Choosing the optimal annuitization time post retirement

(Article begins on next page)

Choosing the optimal annuitization time post retirement*

Russell Gerrard† Bjarne Højgaard‡ Elena Vigna§

September 18, 2009

Abstract

In the context of decision making for retirees of a defined contribution pension scheme in the de-cumulation phase, we formulate and solve a problem of finding the optimal time of annuitization for a retiree having the possibility of choosing her own investment and consumption strategy. We formulate the problem as a combined stochastic control and optimal stopping problem. As criterion for the optimization we select a loss function that penalizes both the deviance of the running consumption rate from a desired consumption rate and the deviance of the final wealth at the time of annuitization from a desired target. We find closed form solutions for the problem and show the existence of three possible types of solutions depending on the free parameters of the problem. In numerical applications we find the optimal wealth that triggers annuitization, compare it with the desired target and investigate its dependence on both parameters of the financial market and parameters linked to the risk attitude of the retiree. Simulations of the behaviour of the risky asset seem to show that under typical situations optimal annuitization should occur a few years after retirement.

Keywords: defined contribution pension scheme, de-cumulation phase, stochastic optimal control, optimal annuitization time.

JEL classification: C61, D91, J26, G11, G23.

*We thank Emanuele Lucà for help with Perl programming, Massimo Marinacci for support and encouragement and two anonymous referees for useful comments that improved the paper. Partially supported by Regione Piemonte.

†Faculty of Actuarial Science and Statistics, Cass Business School, London, UK. Email: r.j.gerrard at city.ac.uk

‡Spar Nord Bank, Denmark. Email: bjh at sparnord.dk

§Corresponding author. Università di Torino and CeRP - Collegio Carlo Alberto, Italy. Email: elena.vigna at econ.unito.it

1 Introduction

In defined contribution pension schemes, the financial risk is borne by the member: contributions are fixed in advance and the benefits provided by the scheme depend on the investment performance experienced during the active membership and on the price of the annuity at retirement, in the case that the benefits are given in the form of an annuity. Therefore, the financial risk can be split into two parts: investment risk, during the accumulation phase, and annuity risk, focused at retirement. In order to limit the annuity risk — which is the risk that high annuity prices (driven by low bond yields) at retirement can lead to a lower than expected pension income — in many schemes the member has the possibility of deferring the annuitization of the accumulated fund. This possibility consists of leaving the fund invested in financial assets as in the accumulation phase, and allows for periodic withdrawals by the pensioner, until annuitization occurs (if ever). In UK this option is named "income drawdown option", in US the periodic withdrawals are called "phased withdrawals".

The current actuarial literature about the financial risk in defined contribution pension schemes is quite rich. Papers dealing with the financial risk in DC schemes in the accumulation phase are, for instance, Battocchio and Menoncin (2004), Blake, Cairns and Dowd (2001), Booth and Yakoubov (2000), Boulier, Huang and Taillard (2001), Devolder, Bosch Princep and Dominguez Fabian (2003), Di Giacinto, Federico and Gozzi (2009a), Gao (2008), Haberman and Vigna (2002), Khorasanee (1998) and Knox (1993). The financial risk in the distribution phase of defined contribution pension schemes has been dealt with in many papers, including: Albrecht and Maurer (2002), Blake, Cairns and Dowd (2003), Gerrard, Haberman and Vigna (2004a), Gerrard, Haberman, Højgaard and Vigna (2004b), Gerrard, Haberman and Vigna (2006), Kapur and Orszag (1999), Khorasanee (1996), Milevsky (2001), Milevsky and Young (2002), Milevsky, Moore and Young (2006).

In this paper we assume that the retiree takes the income drawdown option: she defers the annuitization, meanwhile consumes some income withdrawn from the fund and invests the remainder of the fund. Such a pensioner has three principal degrees of freedom:

1 she can decide what investment strategy to adopt in investing the fund at her disposal;

2 she can decide how much of the fund to withdraw at any time between retirement and ultimate annuitization (if any);

3 she can decide when to annuitize (if ever).

The first two choices represent a classical inter-temporal decision making problem, which can be dealt with using optimal control techniques in the typical Merton (1971) framework (see Gerrard et al. (2006) for an example), whereas the third choice can be tackled by defining an optimal stopping time problem.

In this paper, we formulate a combined stochastic control and optimal stopping problem with the aim of outlining a decision tool that could help members of DC schemes in making their decisions regarding the three choices outlined above. The third choice, when to annuitize, has been analyzed with different approaches, for example, by Blake et al. (2003), Stabile (2006), Milevsky et al. (2006) and Milevsky and Young (2007). In this paper we find closed form solutions in terms of two constants z_0, z^* defined as solutions of given equations. We state and prove an algorithm for numerical solutions for z_0, z^* and apply this algorithm for numerical investigations of the optimization problem and its solution. As far as we know, the problem of optimal annuitization in the presence of quadratic loss functions have not been tackled yet in the literature of defined contribution pension schemes. On the other hand, we regard quadratic loss functions which are target-depending as appropriate for defined contribution pension schemes, for they have proven to produce optimal portfolios that are efficient in the mean-variance setting (see Højgaard and Vigna (2007)). On the contrary, optimal portfolios derived via expected utility maximization of CARA and CRRA utility functions have proved to be not efficient in the mean-variance setting (see Vigna (2009)). The main contribution of this paper to the current actuarial literature on pension funds is the solution in closed form of the optimal annuitization time problem in the presence of quadratic loss functions.

The remainder of the paper is organized as follows: section 2 outlines the general model, section 3 treats the model with quadratic utility functions, in which case a solution is constructed. In section 4 we verify that the constructed solution does solve the optimization problem. Section 5 presents some numerical investigations of the problem and section 6 concludes.

2 The general model

2.1 Basics

A pensioner has a lump sum of size $x(0)$ at time 0, which can be invested either in a riskless asset paying interest at fixed rate r or in a risky asset, whose price evolves as a geometric Brownian motion with parameters λ and σ. We assume that the remaining lifetime T_D of the pensioner is independent of the Brownian motion and exponentially distributed with parameter δ. In other words, the pensioner's force of mortality is supposed constant, equal to δ.

Up until the time of annuitization, the pensioner can choose what proportion of the fund to invest in the risky asset and can choose how much to withdraw from the fund. She is also able to select the time of eventual annuitization. The size of the annuity purchasable with sum x is kx, where $k > r$.

If the amount of money in the fund is ever exhausted, no further investment or withdrawal is permitted.

The pensioner derives utility $U_1(b)$ from a payment of size b before annuitization, $U_2(kx)$ from the same payment after annuitization. The introduction of two utility functions is to account for the fact that she might be wary of withdrawing money from the fund when this will increase the probability of ruin. Both U_1 and U_2 are assumed concave (but not necessarily strictly concave).

Notation:

- T_D is the pensioner's time of death, as measured from the time when the lump sum is received
- T is the time of annuitization
- T_0 is the time when the fund goes below 0
- $x(t)$ is the size of the fund at time t (where $t < \min(T, T_D, T_0)$)
- $y(t)$ is the proportion of the fund invested in the risky asset at time t
- $b(t)\,dt$ is the income withdrawn from the fund between time t and time $t + dt$.

We thus investigate the problem of choosing two continuous control variables, $y(t)$ and $b(t)$, and a stopping time, T, in such a way as to maximise the following expectation

$$E_x\left[\int_0^{T_D\wedge\tau} e^{-\rho t}U_1\big(b(t)\big)\,dt + \mathbf{1}_{\tau<T_D}\int_\tau^{T_D} e^{-\rho t}U_2\big(kx(\tau)\big)\,dt,\right] \tag{2.1}$$

where $E_x(\cdot) = E(\cdot|x(0) = x)$, $\tau = T \wedge T_0$, ρ is a subjective discount factor and the updating equation for x is

$$dx(t) = -b(t)\,dt + y(t)x(t)\big(\lambda\,dt + \sigma\,dB(t)\big) + r\big(1 - y(t)\big)x(t)\,dt, \tag{2.2}$$

where $B(\cdot)$ represents a standard Brownian motion.

Since mortality is assumed to operate independently of the evolution of the fund level, in (2.1) we can integrate out the time of death to obtain the objective function:

$$E_x\left[\int_0^\tau e^{-(\rho+\delta)t}U_1\big(b(t)\big)\,dt + \frac{e^{-(\rho+\delta)\tau}}{\rho+\delta}U_2\big(kx(\tau)\big)\right]. \tag{2.3}$$

The operation of such a scheme may be subject to local regulation:

- $b(t)$ may be restricted to lie in a given range $(b_{\min}, b_{\max})$, with both minimum and maximum values dependent on $x(0)$;
- there may be an upper limit on T, for example if pensioners are required to purchase an annuity by a given age;
- the investment strategy $y(t)$ may be constrained to be non-negative or to be no greater than unity, depending on rules regarding the possibility of short selling of risky assets or borrowing to fund additional equity purchases.

However, in this paper we treat only the situation of unconstrained controls.

Definition 1 (Admissible controls) *A control strategy* $(\{b(t) : t \geq 0\}, \{y(t) : t \geq 0\}, T)$ *is* admissible *if*

a) $\{b(t) : t \geq 0\}$, $\{y(t) : t \geq 0\}$ *and* T *are all adapted to the filtration generated by* $\{\big(x(t), B(t)\big) : t \geq 0\}$*;*

b) There is some constant $C_0 < \infty$ *such that, with probability 1,* $|y(t)x(t)| \leq C_0$ *for all* $t \leq T$*.*

Let $V(x)$ denote the supremal expected reward given that $x(0) = x$, i.e.

$$V(x) = \sup_{b,y,T} E\left[\int_0^\tau e^{-(\rho+\delta)t}U_1\big(b(t)\big)\,dt + \frac{e^{-(\rho+\delta)\tau}}{\rho+\delta}U_2\big(kx(\tau)\big)\right]. \tag{2.4}$$

It is well known from the theory of optimal stopping time combined with stochastic control (see for instance Øksendal (1998), chapter 10, or Peskir and Shiryaev (2006)) that the value function must satisfy the following Hamilton-Jacobi-Bellman equation:

$$V(x) = \max\left\{\frac{U_2(kx)}{\rho+\delta}, V(x) + \sup_{b,y}\Big[U_1(b) - (\rho+\delta)V + \mathcal{L}^{b,y}V\Big]\right\}, \tag{2.5}$$

where

$$\mathcal{L}^{b,y}V = [-b + rx + (\lambda - r)xy]V'(x) + \tfrac{1}{2}\sigma^2x^2y^2V''(x). \tag{2.6}$$

From (2.5) one can prove that for each $x \geq 0$ satisfying (2.5) is equivalent to satisfying the following variational inequalities:

$$V(x) \geq \frac{U_2(kx)}{\rho+\delta} \quad \text{and} \quad \sup_{b,y}\Big[U_1(b) - (\rho+\delta)V + \mathcal{L}^{b,y}V\Big] = 0 \tag{2.7}$$

or

$$V(x) = \frac{U_2(kx)}{\rho+\delta} \quad \text{and} \quad \sup_{b,y}\left[U_1(b) - (\rho+\delta)V + \mathcal{L}^{b,y}V\right] \leq 0 \tag{2.8}$$

A point x will be said to be in the continuation region if the first of these is the case, or in the stopping region if the second is true. The intuition behind (2.5) and related (2.7) and (2.8) is the following. When the level of wealth x is in the continuation region the reward in case of investment-consumption optimization is higher than that achieved in case of exit from the optimization plan and therefore it is convenient to keep playing the game, with value function satisfying the usual HJB equation derived via stochastic control theory (rhs of 2.7). On the other hand, when x is in the stopping region, the reward in case of exit from the optimization plan is higher and coincides with the value function so that it is convenient to stop the game.

The compulsory termination of activity in the event of ruin implies that $V(0-) = U_2(0)/(\rho+\delta)$.

It can be proved rigourously that the value function must satisfy (2.7) and (2.8) but the standard approach is to use the verification techniques where one shows that any function W that satisfies (2.7) and (2.8) under weak assumptions is equal to the value function V. In the following section we prove the verification theorem.

2.2 The verification theorem

Lemma 2 *Assume there exists a C^2 function W that satisfies (2.7) and (2.8) and that for all admissible controls*

$$E\left[\int_0^t y(s)x(s)W'\big(x(s)\big)e^{-(\rho+\delta)s}\,B(ds)\right] = 0. \tag{2.9}$$

for all t. Then $W(x) \geq V(x)$ for all x.

Proof. By Dynkin's Formula and (2.9) we have for any control and stopping time T, that

$$E[e^{-(\rho+\delta)\tau}W\big(x(\tau)\big) - W(x)] = E\left[\int_0^\tau e^{-(\rho+\delta)s}\left[\mathcal{L}^{b(s),y(s)}W\big(x(s)\big) - (\rho+\delta)W\big(x(s)\big)\right]\,ds\right], \tag{2.10}$$

where $\tau = T \wedge T_0$. From (2.7) and (2.8) the integrand on the right hand side is smaller than $-e^{-(\rho+\delta)s}U_1\big(b(s)\big)$. Hence we obtain

$$W(x) \geq E\left[\int_0^\tau e^{-(\rho+\delta)s}U_1(b(s))ds + e^{-(\rho+\delta)\tau}W(x(\tau))\right].$$

From (2.7) and (2.8)

$$W(x(\tau)) \geq \frac{U_2(kx(\tau))}{\rho+\delta}$$

and it follows that $W(x) \geq V(x)$.

Theorem 3 (Verification theorem) *Let W be as in Lemma 2. Assume that $[0,\infty)$ can be split into two regions A and B such that (2.8) is satisfied in A and (2.7) in region B. Let $B^*(x), Y^*(x)$ be the maximizers of the second term in (2.7)and define the controls $y^*(t) = Y^*(x^*(t))$ and $b^*(t) =$*

$B^*(x^*(t))$, where $x^*(t)$ is the solution of (2.2) with $y(t), b(t)$ replaced by $y^*(t), b^*(t)$. Define $T^* = \inf\{t > 0 | x^*(t) \in A\}$ and $\tau^* = T^* \wedge T_0$. Assume that

$$E[e^{-(\rho+\delta)t} W(x^*(t)) 1_{\tau^*=\infty}] \to 0 \tag{2.11}$$

when $t \to \infty$. Then $y^*(t), b^*(t)$ are optimal controls and T^* the optimal stopping time. Furthermore, the function $W(x) = V(x)$.

Proof. Consider the controls $y^*(t), b^*(t)$. Since on $t < \tau^*$, $x^*(t) \in B$ and we get

$$\begin{aligned} & E[e^{-(\rho+\delta)(\tau^*\wedge t)} W(x^*(\tau^* \wedge t)) - W(x)] \\ = \ & E\left[\int_0^{\tau^*\wedge t} e^{-(\rho+\delta)s} (\mathcal{L}^{b^*(s), y^*(s)} W(x^*(s)) - (\rho+\delta) W(x^*(s))) ds\right] \\ = \ & -E\left[\int_0^{\tau^*\wedge t} e^{-(\rho+\delta)s} U_1(b^*(s)) ds\right] \end{aligned}$$

Letting $t \to \infty$, we get by (2.11) that

$$\begin{aligned} W(x) &= E\left[\int_0^{\tau^*\wedge t} e^{-(\rho+\delta)s} U_1(b^*(s)) ds + e^{-(\rho+\delta)\tau^*} W(x^*(\tau^*)) 1_{\tau^*<\infty}\right] \\ &= E\left[\int_0^{\tau^*\wedge t} e^{-(\rho+\delta)s} U_1(b^*(s)) ds + e^{-(\rho+\delta)\tau^*} \frac{U_2(x^*(\tau))}{\rho+\delta} 1_{\tau^*<\infty}\right]. \end{aligned}$$

Now the result follows by applying Lemma 2.

Corollary 4 *Assume that $(\rho+\delta)^{-1} U_2(x)$ satisfies (2.9) and*

$$\sup_{b,y} [U_1(b) - U_2(x) + \mathcal{L}^{b,y} (\rho+\delta)^{-1} U_2(x)] \leq 0 \tag{2.12}$$

for all x. Then $T^* = 0$ and $V(x) = (\rho+\delta)^{-1} U_2(x)$.

Proof. The proof follows easily from Theorem 3. That (2.11) is satisfied in this case is obvious.

2.3 Solution within the continuation region

If x is in the continuation region, then

$$\sup_{b,y} \left[U_1(b) - (\rho+\delta)V(x) + [-b + rx + (\lambda - r)xy]V'(x) + \tfrac{1}{2}\sigma^2 x^2 y^2 V''(x)\right] = 0.$$

We assume that there are no restrictions on y and b. The optimizing value of y is therefore

$$y^* = y^*(x) = -\frac{(\lambda - r)V'(x)}{\sigma^2 x V''(x)}, \tag{2.13}$$

as long as $V''(x) < 0$. We shall be assuming that this holds for all x, i.e. that V is concave throughout the continuation region, otherwise there is no finite maximum for y.

The optimal value of b depends on the form of U_1, but we can write

$$b^* = b^*(x) = \arg\sup_b [U_1(b) - bV'(x)]. \tag{2.14}$$

Therefore

$$U_1\big(b^*(x)\big) - (\rho+\delta)V(x) - \big(b^*(x) - rx\big)V'(x) - \tfrac{1}{2}\beta^2 \frac{V'(x)^2}{V''(x)} = 0, \tag{2.15}$$

where β denotes the Sharpe ratio, $\beta = (\lambda - r)/\sigma$.

We make use of a method illustrated by Karatzas, Lehoczky, Sethi and Shreve (1986) (see also Xu and Shreve (1992) and references therein): we introduce the dual problem by defining a function $X(z)$ to be the inverse of V', so that

$$V'\big(X(z)\big) = z \quad \text{and} \quad V''\big(X(z)\big) = 1/X'(z).$$

The concavity of V implies that X is a decreasing function of z. We may then rewrite **(2.15)** as

$$U_1\big(b^*\big(X(z)\big)\big) - zb^*\big(X(z)\big) + rzX(z) - (\rho+\delta)V\big(X(z)\big) - \tfrac{1}{2}\beta^2 z^2 X'(z) = 0. \tag{2.16}$$

The next step is to differentiate this equation with respect to z to obtain

$$\tfrac{1}{2}\beta^2 z^2 X''(z) + (\rho+\delta+\beta^2-r)zX'(z) - rX(z) = -b^*\big(X(z)\big). \tag{2.17}$$

The complementary function is of the form

$$X(z) = C_1 z^{\alpha_1} + C_2 z^{\alpha_2}, \tag{2.18}$$

where $\alpha_1 > \alpha_2$ are the two roots of the quadratic

$$P(\alpha) = \tfrac{1}{2}\beta^2\alpha^2 + (\rho+\delta+\tfrac{1}{2}\beta^2 - r)\alpha - r. \tag{2.19}$$

Observe that the coefficient of α^2 in $P(\alpha)$ is positive and that $P(0) < 0$, $P(-1) = -(\rho+\delta) < 0$. Therefore one root is positive, the other below -1. We assume that $\alpha_1 > 0 > -1 > \alpha_2$.

Let us denote the particular solution by $\xi(z)$. Thus the general solution takes the form

$$X(z) = \xi(z) + C_1 z^{\alpha_1} + C_2 z^{\alpha_2}, \tag{2.20}$$

$$V\big(X(z)\big) = \frac{1}{\rho+\delta}\left[\eta(z) + C_1\left(r - \tfrac{1}{2}\beta^2\alpha_1\right)z^{\alpha_1+1} + C_2\left(r - \tfrac{1}{2}\beta^2\alpha_2\right)z^{\alpha_2+1}\right], \tag{2.21}$$

where

$$\eta(z) = U_1\big(b^*\big(X(z)\big)\big) - zb^*\big(X(z)\big) + rz\xi(z) - \tfrac{1}{2}\beta^2 z^2 \xi'(z).$$

In the following sections we consider a special case, minimizing quadratic disutility functions, as treated in Gerrard et al. (2004b).

3 Quadratic model

3.1 Basics

In the formulation of the problem and the choice of the disutility function, we follow Gerrard et al. (2004b). We investigate the problem of choosing two continuous control variables, $y(t)$ and $b(t)$, and a stopping time, T, in such a way as to minimise

$$E_x\left[v\int_0^\tau e^{-(\rho+\delta)t}\,(b_0 - b(t))^2\,dt + \frac{we^{-(\rho+\delta)\tau}}{\rho+\delta}\,(b_1 - kx(\tau))^2\right],$$

where $\tau = \min(T, T_0)$, v and w are positive weights, k is the amount of annuity which can be purchased with one unit of money, and the updating equation for x is

$$dx(t) = -b(t)\,dt + y(t)x(t)(\lambda\,dt + \sigma\,dB(t)) + r(1 - y(t))x(t)\,dt.$$

This choice corresponds to $U_1(b) = v(b_0 - b)^2$ and $U_2(kx) = w(b_1 - kx)^2$.

The amount b_0, the income target until the annuity is purchased, will in many cases be equal to kx_0, the size of the annuity which could have been purchased if the retiree had annuitised immediately on retirement. This choice is reasonable, for UK regulations specify that the income drawn down from the fund before annuitisation cannot exceed kx_0.

The process evolves until either it is advantageous to annuitise or the fund falls to a negative value, in which case no further trading is permitted. The loss associated with annuitisation when the level of the fund is x, so that the annuity pays kx per unit time, is

$$K(x) = \frac{w}{\rho + \delta}(b_1 - kx)^2. \tag{3.1}$$

Remark 1

- The fact that annuitisation is compulsory when the fund level goes below zero implies that $V(0-) = K(0) = wb_1^2/(\rho + \delta)$.
- It is always possible to consume the interest received on the fund without investing in the risky asset. Therefore $V(x) \leq v(b_0 - rx)^2/(\rho + \delta)$.
- It is always optimal to purchase an annuity if the fund level reaches b_1/k, since no further losses will be incurred in this case. If the fund level is above b_1/k, the investor can consume at rate b_0 then purchase an annuity if the fund level ever falls to b_1/k. Similarly, if the fund level is above b_0/r, she can consume b_0 without diminishing her fund. Therefore

$$V(x) = 0 \text{ for } x \geq \min\left(\frac{b_0}{r}, \frac{b_1}{k}\right).$$

Since the difference between b_0/r and b_1/k appears often, we define it:

$$D \stackrel{def}{=} \frac{b_0}{r} - \frac{b_1}{k}. \tag{3.2}$$

The formulation of the problem makes the possibility that $D < 0$ very atypical. In fact, typically the starting wealth is $x_0 = \frac{b_0}{k} < \frac{b_0}{r}$. In other words, the initial fund gives the possibility to buy a lifetime annuity of size b_0 which costs less than a perpetuity of size b_0. If $\frac{b_0}{r} < \frac{b_1}{k}$, the fund should cross $\frac{b_0}{r}$ before hitting the desired level $\frac{b_1}{k}$. If the fund reaches $\frac{b_0}{r}$, then, as noted above, it is optimal to invest the whole portfolio in the riskless asset and consume b_0, which gives to the pensioner the same outcome of immediate annuitization at retirement. Therefore, it would be impossible to reach the real goal which is being able to afford an annuity of size $b_1 > b_0$. Considering the fact that the utility from bequest in case of death before annuitization is here disregarded, immediate annuitization would then be preferable to the optimization program because it would avoid the ruin possibility. Thus the choice $D < 0$, although perfectly admissible from a mathematical point of view, is not realistic in this context. For this reason, we will henceforth assume that $D > 0$.

Remark 2
On the interpretation of D, we can write D as[1] $D = \frac{b_0}{k}(\frac{k}{r} - \frac{b_1}{b_0})$. Both terms in parentheses have a meaning. In fact, while $\frac{b_1}{b_0}$ measures the risk aversion of the pensioner (the higher the ratio the lower the risk aversion), $\frac{k}{r}$ is the ratio between the price of a 1\$ perpetuity, $\frac{1}{r}$, and the price of a 1\$ lifetime annuity, $\frac{1}{k} = \frac{1+L}{r+\delta}$, where L is the life office profit loading and δ is the (constant) force of mortality. It is well known that $k > r$, and in fact the price of perpetuity is higher than the price of lifetime annuity. Since, as explained above, an interesting problem must have $D > 0$ and since the parameters r, δ and L are fixed exogenously by the market and by the individual's mortality, we deduce that the risk aversion parameter $\frac{b_1}{b_0}$ can be chosen only such that $1 < \frac{b_1}{b_0} < \frac{(r+\delta)}{r(1+L)}$. From this, we can deduce some important relationships:

- the higher δ – which in turn implies the older the pensioner, given that her expected remaining lifetime is $(\delta)^{-1}$ – the larger the range where she can choose her own ratio $\frac{b_1}{b_0}$, and this is reasonable because her shorter remaining lifetime allows her to aim for a significant improvement in her expected income by taking the income drawdown option;
- vice versa, the higher the profit loading L, the narrower the range where she can choose her own ratio $\frac{b_1}{b_0}$, and this also is intuitive, as part of the targeted improvement is eroded by the insurance profit;
- moreover, the lower r the higher $\frac{r+\delta}{r(1+L)}$, i.e. the larger the range where to choose the ratio $\frac{b_1}{b_0}$, and this is intuitive, because low bond yields enhance the convenience of the income drawdown option.

3.2 The value function

The continuation region U is defined by

$$U := \{x \in \mathbf{R} : V(x) < K(x)\}$$

The variational inequalities (2.7) and (2.8) can then be written as:

$$\begin{array}{llll} LV(x) = 0 & \text{and} & V(x) \le K(x) & \text{for } x \in U \\ LV(x) \ge 0 & \text{and} & V(x) = K(x) & \text{for } x \in U^c \end{array} \tag{3.3}$$

where

$$LV(x) = \inf_{b,y}[v(b_0 - b)^2 - (\rho + \delta)V(x) + \mathcal{L}^{b,y}V(x)] \tag{3.4}$$

and where $\mathcal{L}^{b,y}$, as before, is the linear differential operator

$$\mathcal{L}^{b,y}V(x) = [-b + (\lambda - r)yx + rx]V' + \tfrac{1}{2}\sigma^2 y^2 x^2 V''.$$

In that part of the continuation region that lies between 0 and $\frac{b_1}{k}$ the optimal controls are given by

$$y^*(x) = -\frac{(\lambda - r)V'(x)}{\sigma^2 x V''(x)}, \tag{3.5}$$

[1]We thank an anonymous referee for leading us to this alternative interpretation.

$$b^*(x) = b_0 + \frac{1}{2v} V'(x), \tag{3.6}$$

and the optimal stopping time τ^* is given by

$$\tau^* = \inf\{t \geq 0 : x(t) \notin U\}.$$

One of the difficult tasks consists in finding the continuation region U. However, exploiting remark 1, we can prove the following:

Lemma 5 *The continuation region U contains the set $(\frac{b_1}{k}, +\infty)$, but $\frac{b_1}{k} \in U^c$.*

Therefore the only region where the problem is interesting is $\left[0, \frac{b_1}{k}\right)$.

Lemma 6 *If the set U_0 is defined by*

$$U_0 = \{x \in \mathbf{R} \ : \ LK(x) < 0\} \tag{3.7}$$

then $U_0 \subseteq U$.

Proof. If $x \in U^c$ then $V(x) = K(x)$ and $LV(x) \geq 0$, from which it follows that $LK(x) \geq 0$, i.e., $x \in U_0^c$.

Typically, one obtains information on the continuation region U by first analyzing the set U_0.

3.3 The analysis of the set U_0

The set U_0 under study is:

$$U_0 = \{x \ : LK(x) < 0\}$$

$$LK(x) = \inf_{b,y} \left\{v(b_0 - b)^2 - (\rho + \delta)K + [-b + (\lambda - r)yx + rx]K' + \tfrac{1}{2}\sigma^2 y^2 x^2 K''\right\} \tag{3.8}$$

Given the form (3.1) of $K(x)$, the minimising values of (3.8) are:

$$\hat{b}(x) = b_0 - \frac{kw}{v(\rho + \delta)}(b_1 - kx)$$

$$\hat{y}(x) = \beta \frac{b_1 - kx}{\sigma kx}.$$

By substitution, after some algebra, we obtain:

$$U_0 = \{x : w(b_1 - kx)\,[2krD - \phi(b_1 - kx)] < 0\}, \tag{3.9}$$

where D is given by (3.2), and

$$\phi = \rho + \delta + \beta^2 - 2r + k^2 \frac{w}{v(\rho + \delta)}. \tag{3.10}$$

Lemma 7 *Assume that $\phi \leq 2krD/b_1$. Then, for any $x(0) \in [0, b_1/k]$, the optimal behaviour is to annuitise immediately, implying that $V(x) = K(x)$.*

Proof. The proof follows from Corollary 4. Under the condition $\phi < 2krD/b_1$, (2.12) is satisfied for all $x > 0$.

Remark 3
On the interpretation of Lemma 7, we notice that $\phi < 2krD/b_1$ is equivalent to

$$\frac{w}{v} < \frac{2b_0(\rho+\delta)}{kb_1} - \frac{(\rho+\delta+\beta^2)(\rho+\delta)}{k^2}. \tag{3.11}$$

If we now consider the expression of k as given in Remark 2, $k = (\rho+\delta)/(1+L)$ with L insurance loading factor, and we set the simplifying and not unreasonable values $\rho = r, L = 0$ (notice that the former is a typical assumption in this kind of literature, see e.g. Milevsky and Young (2007)), we have the simple condition:

$$\frac{w}{v} < \frac{2b_0}{b_1} - \left(1 + \frac{\beta^2}{\rho+\delta}\right). \tag{3.12}$$

If condition (3.12) is satisfied, then immediate annuitization turns out to be optimal. From (3.12) one can gather some intuitive rules. It is clear that optimal immediate annuitization will happen when the weight given to loss in case of annuitization is sufficiently low with respect to that given to loss for running consumption, or, equivalently, when the weight given to loss in case of running consumption is sufficiently high with respect to that given to annuitization. However, also the risk aversion of the pensioner plays a role. In fact, if the targeted pension income b_1 is higher than twice the annuity b_0 purchasable at retirement, then the rhs of (3.12) is negative, rendering the above inequality impossible to hold. Not surprisingly, also the financial market has an effect: with too high values of the Sharpe ratio β the rhs can become very small and even negative. All these features are intuitive and desirable in the model considered.

Due to Lemma 7, from now on we restrict attention to the case $\phi > 2krD/b_1$.

In this case,

$$U_0 = \left(-\infty, \frac{b_1}{k} - \frac{2rD}{\phi}\right) \cup \left(\frac{b_1}{k}, +\infty\right)$$

and therefore

$$U \supseteq \left[0, \frac{b_1}{k} - \frac{2rD}{\phi}\right) \cup \left(\frac{b_1}{k}, +\infty\right) \tag{3.13}$$

3.4 Solution within the continuation region

In the continuation region, the value function satisfies (see (3.3)):

$$\tfrac{1}{2}\beta^2 \frac{(V')^2}{V''} + \frac{1}{4v}(V')^2 + (b_0 - rx)V' + (\rho+\delta)V = 0. \tag{3.14}$$

The optimal proportion of the fund to invest in the risky asset and optimal income to draw down are given by (3.5) and (3.6), respectively. By application of the methodology illustrated in the general case, we define in this case X to be the negative of the inverse of V', so that

$$V'\big(X(z)\big) = -z.$$

The corresponding wealth function is:

$$X(z) = \frac{b_0}{r} - \frac{z}{2v(r-\gamma)} + C_1 z^{\alpha_1} + C_2 z^{\alpha_2}, \tag{3.15}$$

where γ is given by

$$\gamma = \rho + \delta + \beta^2 - r. \tag{3.16}$$

and C_1 and C_2 are constants to be determined by the boundary conditions. The corresponding value function is:

$$V\big(X(z)\big) = \frac{z^2}{4v(r-\gamma)} - \frac{1}{\rho+\delta}\left[A_1 C_1 z^{1+\alpha_1} + A_2 C_2 z^{1+\alpha_2}\right], \tag{3.17}$$

where

$$A_1 = r - \tfrac{1}{2}\beta^2\alpha_1, \qquad A_2 = r - \tfrac{1}{2}\beta^2\alpha_2. \tag{3.18}$$

Notice that the coefficients A_1 and A_2 are both positive. In fact, the polynomial P given by (2.19) satisfies $P(2r/\beta^2) > 0$, so that $\alpha_i < 2r/\beta^2$ for $i = 1, 2$, thus $A_i = r - \frac{1}{2}\beta^2\alpha_i > 0$ for both i.

The optimal control functions can then be written as

$$y^*(X(z)) = -\frac{\beta}{\sigma}\frac{zX'(z)}{X(z)} \tag{3.19}$$

$$b^*(X(z)) = b_0 - \frac{z}{2v} \tag{3.20}$$

3.5 The boundary of the continuation region

According to (3.13), the form of the continuation region U is

$$U = [0, x^*) \cup (\tilde{x}, +\infty),$$

where $x^* \geq \frac{b_1}{k} - \frac{2rD}{\phi}$ and $\tilde{x} \leq \frac{b_1}{k}$. We begin by investigating $\tilde{x}$.

Lemma 8 $\tilde{x} = \frac{b_1}{k}$.

Proof. Since $V(x) \leq K(x)$, from $K(b_1/k) = K'(b_1/k) = 0$ it follows that $V(b_1/k) = V'(b_1/k) = 0$. Suppose that every interval of the form $(b_1/k - \epsilon, b_1/k)$ (for $\epsilon > 0$) contains an element of U. Then letting $\epsilon \to 0$ implies the existence of a z such that $X(z) = b_1/k$ satisfying $z = -V'(b_1/k) = 0$. However, if $z = 0$, then $X(z)$, which is given by (3.15), cannot be equal to b_1/k. This contradiction shows that the assumption was false. Therefore, for sufficiently small ϵ,

$$\left(\frac{b_1}{k} - \epsilon, \frac{b_1}{k}\right) \subset U^c,$$

and we conclude that $\tilde{x}$ cannot be less than b_1/k. $\square$

Intuitively, this result can be explained by observing that if $\tilde{x}$ were strictly lower than $\frac{b_1}{k}$, then $\frac{b_1}{k}$ would stay in U, which is absurd, since it is clear that when reaching $\frac{b_1}{k}$ one should stop investing and annuitize to get zero loss.

It remains to determine x^*. One obvious characteristic is that

$$V(x^*) = K(x^*). \tag{3.21}$$

In addition, we may apply the "smooth fit principle" (see Shiryaev (2008)) to obtain the further condition that

$$V'(x^*) = K'(x^*). \tag{3.22}$$

If we define z_* by $z_* = -V'(x^*)$, so that $X(z_*) = x^*$, then these two boundary conditions (3.21) and (3.22) can be written in the form

$$\begin{aligned} -z_* &= -\tfrac{2kw}{\rho+\delta}(b_1 - kx^*) \\ \tfrac{w}{\rho+\delta}(b_1 - kx^*)^2 &= \tfrac{z_*^2}{4v(r-\gamma)} - \tfrac{1}{\rho+\delta}\left[A_1 C_1 z_*^{1+\alpha_1} + A_2 C_2 z_*^{1+\alpha_2}\right] \\ x^* &= \tfrac{b_0}{r} - \tfrac{z_*}{2v(r-\gamma)} + C_1 z_*^{\alpha_1} + C_2 z_*^{\alpha_2} \end{aligned} \tag{3.23}$$

In addition, we require a boundary condition at $x = 0$. Since the pensioner is forced to purchase an annuity as soon as the fund becomes negative, one possible boundary condition is that $V(0) = K(0)$. A solution to the problem which satisfies this boundary condition will be called a RP (Ruin Possibility) solution.

However, this is not the only possibility, since there exist strategies which ensure that the fund level never falls below 0. For example, the pensioner could stop investing in the risky asset as soon as x falls below ϵ, and instead consume only the interest on the fund. This leads to a penalty equal to $vb_0^2/(\rho+\delta)$ when $x = 0$, which may be strictly less than $K(0)$. Such a solution will be called a NR (No Ruin) solution, and is characterized by the condition $\lim_{x\to 0} xy^*(x) = 0$, or, in other words, due to (3.19), there exists a value of z such that both $X(z) = 0$ and $X'(z) = 0$.

It is then clear that if

$$\frac{v}{w} < \left(\frac{b_1}{b_0}\right)^2 \tag{3.24}$$

then RP solution will not be feasible.

Although in general the solution $X(z)$ of (3.15) might not hit zero, any version of X which might be considered as a solution to the current problem must hit 0 at some point. We therefore define $z_0 = \inf\{z > 0 : X(z) = 0\}$, so that

$$\frac{b_0}{r} - \frac{z_0}{2v(r-\gamma)} + C_1 z_0^{\alpha_1} + C_2 z_0^{\alpha_2} = 0. \tag{3.25}$$

Then the boundary condition at z_0 corresponding to a RP solution, $V(0) = K(0)$ is

$$\frac{z_0^2}{4v(r-\gamma)} - \frac{1}{\rho+\delta}\left[A_1 C_1 z_0^{1+\alpha_1} + A_2 C_2 z_0^{1+\alpha_2}\right] = \frac{wb_1^2}{\rho+\delta}, \tag{3.26}$$

while for a NR solution the appropriate requirements, $X'(z_0) = 0$ and $V(0) \le K(0)$, are

$$\begin{aligned} \alpha_1 C_1 z_0^{\alpha_1 - 1} + \alpha_2 C_2 z_0^{\alpha_2 - 1} &= \tfrac{1}{2v(r-\gamma)} \\ &\text{and} \\ \tfrac{z_0^2}{4v(r-\gamma)} - \tfrac{1}{\rho+\delta}\left[A_1 C_1 z_0^{1+\alpha_1} + A_2 C_2 z_0^{1+\alpha_2}\right] &\le \tfrac{wb_1^2}{\rho+\delta} \end{aligned} \tag{3.27}$$

3.6 Construction of a solution

The method of construction is to start with a candidate value z_c for z_*, to derive appropriate values of C_1, C_2 and z_0 and to check whether this constitutes a solution to the problem.

Since $\frac{b_1}{k} > x^* \geq \frac{b_1}{k} - \frac{2rD}{\phi}$, we see from (3.23) that any solution z_* must satisfy

$$0 < z_* \leq \frac{4k^2 wrD}{\phi(\rho+\delta)} = z_U, \text{ say,}$$

and so we choose z_c in this range.

3.6.1 Signs of C_1 and C_2

From (3.23) it follows that the corresponding values of C_1 and C_2 must be

$$C_1(z_c) = \frac{2{z_c}^{-\alpha_1}}{\beta^2(\alpha_1-\alpha_2)}\left[-A_2 D - \phi\frac{\big(r-\gamma+\beta^2(1-\alpha_2)\big)(\rho+\delta)}{4k^2 w(\gamma-r)} z_c\right]. \tag{3.28}$$

$$C_2(z_c) = \frac{2{z_c}^{-\alpha_2}}{\beta^2(\alpha_1-\alpha_2)}\left[A_1 D + \phi\frac{\big(r-\gamma+\beta^2(1-\alpha_1)\big)(\rho+\delta)}{4k^2 w(\gamma-r)} z_c\right] \tag{3.29}$$

After some algebra, one can prove that $r-\gamma+\beta^2(1-\alpha_2) > 0$ and that $r-\gamma+\beta^2(1-\alpha_1) > 0$ if and only if $r > \gamma$. From this, it is possible to prove that $C_2(z_c) > 0$ if and only if

$$z_c < \frac{4k^2 wD}{\phi(\rho+\delta)}\left(r + \tfrac{1}{2}\beta^2\alpha_1\right) = z_U\left(1 + \frac{\beta^2}{2r}\alpha_1\right), \tag{3.30}$$

so this is always true for the range of values of z_c under consideration.

By means of a similar argument we find that $C_1(z_c) > 0$ if and only if

$$r > \gamma \quad \text{and} \quad z_c > z_U\left(1 + \frac{\beta^2}{2r}\alpha_2\right). \tag{3.31}$$

3.6.2 Behaviour of the function $X(z)$

Recall equation (3.15) giving the solution for $X(z)$. Since $X(z)$ depends on C_1 and C_2, we can regard it, too, as a function of z_c, denoted as $X(z; z_c)$. Notice that $\lim_{z\to 0} X(z; z_c) = +\infty$, as $\alpha_2 < 0$ and $C_2(z_c) > 0$.

Now observe that

$$\frac{\partial^2 X}{\partial z^2}(z; z_c) = \alpha_1(\alpha_1-1)C_1(z_c)z^{\alpha_1-2} + \alpha_2(\alpha_2-1)C_2(z_c)z^{\alpha_2-2}. \tag{3.32}$$

By investigating $P(1)$ we see that $\alpha_1 > 1$ if and only if $r > \gamma$. Combining this result with (3.31), we notice that we have to consider three possible situations.

Situation 1: If $r < \gamma$ then $0 < \alpha_1 < 1$ and $C_1(z_c) < 0$, so the right hand side of (3.32) is positive, implying that X is convex, viewed as a function of z. In addition, $X(z; z_c) = \frac{z}{2v(\gamma-r)}(1+o(1))$ as $z \to \infty$. Therefore X has a unique minimum value for each fixed z_c.

Situation 2: If $r > \gamma$ and $z_c > z_U(1 + \frac{1}{2}\beta^2\alpha_2/r)$ then $\alpha_1 > 1$ and $C_1(z_c) > 0$, again implying that X is convex. In this case $X(z; z_c) = C_1(z_c)z^{\alpha_1}(1+o(1))$ as $z \to \infty$. Therefore X again has a unique minimum value for each z_c.

Situation 3: If $r > \gamma$ and $z_c < z_U(1 + \frac{1}{2}\beta^2\alpha_2/r)$. In this case $C_1(z_c) < 0$ and we conclude that $\frac{\partial X}{\partial z}(z; z_c) < 0$ for all z; indeed, as $z \to \infty$, $X(z; z_c) = C_1(z_c)z^{\alpha_1}\big(1 + o(1)\big) \to -\infty$.

Notice that in situation 3 one can only have RP solution, whereas situations 1 and 2 allow for both types of solution.

On differentiating (3.28) and (3.29), we find that

$$\frac{dC_1}{dz_c} = \frac{2(1+\alpha_1)}{\beta^2(\alpha_1-\alpha_2)} \cdot \frac{\phi(\rho+\delta){z_c}^{-(1+\alpha_1)}}{4k^2 w}(z_U - z_c) \tag{3.33}$$

$$\frac{dC_2}{dz_c} = -\frac{2(1+\alpha_2)}{\beta^2(\alpha_1-\alpha_2)} \cdot \frac{\phi(\rho+\delta){z_c}^{-(1+\alpha_2)}}{4k^2 w}(z_U - z_c). \tag{3.34}$$

For a fixed value of z we obtain

$$\frac{\partial}{\partial z_c}X(z; z_c) = \frac{2}{\beta^2(\alpha_1-\alpha_2)} \cdot \frac{\phi(\rho+\delta){z_c}^{-1}}{4k^2 w}(z_U - z_c)\left\{(1+\alpha_1)\left(\frac{z}{z_c}\right)^{\alpha_1} - (1+\alpha_2)\left(\frac{z}{z_c}\right)^{\alpha_2}\right\}.$$

Every term is positive. So, as we decrease z_c, the value of $X(z; z_c)$ also decreases for each fixed z. We can conclude that $\inf_{z\geq 0} X(z; z_c)$ decreases as z_c decreases.

Proposition 9 *For z_c sufficiently small,* $\inf_{z>0} X(z; z_c) < 0$.

Proof. We can write

$$\begin{aligned}
X(z; z_c) &= \frac{b_0}{r} - \frac{z}{2v(r-\gamma)} + C_1(z_c)z^{\alpha_1} + C_2(z_c)z^{\alpha_2} \\
&= \frac{b_0}{r} - \frac{z}{2v(r-\gamma)} + \frac{2}{\beta^2(\alpha_1-\alpha_2)}\left(\frac{z}{z_c}\right)^{\alpha_1}\left[-A_2 D - \phi\frac{\big(r-\gamma+\beta^2(1-\alpha_2)\big)(\rho+\delta)}{4k^2 w(\gamma - r)}z_c\right] \\
&\quad + \frac{2}{\beta^2(\alpha_1-\alpha_2)}\left(\frac{z}{z_c}\right)^{\alpha_2}\left[A_1 D + \phi\frac{\big(r-\gamma+\beta^2(1-\alpha_1)\big)(\rho+\delta)}{4k^2 w(\gamma - r)}z_c\right] \\
&= \frac{b_0}{r} - \frac{\zeta z_c}{2v(r-\gamma)} + \frac{2}{\beta^2(\alpha_1-\alpha_2)}\left\{\zeta^{\alpha_1}[-A_2 D + d_2 z_c] + \zeta^{\alpha_2}[A_1 D + d_1 z_c]\right\},
\end{aligned}$$

where $\zeta = z/z_c$, and d_1, d_2 are constants. Since $A_2 > 0$, we can choose ζ sufficiently large that

$$\frac{2D}{\beta^2(\alpha_1-\alpha_2)}[-A_2\zeta^{\alpha_1} + A_1\zeta^{\alpha_2}] < -4\frac{b_0}{r}.$$

Now choose z_c sufficiently small that

$$\max\left\{\frac{2|d_2|z_c}{\beta^2(\alpha_1-\alpha_2)}\zeta^{\alpha_1}, \frac{2|d_1|z_c}{\beta^2(\alpha_1-\alpha_2)}\zeta^{\alpha_2}\right\} < \frac{b_0}{r}.$$

If $r - \gamma > 0$ then it is easily seen that $X(z; z_c) < 0$. If, on the other hand, $r - \gamma < 0$ then choose z so small that $\frac{\zeta z_c}{2v(r-\gamma)} < \frac{b_0}{r}$. Then $X(\zeta z_c; z_c) < 0$, as required.□

To begin the construction process, we set $z_c = z_U$, so that $X(z; z_U)$ is a convex function of z and has a unique minimum. Depending on the sign of $r-\gamma$, we are then either in situation 1 or 2. What happens next depends on whether $\inf_{z\geq 0} X(z; z_U)$ is positive or negative.

Case 1: $\inf_{z\geq 0} X(z; z_U) \geq 0$

In this case we can progressively reduce z_c, which in turn reduces the minimum value of $X(z; z_c)$, until z_c is just large enough that $\inf_z X(z; z_c) = 0$, in other words, that $\frac{\partial}{\partial z} X(z; z_c) = 0$ at exactly the point when $X(z; z_c) = 0$: let z_M denote the value of z_c when this occurs. If, in this case, $V(0; z_M) \leq K(0)$, then the boundary conditions (3.27) are satisfied, so we have a NR solution and the problem is solved: z_* is equal to z_M and z_0 is $\arg\min X(z; z_M)$.

If, however, $V(0; z_M) > K(0)$, then no NR solution is possible, but we can still seek a RP solution (notice that in this case (3.24) is violated). To this end, we continue to reduce z_c. For each z_c, define $z_0(z_c) = \inf\{z \geq 0 : X(z; z_c) \leq 0\}$. Then z_0 is a decreasing function of z_c. Consider $V(0; z_c) - K(0)$: by assumption this is positive when $z_c = z_M$.

z_0 is given by $0 = \frac{b_0}{r} - \frac{z_0}{2v(r-\gamma)} + C_1 z_0^{\alpha_1} + C_2 z_0^{\alpha_2}$. This implies that

$$\begin{aligned} 0 &= \left[-\frac{1}{2v(r-\gamma)} + \alpha_1 C_1 z_0^{\alpha_1 - 1} + \alpha_2 C_2 z_0^{\alpha_2 - 1}\right]\frac{\partial z_0}{\partial z_c} + z_0^{\alpha_1}\frac{\partial C_1}{\partial z_c} + z_0^{\alpha_2}\frac{\partial C_2}{\partial z_c} \\ &= \left[-\frac{1}{2v(r-\gamma)} + \alpha_1 C_1 z_0^{\alpha_1 - 1} + \alpha_2 C_2 z_0^{\alpha_2 - 1}\right]\frac{\partial z_0}{\partial z_c} \\ &\quad + \frac{2\phi(\rho+\delta)}{4k^2 w\beta^2(\alpha_1-\alpha_2)z_c}(z_U - z_c)\left\{(1+\alpha_1)\left(\frac{z_0}{z_c}\right)^{\alpha_1} - (1+\alpha_2)\left(\frac{z_0}{z_c}\right)^{\alpha_2}\right\} \end{aligned} \tag{3.35}$$

In addition, $V(0; z_c) = \frac{z_0^2}{4v(r-\gamma)} - (\rho+\delta)^{-1}[A_1 C_1 z_0^{1+\alpha_1} + A_2 C_2 z_0^{1+\alpha_2}]$. This implies that

$$\begin{aligned} \frac{\partial}{\partial z_c} V(0; z_c) &= \left\{\frac{z_0}{2v(r-\gamma)} - (\rho+\delta)^{-1}[(1+\alpha_1)A_1 C_1 z_0^{\alpha_1} + (1+\alpha_2)A_2 C_2 z_0^{\alpha_2}]\right\}\frac{\partial z_0}{\partial z_c} \\ &\quad - (\rho+\delta)^{-1}\left[A_1 z_0^{1+\alpha_1}\frac{\partial C_1}{\partial z_c} + A_2 z_0^{1+\alpha_2}\frac{\partial C_2}{\partial z_c}\right] \\ &= \left\{\frac{z_0}{2v(r-\gamma)} - \alpha_1 C_1 z_0^{\alpha_1} - \alpha_2 C_2 z_0^{\alpha_2}\right\}\frac{\partial z_0}{\partial z_c} \\ &\quad - \frac{2\phi(\rho+\delta)}{4k^2 w\beta^2(\alpha_1-\alpha_2)}(z_U - z_c)\left\{\alpha_1\left(\frac{z_0}{z_c}\right)^{1+\alpha_1} - \alpha_2\left(\frac{z_0}{z_c}\right)^{1+\alpha_2}\right\} \end{aligned} \tag{3.36}$$

Putting these together gives

$$\frac{\partial}{\partial z_c} V(0; z_c) = \frac{2\phi(\rho+\delta)}{4k^2 w\beta^2(\alpha_1-\alpha_2)}(z_U - z_c)\left\{\left(\frac{z_0}{z_c}\right)^{1+\alpha_1} - \left(\frac{z_0}{z_c}\right)^{1+\alpha_2}\right\}$$

Every term on the right hand side is positive, so $V(0; z_c)$ is an increasing function of z_c: as z_c decreases, $V(0; z_c)$ also decreases. There may come a value of z_c at which $V(0; z_c) = K(0)$. If so, the boundary condition (3.26) is satisfied and we have a RP solution to the problem, since by construction $X'(z) < 0$ for all $z < z_0$.

We should check that $V(0; z_c)$ really does reach $K(0)$ eventually. Let us consider what happens when z_c is close to 0. In this case

$$C_2(z_c) = \frac{2A_1 D z_c{}^{-\alpha_2}}{\beta^2(\alpha_1-\alpha_2)}(1 + O(z_c)), \quad C_1(z_c) = -\frac{2A_2 D z_c{}^{-\alpha_1}}{\beta^2(\alpha_1-\alpha_2)}(1 + O(z_c))$$

and therefore

$$X(z;z_c) = \frac{b_0}{r} - \frac{\zeta z_c}{2v(r-\gamma)} + \frac{2D}{\beta^2(\alpha_1-\alpha_2)}\left[A_1\zeta^{\alpha_2} - A_2\zeta^{\alpha_1}\right] + O(z_c),$$

where $\zeta = z/z_c$. This implies that $z_0(z_c) = \zeta_0 z_c(1+O(z_c))$, where ζ_0 is the solution to

$$\frac{2D}{\beta^2(\alpha_1-\alpha_2)}\left[A_2\zeta^{\alpha_1} - A_1\zeta^{\alpha_2}\right] = \frac{b_0}{r}.$$

(This definitively does have a solution $\zeta_0 > 1$ because putting $\zeta = 1$ on the left hand side gives D, which is less than b_0/r, whereas when $\zeta \to \infty$ the left hand side diverges to $+\infty$.)

Now

$$V(0;z_c) = V \circ X(z_0(z_c);z_c) = \frac{\zeta_0^2 {z_c}^2}{4v(r-\gamma)} + \frac{2A_1A_2D\zeta_0 z_c}{\beta^2(\alpha_1-\alpha_2)(\rho+\delta)}\left[\zeta_0^{\alpha_1} - \zeta_0^{\alpha_2}\right] + O(z_c)$$

Therefore $\lim_{z_c\to 0} V(0;z_c) = 0 < K(0)$, as required.

Case 2: $\inf_{z\geq 0} X(z;z_U) < 0$

In this case no NR solution is possible. We define $z_0(z_c)$ as above. If $V \circ X(z_0(z_U);z_U) < K(0)$, then no RP solution is possible either, since reducing the value of z_c below z_U will only have the effect of decreasing $V(0;z_c)$, and there will be no value of z_c which gives $V(0;z_c) = K(0)$. If, however, $V \circ X(z_0(z_U);z_U) \geq K(0)$, then progressively reducing z_c will eventually result in a value such that $V \circ X(z_0(z_c);z_c) = K(0)$, which corresponds to a RP solution.

3.7 Optimal consumption at z_0

What is the optimal consumption whenever ruin occurs? The answer is different depending on whether we have a NR or a RP solution.

Let us define

$$z_{neg} := 2vb_0$$

It is clear from (3.20) that $b^*(X(z_{neg})) = 0$ and $b^*(X(z)) < 0$ for $z > z_{neg}$, i.e. the optimal consumption is negative for $z > z_{neg}$. This in turn implies that in a NR solution it must be

$$z_{neg} \leq z_0$$

To show this, let us recall that

$$V(x) = \min_{\pi(\cdot)} J(x;\pi(\cdot)) \tag{3.37}$$

where $J(x;\pi(\cdot))$ is the optimality criterium under strategy $\pi(\cdot)$. Call $\tilde{\pi}(\cdot)$ the null strategy, i.e. under $\tilde{\pi}(\cdot)$ the portfolio is invested entirely in the riskless asset and the consumption is null. If we are at 0 at time t in a NR solution problem, then it cannot be $b^*(0) > 0$. In fact, assume that $b^*(0) > 0$. Since we know by construction that the portfolio is entirely invested in the riskless asset, we would have immediate ruin, implying optimal annuitization:

$$V(0) = K(0)$$

However,

$$K(0) = \frac{wb_1^2}{\rho+\delta} > \frac{vb_0^2}{\rho+\delta} = J(0;\tilde{\pi}(\cdot)),$$

in contradiction with (3.37). Therefore in a NR solution optimal consumption at z_0 cannot be strictly positive.

Alternatively, we can notice that if we have a NR solution, we are either in situation 1 or situation 2, which means that $X(z)$ is convex in z, tends to infinity when z goes to 0 and to infinity, and the minimum of $X(z)$ is zero and is reached in $z = z_0$, i.e.

$$\min_{z \geq 0} X(z) = X(z_0) = 0. \tag{3.38}$$

If it was $z_{neg} > z_0$, then at z_0 the positive consumption, coupled with the fund equal to 0 and the optimal portfolio entirely invested in the riskless asset, would push the fund below zero, contrary to (3.38).

While in a NR solution at z_0 the optimal consumption is bound to be either negative or null, this does not apply to RP solution, when the optimization program stops and optimal annuitization occurs as soon as the fund goes below zero. In RP solution, then, optimal consumption at z_0 can be positive, in which case is positive for all permitted values of z.

Finally, let us remark that in a problem with NR solution the optimal consumption is negative for fund size lower than $X(z_{neg})$, i.e. $b^*(x) < 0$ for $x < X(z_{neg})$. Then, $X(z_{neg})$ acts as a sort of undesirable barrier for the fund, below which the optimal consumption rule states to pay money into the fund instead of withdrawing it. Optimal negative consumption in the de-cumulation phase of DC schemes was already observed in Gerrard et al. (2006).

4 Application of the verification theorem

We are now in a position to state and prove a theorem showing that the constructed solution satisfies the verification theorem (Theorem (3)).

Theorem 10 *Assume that $D > 0$ and that $\phi \geq 2krD/b_1$. Suppose that there exist constants C_1, C_2, z_0 and z_* with $0 < z_* < z_0 < \infty$, such that the function $X(z)$ given by (3.15) satisfies the boundary conditions (3.23), (3.25) and either (3.26) or (3.27).*

Then

(i) For each $z \in (z_*, z_0)$ there is a corresponding $x \in (0, x^*)$ such that $X(z) = x$;

(ii) the function V given by

$$\begin{array}{ll} V(x) = 0 & \textit{for } x \geq \frac{b_1}{k} \\ V(x) = K(x) & \textit{for } x^* \leq x \leq \frac{b_1}{k} \\ V\big(X(z)\big) \textit{ is given by (3.17)} & \textit{for } z_* \leq z \leq z_0 \end{array} \tag{4.1}$$

is the optimal value function;

(iii) the optimal time to annuitise is $\tau^* = \inf\{t : x(t) \in U^c\}$, where the continuation set U is given by

$$U = [0, x^*) \cup \left(\frac{b_1}{k}, \infty\right);$$

(iv) for values of x belonging to $[0, x^*)$, the optimal controls are given by

$$y^*(t) = -\frac{\lambda - r}{\sigma^2} \cdot \frac{V'\big(x(t)\big)}{x(t)\, V''\big(x(t)\big)}, \quad b^*(t) = b_0 + \frac{1}{2v} V'\big(x(t)\big).$$

In order to prove the theorem we need to prove the following proposition.

Proposition 11 *Suppose $(C_1, C_2, z_0, z_*, x^*)$ constitutes either a RP solution or a NR solution constructed as above. Then*

a) $-\infty < X'(z) < 0$ *for* $z_* \le z < z_0$;

b) $V(x) - K(x) \le 0$ *for* $0 \le x \le x^*$.

Proof.
a) This follows directly by the construction of the solution. In fact, if we are in situation 1 or 2 and we have a NR solution, then $X(z)$ is decreasing until the minimum reached in $z = z_0$. If we have a RP solution, then the minimum of $X(z)$ is negative and z_0 is defined as the minimum point z where $X(z)$ crosses 0. Therefore, $X(z)$ is decreasing between 0 and z_0. Finally, if we are in situation 3, then $X'(z) < 0$ for all $z > 0$.

b) The proof consists of a series of lemmas.

Lemma 12 *Suppose there exists $\tilde{x} \in (0, x^*)$ such that*

$$\begin{array}{ll} V'(x) \le K'(x) & \text{for } 0 < x < \tilde{x} \\ V'(x) \ge K'(x) & \text{for } \tilde{x} < x < x^*. \end{array} \tag{4.2}$$

Then

$$V(x) - K(x) \le 0 \qquad \text{for} \qquad 0 \le x \le x^*.$$

Proof. We know that $V(0) - K(0) \le 0$ and $V(x^*) = K(x^*)$. For any $x \in (0, \tilde{x}]$,

$$V(x) - K(x) = V(0) - K(0) + \int_0^x (V'(s) - K'(s))\, ds \le 0;$$

similarly, for any $x \in (\tilde{x}, x^*)$,

$$V(x) - K(x) = -\int_x^{x^*} (V'(s) - K'(s))\, ds \le 0$$

□

Lemma 13 *Define $F(z) = V'\big(X(z)\big) - K'\big(X(z)\big)$ for $z \in (z_*, z_0)$.*
(a) If $F(z) > 0$ for $z_* < z < z_0$, then $V(x) \le K(x)$ for $0 < x < x^*$.
(b) If F is concave on (z_*, z_0), then either there exists $\tilde{x} \in (0, x^*)$ such that the condition (4.2) is satisfied or F is strictly positive on (z_*, z_0).

Proof. (a) $V(X(z)) - K(X(z)) = \int_{z_*}^z [V'(X(\zeta)) - K'(X(\zeta))] X'(\zeta)\, d\zeta = \int_{z_*}^z F(\zeta) X'(\zeta)\, d\zeta \le 0$.
(b) Suppose F is concave for $z \in (z_*, z_0)$. Recall that $F(z_*) = V'(X(z_*)) - K'(X(z_*)) = 0$ and $\int_{z_*}^{z_0} F(z) X'(z)\, dz = V(0) - K(0) \le 0$. F cannot be strictly negative throughout (z_*, z_0), since this would violate the integral condition. Therefore either F is strictly positive or there exists some $\tilde{z} \in (z_*, z_0)$ such that $F(z)$ is positive for $z_* < z < \tilde{z}$ and negative for $\tilde{z} < z < z_0$. □

Lemma 14 *$V'(X(z)) - K'(X(z))$ is either concave for $z \in (z_*, z_0)$ or strictly positive for $z \in (z_*, z_0)$.*

Proof.

$$F(z) = V'(X(z)) - K'(X(z)) = -z + \frac{2k^2 w}{\rho+\delta}\left(\frac{b_1}{k} - X(z)\right)$$

If either (a) $r < \gamma$ or (b) $r > \gamma$ and $z_U > z_* > z_U(1 + \frac{1}{2}\beta^2\alpha_2/r)$, then

$$F''(z) = -\frac{2k^2 w}{\rho+\delta}X''(z) < 0,$$

proving that F is concave.

If, on the other hand, $r > \gamma$ and $z_* < z_U(1 + \frac{1}{2}\beta^2\alpha_2/r)$, then $C_1 < 0$ and

$$F'(z) = -1 - \frac{2k^2 w}{\rho+\delta}\left[-\frac{1}{2v(r-\gamma)} + \alpha_1 C_1 z^{\alpha_1 - 1} + \alpha_2 C_2 z^{\alpha_2 - 1}\right] = \frac{\phi}{r-\gamma} - \frac{2k^2 w}{\rho+\delta}\left[\alpha_1 C_1 z^{\alpha_1 - 1} + \alpha_2 C_2 z^{\alpha_2 - 1}\right].$$

Every term on the right hand side is positive, so F is strictly increasing on the range (z_*, z_0). As $F(z_*) = 0$, it follows that $F(z) > 0$ for $z_* < z < z_0$. □

The proof of (b) of the proposition is now straightforward by application of the previous lemmas. □

4.1 Proof of Theorem 10

(i) is clear, since the function $X(z)$ given by (3.15) is continuous and, due to Proposition 11, strictly decreasing, hence invertible over the range.

(ii) In order to show that the function V defined in the Theorem is the optimal value function, we need to show, first, that it satisfies (3.3). Second, that the controls specified in the Theorem are admissible.

The first requirement is that

$$\inf_{b,y}\left\{\mathcal{L}^{b,y}K - (\rho+\delta)K + \frac{v}{\rho+\delta}(b_0 - b)^2\right\} \geq 0 \text{ for all } x \in U^c.$$

This is guaranteed by the fact that $U^c \subset U_0^c$ (see 3.7). Furthermore, $V(x) = K(x)$ by definition.

Next we need to show that

$$\inf_{b,y}\left\{\mathcal{L}^{b,y}V - (\rho+\delta)V + \frac{v}{\rho+\delta}(b_0 - b)^2\right\} = 0 \text{ for all } x \in [0, x^*).$$

By construction, the function V does satisfy this condition as long as $V''(x) > 0$ for all $x \in (0, x^*)$, i.e., as long as $V''\big(X(z)\big) > 0$ for all $z \in (z_*, z_0)$. But $V''\big(X(z)\big) = -1/X'(z)$, so Proposition 11 is sufficient to demonstrate that this is true.
Furthermore, again due to Proposition 11, we have $V(x) \leq K(x)$ for $x \in (0, x^*)$.

Next we turn to the proof of admissibility. By construction, $b^*(t)$ and $y^*(t)$ are functions of $x(t)$, and τ^* is adapted to the filtration generated by $x(t)$. It therefore remains only to prove that $|y^*(t)x(t)|$ has a finite bound with probability 1. Under the stated policy, given that $x(t) = x$, we have

$$xy^*(x) = -\frac{\lambda - r}{\sigma^2}\cdot\frac{V'(x)}{V''(x)} = -\frac{\lambda - r}{\sigma^2}zX'(z),$$

Now $|X'|$ is a continuous function on a compact interval, so has a finite maximum, C_0, say, over the interval. Thus $|y^*(t)x(t)| \leq \frac{\lambda - r}{\sigma^2} C_0 z_0$ for all $t \leq \tau^*$.

(iii) follows from Theorem 3. Showing that U takes this shape is rather technical and follows from the analysis contained in section 3.3.

(iv) follows from Theorem 3, by observing that b^* and y^* are the minimizers of $LV(x)$.

This ends the proof. □

5 Numerical applications

In this section we show two numerical applications of the model presented.

Firstly, with the help of a Perl program that finds the solution with the methodology described in section 3.6 above, we have found the triplet solution (z_0, z^*,x^*) with a number of different scenarios for market and demographic conditions as well as risk profiles. Recalling the form of the continuation region $U = [0, x^*)$, where $x^* < b_1/k$, it seems of crucial interest to study the dependence of the width of the continuation region on the parameters of the problem. This is done by analyzing the ratio $x^*/(b_1/k)$. Results are reported in section 5.1.

Secondly, we have chosen a typical scenario for all the parameters and have simulated the behaviour of the risky asset, by means of Monte Carlo simulations. We have then analyzed the optimal investment/consumption strategies and the time of optimal annuitization as well as the size of the annuity upon annuitization. We have also focused on the impact of optimal annuitization rules, by comparing this model with a similar one that allows for fixed annuitization time. Results are reported in section 5.2.

5.1 Dependence of the solution on the scenario

Recall that in a realistic setting some of the parameters are chosen by the retiree and some are given.
The parameters given are r, λ, σ (financial market), δ (demographic assumptions) and k (financial and demographic assumptions).
Parameters that can be chosen are the weights given to penalty for running consumption, v, and to penalty for final annuitization, w. We remark that the relevant quantity is the ratio of these weights, w/v. Another parameter chosen by the retiree is the targeted level of annuity, b_1, while it is reasonable to assume that the level of interim consumption b_0 is given and depends on the size of the fund at retirement. A typical choice for b_0 is the size of annuity purchasable at retirement with the initial fund x_0. Thus, typically b_1 is a multiple of b_0, and the relevant quantity is $(b_1/b_0) > 1$. It is easy to see this ratio as a measure of the risk aversion of the retiree: the higher b_1/b_0, the lower the risk aversion and vice versa.
Reporting an observation made after introducing D, in Remark 2, we notice that the subjective choice of the risk aversion ratio $\frac{b_1}{b_0}$ is limited by the values of the parameters r, δ and k. In particular, since $\frac{1}{k} = \frac{1+L}{r+\delta}$, where L is the life office profit loading, we have $1 < \frac{b_1}{b_0} < \frac{(r+\delta)}{r(1+L)}$. We refer the reader to the above mentioned Remark for a useful interpretation of these limits. From now on we will assume that $L = 0$.
A parameter that is somehow arbitrary and somehow given is ρ, the intertemporal discount factor: although subjective by its own nature, in typical situations cannot differ too much from the riskfree

rate of return r. However, what is relevant in the problem is the sum $\rho + \delta$, which measures the patience of the retiree for future events, affected also by her age.

By varying the values of $r \in (0.03, 0.05)$, $\lambda \in (0.07, 0.12)$, $\sigma \in (0.1, 0.25)$ (with these values, the Sharpe ratio β varies between 0.08 and 0.9), $\rho \in (0.03, 0.05)$, $\delta \in (0.005, 0.02)$, $k \in (0.07, 0.1)$, $b_1/b_0 \in (1.2, 2)$, $w/v \in (0.275, 1.25)$, and combining them in many possible ways, we have observed the following results:

1. with typical values of the market parameters, situation 1 ($r < \gamma$) is the most likely to occur
2. the case of no solution seems to occur only with situation 2 ($r > \gamma$)
3. with typical values, NR solution is the most frequent one
4. everything else being equal, NR solution becomes RP solution when
   (a) decreasing β; furthermore, if β is reduced too much RP solution becomes "no solution"
   (b) decreasing $\frac{w}{v}$
   (c) decreasing $\frac{b_1}{b_0}$ (provided that the values of ρ and $\frac{w}{v}$ are respectively high and low enough to permit RP solution)
   (d) increasing $\rho + \delta$
5. everything else being equal, the ratio $\frac{x^*}{(b_1/k)}$, i.e. the width of the continuation region
   (a) increases by increasing β, in both solutions RP and NR
   (b) increases by increasing $\frac{w}{v}$, in both solutions RP and NR
   (c) increases by increasing $\frac{b_1}{b_0}$, in both solutions RP and NR
   (d) generally slightly decreases by increasing $\rho+\delta$ when the problem has NR solution, slightly increases by increasing $\rho + \delta$ when the problem has RP solution

The results 5a, 5b and 5c for NR solution are illustrated in Figures 1, 2 and 3 respectively (similar figures can be obtained for RP solution). For instance, Figure 1 reports β on the x-axis and the ratio $\frac{x^*}{b_1/k}$ on the y-axis, the legend reports the values of all the other relevant parameters (left constant in order to isolate the effect of β on the width of the continuation region). All the figures show two different lines to report some of the variety of combinations of parameters tested. Similarly, Figure 2 reports w/v on the x-axis, and Figure 3 reports b_1/b_0 on the x-axis.

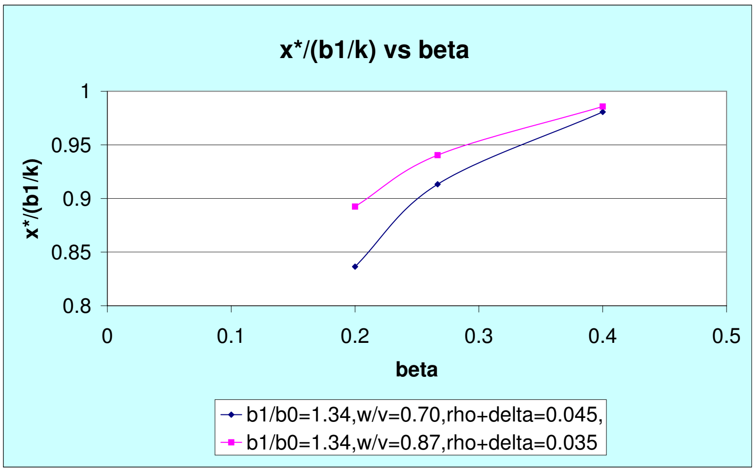

Figure 1.

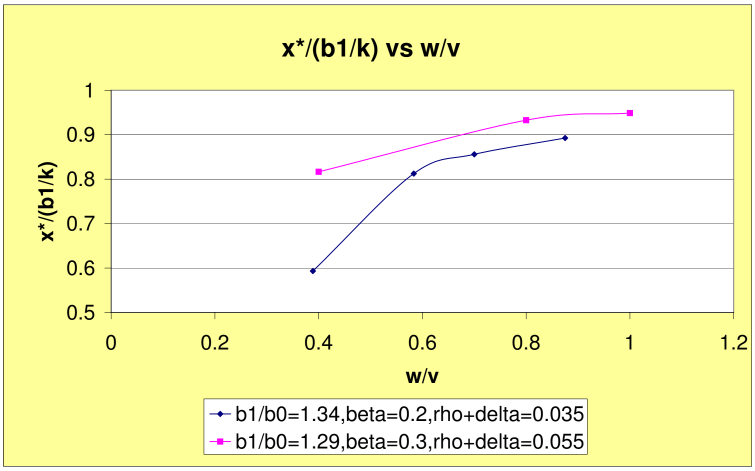

Figure 2.

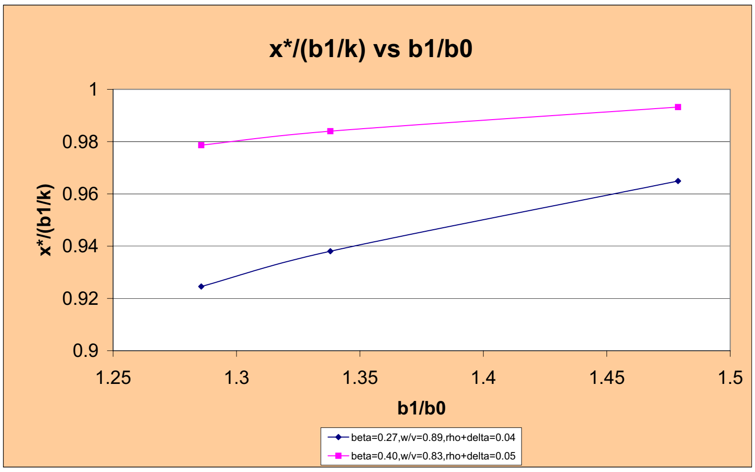

Figure 3.

The observed results can be explained. Due to 4a, 4b, 4c and 5a, 5b, 5c, it is clear that, everything else being equal, by increasing either β or w/v or b_1/b_0 one passes from RP solution to NR solution and the ratio $\frac{x^*}{(b_1/k)}$, that determines the width of the continuation region, increases.

This shows that, in general, the continuation region is larger with NR solution than with RP solution.

The intuition behind this is that it is optimal for longer time to trade your own wealth if you choose and/or are given a set of parameters that lead to NR solution than if you are in a RP solution case. On the other hand, if you choose and/or are given a set of parameters that lead your problem to a RP solution, then you are likely to annuitize earlier than if you are in a NR solution case. This is consistent also with the fact that annuitization occurs in RP solution also in the case of ruin, whereas it does not with a NR solution.

The dependence of the type of solution from the parameters is now easy to understand and explain. In fact, if β is high, the risky asset is good compared to the riskless one, and in this situation it is reasonable to delay annuitization as much as possible (this result was also found in Gerrard et al. (2004a)). If w/v is high, the penalty to be paid in case of annuitization before reaching b_1/k is high compared to that associated to the choice investment-and-consumption, which is then preferable. If b_1/b_0 is high, the retiree has a low risk aversion, thus will be likely to take chances in the financial market instead of locking her position into an annuity. Furthermore, higher values of $\rho + \delta$ are associated to old retirees, who have higher force of mortality and higher subjective discount factor (as they are less patient for future events), and it is reasonable to expect them to be more willing to annuitize rather than continuing investing in the market.

5.2 Simulations

In this application[2] we consider the position of a male retiree aged 60, who retires with initial fund $x_0 = 1000$. We have selected the following values of the parameters:

$$r = 0.04,\ \lambda = 0.08,\ \sigma = 0.1,\ \rho + \delta = 0.045,\ w = v = 0.04,\ b_0 = 69.95,\ b_1 = 120,\ k = 0.095$$

This implies

$$\beta = 0.4 \qquad \frac{w}{v} = 1, \qquad \frac{b_1}{b_0} = 1.72, \qquad \frac{b_1}{k} = 1263.16,$$

In turn, the solution (NR) is

$$x^* = 1257.14, \qquad \frac{x^*}{b_1/k} = 0.995$$

We have simulated the behaviour of the risky asset with Monte Carlo simulations in 1000 scenarios, and in each scenario we have adopted the optimal investment and consumption strategies until the minimum between time of annuitization and 15 years. The choice of a terminal time of the optimization program is consistent with current regulation in UK, whereby annuitization becomes compulsory at age 75.

An interesting result is that the probability of annuitization within 15 years from retirement is 74.50% and on average optimal annuitization occurs after 6.05 years after retirement. The mean size of annuity is 78.18 and in almost 50% of the cases the annuity value lies between 90 and 110. More detailed information can be gathered from the histograms of Figures 4 and 5 that report, respectively, the distribution of time of optimal annuitization (measured in years from retirement) and the distribution of size of final annuity. In both Figures, the first column labelled by "NA" (No Annuitization) report the cases in which optimal annuitization does not occur within the time frame (255 cases out of 1000).

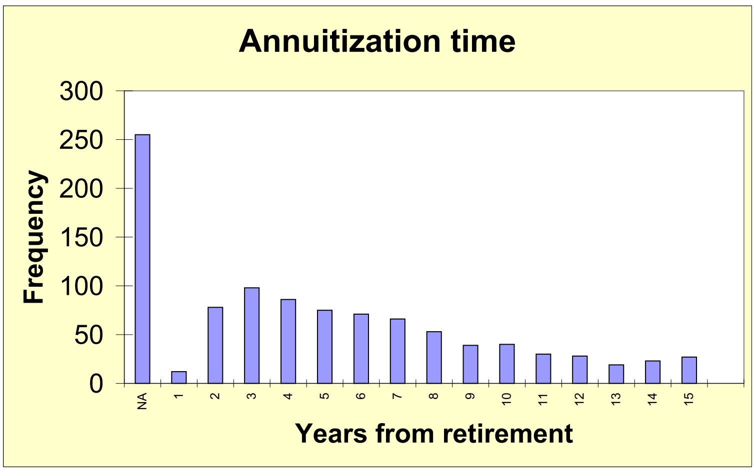

Figure 4.

[2]We thank Marina Di Giacinto for spotting a forgotten term in the simulation program, that improved accurateness of results.

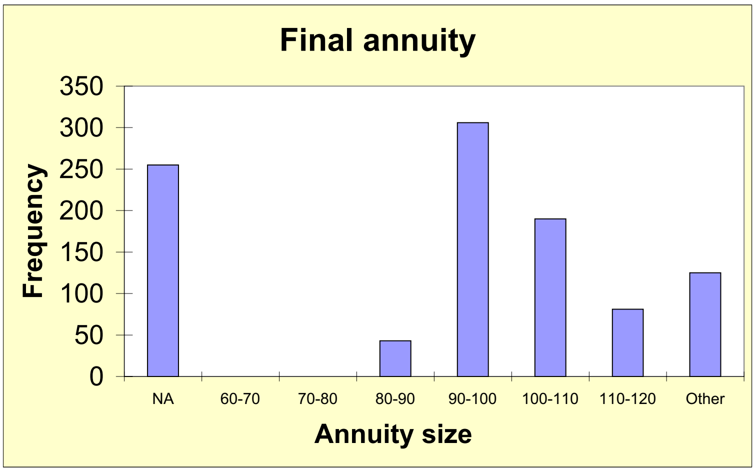

Figure 5.

Figure 6 reports some statistics (mean, standard deviation, 5^{th} and 95^{th} percentiles) of the optimal consumption in the 15 years after retirement. The interim target consumption, b_0 is reported for comparison. We notice from (3.20) that optimal consumption cannot exceed b_0. However, in Figure 6 we see that on average optimal consumption is higher than b_0. This is due to the fact that here in the 886 scenarios in which optimal annuitization occurs before age 75, the consumption reported after annuitization time is the annuity value, which is always higher than b_0. This highlights the financial convenience for the retiree of deferment of annuitization until a more propitious time, in this particular example. The event of negative consumption occurs in 7 cases out of 1000 simulations: in those unfavourable cases the fund falls below the undesirable level of fund below which optimal consumption is negative (here equal to $X(z_{neg}) = 69.5$). It is worth mentioning that in 2 cases out of 1000 the fund falls very close to 0, but then, due to null investment in the risky asset and negative consumption, goes up again. Both events of negative consumption and almost-ruin happen on average at age 73 and always between ages 70 and 75, that is not surprising.

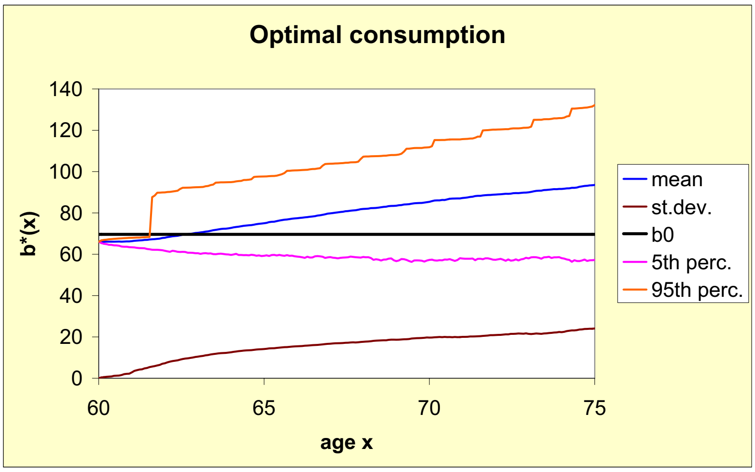

Figure 6.

Similarly, Figure 7 reports some statistics (mean, standard deviation, 5^{th} and 95^{th} percentiles) of the optimal fraction of portfolio to be invested in the risky asset over time. We notice that although the control is not constrained between 0 and 1, the optimal $y^*(t)$ is never negative. This interesting and desirable feature comes directly from the form of the optimal control (3.19), observing that the fund is by construction always non-negative and the function $X(z)$ is decreasing. Let us notice that this characteristic was also present in the model by Gerrard et al. (2006). Furthermore, on average $y^*(t)$ is decreasing from one to zero, and deviations above one, though undesirable, are not huge.

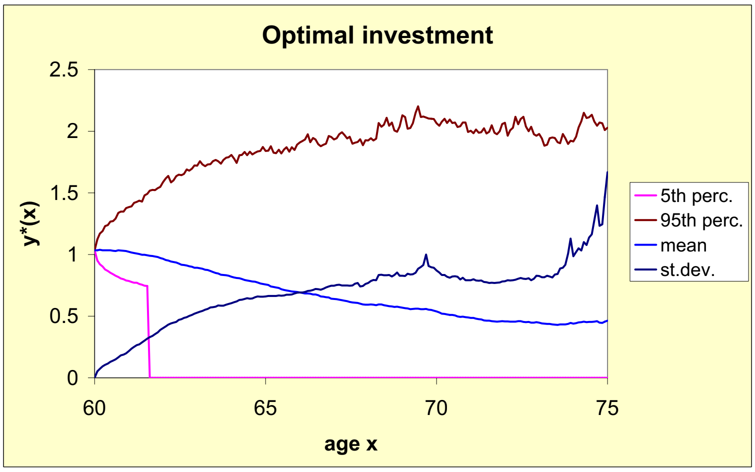

Figure 7.

Comparison with a model without optimal annuitization time.

It is now our aim to try to compare the results obtained in our model with a model that allows for compulsory annuitization at terminal date without possibility of earlier annuitization. We acknowledge that the comparison is very hard if not impossible to make because two different models are considered. However, we think that an attempt to compare different choice models available to the retiree can be useful to help the member in the decision of what model should be adopted. The comparison is done with the model introduced by Gerrard et al. (2006), because this is the most similar to the one presented here and the comparison allows us to isolate and measure the effect of adopting optimal annuitization rules. In particular, by choosing, in the mentioned paper, $u = 0$ we have the same loss function. All the other parameters either have equal meaning or play similar roles in the definition of the model. The main difference is, clearly, the absence in the mentioned paper of an optimal exit from the optimization program, which is run until terminal time T, when the fund is annuitized.

From now onwards, we will call the model of this paper "model A", and the model of the paper without optimal annuitization time "model B". In order to compare the results in a consistent way, we have run the same simulations for the risky asset used above and in each of the 1000 scenarios we have applied the optimal investment and consumption rules indicated by the mentioned paper. Annuitization occurs at age 75, and the values of the parameters have been chosen all equal, apart from the value of k, here chosen equal to 0.11 (the value of k has to be different from the one chosen before for this formulation of the problem to make sense). As already noted in Gerrard et al. (2006), in model B the final target b_1 is approached in a very satisfactory way after 15 years. This is highlighted by the distribution of the final annuity in this case, reported in Figure 8. In 76% of the cases the final annuity amounts between 115 and 120, in 11% of the cases between 110 and 115.

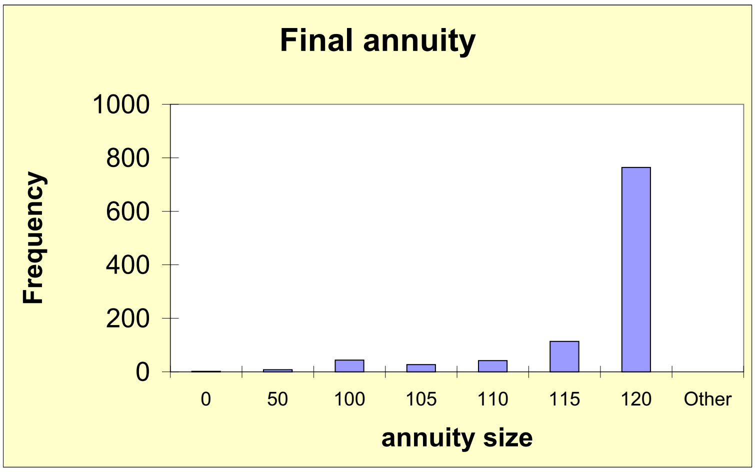

Figure 8.

In fact, in most of the cases (82.70%) the ultimate annuity received by the retiree is higher in model B than in model A. This is due to the fact that in the model without optimal annuitization, the fund approaches the target b_1/k very closely, whereas in model A annuitization occurs whenever the fund reaches x^*, that is lower than b_1/k. However, the fact that optimal annuitization occurs before T implies that optimal consumption before T is generally higher with model A than with model B. Therefore, the comparison has to be done between different paths of consumption, ideally from retirement up to time of death. There are many possible ways to make a comparison of different streams of money at different times, and a thorough discussion about appropriateness of different methods is beyond the scope of this paper. Here, with illustrative purpose only, we choose the criterium of the expected present value (EPV) of the streams. In particular, we discount flows from retirement to T with the rate $\rho+\delta$ and we then add the expected present value of the actuarial value of the annuity achieved from T until death. In Table 1 we separate the expected present values of the consumption streams from 60 to 75 and from 75 to death, in order to show the different effect of the two different periods on the total expected present value (reported in the last two columns).

	A: EPV cons. ages 60-75	B: EPV cons. ages 60-75	A: EPV cons. age 75-death	B: EPV cons. age 75-death	A: EPV cons. age 60-death	B: EPV cons. age 60-death
min	18377	35556	1	-539	18427	35468
5th perc.	34370	38569	630	826	34702	38987
25th perc.	37747	39110	776	961	38256	39598
50th perc.	47584	39300	843	985	48030	39799
75th perc.	49148	39391	963	995	49544	39896
95th perc.	49880	39463	1151	999	50281	39970
max	51166	39533	1272	1000	51558	40042
mean	44441	39194	866	955	44882	39681
st.dev.	5879	339	167	111	5874	382

Table 1.

According to Table 1 and to the criterium of EPV, model A seems to perform better than model B. In fact, the generally higher values of the EPV of consumption streams from retirement to death in A-model w.r.t. B-model indicate that the generally higher income received in model B from 75 to death does not compensate the reduced income in years between optimal annuitization and 75. However, the dramatically lower standard deviation of the EPV in model B, indicating a much higher stability of final outcomes, could be more appealing to a risk averse retiree and push her to choose model B.

As noted above, here there are only some guidelines for possible comparisons between different decision making models to be adopted after retirement, and a full discussion of pro and contra of these models is left to further research. Beyond the scope of this paper, but certainly interesting and left to future research is the comparison of our model with one giving optimal annuitization rules driven by optimization criteria different from ours (such as e.g. in Milevsky et al. (2006), where the criterium is the minimization of the probability of financial ruin, or Milevsky and Young (2007), where the agent maximizes expected utility of lifetime consumption and bequest).

6 Conclusions and further research

In this paper we have considered the problem faced by a retiree of a defined contribution pension scheme who defers annuitization of the fund and has to decide about investment allocation, consumption strategy and time of annuitization.

The problem is naturally formulated as a combined stochastic control and optimal stopping problem. The optimization criterion consists of a quadratic running cost penalizing deviations of interim consumption from a target and a quadratic final cost penalizing the deviation of the annuity size achieved from a certain desired level of annuity. We tackle the delicate issue of ruin by imposing the constraint that the optimization program stops whenever the fund becomes negative. This implies that, depending on the values of the parameters of the model, the problem either has no solution or has a solution that can be of two different types. By construction we find closed form solutions to the HJB equation which, by means of the verification technique, is shown to satisfy the optimization problem. The construction leads to an algorithm that is applied for numerical investigations of the solution.

The numerical applications presented are twofold. Firstly, we investigate the dependence of the type of solution and of the width of the continuation region on the values of the parameters of the model. In particular, we find that the key values in determining type of solution and width of continuation region are the Sharpe ratio of the risky asset, the importance given by the retiree to the loss associated to running consumption relative to that associated to final cost, the ratio between desired annuity size and that purchasable at retirement (i.e. the risk attitude of the retiree). This investigation shows the reasonable result that, ceteris paribus, it is optimal to defer annuitization for longer time if either the Sharpe ratio is high, or the penalty paid in case of annuitization is high with respect to that paid for low running consumption, or the pensioner has a low risk aversion.

Secondly, we select a particular scenario for market and demographic conditions and risk profile of the retiree, find the solution and simulate the behaviour of the risky asset via Monte Carlo method. Simulation results indicate that in the particular scenario chosen optimal annuitization occurs within 15 years from retirement in most (75%) of the cases and on average should occur a few years after retirement. Furthermore, the event of ruin never occurs and optimal consumption is negative in less than 1% of the cases. A few guidelines are given for possible comparison of results with a model without optimal annuitization and, based on the criterium of expected present value of consumption streams from retirement to death, we find that a model with optimal annuitization should be preferred to one with fixed annuitization time.

We believe that this paper leaves scope for further research in many directions, both on the applicative side and on the theoretical one.
Due to space constraints, we have not carried out analysis of robustness for the numerical investigations. A greater variety of scenarios for the market and demographic assumptions and for the decision maker's risk profile would certainly add more useful insight for practical applications of the model, as well as an accurate comparison with different models of optimal annuitization time.
The strong assumption of constant force of mortality could be relaxed and stochastic mortality might be introduced in the model. Finally, we have considered unconstrained controls. Nevertheless, our optimal investment in the risky asset and optimal consumption are naturally constrained to be greater than 0 and lower than the targeted consumption, respectively. The addition of bilateral restrictions on the investment allocation is subject of ongoing research (see Di Giacinto, Federico, Gozzi and Vigna (2009b)) and the introduction of further restriction on the consumption is in the agenda for future research.

References

Albrecht, P. and Maurer, R. (2002). Self-annuitization, consumption shortfall in retirement and asset allocation: The annuity benchmark, *Journal of Pension Economics and Finance* **1**: 269–288.

Battocchio, P. and Menoncin, F. (2004). Optimal pension management in a stochastic framework, *Insurance: Mathematics and Economics* **34**: 79–95.

Blake, D., Cairns, A. J. G. and Dowd, K. (2001). Pensionmetrics: Stochastic pension plan design and value-at-risk during the accumulation phase, *Insurance: Mathematics and Economics* **29**: 187–215.

Blake, D., Cairns, A. J. G. and Dowd, K. (2003). Pensionmetrics 2: Stochastic pension plan design during the distribution phase, *Insurance: Mathematics and Economics* **33**: 29–47.

Booth, P. and Yakoubov, Y. (2000). Investment policy for defined contribution pension scheme members close to retirement: an analysis of the 'lifestyle' concept, *North American Actuarial Journal* **4**: 1–19.

Boulier, J.-F., Huang, S. and Taillard, G. (2001). Optimal management under stochastic interest rates: the case of a protected defined contribution pension fund, *Insurance: Mathematics and Economics* **28**: 173–189.

Devolder, P., Bosch Princep, M. and Dominguez Fabian, I. (2003). Stochastic optimal control of annuity contracts, *Insurance: Mathematics and Economics* **33**: 227–238.

Di Giacinto, M., Federico, S. and Gozzi, F. (2009a). Pension funds with a minimum guarantee: a stochastic control approach, *to appear in Finance and Stochastic.*

Di Giacinto, M., Federico, S., Gozzi, F. and Vigna, E. (2009b). Constrained portfolio choices in the decumulation phase of a pension plan, *mimeo.*

Gao, J. (2008). Stochastic optimal control of dc pension funds, *Insurance: Mathematics and Economics* **42**: 1159–1164.

Gerrard, R., Haberman, S. and Vigna, E. (2004a). Optimal investment choices post retirement in a defined contribution pension scheme, *Insurance: Mathematics and Economics* **35**: 321–342.

Gerrard, R., Haberman, S. and Vigna, E. (2006). The management of de-cumulation risks in a defined contribution pension scheme, *North American Actuarial Journal* **10 (1)**: 84–110.

Gerrard, R., Haberman, S., Højgaard, B. and Vigna, E. (2004b). The income drawdown: quadratic loss, *Actuarial Research Paper No. 155, Cass Business School, London.*

Haberman, S. and Vigna, E. (2002). Optimal investment strategies and risk measures in defined contribution pension schemes, *Insurance: Mathematics and Economics* **31**: 35–69.

Højgaard, B. and Vigna, E. (2007). Mean-variance portfolio selection and efficient frontier for defined contribution pension schemes, *Technical report R-2007-13. Department of Mathematical Sciences. Aalborg University.*

Kapur, S. and Orszag, J. M. (1999). A portfolio approach to investment and annuitization during retirement, *Proceedings of the Third International Congress on Insurance: Mathematics and Economics, London.*

Karatzas, I., Lehoczky, J. P., Sethi, S. P. and Shreve, S. E. (1986). Explicit solution of a general consumption/investment problem, *Mathematics of Operations Research* **11**: 261–294.

Khorasanee, M. Z. (1996). Annuity choices for pensioners, *Journal of Actuarial Practice* **4**: 229–255.

Khorasanee, M. Z. (1998). Deterministic modelling of defined contribution pension funds, *North American Actuarial Journal* **1**: 83–103.

Knox, D. M. (1993). A critique of defined contribution plans using a simulation approach, *Journal of Actuarial Practice* **1**: 49–66.

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

Milevsky, M. (2001). Optimal annuitization policies: Analysis of the options, *North American Actuarial Journal* **5**: 57–69.

Milevsky, M. A. and Young, V. R. (2002). Optimal asset allocation and the real option to delay annuitization: It's not now-or-never, working paper.

Milevsky, M. A. and Young, V. R. (2007). Annuitization and asset allocation, *Journal of Economic Dynamics and Control* **31**: 3138–3177.

Milevsky, M. A., Moore, K. S. and Young, V. R. (2006). Asset allocation and annuity-purchase strategies to minimize the probability of financial ruin, *Mathematical Finance* **16**: 647–671.

Øksendal, B. (1998). *Stochastic Differential Equations*, Springer Verlag, Berlin.

Peskir, G. and Shiryaev, A. N. (2006). *Optimal stopping and free-boundary problems*, Springer.

Shiryaev, A. N. (2008). *Optimal Stopping Rules*, Stochastic Modelling and Applied Probability Vol. 8, Springer Verlag, Berlin.

Stabile, G. (2006). Optimal timing of the annuity purchase: combined stochastic control and optimal stopping problem, *International Journal of Theoretical and Applied Finance* **9(2)**: 151–170.

Vigna, E. (2009). Mean-variance inefficiency of CRRA and CARA utility functions for portfolio selection in defined contribution pension schemes, *Carlo Alberto Notebook 108/09 and CeRP WP N. 89/09.*

Xu, G.-L. and Shreve, S. E. (1992). A duality method for optimal consumption and investment under short-selling prohibition. I. General market coefficients, *The Annals of Applied Probability* **2**: 87–112.