x) \tag{8}$$

$$\text{and } \frac{\partial}{\partial p}\Big(\cdot\Big) = \lambda x V_x + Dpx^2 V_{xx} = 0 \tag{9}$$

$$\Rightarrow p^*(t,x) = -\left(\frac{V_x}{xV_{xx}}\right)D^{-1}\lambda \tag{10}$$

We see from the form of p^* that the amounts invested in each of the risky assets always stay in the same proportion. Thus we may define a special portfolio, A, which is a mixture of assets 1 to n in the same proportions (in market value terms) as $D^{-1}\lambda$. Then for any x we hold a proportion $\tilde{p}(x)$ (which depends upon $V(t,x)$) in portfolio A and $1-\tilde{p}(x)$ in cash. This result has obvious parallels in modern portfolio theory where the combination here of cash and portfolio A mimics movement along the capital market line. However, here we have not yet specified any form for the loss function $L(t,c,x)$ whereas modern portfolio theory (which works in discrete time) relies upon the use of a quadratic loss function. Further consideration of the model shows that portfolio A (which is efficient in the sense of minimising the value function) is also efficient in the sense of modern portfolio theory: that is, it has the lowest instantaneous volatility for a given rate of return.

Classical portfolio theory has been extended to include liabilities by Wise (1984), Wilkie (1985), Sharpe and Tint (1990) and Keel and Müller (1995). Working in discrete time and using a quadratic loss function Keel and Müller (1995) find that the composition of efficient portfolios can be altered by the inclusion of liabilities: in particular, where liabilities are random and not independent of the asset returns.

The precise form for $V(t,x)$ is, of course, still not yet known: we only have expressions for c^* and p^* involving $V(t,x)$.

It is necessary that the loss function is a strictly convex function of c. This ensures that the inverse of L_c exists. This requirement excludes, for example, downside loss functions which are convex but not strictly convex.

Here we restrict ourselves to the following quadratic loss function:

$$L(t,c,x) = (c-c_m)^2 + 2\rho(c-c_m)(x-x_p) + (k+\rho^2)(x-x_p)^2 \tag{11}$$

$$\text{where } k \geq 0.$$

$$\text{Thus } L_c^{-1}(-e^{\beta t}V_x) = c_m - \rho(x-x_p) - \frac{1}{2}e^{\beta t}V_x \tag{12}$$

(that is, if $c = c_m - \rho(x-x_p) - \exp(\beta t)V_x/2$ we have $L_c(t,c,x) = -\exp(\beta t)V_x$).

A special case of this loss function is the one suggested by Haberman and Sung (1994) (in a discrete-time framework).

We apply this to the Hamilton-Jacobi-Bellman equation to give:

$$
\begin{aligned}
0 &= e^{-\beta t}\left[\left(-\frac{1}{2}e^{\beta t}V_x\right)^2 + k(x-x_p)^2\right] + V_t + (\delta_0 x - B)V_x - \lambda^T D^{-1}\frac{V_x}{xV_{xx}}\lambda x V_x \\
&\quad + \left(c_m - \rho(x-x_p) - \frac{1}{2}e^{\beta t}V_x\right)V_x + \frac{1}{2}V_{xx}\left[x^2\left(\frac{V_x}{xV_{xx}}\right)^2\lambda^T D^{-1}DD^{-1}\lambda + \sigma_b^2\right]
\end{aligned} \tag{13}
$$

Given the form of the objective function (Markov and time-homogeneous) it is clear that the optimal strategies c^* and p^* depend only upon x and not upon t. Thus $V(t,x)$ will be of the form $e^{-\beta t}F(x)$ and therefore

$$
\begin{aligned}
0 &= \frac{1}{4}F_x^2 + k(x-x_p)^2 - \beta F + (\delta_0 x - B)F_x - \lambda^T D^{-1}\lambda\frac{F_x^2}{F_{xx}} \\
&\quad + (c_m - \rho(x-x_p) - \frac{1}{2}F_x)F_x + \frac{1}{2}F_{xx}\left[\frac{F_x^2}{F_{xx}^2}\lambda^T D^{-1}\lambda + \sigma_b^2\right]
\end{aligned} \tag{14}
$$

Try $F(x) = Px^2 + Qx + R$, and write $\varepsilon = \lambda^T D^{-1}\lambda$. Then

$$
\begin{aligned}
0 &= -\frac{1}{4}(2Px+Q)^2 + k(x-x_p)^2 - \beta(Px^2+Qx+R) + (\delta_0 x - B)(2Px+Q) \\
&\quad -\frac{1}{2}\varepsilon\frac{(2Px+Q)^2}{2P} + (c_m - \rho(x-x_p))(2Px+Q) + P\sigma_b^2 \qquad (15) \\
\Rightarrow 0 &= x^2[-P^2 + k - \beta P + 2P\delta_0 - P\varepsilon - 2\rho P] \\
&\quad + x[-PQ - 2kx_p - \beta Q - 2PB + Q\delta_0 - Q\varepsilon - \rho Q + 2P(c_m + \rho x_p)] \\
&\quad + [-\frac{1}{4}Q^2 + kx_p^2 - \beta R - BQ - \frac{Q^2\varepsilon}{4P} + (c_m + \rho x_p)Q + P\sigma_b^2] \qquad (16)
\end{aligned}
$$

Define $\hat{P} = 2\delta_0 - \beta - \varepsilon - 2\rho$. Then we find that

$$
\begin{aligned}
P(k) &= \frac{\hat{P} + \sqrt{\hat{P}^2 + 4k}}{2} && (17) \\
Q(k) &= \frac{2[P(k)(B - c_m - \rho x_p) + kx_p]}{-P(k) + \delta_0 - \beta - \varepsilon - \rho} && (18) \\
R(k) &= \frac{1}{\beta}\left[-\frac{1}{4}Q(k)^2 + kx_p^2 - BQ(k) - \frac{Q(k)^2\varepsilon}{4P} + (c_m + \rho x_p)Q(k) + P(k)\sigma_b^2\right] && (19)
\end{aligned}
$$

This is an admissible solution provided $\hat{P} > 0$.

We find then that

$$
\begin{aligned}
c^*(x) &= c_m - \rho(x - x_p) - \frac{1}{2}(2P(k)x + Q(k)) && (20)\\
\text{or } c^*(x) &= c_0^* - c_1^* x && (21)\\
p^*(x) &= -\frac{(2P(k)x + Q(k))}{2P(k)x} D^{-1}\lambda && (22)\\
\text{or } p^*(x) &= \frac{p_0^* + p_1^* x}{x} && (23)
\end{aligned}
$$

where p_0^* and p_1^* are both $n \times 1$ vectors which are proportional to $D^{-1}\lambda$.

Note that when $x = -Q(k)/2P(k)$, $p^*(x) = 0$: that is, we are invested entirely in the risk-free asset. Furthermore if a portfolio, A, is synthesised from the n risky assets in the proportions $D^{-1}\lambda$ as described earlier, then, given a funding level of x, we should hold a proportion of the fund

$$
\tilde{p}(x) = e^T p^*(x) = -\frac{2P(k)x + Q(k)}{2P(k)x} e^T D^{-1}\lambda \qquad (24)
$$

in portfolio A and $1 - \tilde{p}(x)$ in cash. ($e^T = (1, \ldots, 1)$ is the unit vector.)

We can also note that $F(x)$ is minimised at $x = -Q(k)/2P(k)$, which we will denote by x_{min} say. As discussed in Section 3.6 this presents, to a certain extent, a barrier through which it is difficult for the funding level, $X(t)$, to pass. Depending upon the relationship between c_m, k and x_p this could take the form of a ceiling or a floor.

It is important to note that $P(k)$ and $Q(k)$ do not depend upon σ_b. It follows, therefore, that the optimal control strategy (both contributions and investments) do not depend upon σ_b. Thus, demographic variability is a factor which affects the value function $V(t,x)$ only and we should treat small funds in the same way as large funds.

It is also important to note that the precise proportions of each asset held in portfolio A do not depend upon the form of the loss function, nor does it depend upon σ_b.

Remark

The non-linear ordinary differential equation (14) is subject to the boundary condition $0 \leq F(x)$ for all x. We have two degrees of freedom in how we solve this equation. Numerical work suggests that there are also solutions to (14) which either have singularities (which we regard as an inadmissible solution) or which are asymptotically linear as $x \to \pm\infty$. Now if $F(x) \sim a + bx$ as $x \to +\infty$, $c^*(x) \sim c_m - b$ as $x \to +\infty$. With such a solution we may find that $X(t)$ will drift off to infinity. This drift, however, is countered by the asset-allocation strategy which is quite extreme:

- As $X(t)$ gets very large the fund goes very long in cash and very short in risky assets. This ensures that there is a very inefficient strategy which more-or-less throws away money in order to get back to the target funding level x_p.
- As $X(t)$ gets very small the fund goes very long in risky assets and very short in cash to get a high expected return to help us get back to a better funded position as quickly as possible.

In the quadratic-$F(x)$ case these problems with the asset-allocation strategy also apply but they are much less extreme. Furthermore, the optimal contribution rate is a linear function of $X(t)$. It is a necessary condition for stationarity that the contribution rate is at least linear. (Note, however, that linearity is sufficient only when the slope c_1^* is greater than a certain minimum level described later in this paper.)

Thus we can reasonably put in the further boundary condition that $F(x)/x^2 \to$ constant as $x \to \pm\infty$.

3.2 Constraints on cash

We have up until now assumed that the amount of money invested in cash could vary without bound. Here we go to the other extreme and assume that we invest a proportion p_m of the fund in risky assets and $1-p_m$ in cash, where p_m is fixed. It is reasonable that $p_m < 1$ allowing for a small but fixed amount in cash to provide short-term liquidity for the fund to cover immediate benefit payments. (A typical figure for a UK pension fund in the UK is 5% cash and $p_m = 95\%$ risky assets.) Subject to this constraint, there is total freedom in the proportions invested in the n risky assets.

Recall the Hamilton-Jacobi-Bellman equation:

$$0 = \inf_{\substack{c,p \\ e^T p = p_m}} \left(e^{-\beta t} L(t,c,x) + V_t + [(\delta_0 + p^T \lambda)x + c - B]V_x + \frac{1}{2} V_{xx}(x^2 p^T D p + \sigma_b^2) \right) \tag{25}$$

where $e = (1,\ldots,1)^T$

We differentiate the expression in brackets with respect to c as before to get:

$$\frac{\partial}{\partial c}\Big(\cdot\Big) = e^{-\beta t} L_c + V_x = 0, \qquad \text{where } L_c = \partial L/\partial c \tag{26}$$

$$\Rightarrow c^*(t,x) = L_c^{-1}(-e^{\beta t} V_x) \tag{27}$$

To minimise over p subject to the constraint we use the method of Lagrangians. Thus we minimise the function

$$G(p,\gamma) = x V_x \lambda^T p + \frac{1}{2} x^2 V_{xx} p^T D p + \gamma(e^T p - p_m) \tag{28}$$

over p and γ.

$$\frac{\partial G}{\partial p} = x V_x \lambda + x^2 V_{xx} D p + \gamma e = 0 \tag{29}$$

$$\frac{\partial G}{\partial \gamma} = e^T p - p_m = 0 \tag{30}$$

for which the solution is

$$
\begin{aligned}
p = p(x) &= \left(p_m + \frac{V_x}{xV_{xx}} e^T D^{-1}\lambda\right) \frac{1}{e^T D^{-1} e} D^{-1} e - \frac{V_x}{xV_{xx}} D^{-1}\lambda \qquad (31)\\
&= \left(d_0 + d_1 \frac{V_x}{xV_{xx}}\right) D^{-1} e - \frac{V_x}{xV_{xx}} D^{-1}\lambda \qquad (32)\\
\text{where } d_0 &= \frac{p_m}{e^T D^{-1} e}\\
\text{and } d_1 &= \frac{e^T D^{-1}\lambda}{e^T D^{-1} e}
\end{aligned}
$$

We note, as in the previous section, the connection with modern portfolio theory. We have already discussed the relevance of $D^{-1}\lambda$. Here we note that portfolios which invest in the same proportion as $D^{-1}e$ have the minimum variance given that there are to be no investments in cash. Furthermore, all efficient portfolios are linear combinations of $D^{-1}\lambda$ and $D^{-1}e$.

Again because of the form of the value function we substitute $V(t,x) = e^{-\beta t}F(x)$.

$$
\begin{aligned}
\text{Now } L(t,c,x) &= (c-c_m)^2 + 2\rho(c-c_m)(x-x_p) + (k+\rho^2)(x-x_p)^2 \qquad (33)\\
\Rightarrow L_c^{-1}(-e^{\beta t}V_x) &= c_m - \rho(x-x_p) - \frac{1}{2}F_x \qquad (34)
\end{aligned}
$$

We apply this to the Hamilton-Jacobi-Bellman equation to give:

$$
\begin{aligned}
0 &= \frac{1}{4}F_x^2 + k(x-x_p)^2 - \beta F\\
&\quad + \left[\left(\delta_0 + \lambda^T D^{-1}\left(d_0 e + (d_1 e - \lambda)\frac{F_x}{xF_{xx}}\right)\right)x + c_m - \rho(x-x_p) - \frac{1}{2}F_x - B\right]F_x\\
&\quad + \frac{1}{2}F_{xx}\left[x^2\left(d_0 e + (d_1 e - \lambda)\frac{F_x}{xF_{xx}}\right)^T D^{-1}\left(d_0 e + (d_1 e - \lambda)\frac{F_x}{xF_{xx}}\right) + \sigma_b^2\right] \qquad (35)
\end{aligned}
$$

As in the unconstrained case this has a quadratic solution $F(x) = P(k)x^2 + Q(k)x + R(k)$. The form of $p(x)$ indicates that we require two portfolios A and B. Portfolio A is made up of fixed proportions of assets 1 to n in proportion to the vector $D^{-1}\lambda$, while portfolio B is synthesised similarly but in proportion to the vector $D^{-1}e$.

As in Section 3.1 portfolios A and B are independent of the form of the loss function.

3.3 Further discussion of the general model

We now consider the optimal asset-allocation strategy in more detail. In particular, consider the instantaneous rate of return on the investments: that is, $\delta_0 + \lambda^T p^*(x)$. Consider the unconstrained case first.

$$\delta_0 + \lambda^T p^*(x) = \delta_0 - \frac{2P(k)x + Q(k)}{2P(k)x}\lambda^T D^{-1}\lambda \quad (36)$$

$$= \delta_0 - \lambda^T D^{-1}\lambda - \frac{Q(k)}{2P(k)x}\lambda^T D^{-1}\lambda \quad (37)$$

Now D is positive definite so that $\lambda^T D^{-1}\lambda$ is positive. Furthermore, $P(k)$ is positive and $Q(k)$ is normally negative. Hence $\delta_0 + \lambda^T p^*(x)$ is normally a decreasing function of x.

Similarly consider the constrained case.

$$\lambda^T p^*(x) = \lambda^T \left[d_0 D^{-1} e + \frac{2P(k)x + Q(k)}{2P(k)x}(d_1 D^{-1} e - \lambda) \right] \quad (38)$$

$$= d_0 \lambda^T D^{-1} e + d_1 \lambda^T D^{-1} e - \lambda^T D^{-1}\lambda + \frac{Q(k)}{2P(k)x}\lambda^T D^{-1}(d_1 e - \lambda) \quad (39)$$

Note that $e^T D^{-1}(d_1 e - \lambda) = 0$. Hence $\lambda^T D^{-1}(d_1 e - \lambda) = -(d_1 e - \lambda)^T D^{-1}(d_1 e - \lambda) < 0$ since D is positive definite. Again $P(k) > 0$ and normally $Q(k) < 0$ so that $\lambda^T p^*(x)$ is a decreasing function of x.

Furthermore an analysis of the instantaneous variance of the investments confirms that as the instantaneous rate of return decreases, the instantaneous variance decreases also and then starts to increase as we go long (effectively) in low-risk assets and short in high-risk assets.

Thus we find that when the funding level is low we invest more in high-risk assets and as the funding level rises we shift from high-risk into low-risk assets. This is a rather counterintuitive investment strategy. We would expect that as the funding level falls that we might shift into lower-risk assets to protect our position. The strategy we have found here does the opposite. The reason for this is because of the quadratic form of the objective function. This, in a sense, defines an ideal funding level x_p and an ideal contribution rate c_m. If the funding level is below this then we invest in high-return, high-risk assets to increase the chance of getting quickly back to the ideal level. Conversely if the funding level is too high then we are prepared to invest in what is effectively an inefficient, high-risk, low-return investment strategy in order to get back to the ideal level. Indeed the fund will go long in cash and short in equities. In effect the scheme would be throwing money away since, for the same level of risk (that is, volatility of asset returns) it could have a higher expected return. The inefficiency here turns out, with hindsight, to be a result of the quadratic loss function. This actually prefers the positive target contribution rate, c_m, to refunds. In other words, it is better to throw money away than to take a refund. (There is nothing new in this observation. Related problems in other branches of financial economics come to the same counter-intuitive conclusions where, for example, quadratic utility functions are employed.)

Now consider the optimal contribution rate. Sometimes this is written in the form $(c_0 - c_1 x_p) - c_1(x - x_p)$ where $x - x_p$ is the surplus relative to the target fund size x_p. c_1 is the rate at which we try to remove surplus or *amortize* this surplus. It can be noted that the

optimal amortization rate, $c_1^* = \hat{P}(k)$, depends on k, δ_0, λ and D but not on c_m, x_p or σ_b^2. On the other hand, c_0^* also depends on x_p and c_m but again not on σ_b^2.

Similarly it can be seen that $p^*(x)$ does not depend upon σ_b^2. Thus, it has been demonstrated that for such a quadratic loss function $L(\cdot)$ the optimal contribution and asset-allocation strategies do not depend in any way upon the randomness in the level of benefit outgo (at least where this uncertainty is uncorrelated with investment returns).

Later in this paper we will return to the dynamic optimisation problem where we have a different objective function and where there are constraints on the investment strategy and of the funding level.

3.4 Optimal strategy when p is fixed

Suppose instead that the asset-allocation strategy is static: that is, $p(t,x) = p$ for all t, x, for some p. We can still apply the Bellman equation but minimise over $c(t,x)$ only. Thus we find that

$$\begin{aligned} 0 &= \left[\frac{1}{4}F_x^2 + k(x-x_p)^2\right] - \beta F \\ &\quad + \left[\delta_0 x - B - p^T\lambda x + \left(c_m - \rho(x-x_p) - \frac{1}{2}F_x\right)\right]F_x + \frac{1}{2}F_x x(p^T D p x^2 + \sigma_b^2) \end{aligned} \quad (40)$$

Again we try to find a solution of the form $F(x) = Px^2 + Qx + R$ and we find that

$$P = P(k) = \frac{\hat{P} + \sqrt{\hat{P}^2 + 4k}}{2} \quad (41)$$

$$\text{where } \hat{P} = 2\delta_0 - \beta - 2p^T\lambda - 2\rho + p^T D p \quad (42)$$

$$Q = Q(k) = -\frac{2\left[kx_p + P(k)\left(B - c_m - \rho x_p\right)\right]}{P(k) + \beta - \delta_0 + p^T\lambda + \rho} \quad (43)$$

$$R = R(k) = \frac{1}{\beta}\left[-\frac{1}{4}Q(k)^2 + kx_p^2 - BQ(k) + (c_m + \rho x_p)Q(k) + P(k)\sigma_b^2\right] \quad (44)$$

The question now arises: how do we choose the optimal static p?

We will consider one option here: minimise $P(k)$ over p. This means that the optimal curve $F(x)$ will be as close as possibile in the limit as x tends to $\pm\infty$ to the superior solution derived in Section 3.1. Clearly the solution derived in Section 3.1 will be lower for all x regardless of the value of p. (Other possibilities include minimising $F(x)$ over p for a specific value of x, or minimising the minimum of $F(x)$ over p.)

To minimise $P(k)$ over p we differentiate:

$$\frac{dP}{dp} = \frac{dP}{d\hat{P}}\frac{d\hat{P}}{dp} \quad (45)$$

$$= \left(\frac{1}{2}+\frac{1}{2}\hat{P}\left(\hat{P}^2+4k\right)^{-1/2}\right)(-2\lambda+2Dp) \quad (46)$$

$$\Rightarrow \hat{p} = D^{-1}\lambda \quad (47)$$

A consideration of the form of $\hat{P}$ as a function of p shows that this is a minimum at $\hat{p}$.

We note that the proportions in the risky assets as given in $\hat{p}$ are the same as those derived in Section 3.1. Furthermore, we find that, given $p=\hat{p}$, $\hat{P}=2\delta_0-\beta-2\rho-\lambda^T D^{-1}\lambda$ (again the same as in Section 3.1). This means that:

- for large or small values of x the loss of optimality as a result of fixing p does not become too great;
- if we write $c^*(t,x)=c_0^*-c_1^*x$ then c_1^* is not affected by the restriction on the investment strategy (that is, the rate of amortisation of surplus or deficit is not affected).

3.5 Comparison of the strategies

Let us consider a specific example to compare the effectiveness of the optimal strategies derived in Sections 3.3 and 3.5 compared to that in Section 3.1. The fixed parameters are as follows:

$$\delta_0=0.03,\ \delta=\begin{pmatrix}0.04\\0.06\end{pmatrix},\ S=\begin{pmatrix}0.05 & 0.05\\0.05 & 0.2\end{pmatrix},\ B=1,\ \sigma_b=0.1 \quad (48)$$

The control parameters are:

$$c_m=0.6,\ k=0.001,\ x_p=10,\ \beta=0.03,\ \rho=0 \quad (49)$$

In this and in subsequent sections we define the funding level, $X(t)$, as the value of the assets divided by the expected rate of benefit outgo. Alternatively, if expected benefit outgo is defined as it is here as $B=1$ then $X(t)$ is also the fund size.

The optimal value functions $F(x)$ are plotted in Figure 1 and their stationary distributions (as derived later on in Section 4) are plotted in Figure 2.

Selected statistics are given in Table 1. From Table 1 and Figure 1 we can see that the unconstrained solution is significantly better that the other two. The unconstrained and static cases are quite similar in some ways (shape and contribution strategy) but the lack of flexibility in the investment strategy adds on a fixed and substantial penalty. The constrained (no cash) case looks much more different. By reference to Figure 1 it performs well in the middle of the range and, indeed, attempts to stay there by applying a more aggressive amortization strategy. For more extreme values of x this strategy is much poorer than the static case. However, by looking also at the stationary densities of the funding level under the three strategies (Figure 2) we can see that such extreme values will occur very rarely indeed.

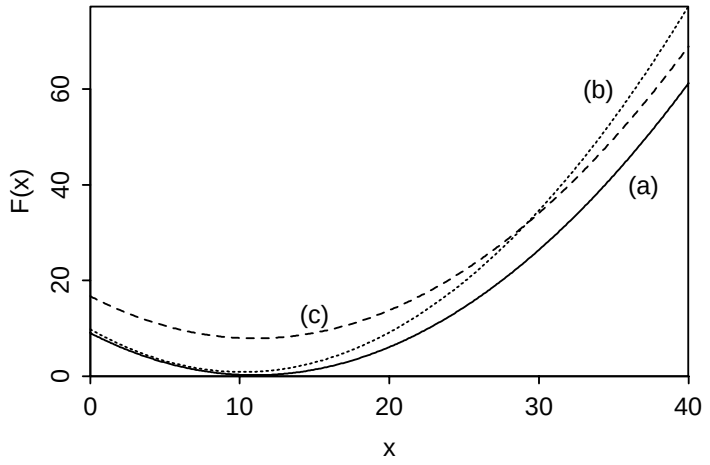

Figure 1: Comparison of value functions for three investment/contribution strategies. (a) (solid line) unconstrained optimum. (b) (dotted line) optimum under the constraint of no cash ($p_m = 1$). (c) (dashed line) optimum under a static investment strategy.

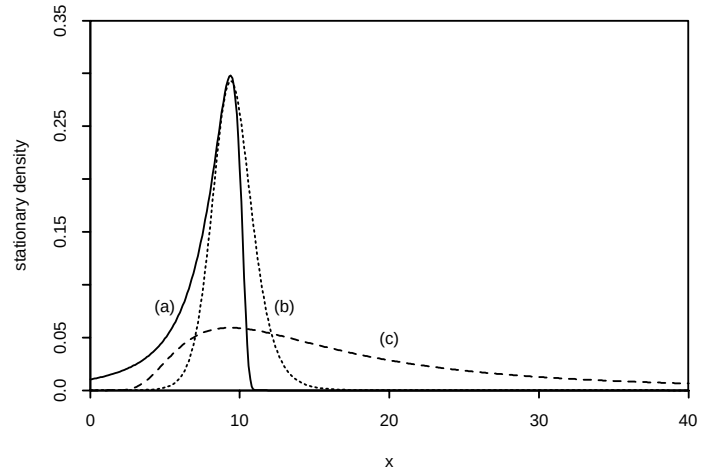

Figure 2: Comparison of stationary densities for three investment/contribution strategies. (a) (solid line) unconstrained optimum. (b) (dotted line) optimum under the constraint of no cash ($p_m = 1$). (c) (dashed line) optimum under a static investment strategy. (The funding level is defined here as fund size divided by the expected rate of benefit outgo.)

Table 1: Comparison of optimal strategies with and without constraints.

	P	Q	R	Minimum $F(x)$	c_0^*	c_1^*
Unconstrained	0.073	-1.60	9.00	0.19	1.40	0.073
Constrained ($p_m = 1$)	0.086	-1.75	9.79	0.85	1.48	0.086
Static	0.073	-1.60	16.70	7.89	1.40	0.073

As mentioned in Section 3.1 we can look at Figures 1 and 2 and see that under the unconstrained investment strategy the value function is minimised at $x_{min} = -Q/2P = 11.01$ and that this, in effect, turns out to be a ceiling (although $X(t)$ can have brief excursions above this value because of volatility in the benefit outgo). Under such circumstances (that is the existence of, effectively, a ceiling) some of the criticisms of the approach with the quadratic loss function become somewhat irrelevant since we are practically never at a funding level where we choose effectively to throw away money (in the sense described in Section 3.4).

Under other circumstances (for example, here if we took $k = 0.005$ and $x_p = 10$ as before but changed c_m from 0.6 to 0.8) the ceiling would turn into a floor at 8.99 and the funding level would spend most of the time above this floor. While this appears to be an appealing strategy the reservations about the investment strategy discussed in Section 3.4 are well founded here.

Under the constrained strategy the value of x_{min} is 10.19 but we can see that the funding level can frequently go above this level. At x_{min} note that the fund here is invested in proportion to the minimum variance portfolio $D^{-1}e$.

Finally we can see from Figure 2 that the static investment strategy leads to much wider fluctuations in the funding level which could only be reduced by increasing the value of k in the loss function.

We will return to this example in Section 5.

3.6 Power and exponential loss functions

3.6.1 Power loss function

Let us complete this section now with a short analysis of the special case where $\sigma_b = 0$ and

$$L(t,c,x) = \begin{cases} -\frac{1}{\gamma}(c_m - c)^\gamma & \text{for } c \le c_m \\ +\infty & \text{for } c > c_m \end{cases} \qquad \text{for } 0 < \gamma < 1. \tag{50}$$

Again we assume that the optimal value function takes the form $V(t,x) = \exp(-\beta t)F(x)$. Then the form of the Hamilton-Jacobi-Bellman equation takes the form:

$$\inf_{c,p}\left\{ -\frac{1}{\gamma}(c_m - c)^\gamma - \beta F + \left[\left(\delta_0 + p^T\lambda\right)x + c - B\right]F_x + \frac{1}{2}F_{xx}x^2 p^T D p \right\} \;=\; 0 \tag{51}$$

(We restrict optimisation to strategies which keep the fund size positive. Without this condition it is clearly optimal to take contribution refunds of infinite size.)

Now

$$\frac{\partial}{\partial c}\left(\ \cdot\ \right) \;=\; 0 \tag{52}$$

$$\Rightarrow (c_m - c)^{\gamma-1} + F_x = 0 \quad (53)$$

$$\Rightarrow c^*(x) = c_m - (-F_x)^{1/(\gamma-1)} \quad (54)$$

$$\frac{\partial}{\partial p}\left(\cdot\right) = 0 \quad (55)$$

$$\Rightarrow p^*(x) = -D^{-1}\lambda\frac{F_x}{xF_{xx}} \quad (56)$$

Inserting $c^*(x)$ and $p^*(x)$ into equation (51) we get

$$(-F_x)^{\gamma/(\gamma-1)}\left(\frac{\gamma-1}{\gamma}\right) - \beta F + (\delta_0 x + c_m - B)F_x - \frac{1}{2}\frac{F_x^2}{F_{xx}}\lambda^T D^{-1}\lambda = 0 \quad (57)$$

We try for a solution of the form $F(x) = -k(x - x_m)^\alpha$.
Inserting this into equation (57) we get, for all x:

$$\begin{aligned}(k\alpha)^{\gamma/(\gamma-1)}(x-x_m)^{\gamma(\alpha-1)/(\gamma-1)}&\left(\frac{\gamma-1}{\gamma}\right)\\ +\beta k(x-x_m)^\alpha - \delta_0&\left(x - \frac{B-c_m}{\delta_0}\right)k\alpha(x-x_m)^{\alpha-1}\\ &+\frac{1}{2}\frac{k^2\alpha^2(x-x_m)^{2\alpha-2}}{k\alpha(\alpha-1)(x-x_m)^{\alpha-2}}\lambda^T D^{-1}\lambda = 0\end{aligned} \quad (58)$$

$$\Rightarrow x_m = \frac{B-c_m}{\delta_0} \quad (59)$$

$$\alpha = \gamma \quad (60)$$

$$\text{and } k = \frac{1}{\gamma}c_1^{\gamma-1} \quad (61)$$

$$\text{where } c_1 = \left(\frac{\beta - \delta_0\gamma + \frac{1}{2}\frac{\gamma}{\gamma-1}\lambda^T D^{-1}\lambda}{1-\gamma}\right) \quad (62)$$

Hence

$$c^*(x) = c_m - c_1(x - x_m) \quad (63)$$

$$p^*(x) = D^{-1}\lambda\frac{(x-x_m)}{(1-\gamma)x} \quad (64)$$

We note the similarity of the problems and solutions here with a well-known optimal-consumption problem described by Merton (1971, 1990). Equivalence is achieved by equating the controllable level of consumption with $B - c^*(t)$ in the current model. This

enables us to speculate that recent extensions of this work to include the effects of transactions costs can be applied to the present problem. For example, the problem of proportional transactions costs has been considered by, amongst others, Davis & Norman (1990) and Shreve & Soner (1994).

As with the quadratic loss function, contributions decrease linearly with x with the amortisation rate c_1 being determined by the discount rate β and the risk-aversion parameter γ (but not the maximum acceptable contribution rate, c_m).

Investment in risky assets, $p^*(x).x$ increases linearly in x above the minimum x_m and therefore appears to conform better with conventional wisdom. However, it turns out that this solution gives rise to one of two trivial stationary solutions for $X(t)$: that is, $X(t) \to x_m$ or $+\infty$ depending upon the value of β.

Returning to the dynamics of the funding level $X(t)$ we find that $c^*(x)$ and $p^*(x)$ give rise to

$$dX(t) = (X(t)-x_m)\left[\left(\delta_0 + \frac{1}{1-\gamma}\lambda^T D^{-1}\lambda - c_1\right)dt + \frac{1}{1-\gamma}\lambda^T S^{-1} dZ(t)\right] \quad (65)$$

$$\stackrel{D}{=} (X(t)-x_m)\left[\left(\delta_0 + \frac{1}{1-\gamma}\lambda^T D^{-1}\lambda - c_1\right)dt + \frac{1}{1-\gamma}\sqrt{\lambda^T D^{-1}\lambda}\, d\tilde{Z}(t)\right] \quad (66)$$

where $\tilde{Z}(t)$ is another Brownian motion. It follows that (and inserting the known form of c_1):

$$X(t)-x_m = (X(0)-x_m)\exp\left[\left(\delta_0 + \frac{1}{2(1-\gamma)}\lambda^T D^{-1}\lambda - \beta\right)t + \frac{1}{1-\gamma}\sqrt{\lambda^T D^{-1}\lambda}\,\tilde{Z}(t)\right] \quad (67)$$

That is, $X(t)-x_m$ is a geometric Brownian motion which tends to zero if $\beta > \delta_0 + \lambda^T D^{-1}\lambda/2(1-\gamma)$ and to $+\infty$ if $\beta < \delta_0 + \lambda^T D^{-1}\lambda/2(1-\gamma)$.

There are some similarities between this solution and that of Boulier *et al.* (1995) under which $x_m - X(t)$ is also a geometric Brownian motion.

With either the introduction of volatility in benefit outgo ($\sigma_b > 0$) or with restrictions on the amount of cash we cannot have both a lower bound on the funding level and an upper bound on the contribution rate.

The loss function $L(c) = c^\gamma/\gamma$ for $c > 0$ and $\gamma > 1$ has been considered by Siegmann & Lucas (1999). They obtain similar results to those described above, except that x_m becomes a maximum, and contributions are bounded below by 0 rather than above by c_m.)

3.6.2 Exponential loss function

Similarly, we can consider the exponential loss function (for example, see Siegmann & Lucas, 1999):

$$L(t,c,x) = \exp(\gamma c - \theta x) \tag{68}$$

where $\gamma > 0$ and $\theta > 0$. Here the relationship between γ and θ determines the relative emphasis on the employer and the members.

This gives us the solution:

$$\begin{aligned}
F(x) &= \exp(a - bx) && (69)\\
\text{where } b &= \gamma\delta_0 + \theta \\
\text{and } a &= \log\frac{\gamma}{\gamma\delta_0+\theta} + \frac{-\beta + (\gamma\delta_0+\theta)B + (\gamma\delta_0+\theta)/\gamma - \frac{1}{2}\lambda^T D^{-1}\lambda}{(\gamma\delta_0+\theta)/\gamma} \\
\Rightarrow c^*(t) &= c_0 - c_1 x && (70)\\
p^*(t) &= \frac{1}{x}p_0 && (71)\\
\text{where } c_0 &= \frac{-\beta + (\gamma\delta_0+\theta)B + (\gamma\delta_0+\theta)/\gamma - \frac{1}{2}\lambda^T D^{-1}\lambda}{\gamma\delta_0+\theta} \\
c_1 &= \delta_0 \\
\text{and } p_0 &= \frac{D^{-1}\lambda}{\gamma\delta_0+\theta}
\end{aligned}$$

This solution is more like the quadratic loss function considered in earlier sections: that is, the proportion of the fund invested in risky assets decreases as x increases. If we increase θ then p_0 decreases. This reflects the fact that there is a greater degree of risk aversion when we consider the interests of the members, so we invest less in risky assets.

With a little algebra we can see that $X(t)$ follows a Brownian motion with drift $\mu = (-\beta + (\gamma\delta_0+\theta)/\gamma + \frac{1}{2}\lambda^T D^{-1}\lambda)/(\gamma\delta_0+\theta)$ and volatility $\sqrt{\alpha} = \sqrt{\lambda^T D^{-1}\lambda}/(\gamma\delta_0+\theta)$. This means that the solution is unsatisfactory because it is both non-stationary and because it gives rise to a 'counterintuitive' investment strategy.

4 The stationary distribution of $X(t)$

4.1 General model

Assume now that

$$\begin{aligned}
c(t,x) = c(x) &= c_0 - c_1 x && (72)\\
p(t,x) = p(x) &= \frac{p_0 + p_1 x}{x} && (73)
\end{aligned}$$

where p_0 and p_1 are $n \times 1$ vectors.

The reason for assuming a linear form for $c(x)$ and $xp(x)$ is simple. They are consistent with the optimal dynamic controls derived in Section 3 when we use a quadratic loss function. Furthermore let us recall the value function

$$W(t,x)(c,p) = E\left[\int_t^\infty \exp(-\beta s)L(s,c(s,X(s)),X(s))ds \mid X(t)=x\right] \tag{74}$$

As $\beta \to 0$, $\beta W(t,x) \to E[L(s,c,X)]$: that is, the limiting optimal dynamic controls are also optimal in the static case if we use the same quadratic loss function.

The dynamics of the fund size, $X(t)$, are then

$$\begin{aligned} dX &= X\left[\left(\delta_0 + p(x)^T\lambda\right)dt + p(x)^T S d\tilde{Z}\right] + (c_0 - c_1 X - B)dt + \sigma_b dZ_b & (75)\\ &= \left[\left(\delta_0 X + (p_0 + p_1 X)^T\lambda\right)dt + (p_0 + p_1 X)^T S d\tilde{Z}\right] + (c_0 - c_1 X - B)dt + \sigma_b dZ_b & (76)\\ &\stackrel{D}{=} \mu.dt - \nu X.dt + (\alpha + \beta X + \gamma X^2)^{1/2} dZ & (77)\end{aligned}$$

where $\tilde{Z}(t)$ is a standard n-dimensional Brownian Motion, and $Z(t)$ is a standard Brownian motion which depends upon $\tilde{Z}(t)$ and $Z_b(t)$,

$$\begin{aligned} \mu &= c_0 - B + p_0^T\lambda & (78)\\ \nu &= c_1 - \delta_0 - p_1^T\lambda & (79)\\ \alpha &= p_0^T D p_0 + \sigma_b^2 & (80)\\ \beta &= 2p_0^T D p_1 & (81)\\ \gamma &= p_1^T D p_1 \ . & (82)\end{aligned}$$

In order to discuss the properties of this model we state the following theorem:

Theorem 4.1.1

Let the continuous-time stochastic process X_t satisfy the stochastic differential equation

$$dX_t = (\alpha + \beta X_t + \gamma X_t^2)^{1/2} dZ + \mu dt - \nu X_t dt \tag{83}$$

subject to the constraints on the parameters $\alpha > 0$, $\gamma > 0$, $\beta^2 - 4\alpha\gamma \le 0$, $\mu > 0$ and $\nu > 0$.
(a) If $\beta^2 - 4\alpha\gamma < 0$, the stationary density function of X_t is

$$\begin{aligned} f_X(x) &= K\exp\left[2a\tan^{-1}\frac{x+b}{c}\right](\alpha + \beta x + \gamma x^2)^{-1-\nu/\gamma} & (84)\\ &\qquad\qquad \textit{for } -\infty < x < \infty \\ \textit{where } a &= \frac{1}{\sqrt{4\alpha\gamma - \beta^2}}\left(\frac{\nu\beta}{\gamma} + 2\mu\right)\end{aligned}$$

$$
\begin{aligned}
b &= \frac{\beta}{2\gamma} \\
c &= \frac{\sqrt{4\alpha\gamma-\beta^2}}{2\gamma}.
\end{aligned}
$$

(b) If $\beta^2 - 4\alpha\gamma = 0$ *and* $X_0 > -b$*, the stationary density function of* X_t *is*

$$
f_X(x) = \begin{cases} K(x+b)^{-\theta}\exp[-\phi/(x+b)] & \begin{cases} \text{for } x > -b \text{ if } \theta > 0 \\ \text{for } x < -b \text{ if } \theta < 0 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{85}
$$

$$
\begin{aligned}
\text{where } b &= \frac{\beta}{2\gamma} \\
\theta &= 2\left(1+\frac{\nu}{\gamma}\right) \\
\phi &= \frac{\nu\beta+2\mu\gamma}{\gamma^2}
\end{aligned}
$$

that is, the Translated-Inverse-Gamma distribution with parameters $-b$*,* $\theta - 1 > 0$ *and* $\phi > 0$ *(*$TIG(-b, \theta-1, \phi)$*). (If* $X \sim TIG(k, \alpha, \beta)$ *then* $(X-k)^{-1} \sim$ *Gamma*(α, β)*.)*
In each case K is a normalizing constant.

Proof: Proofs of these two results have been provided before by a number of authors. Distribution (b) was first derived in the context of pension funding by Dufresne (1990) in the case where there is one risky asset, no cash and no demographic volatility ($\sigma_b = 0$). Dufresne also noted that the stationary distribution of the funding level was the same as the distribution of a perpetuity. An alternative proof for the distribution of the present value of a perpetuity was also shown to have distribution (b) by Yor (1992) and by De Schepper *et al.* (1994). Föllmer and Schweizer (1993, Theorem 5.1 and erratum) considered the diffusion process defined above as underlying a model for stock prices. They derived both of the limiting distributions given in (a) and (b).

The two distributions above are also known as Pearson type IV and type V distributions respectively (for example, see Johnson *et al.*, 1994).

Let us consider the Pearson Type IV distribution. This distribution has four degrees of freedom. The fifth degree of freedom used in the dynamics of the fund size determines the speed of the process.

Following the notation of Johnson *et al.* (1994) we define $\mu'_r = E[X^r]$. We define $\mu'_{-1} = 0$ and have $\mu'_0 = 1$. It is easy to show that the μ'_r satisfy the following recursive relationship:

$$
-k_0 r\mu'_{r-1} + (k_3 - (r+1)k_1)\mu'_r + (1-(r+2)k_2)\mu'_{r+1} = 0 \tag{86}
$$

$$
\text{where } k_0 = \frac{\alpha}{2(\gamma+\nu)}, \quad k_1 = \frac{\beta}{2(\gamma+\nu)}, \quad k_2 = \frac{\gamma}{2(\gamma+\nu)}, \quad k_3 = \frac{\beta-2\mu}{2(\gamma+\nu)} \tag{87}
$$

Hence we have

$$E[X] = \mu_1' = \frac{k_1 - k_3}{1 - 2k_2} = \frac{\mu}{\nu} \tag{88}$$

$$E[X^2] = \mu_2' = \frac{k_0 + (2k_1 - k_3)\mu/\nu}{1 - 3k_2} = \frac{\alpha\nu + \beta\mu + 2\mu^2}{\nu(2\nu - \gamma)} \tag{89}$$

$$\Rightarrow Var[X] = \frac{\alpha\nu + \beta\mu + 2\mu^2}{\nu(2\nu - \gamma)} - \frac{\mu^2}{\nu^2} = \frac{\nu^2\left(\alpha + \beta\frac{\mu}{\nu} + \gamma\frac{\mu^2}{\nu^2}\right)}{\nu^2(2\nu - \gamma)} = \frac{\alpha + \beta E[X] + \gamma E[X]^2}{2\nu - \gamma} \tag{90}$$

$$E[c(X)] = c_0 - c_1 E[X] = c_0 - c_1\frac{\mu}{\nu} \tag{91}$$

$$Var[c(X)] = c_1^2 Var[X] = c_1^2\frac{\alpha + \beta E[X] + \gamma E[X]^2}{2\nu - \gamma} \tag{92}$$

The stationary distribution exists if and only if $2(1 + \nu/\gamma) > 1$: that is $c_1 > \delta_0 + p_1^T\lambda - \frac{1}{2}\gamma$ (in fact, $\gamma^{1/2}$ is the asymptotic volatility as $|x| \to \infty$). Similarly $E[X]$ exists if and only if $c_1 > \delta_0 + p_1^T\lambda$ and $Var[X]$ exists if and only if $c_1 > \delta_0 + p_1^T\lambda + \frac{1}{2}\gamma$.
Coming back to the optimal solution for the dynamic problem we found that $c_1^* = P(k) = P(k, \beta)$. Thus the condition for stationarity is

$$P(k) > \delta_0 + p_1^{*T}\lambda - \frac{1}{2}p_1^{*T}Dp_1^* \tag{93}$$

$$\text{where } p_1^* = -D^{-1}\lambda \quad \text{(unconstrained case).}$$

$$\Rightarrow P(k) > \delta_0 + \frac{3}{2}\lambda^T D^{-1}\lambda \tag{94}$$

Now $P(k, \beta)$ is a decreasing function of β and an increasing function of k, so the condition above is less likely to be satisfied if β is large or k is small. Under such circumstances the funding level will diverge as t tends to infinity. The situation, therefore, is that with a relatively large value of β we pay more attention to control of short-term variability in the contribution rate at the expense of larger fluctuations in the long term. Likewise, if the value of k is too small then we also pay too much attention to short-term contribution-rate stability.

4.2 Continuous proportion portfolio insurance

The idea of continuous proportion portfolio insurance (CPPI) was introduced by Black and Jones (1988) and Black and Perold (1992).

The previous sections in this paper have concentrated upon quadratic loss functions. The motivation behind CPPI is that in certain countries there exist minimum funding constraints: that is, there exists a floor below which the funding level must not fall. CPPI was proposed as a means of reducing the risk that the fund falls below this floor.

Under CPPI if the funding level is low then the fund will be invested more in low-risk assets (in particular, those which will best match variations in the floor). As the funding level improves the fund can be shifted more into risky assets which provide the fund with higher upside potential.

Suppose that the minimum funding level (or floor) is M. We have a low-risk portfolio A with a proportion π_{Ai} of the fund invested in asset i ($i = 1, 2, \ldots, n$). We also have a higher-risk portfolio B which invests in proportion to the vector π_B. At funding level x a proportion $p_A(x)$ of the fund is invested in portfolio A and $p_B(x) = 1 - p_A(x)$ in portfolio B. Since A is less risky we have (normally):

$$\begin{aligned} \pi_A^T \lambda &< \pi_B^T \lambda \text{ (that is, A has a lower expected return)} && (95) \\ \pi_A^T D \pi_A &< \pi_B^T D \pi_B \text{ (that is, A is lower risk).} && (96) \end{aligned}$$

We define $p_B(x)$ in one of the following ways:

$$\begin{aligned} p_B(x) &= \frac{x-M}{M} && (97) \\ \text{or } \bar{p}_B(x) &= \max\left\{\frac{x-M}{M}, 0\right\} && (98) \end{aligned}$$

We will concentrate here on $p_B(x)$ for the sake of mathematical convenience since it is normally the case that the probability that $X(t)$ falls below M under this strategy is very small if A is very low risk.

The vector of proportions invested in each asset under CPPI is thus:

$$\begin{aligned} p_c(x) &= \frac{M}{x}\pi_A + \frac{x-M}{x}\pi_B && (99) \\ &= \frac{p_{c0} + p_{c1} x}{x} && (100) \\ \text{where } p_{c0} &= M(\pi_A - \pi_B) \\ p_{c1} &= \pi_B \end{aligned}$$

We can, therefore, apply all of the results discussed in Section 4.1 to CPPI. For example, it is of interest to compare the effectiveness of CPPI relative to a static investment strategy. Let us look first at the stationary mean and variance of the funding level. In the equations below we use a subscript c for calculations under CPPI and s where we are considering the static strategy. Thus

$$\begin{aligned} E[X_c] &= \frac{\mu_c}{\nu_c} = \frac{c_0 - B + p_{c0}^T \lambda}{c_1 - \delta_0 - p_{c1}^T \lambda} = m_c && (101) \\ \text{and } Var[X_c] &= \frac{\alpha_c + \beta_c m_c + \gamma_c m_c^2}{2\nu_c - \gamma_c} = s_c^2 \text{ say.} && (102) \end{aligned}$$

We will assume that the floor, M, is sufficiently small and that portfolios A and B, and the contribution strategy have been chosen in such a way that $X(t)$ is stationary with $M < m_c < \infty$.

Now suppose that we will employ a static investment strategy under which we hold assets in proportion to the vector $p_s\pi_A + (1-p_s)\pi_B$ for all x where p_s is some scalar quantity. Then we have

$$E[X_s] = \frac{c_0 - B}{c_1 - \delta_0 - (p_s\pi_A + (1-p_s)\pi_B)^T\lambda} = m_s. \tag{103}$$

Now choose p_s in such a way that $m_s = m_c$: that is,

$$p_s = \frac{m_c(c_1 - \delta_0 - \pi_B^T\lambda) - (c_0 - B)}{m_c(\pi_A - \pi_B)^T\lambda} = \frac{M(\pi_A - \pi_B)^T\lambda}{m_c(\pi_A - \pi_B)^T\lambda} = \frac{M}{m_c} \tag{104}$$

Note that $0 < p_s < 1$.

We now claim that $Var[X_s] < Var[X_c]$.

$$\begin{aligned}
& \alpha_c + \beta_c m_c + \gamma_c m_c^2 && (105)\\
= \; & \sigma_b^2 + M^2(\pi_A - \pi_B)^T D(\pi_A - \pi_B) + 2M(\pi_A - \pi_B)^T D\pi_B m_c + \pi_B^T D\pi_B m_c^2 && (106)\\
= \; & \sigma_b^2 + (M(\pi_A - \pi_B) + m_c\pi_B)^T D\,(M(\pi_A - \pi_B) + m_c\pi_B) && (107)\\
= \; & \sigma_b^2 + (p_c\pi_A + (1-p_c)\pi_B)^T D\,(p_c\pi_A + (1-p_c)\pi_B)\, m_c^2 && (108)
\end{aligned}$$

where $p_c = M/m_c$.

We also have

$$\alpha_s + \beta_s m_s + \gamma_s m_s^2 \;=\; \sigma_b^2 + (p_s\pi_A + (1-p_s)\pi_B)^T D\,(p_s\pi_A + (1-p_s)\pi_B)\, m_s^2 \tag{109}$$

But $m_c = m_s$ and $p_c = p_s$ so that $\alpha_c + \beta_c m_c + \gamma_c m_c^2 = \alpha_s + \beta_s m_s + \gamma_s m_s^2$.

Next consider

$$\begin{aligned}
2\nu_s - \gamma_s \;=\; & 2\left(c_1 - \delta_0 - \pi_B^T\lambda - p_S(\pi_A - \pi_B)^T\lambda\right) && (110)\\
& \quad - (p_s\pi_A + (1-p_s)\pi_B)^T D\,(p_s\pi_A + (1-p_s)\pi_B) && (111)\\
= \; & 2\left(c_1 - \delta_0 - \pi_B^T\lambda\right) - \pi_B^T D\pi_B - 2p_S(\pi_A - \pi_B)^T\lambda && (112)\\
& \quad + \left[\pi_B^T D\pi_B - (p_s\pi_A + (1-p_s)\pi_B)^T D\,(p_s\pi_A + (1-p_s)\pi_B)\right] && (113)
\end{aligned}$$

Now $0 < p_s < 1$, $(\pi_A - \pi_B)^T\lambda < 0$ and

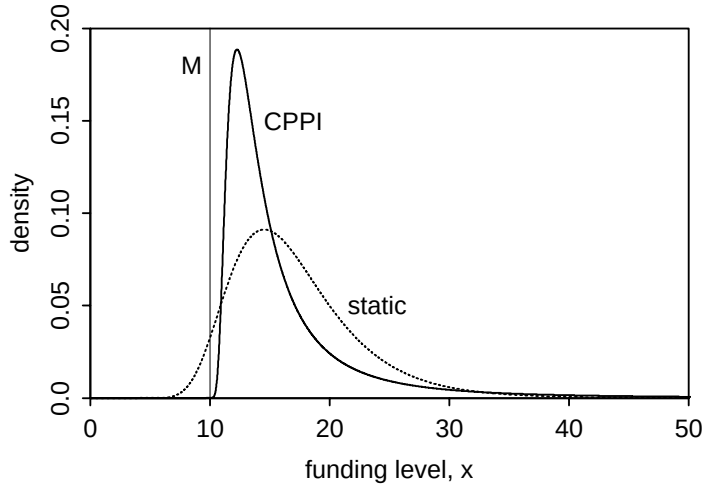

Figure 3: Comparison of stationary distributions for static and CPPI investment strategies. Both distributions have the same mean.

$$\left[\pi_B^T D\pi_B - (p_s\pi_A + (1-p_s)\pi_B)^T D\,(p_s\pi_A + (1-p_s)\pi_B)\right] > 0 \tag{114}$$

(since the expression in square brackets is convex, quadratic in p_s and $\pi_A^T D\pi_A < \pi_B^T D\pi_B$). Therefore

$$2\nu_s - \gamma_s > 2\nu_c - \gamma_c. \tag{115}$$

Hence

$$\frac{\alpha_s + \beta_s m_s + \gamma_s m_s^2}{2\nu_s - \gamma_s} < \frac{\alpha_c + \beta_c m_c + \gamma_c m_c^2}{2\nu_c - \gamma_c} \tag{116}$$

$$\Rightarrow Var[X_s] < Var[X_c] \tag{117}$$

This can be summarised in the following theorem:

Theorem 4.2.1

For any CPPI investment strategy let m_c and s_c^2 be the stationary mean and variance of the funding level X_t. There exists a static investment strategy which invests under which the stationary mean funding level, m_s, is equal to m_c but the stationary variance of the funding level, s_s^2, is less than s_c^2.

Interpretation: In the variance sense, the static strategy is *more efficient* than CPPI: that is, given a CPPI strategy we can always find a static strategy which delivers the same mean funding level but a lower variance.

One example illustrating this result is plotted in Figure 3. Here we use the same fixed parameters as in Section 3.5. In addition we have $c_0 = 1.5$ and $c_1 = 0.07$ for both the

static and CPPI strategies. Under CPPI we have a floor of $M = 10$ with $\pi_A^T = (0,0)$ and $\pi_B^T = (0.2, 0.8)$ (meaning that at the floor ($X(t) = M$) the fund is invested 100% in cash). This gives rise to a mean funding level (that is, assets divided by expected benefit outgo) of 17.1 while the variance of the funding level is, in fact, infinite. Under the matching static investment strategy the expected funding level is also 17.1 while the standard deviation of the funding level is 5.4. This marked difference in the variances is caused by the fatness of the tail of the CPPI distribution although this is not clear from Figure 3. What we can see in Figure 3 is that the two distributions are quite different.

One might ask why would we use CPPI when the static strategy has been shown to be more efficient. The answer to this is that it depends upon the objectives of the pension fund. If the objective is to minimise variance then clearly the static strategy is superior (although we have shown in Section 3 that a form of "inverse" CPPI is better still). On the other hand, if the objective is to minimise the probability that the funding level falls below the floor, M, then CPPI is clearly superior.

5 Numerical examples

We consider now an example in which the following parameters are fixed (as in Section 3.6):

$$\delta_0 = 0.03, \;\; \delta = \begin{pmatrix} 0.04 \\ 0.06 \end{pmatrix}, \;\; S = \begin{pmatrix} 0.05 & 0.05 \\ 0.05 & 0.2 \end{pmatrix}, \;\; B = 1, \;\; \sigma_b = 0.1 \tag{118}$$

Here we consider an analysis of the sensitivity of the optimal control strategies to variation of the input parameters in the value function and the loss function. The central parameter values which we will use are:

$$c_m = 0.6, \;\; k = 0.001, \;\; x_p = 10, \;\; \beta = 0.03 \tag{119}$$

Furthermore, we assume that none of the fund can be invested in cash (as in Section 3.3 with $p_m = 1$). Throughout this analysis we keep the fifth input parameter ρ equal to 0.

In Tables 2 (Dynamic optimisation) and 3 (Stationary optimisation) below we give the values of the input parameters (c_m, k, x_p and β), the optimal values of p_0, p_1, c_0 and c_1, and the mean and standard deviation of the stationary fund size and the contribution rate.

The values given for p_{B0} and p_{B1} relate to the proportion of the fund invested in the more risky but efficient portfolio B: that is, the portfolio in which investments are in proportion to the vector $D^{-1}\lambda$. In particular, the proportion of the fund invested in portfolio B is $p_B(x) = (p_{B0} + p_{B1}x)/x$. Since $p_m = 1$ the remainder of the assets are invested in the minimum variance portfolio A: $D^{-1}e/(e^T D^{-1} e)$.

Table 2: Dynamic Optimisation

Ex.	c_m	k	x_p	β	σ_b	p_{B0}	p_{B1}	c_0	c_1	$E[X]$	$S.D.[X]$	$E[C]$	$S.D.[C]$
1	0.6	0	–	0.02	0.1	20.7	-1.97	1.00	0.038	8.9	3.8	0.660	0.146
2	0.6	0	–	0.03	0.1	20.7	-1.97	0.89	0.028	7.0	∞	0.699	∞
3 (*)	0.6	0.005	10	0.03	0.1	20.1	-1.97	1.48	0.086	9.6	1.5	0.651	0.132
4	0.6	0.05	10	0.03	0.1	19.8	-1.97	3.00	0.238	9.9	0.8	0.647	0.194
5	0.5	0.005	10	0.03	0.1	22.1	-1.97	1.47	0.086	9.7	1.6	0.633	0.141
6	0.6	0.005	15	0.03	0.1	26.0	-1.97	1.74	0.086	14.2	2.3	0.512	0.194
7	0.6	0.005	10	0.03	0.2	20.1	-1.97	1.48	0.086	9.6	1.6	0.651	0.139

Table 3: Stationary Optimisation

Ex.	c_m	k	x_p	β	σ_b	p_{B0}	p_{B1}	c_0	c_1	$E[X]$	$S.D.[X]$	$E[C]$	$S.D.[C]$
1	0.6	0	–	–	0.1	20.7	-1.97	1.21	0.058	9.7	2.2	0.644	0.127
2			–	–									
3 (*)	0.6	0.005	10	–	0.1	20.1	-1.97	1.68	0.105	9.8	1.3	0.647	0.141
4	0.6	0.05	10	–	0.1	19.9	-1.97	3.16	0.254	9.9	0.8	0.647	0.200
5	0.5	0.005	10	–	0.1	22.4	-1.97	1.70	0.105	10.3	1.4	0.619	0.152
6	0.6	0.005	15	–	0.1	25.6	-1.97	1.97	0.105	13.7	1.9	0.526	0.197
7	0.6	0.005	10	–	0.2	20.1	-1.97	1.68	0.105	9.8	1.4	0.647	0.149

5.1 Notes on the numerical examples

- Examples 1 and 2 show the effect of changing the risk-discount rate β when $k = 0$. Note how the variances in the dynamic case become infinite as β increases. However, when $\beta = 0.03$ the dynamic optimum still has a stationary distribution, albeit with infinite variances.

  When $k = 0$ we can also see that the optimal asset-allocation strategies for the dynamic and static cases are the same and do not depend upon β. We also see that $c_1^D = c_1^S - \beta$ if $k = 0$.

  As β tends to 0 the optimal dynamic solutions converge to the same values as the optimal static solution. The effect of β is therefore to suppress variance in the short term through a lower value of c_1. A low value of c_1 may reduce variance in the short term but it increases it in the long run by allowing fluctuations in the fund size to persist.

  In Example 1 we also see that $E[C] > c_m$. This reflects that fact that the minimum variance of C falls as $E[C]$ increases (and $E[X]$ falls).

- In Example 1, the fund is invested 100% in portfolio B when X equals about 7.0. Below this the fund goes long in portfolio B and short in portfolio A. Conversely, when X reaches just above 10.5 the fund has 100% in portfolio A. When X goes above this there is a long position in portfolio A and a short position in portfolio B.

  Similar ranges apply for each of the other examples.

- Examples 2, 3 and 4 show the effect of increasing k. This shifts the emphasis onto reducing the variance of the fund size rather than of the contribution rate. The principle effect is that c_1 increases with k: that is, surplus or deficit is amortised more quickly. The changes in p_0 and c_0 are primarily a knock on effect.

- Examples 3 and 5 demonstrate the consequences of changing c_m. p_1 remains unchanged as it does throughout. The changes in the remaining control parameters have the effect of shifting the mean values principally but also affect the variances.

- Examples 3 and 6 consider the effect of changing the target fund size x_p. There is no change in c_1 or p_1. p_0 and c_0 change in order to shift the mean fund size. The variance rises because the target fund size is being moved away from the more natural mean observed in Example 3. This increases the tension on the mean contribution rate since a target fund size of 15 is not entirely consistent with a target contribution rate of 0.6.

- Examples 3 and 7 show the influence of the uncertainty in the level of benefit outgo. As was remarked in Section 3, σ_b has no effect on the optimal values of p_0, p_1, c_0 and c_1. Furthermore, the increases in the variances are small indicating that at this level ($\sigma_b = 0.1$ or 0.2) the main source of variability in the contribution rate is due to investment risk.

- The stationary distributions for the fund size for the dynamic and the stationary optima in Example 3 are plotted in Figure 4. It can be seen that the results are similar although the dynamic optimum gives rise to a stationary distribution which is less peaked and which has fatter tails. In other cases (for example, Example 6) if there is some tension between the target funding level, x_p, and the target contribution rate, c_m, there will be more of a difference between the two stationary distributions.

6 Constraints and discontinuities

A number of possible constraints can be put in place which complicate considerably the preceeding analyses. These are:

- upper and lower barriers for the funding level, $X(t)$. These might be legislative requirements or self-imposed by the fund sponsor and the trustees.

- an upper limit set by the fund sponsor on the contribution rate.

- restrictions on the short-selling of assets.

Further discontinuities might exist where the objective function has a non-standard form. For example, we may have

$$L(t,c,x) = (c-c_m)^2 + k\,(\max\{x_p - x, 0\})^2 \tag{120}$$

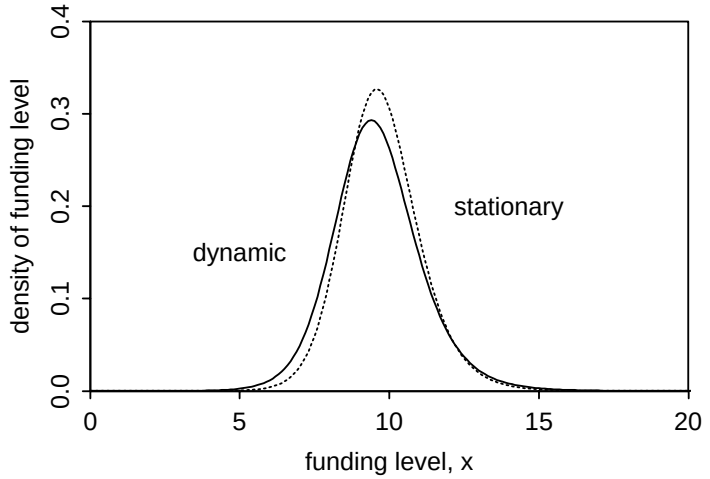

Figure 4: Example 3: Comparison of the stationary distribution of the funding levels for the dynamic and stationary optimal solutions.

where the second term only introduces a penalty when the funding level drops below x_p. Such a function can also be used as a means of investigating the effects of a barrier since as k gets larger and larger the optimal contribution rate below x_p will increase in an effort to raise the funding level above x_p as quickly as possible. For large k this will have the effect of looking like a reflection off the barrier.

Analysis of many of these problems is under way but there are only a few interesting results to discuss at this stage.

6.1 Dynamics in the presence of a minimum barrier

A much simplified version of the minimum funding requirement in the UK is as follows. There is a floor M below which the funding level should not fall. If $X(t)$ does drop below M then it is immediately increased to M by a special contribution.

This problem can be approached by modifying the original setup described in Section 2 by adding an additional contribution rate $c^+.\max\{M-X(t),0\}$: that is, when the funding level is below M. As c^+ tends to infinity the dynamics of the model approach that described above and the process reflects off the barrier M. In this limit the process can be written as follows:

$$dX = X\left[\left(\delta_0 + p(X)^T\lambda\right)dt + p(X)^T S dZ\right] + c(X).dt - B.dt - \sigma_b.dZ_b + dL_t^M \ . \quad (121)$$

The new term in this formula, dL_t^M, is called the *local time* of the process, $X(t)$, at M and is defined as

$$L_t^M \ = \ \int_0^t dL_s^M \quad (122)$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t I(M \leq X(t) < M + \varepsilon) ds \tag{123}$$

where $I(\cdot)$ is equal to 1 when $X(t)$ lies between M and $M+\varepsilon$ and 0 otherwise.

L_t^M is a measure of how much time the process spends in the vicinity of M.

The dL_t^M term represents the additional contributions required when the process hits the barrier to keep $X(t)$ above M. In a sense it gives the process a small upwards 'kick' every time it hits the barrier.

It is possible to analyse the stationary distribution of such a process when $c(x)$ and $p(x).x$ are linear in x away from the barrier: it has a truncated Pearson type IV distribution. However, $Var[c(X)dt + dL_t^M]/dt^2$ is infinite whereas the variance is finite when there is no barrier. This inhibits the optimisation of, for example, quadratic objective functions. For such problems it is easier to replace dL_t^M by $c^+ \max\{M - X(t), 0\}dt$ and consider what happens as c^+ tends to infinity.

A more suitable loss function which accommodates local time as a result of the existence of upper and lower barriers is:

$$L(c) = l_0 c + \sqrt{1 + l_1^2 (c - l_2)^2}. \tag{124}$$

Note that as $c \to +\infty$, $L(c) \sim (l_0 + l_1)c$, while as $c \to -\infty$, $L(c) \sim (l_0 - l_1)c$. This asymptotic linearity is required to ensure that the expected value of the loss function does not become infinite when a reflecting barrier and local time is introduced. If $l_0 < l_1$ then $L(c)$ is increasing and convex. In other words, the fund sponsor prefers to pay less rather than more and prefers stability to instability. Furthermore, the employer will be prepared to pay a higher average contribution rate in the long run in return for lower volatility in the contribution rate.

6.2 No short-selling of assets

Suppose that the holdings in each asset must be non-negative: that is, $0 \leq p(t,x) \leq 1$ for all t, x.

Let us consider the following piecewise linear model for the proportion of the fund invested in the more risky asset 2 in a 2-asset model:

$$c(x) = c_0 - c_1 x \text{ for all } x \tag{125}$$

$$p_2(x) = \begin{cases} 1 & \text{if } x < x_0 \\ \frac{x_0}{(x_1 - x_0)} \frac{(x_1 - x)}{x} & \text{if } x_0 \leq x < x_1 \\ 0 & \text{if } x_1 \leq x \end{cases} \tag{126}$$

In the unconstrained case

Table 4: Stationary optimisation under constraints

Case	x_0	x_1	c_0	c_1	$E[X]$	$S.D.[X]$	$E[C]$	$S.D.[C]$	E[L(C,X)]
Constrained	3.87	7.40	1.77	0.123	9.29	1.50	0.628	0.184	0.0485
Unconstrained	3.96	7.89	1.68	0.105	9.85	1.35	0.645	0.142	0.0315

$$p_2(x) = \frac{x_0}{(x_1 - x_0)} \frac{(x_1 - x)}{x} \tag{127}$$

for all x. The use of x_0 and x_1 here makes it easier to see where the constraints lock in. For asset 1 and cash we have:

$$p_1(x) = 1 - p_2(x) \tag{128}$$
$$p_0(x) = 0 \tag{129}$$

Over each interval $[0, x_0)$, $[x_0, x_1)$ and $[x_1, \infty)$ the stationary distribution function is a scaled Pearson type IV with different parameters over each interval. Since $p(x)$ is continuous the stationary density function is continuous. This allows numerical evaluation of an objective function and hence optimisation over c_0, c_1, x_0 and x_1.

Let us consider a numerical example. We use the same model parameter values and objective function as in Section 5. In the unconstrained problem the optimal solution is linear in x as usual. In the constrained problem the optimal solution will not be linear or piecewise linear in x, but here we optimise only over piecewise linear strategies.

It can be seen by referring to Table 4 that the effects of the constraints in this example are fairly small but, nevertheless, significant. The size of the effect of the constraint depends upon to what extent the interval $[x_0, x_1)$ comes into play in the unconstrained case. If $X(t)$ falls into $[x_0, x_1)$ most of the time then the effect of the constraint will be small. Here, in the unconstrained case, most of the time the fund is invested long in the low-risk asset 1 and short in the high-risk asset 2.

6.3 Upper limit on the contribution rate

Boulier *et al.* (1996) considered the effect of an upper bound on the optimal contribution rate. This resulted in a nearly linear form for $c(x)$ and a bell-shaped curve for $p(x)$. Their solution required the existence of a risk-free asset for the fund and zero volatility in the benefit outgo: otherwise the dynamics of the model would be non-stationary since a sufficiently large deficit will eventually build up which cannot be eradicated.

7 Conclusions

This paper has considered the optimal control of a pension fund using the asset-allocation strategy and the contribution strategy.

Optimal solutions have been derived for power and exponential loss functions (with no demographic risk – $\sigma_b = 0$) and, in more detail, for a quadratic loss function. In most cases the contribution strategy appears to be sensible and conforms with current practice. In each case aspects of the solution were not completely satisfactory. First power and exponential loss functions were found to give rise to non-stationary solutions. Second, when we considered the quadratic and exponential loss functions, the optimal asset-allocation strategy derived was rather counterintuitive: moving, say, out of equities into bonds when the level of surplus is growing.

This has one of two explanations. Funds may be operating in a very non-optimal way. Alternatively, they may be operating optimally but with different objectives. For example, in the UK, the government has recently introduced minimum funding legislation. This should lead to loss functions which heavily penalise events when the fund size falls below the legal minimum. Boulier *et al.* (1996) considered a related problem in which the contribution rate was subject to an upper constraint (say, twice the target rate). However, in the present framework (in which all assets are risky and where there is volatility in the benefit outgo) it is not possible to constrain the contribution rate in this way, for otherwise the fund size would ultimately drift off to minus infinity.

There is, however, some sense in a shift out of equities if the fund size is well above its target level. First if there is too much surplus then there will be pressure on the sponsoring employer to use this surplus to pay for discretionary pension increases which, perhaps, had not been promised. In any event the members would be benefitting from good investment returns while the employer has to pay when things go badly. Second if the employer is able to take a refund, the refund may be liable to tax (for example, in the UK this is 40% with the aim of inhibiting exploitation of the tax advantages enjoyed by a pension fund). Third, too much surplus may lead to the removal of part or all of the fund's special tax status (again this is the case in the UK). All of these reasons mean that it should be advantageous to put a bigger proportion of the fund into low-risk assets when the fund has a large surplus. The results described in this paper back up this viewpoint.

It is clear from the results contained in this paper that we must look for alternative loss functions. The target are ones which give rise to stationary solutions and sensible asset-allocation strategies.

The results presented in this paper and in that of Boulier *et al.* (1995) also draw attention to the following issues:

- what objective functions (if any) are used by pension funds? Are pension funds currently operating in a sub-optimal way or do they have different objective functions from the one considered here?
- what constraints (if any) on contributions and investments are appropriate? Can investment constraints be circumvented by prudent use of derivatives?

- is too much emphasis placed on the calculation of the so-called actuarial liability when this may have no relationship to the target funding level under the optimised objective function? Here the problems have been analysed solely with reference to the objective function. A framework which relies heavily on the actuarial liability might result in a solution which is sub-optimal with reference to this stated objective function.

It must be stressed again that we have assumed a stable membership structure in the pension plan. In many problems there may be a reason, for example, to incorporate changes in the membership as a sponsoring company evolves or restructures. Such situations would require an adaptation of, for example, the use of $X(t) = F(t)/W(t)$ as the key process and of the objective function. The findings in this paper suggest that the interests of the employer and fund members might be served better by a combination of dynamic control theory and more traditional actuarial valuation techniques. In this respect, theoretical solutions to simplified problems give us a basis for investigations of more complex situations.

Acknowledgements

I would like to thank the many people who have commented on earlier versions of this paper and at various talks. In particular, I have had useful conversations with or suggestions from Jean-Francois Boulier, Daniel Dufresne, Martin Schweizer, Wolfgang Rungaldier, Michael Taksar, Hanspeter Schmidli, Ravi Mazumdar, Paul Embrechts, Gary Parker, Martino Graselli, Mary Hardy, David Wilkie, Terence Chan, Iain Currie, Jack Carr, Ken Brown, and Andrew Lacey.

References

Black, F. and Jones, R. (1988) Simplifying portfolio insurance for corporate pension plans. *Journal of Portfolio Management* **14(4)**, 33-37.

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

Boulier, J-F., Trussant, E. and Florens, D. (1995) A dynamic model for pension funds management. *Proceedings of the 5th AFIR International Colloquium* **1**, 361-384.

Boulier, J-F., Michel, S., and Wisnia, V (1996) Optimizing investment and contribution policies of a defined benefit pension fund. *Proceedings of the 6th AFIR International Colloquium* **1**, 593-607.

Cairns, A.J.G. (1995) Pension funding in a stochastic environment: the role of objectives in selecting an asset-allocation strategy. *Proceedings of the 5th AFIR International Colloquium* **1**, 429-453.

Cairns, A.J.G. (1996) Continuous-time stochastic pension fund modelling. *Proceedings of the 6th AFIR International Colloquium* **1**, 609-624.

Cairns, A.J.G., Blake, D., and Dowd, K. (1999) Optimal dynamic asset allocation for defined contribution pensions. Preprint.

Cairns, A.J.G., and Parker, G. (1997) Stochastic pension fund modelling. *Insurance Mathematics and Economic* **21**, 43-79.

Davis, M.H.A., and Norman, A.R. (1990) Portfolio selection with transaction costs.

Mathematics of Operations Research **15**, 676-713.

Deelstra, G., Grasselli, M., and Koehl, P-F. (1999) Optimal investment strategies in a CIR framework.. Preprint.

De Schepper, A., Teunen, M., and Goovaerts, M. (1994) An analytical inversion of a Laplace transform related to annuities certain. *Insurance: Mathematics and Economics* **14**, 33-37.

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

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

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

Fleming, W.H., and Rishel, R.W. (1975) *Deterministic and stochastic optimal control.* Springer-Verlag, New York.

Föllmer, H., and Schweizer, M. (1993) A microeconomic approach to diffusion models for stock prices. *Mathematical Finance* **3**, 1-23. (Erratum: **4:**285.)

Gerber, H.U., and Shiu, E.S. (2000) Investing for retirement: optimal capital growth and dynamic asset allocation. To appear in North American Actuarial Journal **4** (2).

Haberman, S. and Sung, J-H. (1994) Dynamic approaches to pension funding. *Insurance: Mathematics and Economics* **15**, 151-162.

Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994) *Continuous Univariate Distributions: Volume 1, 2nd Edition.* Wiley, New York.

Keel, A., and Müller, H.H. (1995) Efficient portfolios in the asset liability context. *ASTIN Bulletin* **25**, 33-48.

Merton, R.C. (1971) Optimal consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* **3**, 373-413.

Merton, R.C. (1990) *Continuous-Time Finance.* Blackwell, Cambridge, Mass..

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.

Øksendal, B. (1998) *Stochastic Differential Equations, 5th Edition.* Springer-Verlag, Berlin.

Sharpe, W.F., and Tint, L.G. (1990) Liabilities – a new approach. *Journal of Portfolio Management* **Winter 1990**, 5-10.

Siegmann, A.H., and Lucas, A. (1999) Continuous-time dynamic programming for ALM with risk-averse loss functions. *Proceedings of the 9th AFIR International Colloquium* **2**, 183-193.

Shreve, S.E., and Soner, H.M. (1994) Optimal investment and consumption with transaction costs. *Annals of Applied Probability* **4**, 609-692.

Wilkie, A.D. (1985) Portfolio selection in the presence of fixed liabilities. *Journal of*

the Institute of Actuaries **112**, 229-277.

Wise, A. (1984) The matching of assets to liabilities. *Journal of the Institute of Actuaries* **111**, 445-501.

Yor, M. (1992) Sur certaines foncionnelles exponentielles du mouvement brownien réel. *Journal of Applied Probability* **29**, 202-208.