Annuities and Aggregate Mortality Risk*

Justin van de Ven and Martin Weale
National Institute of Economic and Social Research,
2, Dean Trench Street,
London SW1P 3HE
Email: mweale@niesr.ac.uk

31st December 2007

Abstract

This paper explores the effect of aggregate mortality risk on the pricing of annuities. It uses a two-period OLG model; in the first period, 'young' people have a zero probability of death, and in the second period 'old' people face an initially unknown risk of death. Old people can either carry their aggregate mortality risk, or buy annuities which are sold by young people. A market-clearing price for such annuities is established. The alternative where annuities are purchased from the government is also explored, and this is found to dominate the private market solution in welfare terms.

KEYWORDS: Annuity, asset-pricing, mortality risk, overlapping generations
JEL Code: C610, C680, D530, D580, D910, J100, J180

---

[*] This work has been supported by the ESRC under its World Economy and Finance Programme.

1 Introduction

Many of the issues associated with annuity markets have been well-researched in recent years, as reported in the comprehensive survey by Davidoff et al. (2005). Much of this research has focussed upon the merits of conventional annuities, with the general finding that there is likely to be a substantial role for annuitisation even in context of heterogeneous preferences for investment, the time stream of consumption, and bequests.

In contrast, very little work has been conducted to understand the workings of conventional annuities markets with regard to aggregate mortality risk.[1] The topic of aggregate mortality risk was not discussed in the survey of related literature by Davidoff et al. (2005). While Bohn (2005) does air the question of uncertain longevity, he limits himself to scenario analysis – the effect of a permanent increase in longevity during retirement – and does not address the question of risk in its more conventional sense. The problem is that the true mortality rate of any particular cohort is known only *ex post*. When an insurance company sells a conventional annuity it contracts to pay a regular income until the death of the annuitant, calculated on the basis of a forecast of the respective cohort's mortality rate. Like any forecast, projections of mortality rates are subject to error, and an insurance company selling conventional annuities therefore carries aggregate mortality risk. Casual discussion in the United Kingdom has suggested that this risk may be substantial. Banks & Blundell (2005) make the point that "unanticipated increases in longevity put pressure on all forms of pension systems" although they do not offer any view on the scale of this pressure.[2] Hardy (2005) states that "Even actuaries recognize that longer life is a good thing – but, to the extent that it is unanticipated, it is also an enormous problem for the managers of annuity portfolios", with the implication that an annuitant should expect to pay a substantial risk premium. On the other hand, studies such as Finkelstein & Poterba (2002) argue that, after allowing for adverse selection, annuity pricing is close to actuarially fair before making any allowance for a risk premium. This last observation implies that either the risk premium cannot be very

---

[1] To clarify terms, *aggregate mortality risk* refers to the uncertainty that is associated with the mortality rate of a cohort that is comprised of sufficient members to ignore small-sample effects. Individual mortality risk, by contrast, is the uncertainty of the precise time of death of any individual. A *conventional annuity* pays a guaranteed level income until time of death of the annuitant, thereby insuring against both forms of mortality risk defined here.

[2] In the United Kingdom there is a separate problem that mortality rates have been systematically over-predicted by the Government Actuary. Here, however, we focus on shocks relative to unbiased forecasts.

large, or that – once it is taken into account – conventional annuities are substantially mis-priced.

One clue to the pricing of mortality risk might be offered by financial markets. In November 2004 the European Investment Bank (EIB) issued a mortality bond following an earlier issue in December 2003 by the insurance company Swiss Re. A mortality bond is a loan stock whose payout depends on the mortality rate of a specific cohort and in the case of the EIB issue, the mortality of the cohort of men in the United Kingdom aged 65 in 2003. The pay-out received by purchasers of the stock increases proportionately to the longevity of the defined cohort, thus providing a hedge against aggregate mortality risk. Friedberg & Webb (2006) explore the pricing of the EIB mortality bond in the context of the consumption capital asset pricing model (CCAPM). Despite their observation that conventional annuities expose insurance companies to "substantial risk", they find on the basis of their model that the cost of hedging aggregate mortality risk should be very low.

It is, however, questionable whether the conventional CCAPM is the most appropriate tool for analysing the pricing of annuities. The assumption of an infinitely lived representative consumer by the standard CCAPM sits uncomfortably alongside the uncertain mortality that influences demand for conventional annuities in practice.[3] Annuities represent a transaction, where one group of consumers (the elderly) divest themselves of their aggregate mortality risk by purchasing insurance from another group (the young).

In this paper we consequently explore the pricing of aggregate mortality risk in an overlapping-generations model. In our model, young people have no risk of death while young but, beyond a threshold age when they become old, they face a constant risk of death. Old people can buy conventional annuities from young people (or from insurance companies whose shares are owned by young people) at a price which balances the willingness of the young to carry aggregate mortality risk with the desire of the old to divest themselves of it. The old and the young have different attitudes to the mortality risk of the old because they are of different ages; young people can adjust their consumption while still young in response to the gains or losses that they experience from carrying the mortality risk of the old. The fact that different generations are affected differently by shocks to the mortality rate of any particular cohort makes transactions in

---

[3] The model could be extended to address aggregate mortality risk using the approach suggested by Yaari (1965) in which all consumers have the same mortality rate independent of their age and where, therefore, the notion of a representative consumer can be retained, as Blanchard (1985) does in his analysis of fiscal policy with finite horizons.

mortality risk possible. The model is therefore a member of the family of heterogeneous agent CCAPMs. Other members of this family have been put to a variety of uses – see, for example, Constantinides et al. (2002) who use a heterogeneous agent model to explore the equity risk premium.

An important feature of our analysis is the assumption that *mortality-adjusted annuities* can be purchased as a substitute for conventional annuities. A mortality-adjusted annuity is considered to pay an annual dividend based upon the realised mortality rate of a cohort.[4] An individual who purchases a mortality adjusted annuity with a dividend determined by their own cohort's mortality rate is consequently insured against individual specific mortality risk, but remains exposed to aggregate mortality risk, to the extent that their cohort's mortality rate is uncertain. This differs from a conventional annuity, which pays a guaranteed income to annuitants until death, and consequently insures against both forms of mortality risk. Plainly, if the cost imposed by sellers of conventional annuities for carrying aggregate mortality risk is substantial, then, given the choice, annuitants would prefer to carry some of their own aggregate mortality risk. This generates demand for mortality-adjusted annuities in the current study.[5] Circumstances in which mortality-adjusted annuities are not available then appear as a special case of our results.

The next section discusses the difficulties associated with forecasting contemporary mortality rates, and describes the process that is assumed to govern mortality rate variation for the analysis presented here. We then set out a modelling framework to assess the impact of aggregate mortality risk on the pricing of conventional annuities, and use this to explore the risk margins which we might expect to observe. We also show how the framework can be used to explore the consequences of relying on the government instead of the private market to insure against aggregate mortality risk. In such a situation the government announces the terms on which it is prepared to sell conventional annuities to households; it also makes a lump sum payment or collects a lump sum tax to ensure that its budget is balanced inter-temporally. A summary is

---

[4] Conceptually a mortality-adjusted annuity is not very different from a tontine. Historically, however, tontines have operated by pooling the contributions of the participants with the last survivor receiving the whole amount and the other participants receiving nothing.

[5] We are not aware of any contemporary sellers of mortality-adjusted annuities in the United Kingdom but it is likely that there would be benefits to filling this gap in financial markets; certainly, if the premium associated with aggregate mortality risk is large, then one might expect that annuitants would value the opportunity to carry at least some of the aggregate mortality risk themselves. Annuities are available which invest the funds in the stock market and which therefore pay out on a with profits basis; typically with these the annuitants do carry some aggregate mortality risk.

provided in a concluding section.

2 Forecasting Mortality Rates

2.1 Statistical Difficulties in Forecasting Mortality Rates in the UK

In the United Kingdom, recent experience suggests that mortality projections systematically have over-stated the practical reality with forecasts of life expectancy being too low in consequence. As pointed out by the Pensions Commission, for example, “in the early 1980s public pension policy and private pension provision decisions were based on the assumption that average male life expectancy at 65 in 2010 would be 15.1 years: the best estimate is now 20.1 years (Pensions Commission 2005, p. 90). The consequent under-prediction of life expectancies has contributed to the sizeable pension fund deficits that have been identified during the last decade. A time-series statistician might think that there are actually two separate issues present here. On the one hand is the uncertainty that underlies any forecast. On the other, the systematic over-prediction of mortality rates in recent years may be taken as indicative of bias, which could be eliminated by adopting a more appropriate forecasting model. Unfortunately, the practical reality is less clear.

A standard and fairly simple framework for forecasting mortality is provided by Lee & Carter (1992). They assume that the change in mortality rates at different ages can be largely accounted for by the movement in a single principal component of the historic mortality data. They recommend forecasting using the assumption that this principle component follows a random walk with drift.

In a single factor model of mortality – as in any reduced form regression – the uncertainty surrounding mortality projections depends on the uncertainty of the model parameters and the uncertainty in the error term. Lee & Carter (1992) show that, for long-term projections of the type which concern us here, it is the uncertainty in the error term which dominates. Lee & Miller (2002) assess the performance of the Lee-Carter model out of sample for the Untied States, looking at recursive forecasts for period (not cohort) life expectancies at birth for men and women combined produced from 1920 to 1998 and making the assumption that the mortality factor evolves as a random walk with drift. In the 1920s, when life expectancy was between 55 and 60 years, the median forecasts for mortality in 1998 were between 70 to 75 years, as compared to an out-turn of 76.7.

If contemporary trends in mortality rates were to conform to the stylised observations that are reported above, then there would appear to be strong evidence in support of a single factor forecasting model, like the one suggested by Lee and Carter. Unfortunately, however, the relationship between mortality rates and time appears to have suffered a structural break – or at least to be changing over time. Mortality rates are currently falling much faster for people aged 65-85 than would have been expected either from the single factor model or from other analysis of mortality rates. This pattern has led some authors (see, for example, Willets (2004)) to suggest that the much lower mortality rates of people currently in their seventies and early eighties will spread into older age groups as the survivors of this cohort age (the cohort effect).

In forecasting mortality rates actuaries face the problem of deciding how long recently observed trends will last – and the historical experience cannot be said to offer much of a guide to this. If one looks further back – to 1840 – the experience of the last twenty years stands out very markedly. Not surprisingly views on this issue are divided, with some authors arguing that there is an upper limit to life expectancy (Wilmoth & Horiuchi 1999),[6] while others point out that there is no sign of a life expectancy limit having been reached, and plenty of evidence of assumed limits being broken (Oeppen & Vaupel 2002).

The Pensions Commission (2005) explored the avaliable evidence on aggregate mortality uncertainty in the United Kingdom with the aim of coming to a balanced view. It suggested that, for men aged 65 in 2004 life expectancy lay in the range 17.6-19.9 years, while for a man aged 19 in 2004 it lay in the range 20-29 years. These ranges for uncertainty, the Commission suggests, could be regarded as 90% confidence intervals. They form the backdrop to this paper.

3 Mortality Risk and the Decision Framework

We consider the savings and investment decisions implied by a rational agent life-cycle model in the context of aggregate mortality risk. In our model, heterogeneity is age and time specific, so that the behaviour of any given birth cohort can be explored with

---

[6] An upper limit to life expectancy does not imply an upper limit to human life. If the mortality rate for people aged 100+ settles at some figure below one, then an infinitely small number of people will live infinitely long. With a rate of 0.5, only one in a thousand centenarians would reach one hundred and ten, so it is not surprising that the statistical evidence is unclear.

reference to a representative agent from that cohort. The discussion that follows consequently makes explicit reference to the behaviour of cohorts, rather than of individuals. Each cohort i, at age t, has an instantaneous utility function described by:

$$u\left(c_i^t\right) = A + \frac{\left(c_i^t\right)^{1-\alpha}}{1-\alpha} \tag{1}$$

where A represents the joy of living, and can be set to ensure that individuals prefer strictly longer lifetimes; $u\left(c_i^t\right) > 0$ for all plausible values of instantaneous consumption, c_i^t.[7] This utility function forms the basis for the recursive solution of our model.

Our analysis focuses upon transactions relating to aggregate mortality risk in two decision making contexts; one in which overlapping generations are considered to trade with each other; and another in which each generation is considered to trade with the government. These two contexts are treated separately below.

As noted above, a key to our analysis is to represent a situation in which there might be both sellers and buyers of mortality risk. We do this with a stylised overlapping generations model that divides life into two periods. Economically active life begins at age twenty-five, and the first period lasts for forty years. In this phase the mortality rate of each cohort is assumed to be zero. For those aged sixty-five or above, in the second period of life, the mortality rate is positive and does not vary with age; life expectancy at age sixty-five is therefore given by the reciprocal of the mortality rate once that is known. In subsequent sections it is convenient to refer to the start of the first period of life as age 0 and the start of the second period of life as age 1. Before explaining the detailed structure of our model we describe how it allows us to represent shocks to aggregate mortality.

3.1 A Model of Mortality Risk

In the absence of a consensus model for forecasting mortality rates, the current study is simplified by assuming a random walk process. Specifically, we assume that the mortality rate, ρ_i of cohort i in the second period of its life, is log-normally distributed. Thus:

$$\log \rho_i = \frac{\gamma_1 + \nu_i}{\zeta}$$

where ν_i is given by:

$$\nu_i = \nu_{i-1} + \upsilon_i \quad \text{where } \upsilon_i \sim N(\gamma_2, 1), \quad Cov(\upsilon_i, \upsilon_j) = 0 \text{ for all } i \neq j$$

---

[7] This condition must be met if consumers are to regard long life as a good thing rather than a burden. Note, however, that A does not influence utility maximising decisions.

This gives the following auto-regressive structure for the mortality rate ρ_i:

$$\log \rho_i = \log \rho_{i-1} + \frac{\upsilon_{i-1}}{\zeta} \tag{2}$$

Setting $\gamma_2 = -1/2\zeta$ ensures that $E\left(\rho_i \mid \rho_{i-1}\right) = \rho_{i-1}$, removing drift. The assumption of log-normality is common in the related literature, and implies that values of ρ_i are positive. This structure provides a formal framework in which the mortality rate of cohort i is not known in advance and any assumptions made about this mortality rate are therefore subject to aggregate mortality risk.

We solve the model for a succession of periods and, noting that each period represents forty years, consider a value of γ_1 so that, in the initial period, the expected mortality rate is set to two (corresponding to life-expectancy of 20 years):

$$E\left\{e^{\frac{\gamma_1+\nu_i}{\zeta}}\right\} = 2$$

We assume a standard deviation of 0.2, roughly equivalent to a standard deviation in life expectancy of two years so that

$$Var\left\{e^{\frac{\gamma_1+\nu_{i-1}}{\zeta}}\right\} = 0.04$$

These conditions are met if $\gamma_1 = 6.89887$, $\gamma_2 = -0.0499$ and $\zeta = 10.0249$[8]. These parameter values imply that the standard deviation of the mortality rate of young men at age twenty-five is 0.285, giving a standard deviation of life expectancy of about 2.8 years. If one assumes that the preferred Pensions Commission figures represent 3.3 standard deviations (i.e. are taken from a normal distribution) then the Commission's estimates imply a standard deviation in life expectancy of 2.7 years for a man aged 19 and 0.7 years for a man aged 65. Our statistics are consequently close to those projected for young men by the Commission, and overstate the uncertainty projected for old men.

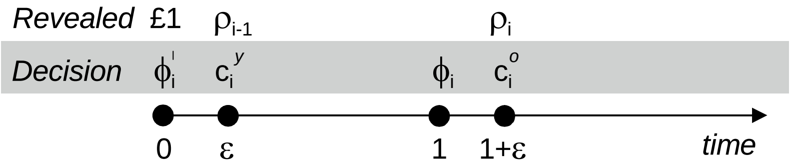
3.2 A Private Market in Annuities

We initially explore market interactions between two cohorts, which are referred to as young and old in the discussion that follows. A cohort begins its economic life at age 0, and is subject to no mortality until age 1 when it becomes old – as noted in section 3.1, one period can be thought of as representing 40 years. Each cohort is subject to a constant cohort-specific mortality rate from age 1. Importantly, the mortality rate of a

---

[8] Assuming that $\nu_{i-1} = 0$ for the first cohort.

cohort is uncertain at the time when investment decisions for old age need to be made. An old cohort is able to insure against the risk represented by the uncertainty of its own mortality rate by purchasing conventional annuities from the co-surviving young cohort. But the young cohort will only be willing to supply annuities at a price that adequately compensates it for the risk that the sale of annuities represents. Mutually beneficial transactions of aggregate mortality risk between the young and the old are made possible by the timing that we assume for the intertemporal decision making problem.



Notes: ϕ_i' denotes value of annuities sold, expressed as a proportion of wealth when young
ϕ_i denotes value of annuities bought, expressed as a proportion of wealth when old
c_i^y denotes instantaneous consumption when young
c_i^o denotes instantaneous consumption when old
ρ_i denotes realised mortality rate of cohort i

Figure 1: Decision Making Time-Line Assumed for Analysis

We assume that cohort i faces four decision making points when conventional annuities are traded in the private market, which are displayed graphically in Figure 1. Each cohort is considered to be endowed with a unit of wealth at age 0. At age 0, the cohort must decide what value of conventional annuities it is willing to sell to the current old cohort, expressed as a proportion of its endowment ϕ_i'. This decision is taken in context of the prevailing annuity price, and the uncertainty surrounding the old cohort's mortality rate. At age ε, the old cohort's mortality rate is revealed, ρ_{i-1}, and the young cohort must choose its consumption rate between ages ε and 1, c_i^y, in light of the capital gains or losses that it has realised on its sale of conventional annuities. At age 1, the cohort becomes old and it must choose the proportion of its remaining wealth to invest in conventional annuities, ϕ_i. Any wealth that is not used to purchase a conventional annuity is used to purchase a mortality-adjusted annuity, which pays an actuarially fair level income based upon the cohort's realised mortality rate: for the reasons given by Davidoff et al. (2005) this clearly dominates holding simple interest-bearing assets. This investment decision is made, given the prevailing annuity price and the uncertainty surrounding the cohort's own mortality rate. At age $1+\varepsilon$, the cohort's own mortality rate is revealed, thereby determining its income in old age. The cohort must then choose an instantaneous consumption rate for the remainder of its life, c_i^o.

We use backward induction to solve for decisions that maximise expected lifetime utility at each of the decision points that are described above, in common with the dynamic programming literature. Furthermore, as the description provided above implies, we assume that ε is so small that we can ignore the consumption that takes place over the intervals $[0,\varepsilon)$ and $[1,1+\varepsilon)$.[9] We now describe these utility maximising problems at greater length.

1. *Age* $1+\varepsilon$. The cohort is subject to no further uncertainty after its mortality rate is revealed. The decision problem at age $1+\varepsilon$ is consequently described by:

$$\begin{aligned} V_{i,1+\varepsilon} = V_{1+\varepsilon}\left(y_i^o\right) &= \max_{c_i^t} \int_{1+\varepsilon}^{\infty} u\left(c_i^t\right) e^{-(r+\rho_i)(t-(1+\varepsilon))} dt \\ \text{subject to} &: \int_{1+\varepsilon}^{M} \left\{y_i^o - c_i^t\right\} e^{-(r+\rho_i)(t-(1+\varepsilon))} dt \geq 0 \qquad (3) \\ \text{for all } M &\in [1+\varepsilon,\infty) \end{aligned}$$

   where r is the interest rate, assumed to be equal to the discount rate, and y_i^o is the level income that is provided by the combination of conventional and mortality-adjusted annuities that the cohort chose to invest in at age 1. Given the assumptions that the interest rate is equal to the discount rate, and that the mortality rate from age 1 is constant, the solution to (3) will involve setting $c_i^t = c_i^o = y_i^o$ for all $t \in [1+\varepsilon,\infty)$.

2. *Age* 1. Cohort i must decide what proportion of its wealth, ϕ_i, to invest in conventional annuities, investing the remainder in a mortality-adjusted annuity.[10] The decision problem is described by:

$$\begin{aligned} V_{i,1} = V_1\left(w_i^1,\pi_i,\rho_{i-1}\right) &= \max_{\phi_i} E\left(V_{1+\varepsilon}\left(y_i^o\right)\right) \\ &= \max_{\phi_i} \int_0^{\infty} \left[\frac{\left(y_i^o\right)^{1-\alpha}}{(1-\alpha)} + A\right] \frac{1}{r+\rho_i} f\left(\rho_i,\rho_{i-1}\right) d\rho_i \\ \text{subject to} &: y_i^o = \left[\phi_i\pi_i\theta_i + (1-\phi_i)(r+\rho_i)\right] w_i^1 \text{ and} \\ & \log\rho_i = \log\rho_{i-1} + \frac{\upsilon_{i-1}}{\zeta}, \upsilon_{i-1} \sim N(\gamma_2,1) \qquad (4) \end{aligned}$$

---

[9] It would be possible to solve for consumption in the interval $[0,\varepsilon)$ at the same time as solving for the portfolio decision at time 0. Similarly we could solve for consumption in the interval $[1,1+\varepsilon)$ at age ε since no information accrues in the interval $[\varepsilon,1+\varepsilon)$. However, the introduction of these extra variables would complicate our analysis without adding to it.

[10] With our assumptions, buying a mortality adjusted annuity with any wealth that is not invested in a conventional annuity is equivalent to buying a conventional annuity in the absence of aggregate mortality risk at time $1+\varepsilon$.

where E is the expectations operator, and w_i^1 is cohort i's wealth at age 1. We denote by θ_i the pay-out rate of the actuarially fair annuity. Given our assumption of constant mortality and interest rates, this is equal to the harmonic mean of the sum of these (Lando 1998, Schrager 2006). π_i is the ratio of the actual pay-out of the conventional annuity to the actuarially fair pay-out; we call this the pay-out ratio. All other variables in (4) are as defined previously. We represent the density function of ρ_i as depending on the values of ρ_{i-1} to take account of the persistence of mortality rates described by the respective relation in the constraint (see section 3.1).

π_i is the market clearing annuity pay-out ratio when cohort i is aged 1, and is taken as given by cohort i when solving the decision making problem described by (4). Note, however, that the demand for conventional annuities, ϕ_i, is determined partly by wealth w_i^1, which depends upon the past decisions of cohort i. Hence, the market clearing pay-out ratio for conventional annuities that is explored here describes a Stackelberg equilibrium in which the old cohort has first mover advantage.

3. *Age* ε. The mortality rate of the old cohort is revealed, which determines the gains or losses that the young cohort realises on their sale of conventional annuities, w_i^ε. The young cohort must then choose their consumption rate over the interval $[\varepsilon, 1)$, c_i^y:

$$\begin{aligned} V_{i,\varepsilon} = V_\varepsilon\left(w_i^\varepsilon, \rho_{i-1}, i\right) &= \frac{A\left(1-e^{-r}\right)}{r} + \max_{c_i^y}\left\{\frac{\left(1-e^{-r}\right)}{r\left(1-\alpha\right)}\left(c_i^y\right)^{1-\alpha} + e^{-r}V_1\left(w_i^1, \pi_i, \rho_{i-1}\right)\right\} \\ \text{subject to} &: \quad w_i^1 = w_i^\varepsilon e^r - \frac{c_i^y}{r}\left(e^r - 1\right) \geq 0 \text{ and} \\ &\quad\; \pi_i = \pi\left(w_i^1, \rho_{i-1}, i\right) \end{aligned} \tag{5}$$

Here the first mover advantage of cohort i is made clear by the specification of π_i as a function of w_i^1, ρ_{i-1}, and i; $\pi_i = \pi\left(w_i^1, \rho_{i-1}, i\right)$. π_i is a function of w_i^1 through its influence on the demand for conventional annuities as described by (4), and is influenced separately by ρ_{i-1} because this affects expectations over ρ_i (as described by equation (2)). Furthermore, the relation between π_i, w_i^1, and ρ_{i-1} is cohort specific due to the shadow effects of the terminal conditions that are assumed for the model. These are discussed at the end of the current subsection.

4. *Age* 0. At the beginning of its economic life a cohort must decide how far it is prepared to sell conventional annuities to the current old cohort, expressed

as a proportion of its wealth endowment, ϕ_i'. In the absence of uncertainty, an actuarially fair annuity bought for £1 would yield a dividend of £$(r+\rho_i)$. The capital value of a conventional annuity once the uncertainty surrounding the old cohort's mortality rate is resolved is consequently £$\pi_i\theta_i/\left(r+\rho_i\right)$ per £1 invested. Hence, ϕ_i' is found by solving:[11]

$$\begin{aligned} V_{i,0} = V_0\left(\pi_{i-1},\rho_{i-2},i\right) &= \max_{\phi_i'} E\{V_\varepsilon\left(w_i^\varepsilon,\rho_{i-1},i\right)\} \\ \text{subject to} &: w_i^\varepsilon = 1+\phi_i'\left(1-\frac{\pi_{i-1}\theta_{i-1}}{r+\rho_{i-1}}\right) \text{ and} \\ & \log\rho_{i-1} = \log\rho_{i-2}+\frac{\upsilon_{i-2}}{\zeta}, \upsilon_{i-2}\sim N(\gamma_2,1) \end{aligned} \tag{6}$$

The homothetic form of

$$V_\varepsilon\left(w_i^\varepsilon,\rho_{i-1},i\right) - A\left[\frac{1-e^{-r}}{r}+\int_0^\infty\frac{f\left(\rho_i,\rho_{i-1}\right)}{r+\rho_i}d\rho_i\right]$$

means that, with

$$U\left(\rho_{i-1}\right) = A\left[\frac{1-e^{-r}}{r}+\int_0^\infty\frac{f\left(\rho_i,\rho_{i-1}\right)}{r+\rho_i}d\rho_i\right]$$

$$\begin{aligned} V_0\left(\pi_{i-1},\rho_{i-2},i\right) &= \max_{\phi_i'} E\{\left[1+\phi_i'\left(1-\frac{\pi_{i-1}\theta_{i-1}}{r+\rho_{i-1}}\right)\right]^{1-\alpha}\left[V_\varepsilon\left(1,\rho_{i-1},i\right)-U\left(\rho_{i-1}\right)\right] \\ &\quad +U\left(\rho_{i-1}\right)\} && (7) \\ &= \max_{\phi_i'}\int_0^\infty\left[1+\phi_i'\left(1-\frac{\pi_{i-1}\theta_{i-1}}{r+\rho_{i-1}}\right)\right]^{1-\alpha} \\ &\quad \left[V_\varepsilon\left(1,\rho_{i-1},i\right)-U\left(\rho_{i-1}\right)\right]f(\rho_{i-1},\rho_{i-2})d\rho_{i-1} && (8) \\ &\quad +\int_0^\infty U\left(\rho_{i-1}\right)f(\rho_{i-1},\rho_{i-2})d\rho_{i-1} \\ \text{subject to} &: \log\rho_{i-1} = \log\rho_{i-2}+\frac{\upsilon_{i-2}}{\zeta}, \upsilon_{i-2}\sim N(\gamma_2,1) && (9) \end{aligned}$$

The above provides a complete description of the lifetime decision making problem considered for any single cohort. In obtaining a solution to the optimisation problems

---

[11] We do not impose a separate constraint on ϕ_i' to ensure that w_i^ε is strictly positive in the context of extreme values of ρ_{i-1}. We note, however, that ρ_{i-1} would need to be more than 9 standard deviations from its expectation to obtain zero wealth in any of the analytical contexts that are considered in the current study.

that are described above, however, we must also take into account conditions at the private market's opening and closure. When the private market for conventional annuities is first introduced, the current old cohort is assumed to have the opportunity to purchase annuities, without having had the opportunity to sell them when young. Denote this cohort $i = 1$, for which $\phi_1' = 0$, so that $V_{1,0} = \int_{\rho_0} \{V_\varepsilon(1, \rho_0, 1) - U(\rho_0)\} f(\rho_0, \rho_{-1}) + U(\rho_0) f(\rho_0, \rho_{-1})\, d\rho_0$. Similarly, when the private market for annuities is closed, the current young cohort is assumed to have the opportunity to sell annuities, without having the opportunity to buy them when old. Denote this cohort $i = T$, for which $\phi_i = 0$, so that:

$$V_{T,1} = V_1\left(w_T^1, \rho_{T-1}\right) = \int_0^\infty \frac{\left[w_T^1 (r + \rho_T)\right]^{1-\alpha}}{(1-\alpha)(r + \rho_T)} + \frac{A}{(r + \rho_T)} f\left(\rho_T, \rho_{T-1}\right) d\rho_T$$

No analytical solutions exist to these optimisation problems, and so numerical methods were employed. These are discussed in section 3.5.

3.3 Compulsory Annuitisation

In the United Kingdom annuitisation of retirement savings is compulsory at the age of seventy-five. In practice, of course, the choice between mortality-adjusted annuities and conventional annuities does not yet exist. Nevertheless, analysis of compulsory annuitisation provides an indication of the welfare and budgetary costs of policies that protect people fully from mortality risk. In our model, compulsory annuitisation is explored by imposing the restriction that $\phi = 1$, holding all other aspects of the model fixed as described above.

3.4 Government Provision of Annuities

The government is an alternative supplier of annuities and it, unlike the private market, is in a position to spread risk across all future cohorts. Gordon & Varian (1988) show that, in a general context, governments can raise welfare by spreading shocks across future cohorts as well as those that are currently alive. We explore the effects of government intervention in the annuities market with reference to the following policy structure. When cohort i is aged ε, the government announces a (possibly negative) dividend D_i, which is paid immediately to cohort i. It also announces the pay-out ratio, π_i, at which the cohort will be able to buy conventional annuities on reaching age 1. Individuals are subject to the same decision making points as described above for the private market in annuities, except that they are not permitted to sell annuities at age 0. The government is considered to select its policy parameters to maximise the expected value of a utilitarian welfare function aggregated over all future cohorts that will be subject to the policy in annuities, described by:

$$\begin{aligned} W_i^{\varepsilon}\left(H_i, \rho_{i-1}\right) &= \max_{D_i, \pi_i} \sum_{j=i}^{T} e^{-\delta(j-i)} E\left(V_{j,\varepsilon}\right) \\ &= \max_{D_i, \pi_i} V_{i,\varepsilon} + e^{-\delta} E\left(W_{i+1}^{\varepsilon}\left(H_{i+1}, \rho_i\right)\right) \\ \text{subject to} &: \; H_{i+1} = e^{r}\left(H_i - D_i\right) + \phi_i w_i^1 \left(1 - \frac{\pi_i \theta_i}{r + \rho_i}\right) \end{aligned} \tag{10}$$

where δ is the government's discount rate, $H_{T+1} = 0$, $\pi_T = 0$ – so that $D_T = H_T$ – and $W_{T+1}^{\varepsilon} = 0$. In selecting its policy parameters, the government is also considered to be subject to an incentive compatability constraint, so that it must take into consideration the optimising reactions of each cohort to the policy environment.

As no information accrues between ages ε and 1, and as the optimising decision at age $1 + \varepsilon$ continues to involve setting instantaneous consumption in old age equal to annuity income, all three decisions that cohorts are considered to make in the current context (c_i^y, ϕ_i, c_i^o) can be resolved at age ε. An alternative analytical scenario involves announcing the policy variables, $[\pi_{i,} D_i]$ at time 0, in which case the government would carry the additional risk arising from the uncertainty surrounding ρ_{i-1}. If the policy is announced at time ε, however, then the structure is more like the private market problem described in the preceding subsection, in which people realise their capital gains at time ε and can then predict the pay-out rate that they will face when aged 1.

Although this policy structure results in transfers between generations, it is different in form from a typical pay as you go system (Demange 2002). With a pay as you go system, no explicit charge is levied for the transfer of risk between generations.

As Gordon & Varian (1988) point out, the policy structure that is considered here may be time-inconsistent, as examples can be constructed whereby a young cohort is made strictly better off by withdrawing its support for the scheme. Varian is optimistic that there are obstacles to default arising from institutional features that are beyond the scope of his model. In our framework we also note that there would appear to be a higher likelihood of the scheme being closed prematurely when it is in surplus than when it is in deficit. This is because closure while in deficit suggests a scenario in which a young cohort appropriatse resources from the older living cohort, which the older cohort is likely to resist. In contrast, closure while in surplus suggests a scenario in which current surviving cohorts appropriate resources held for the benefit of future cohorts. Consideration of these issues could be expected to raise issues similar to those explored in the context of behaviour subject to quasi-hyperbolic discounting; see, for example, Diamond & Koszegi (2003).

3.5 Solution Methods

A solution is required for:

- c_i^o the instantaneous consumption rate from age $1+\varepsilon$ of cohort i
- ϕ_i cohort i's demand for annuities, for all but the youngest cohort, $i=T$
- π_i the pay-out ratio at which cohort i can buy conventional annuities, for all but the youngest cohort, $i=T$
- c_i^y the instantaneous consumption rate between ages ε and 1 of cohort i
- ϕ_i' cohort i's supply of annuities in the case where annuities are provided through a private market, for all but the oldest cohort, $i=1$.
- D_i the lump sum that the government pays to or tax it collects from cohort i where annuities are provided through the government

The uncertainty that is assumed to govern the intertemporal variation in cohort specific mortality rates implies that analytical solutions to the optimisation problems that are described above do not exist. The procedure that we adopt consequently uses backward induction to solve the required inter-temporal Bellman equations.

The first period in the decision making problem relates to the youngest decision making point of the oldest cohort's life – age ε of cohort 1 – and the last period relates to the final decision making point of the youngest cohort's life – age $1+\varepsilon$ of cohort T. Our solution method consequently works backward from age $1+\varepsilon$ of cohort T to age ε of cohort 1, with the decision making problems of individual cohorts linked, either through their respective market interactions in the case of a private market in annuities, or through government reserves, H_i. The objective functions that we are interested in here are all smooth and concave. Numerical approximations to optimising solutions were consequently obtained using the *Matlab* routine *fminsearch.*

At age $1+\varepsilon$ of cohort T, we know that the utility maximising instantaneous consumption rate is equal to the flow of income for that cohort in old age, y_T^o (as described with reference to the optimisation problem (3)). Furthermore, given the assumptions made for the terminal conditions of both models of annuity provision that are considered here, cohort T has no access to conventional annuities at age 1. Hence, $c_T^t = c_T^o = y_T^o = (r+\rho_T)\,w_T^1$ for all $t \in [1+\varepsilon, \infty)$. We can consequently determine the value function – the optimised expected lifetime utility – of cohort T at age $1+\varepsilon$ for any combination of state variables, $V_{1+\varepsilon}(y_T^o) = V_{1+\varepsilon}((r+\rho_T)\,w_T^1) = V_{T,1+\varepsilon}(w_T^1, \rho_T)$.

A two dimensional grid was defined in (w_T^1, ρ_T), comprised of 61 evenly spaced points in each dimension, with w_T^1 in [0.6, 1.4], and ρ_T in [0.9728, 4.0712], with the ρ dimension specified in logs. The value function was evaluated at all of the nodes of this two dimensional grid.

At age ε of cohort T, the decision problem (5) was solved for consumption, c_T^y, at each node of a two dimensional grid. With regard to the private market problem, this two dimensional grid was defined in w_T^ε and ρ_{T-1}, and was specified with the same dimensions as described above for age $1+\varepsilon$. In the case where conventional annuities are considered to be sold by the government, the w axis referred to above was replaced by the size of the social fund, H. For cohort T, the dividend is $D_T = H_T$ by assumption, so that $w_T^\varepsilon = 1 + H_T$. Thirteen equally spaced points were selected for H_T covering the interval $[-0.36, ..., +0.36]$. The expectation over ρ_T associated with $V_{T,1} = E\left(V_{T,1+\varepsilon}(w_T^1, \rho_T)\middle|\, \rho_{T-1}\right)$ that is required for $V_{T,\varepsilon}$was evaluated using a Gaussian quadrature with five abscissae points.[12] At each combination of state variables considered for analysis, the instantaneous consumption rate c_T^y that solves for $V_{T,\varepsilon}$ was numerically approximated using the *fminsearch* routine.[13]

At age 0 of cohort T, in the case of the private market for annuities, the cohort is considered to choose the value of conventional annuities that it is willing to sell to cohort $T-1$(no decision is made at this decision point in the case of government provision). In this case, the assumption of a perfectly competitive market in conventional annuities provides the link between the decision making problems of cohorts T and $T-1$. In the

---

[12] This is a standard approach for numerically approximating the expectation of a continuously distributed random variable. It involves evaluating the objective function at a discrete number of abscissae points, and aggregating subject to a defined weighting schedule. In the current case, the objective function associated with each of the abscissae was evaluated by bi-cubic interpolation of the solutions stored in the grid considered for cohort T at age $1+\varepsilon$, as described above.

[13] To verify the accuracy of the *MATLAB* solution routine considered here, we obtained solutions to the private market problem through separate routines programmed in Fortran.

first instance, we solved for cohort T's supply of annuities, ϕ'_T, for all of the nodes of a two dimensional grid in $\rho_{T-2}$and π_{T-1}. As above, each dimension was divided into 61 evenly spaced points, with ρ specified in logs, and $\pi \in [0.95, 1.0]$. Again, the expectations required to evaluate the value function, $V_{T,0}$ were calculated using a five point Gauss-Hermite quadrature and cubic interpolation of the grid described for age ε. In this regard, it is of note that we experienced some numerical difficulties in extrapolating in the ρ dimension. The grid for age 0 was consequently defined to limit these extrapolations, with $\rho \in [1.3, 3.0769]$ considered for analysis. Note that this range is centered on the expectation $\rho = 2.0$.

At age $1 + \varepsilon$ of cohort $T - 1$, we know – as in the case of cohort T– that the utility maximising instantaneous consumption rate is $c^t_{T-1} = c^o_{T-1} = y^o_{T-1}$ for all $t \in [1+\varepsilon, \infty)$. In contrast to cohort T, however, cohort $T - 1$ can choose to invest some of its wealth at age 1, w^1_{T-1}, in conventional annuities; buying them from cohort T in the context of a private market, or from the government in the context of public provision. For the purposes of computational simplicity, we chose to collapse the decision making problems that are described in section 3.2 for ages ε and 1 into a single decision making point, at ε. This is made possible by the fact that no uncertainty is resolved between ages ε and 1, so that the transition of state variables between the two points is deterministic.

At age ε of cohort $T - 1$, we considered decisions regarding the instantaneous consumption rate between ages ε and 1, and the demand for conventional annuities that maximise expected lifetime utility, $V_{T-1,\varepsilon}$. This was done for the same combinations of state variables – $(w^\varepsilon_{T-1}, \rho_{T-2})$ in the case of private market provision, and $\left(H_{T-1}, \rho_{T-2}\right)$ in the case of government provision – that were considered for cohort T. In the case of cohort $T-1$, however, we considered a three dimensional grid, with the third dimension representing the annuity pay-out rate π_{T-1}. As for the other dimensions considered for analysis, the πaxis was divided into 61 points, evenly spaced between 0.95 and 1.0. Furthermore, the expectation associated with $V_{T-1,1} = E\left(V_{T-1,1+\varepsilon} | \, \rho_{T-2}\right)$ that is required for $V_{T-1,\varepsilon}$ was again evaluated using a Gaussian quadrature with five abscissae points.

In the model of a private market for conventional annuities, we used a simple gradient procedure to search for the value of π_{T-1}that balanced demand for annuities (as described by the grid for cohort $T - 1$ at age ε) against the supply of annuities (as described by the grid for cohort T at age 0).

At age 0 of cohort $T - 1$, in the case of the private market in annuities, the cohort is considered to choose the value of conventional annuities it is willing to sell to cohort

$T-2$. Again, the market in conventional annuities provides the link between the decision making problems of cohorts $T-1$ and $T-2$.

In the model of government provision of annuities, a solution was obtained to ϕ_{T-1}, c^y_{T-1}, π_{T-1}, and D_{T-1}, at given values of $H_{T-1}$and ρ_{T-2}, by a nested application of the *fminsearch* routine. The outer-loop of the nested search procedure was used to solve the inter-cohort Bellman equation described by (10) for the values of $\pi_{T-1}$and D_{T-1}. The inner-loop of the procedure was used to solve (5) for values of $\phi_{T-1}$and c^y_{T-1}.

Decisions for each cohort preceding cohort $T-1$ were solved following repeated application of the steps described above for cohort $T-1$, with the exception that cohort 1 was subject to the further restriction in the private market scenario that it could not sell annuities, $\phi'_1 = 0$, as noted above.

4 The Market Solution

We present an analysis of the market soution, beginning by setting out demand and supply curves for conventional annuities.

4.1 The Demand for Annuities

The demand curve for conventional annuities is given as the value of ϕ_i that optimises equation (4), at a given pay-out ratio π_i. Although no analytical solution exists to this optimisation problem, we can derive an analytical expression that approximates the implied demand function. This subsection begins by deriving analytical approximations to the demand for annuities at two specific pay-out ratios. We then present numerically derived demand functions evaluated at a large number of pay-out ratios, two of which coincide with the pay-out ratios considered for the analytical approximations. The analytical approximations consequently provide a useful benchmark against which to check the accuracy of the numerical methods that are considered for analysis.

Analytical approximations

Consider a general, continuously differentiable instantaneous utility function, $u(.)$. Then the value function at age 1, as in equation (4), is described by:

$$V_{i,1} = \max_{\phi_i} E \frac{u(c^o_i)}{r+\rho_i} \tag{11}$$

$$c^o_i = [\pi_i \theta_i + (1-\phi_i)(r+\rho_i - \pi_i \theta_i)] w^1_i \tag{12}$$

where all variables are defined as stated previously. Taking a Taylor series expansion about $\pi_i\theta_i w_i^1$ of $u(.)$, and substituting into (11):

$$\begin{aligned} V_{i,1} &= \max_{\phi_i}\left\{E\frac{u\left(\pi_i\theta_i w_i^1\right)}{r+\rho_i}+\frac{\left(1-\phi_i\right)\left(r+\rho_i-\pi_i\theta_i\right)w_i^1}{r+\rho_i}u'\left(\pi_i\theta_i w_i^1\right)\right. \\ &\quad\left.+\frac{\left(1-\phi_i\right)^2\left(r+\rho_i-\pi_i\theta_i\right)^2\left(w_i^1\right)^2}{2(r+\rho_i)}u''\left(\left[\pi_i\theta_i+\lambda(1-\phi_i)\left(r+\rho_i-\pi_i\theta_i\right)\right]w_i^1\right)\right\} \end{aligned}$$

Differentiating with respect to ϕ_i:

$$\begin{aligned} \frac{dV_{i,1}}{d\phi_i} &= E\left\{-\frac{(r+\rho_i-\pi_i\theta_i)w_i^1}{r+\rho_i}u'\left(\pi_i\theta_i w_i^1\right)+\right. \\ &\quad -\frac{\left(1-\phi_i\right)\left(r+\rho_i-\pi_i\theta_i\right)^2\left(w_i^1\right)^2}{(r+\rho_i)}u''\left(\left[\pi_i\theta_i+\lambda(1-\phi_i)\left(r+\rho_i-\pi_i\theta_i\right)\right]w_i^1\right) \\ &\quad\left.-\frac{\lambda(1-\phi_i)^2\left(r+\rho_i-\pi_i\theta_i\right)^3\left(w_i^1\right)^3}{2(r+\rho_i)}u'''\left(\left[\pi_i\theta_i+\lambda(1-\phi_i)\left(r+\rho_i-\pi_i\theta_i\right)\right]w_i^1\right)\right\} \end{aligned} \tag{13}$$

Let $\pi_i\theta_i$ denote the harmonic mean, $\left(r+\rho_i\right)^h$, so that the pay-out ratio $\pi_i=1$. Then, from (13):

$$\frac{dV_{i,1}}{d\phi_i} = E\left\{-\frac{\left(1-\phi_i\right)\left(r+\rho_i-\pi_i\theta_i\right)^2\left(w_i^1\right)^2}{(r+\rho_i)}u''\left(\left[\pi_i\theta_i+\lambda(1-\phi_i)\left(r+\rho_i-\pi_i\theta_i\right)\right]w_i^1\right)\right. \tag{14}$$

$$\left.-\frac{\lambda(1-\phi_i)^2\left(r+\rho_i-\pi_i\theta_i\right)^3\left(w_i^1\right)^3}{2(r+\rho_i)}u'''\left(\left[\pi_i\theta_i+\lambda(1-\phi_i)\left(r+\rho_i-\pi_i\theta_i\right)\right]w_i^1\right)\right\} \tag{15}$$

It then immediately follows that $dV_{i,1}/d\phi_i = 0$ if $\phi_i = 1$. Furthermore, this will be a (local) maximum if $u''(.) < 0$.

Now let $\pi_i\theta_i$ denote the arithmetic mean, $(r+\rho_i^e)$, so that $\pi_i = \left(r+\rho_i^e\right)/\left(r+\rho_i\right)^h > 1$. Setting $\lambda = 0$ to make the second-order approximation:

$$\begin{aligned} \frac{dV_{i,1}}{d\phi_i} &\simeq E\left\{-\frac{(r+\rho_i-\pi_i\theta_i)w_i^1}{r+\rho_i}u'\left(\pi_i\theta_i w_i^1\right)+\right. \\ &\quad\left.-\frac{\left(1-\phi_i\right)\left(r+\rho_i-\pi_i\theta_i\right)^2\left(w_i^1\right)^2}{(r+\rho_i)}u''\left(\pi_i\theta_i w_i^1\right)\right\} \end{aligned} \tag{16}$$

And taking another second-order Taylor expansion about ρ_i^e to evaluate the expectation term:

$$\frac{dV_{i,1}}{d\phi_i} \simeq \frac{\sigma_i^2 w_i^1}{\left(\pi_i\theta_i\right)^2}u'\left(\pi_i\theta_i w_i^1\right)-\frac{\sigma_i^2\left(1-\phi_i\right)\left(w_i^1\right)^2}{\pi_i\theta_i}u''\left(\pi_i\theta_i w_i^1\right)$$

If $u\left(\pi_i\theta_i w_i^1\right) = \frac{\left(\pi_i\theta_i w_i^1\right)^{1-\alpha}}{1-\alpha}$ then it follows that $dV_{i,1}/d\phi_i \simeq 0$ if:

$$\phi_i \simeq \frac{\alpha+1}{\alpha}$$

Furthermore, this will be a (local) maximum if $u''(.) < 0$.

Hence, for any differentiable instantaneous utility function with expected utility discounted at the real rate of interest, if $\pi_i\theta_i$ is equal to the harmonic mean, $\left(r+\rho_i\right)^h$, then expected lifetime utility when old will obtain a (local) maximum when wealth is fully annuitised, $\phi_i = 1$. Alternatively, if $\pi_i\theta_i$ is equal to the arithmetic mean, $r + \rho_i^e$, then to a second-order approximation, and with a constant elasticity of substitution utility function with elasticity of substitution α, expected lifetime utility when old will obtain a (local) maximum when wealth is over-annuitised to the degree, $\phi_i = \frac{\alpha+1}{\alpha}$. These two special cases allow us to identify the demand curve to the extent that it can be approximated by a straight line.

Numerical approximations

The homotheticity of the utility function implies that the solution is independent of w_i^1. We illustrate demand curves for $r = 1\%$ p.a. and $r = 2\frac{1}{2}\%$ p.a. in figure 2. We use a value of $\alpha = 4$. This is towards the upper end of the range of plausible estimates for this value; we choose a value which implies a high degree of risk aversion because the aim of this paper is, within the framework of a structural model, to look at the effects of risk aversion on the supply of and demand for conventional annuities. We also show in figure 2 the demand curves constructed by extrapolating the lines that join the analytical approximations derived above to the solutions for the harmonic mean (which is independent of r) and the arithmetic mean (which is a function of r).

The proportionate uncertainty in consumption arising from uncertainty in ρ_i is decreasing as a function of r. If the interest rate is low then the overall return on a mortality-adjusted annuity, or on funds annuitised at age $1+\varepsilon$ is dominated by ρ_i; it follows that shocks to ρ_i will have a large proportionate impact and people will be keen to buy conventional annuities even if the pay-out is poor. If r is large, then both ρ_i and shocks to ρ_i are likely to be less important. As a consequence, at any given value of $\pi_i < 1$ the proportion of w_i^1 which old people wish to annuitise will be a declining function of r.

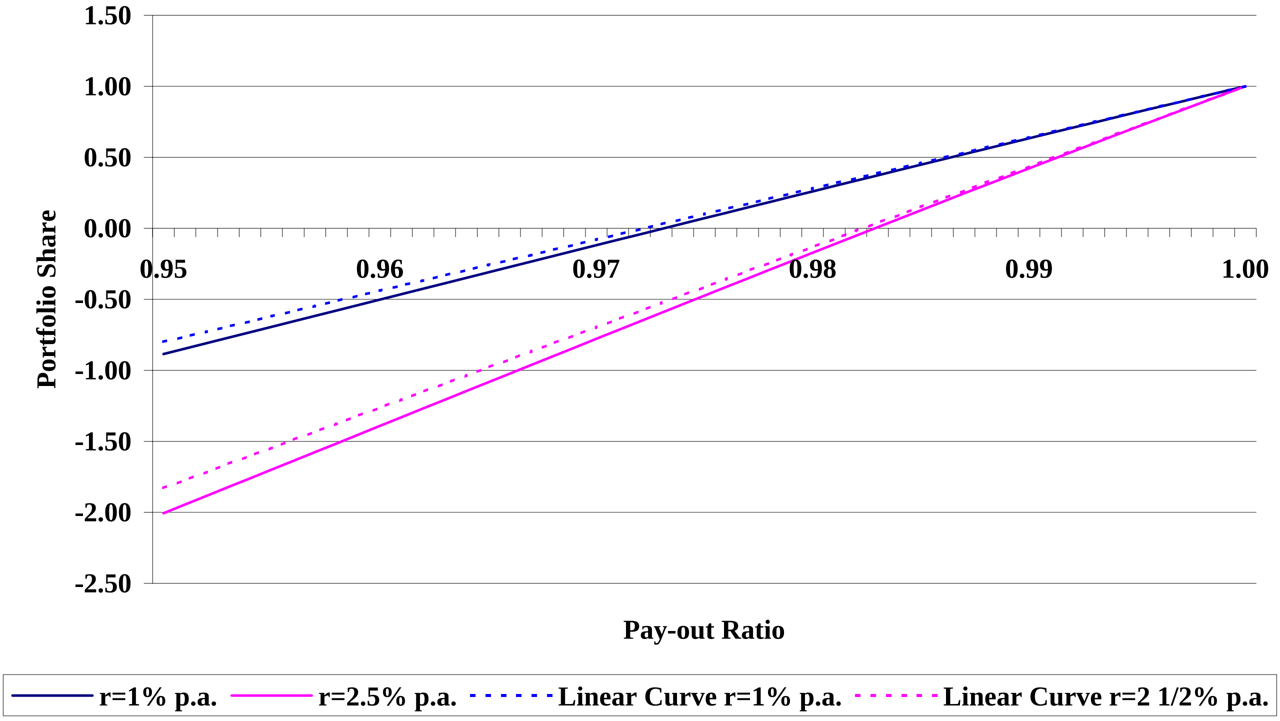


Figure 2: Demand for Annuities, $r = 1\% p.a.$ and $r = 2.5\% p.a.$

The curves confirm that the steepness of the demand curve is a function of the interest rate, for the reasons that we discuss above. With $\pi_i = 1$ our simulations confirm that people are fully invested in conventional annuities. At low annuity rates it is not surprising that old people would like to be net sellers rather than net buyers of conventional annuities. But the graph suggests that, for an annuity market to exist when investors can choose between conventional annuities and mortality-adjusted annuities, the value of π_i cannot be very far below 1. If sellers of conventional annuities need to make a substantial charge for risk, then there should not be an annuity market, and its existence can be explained only as a result of the legal requirement for pension funds to be used to purchase conventional annuities rather than mortality-adjusted annuities. If mortality-adjusted annuities are not available, the demand curves for conventional annuities reported in figure 2 are horizontal.

It should be remembered that these demand curves show the proportion of wealth, at age 1, that people would like to annuitise. The market-clearing equilibrium is, of course, also determined by the amount of wealth that people hold at this age, and the associated supply of annuities by the young.

4.2 The Supply of Conventional Annuities

As the description of the solution method provided in section 3.5 makes clear, the supply of conventional annuities depends on the uncertainty that is associated with future annuity pay-out rates. The supply of conventional annuities consequently alters as one approaches the terminal period.

In constructing the supply curves that are reported here, the model was solved recursively for twenty cohorts, $T = 20$, with the supply curves that are reported here being calculated for the oldest cohort, cohort $i = 1$ (ignoring the assumption that the oldest cohort cannot supply annuities in the model). The supply curves are based upon the assumptions that $\rho_{-1} = 2$ (the assumed expectation), and that cohort 1 takes into account the market-clearing pay-out ratio, $\pi_1 = \pi\left(w_1^{\varepsilon}, \rho_0, 1\right)$, determined by the interaction of cohort $i = 1$'s demand for annuities and cohort $i = 2$'s supply. Working to four decimal places we found that, for any given set of state variables, $\left(w_j^{\varepsilon}, \rho_{j-1}\right)$, the market-clearing pay-out ratio $\pi_j = \pi\left(w_j^{\varepsilon}, \rho_{j-1}, j\right)$ did not change between j and $j - 1$, for all $j < T - 1$. The supply curves that we report here can consequently be regarded as representative of the market steady state.

We find again that the steepness of the curve is increasing in the interest rate, showing in figure 3 curves constructed when mortality-adjusted annuities are assumed to be available. These curves once more reflect the fact that, with high interest rates mortality risk is quantitatively less important than with low interest rates. In consequence, the proportion of wealth people are prepared to annuitise rises more steeply as the pay-out on the annuity falls. With a pay-out of 1 young people want to be buyers rather than sellers of conventional annuities (selling a negative proportion of initial wealth). This is a consequence of the persistence in the mortality rate. When they they can trade in annuities at a price of 1 they are buying some insurance against their own risk of collective longevity. Once again the higher is the interest rate, the less important is mortality uncertainty and thus the fair pay-out ratio is closer to 1. Compulsory purchase of conventional annuities in old age has very little effect on the position of the supply curve; we find it changes only the fourth decimal place of the amount of annuities that young people are prepared to sell.

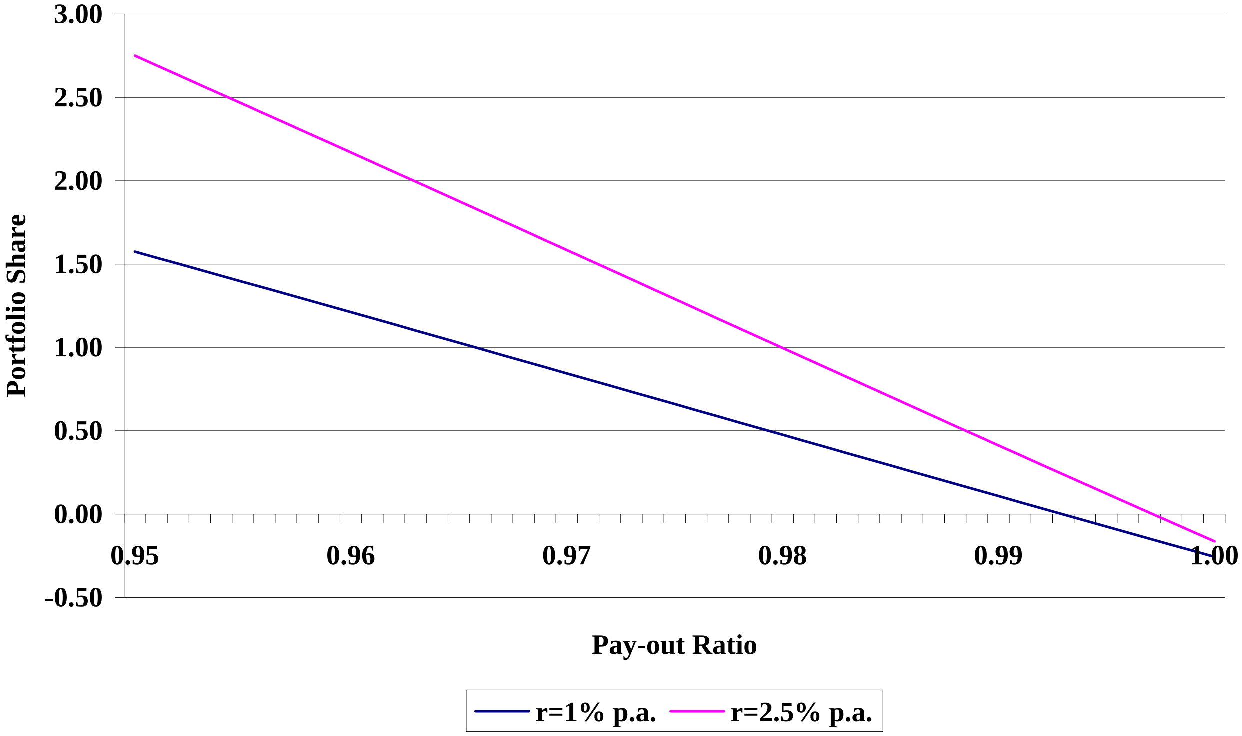


Figure 3: Supply of Annuities, $r = 1\% p.a.$ and $r = 2.5\% p.a.$

4.3 Market Equilibrium

Here we report the properties of market equilibria that were calculated from the same model solutions that were used to compute the supply functions reported in the preceding subsection. Table 1 displays stastitics obtained where the real rate of interest is assumed to be 1% p.a., $\alpha = 4$, and the out-turn of ρ_i is considered to repeatedly take its expected value of 2. The choice of a low real interest rate and a high value of α is deliberate, as these exaggerate the incentive effects of aggregate mortality risk on cohort behaviour. In the solutions that we present below, cohort i sells conventional annuities to cohort $i+1$. The table makes it possible to follow each cohort through its life-course. The first three columns show the decisions made be each cohort at age 0. They indicate the sale of annuities ϕ' as a fraction of initial wealth, 1, the market clearing pay-out ratio for those annuities faced by the cohort at age 0, π_0, and the cohort's value function at age 0, V_0. The next three columns show the state of affairs at time ϵ; here the cohort choses its consumption when young, c^y in the light of its wealth after realising gains on sales of annuities to its parent's cohort, W^ϵ and the value of its welfare function, V_ϵ at the same time. The last two columns show the state of affairs when the cohort is aged 1, indicating the fraction of its wealth which it annuitises, ϕ and the pay-out ratio, π_1 at which it does so. This pay-out ratio is obviously the same as that, shown in column 3, at which the next cohort sells annuities

The results confirm what is alluded to by the demand and supply curves that are reported above; that *market-clearing pay-out rates are not very far below 1.* When people can choose between mortality-adjusted and conventional annuities, the steady-state pay-out is 98.73% of the actuarily fair rate if the real interest rate is 1% p.a.[14]. The results reported in table 1 are striking in the stability that they indicate between alternative cohorts. Nevertheless, there are small variations over time which merit some comment.

The first cohort cannot trade annuities when young, although it can take advantage of the market once it reaches old age. This means that, if mortality rates conform to expected values, it does not make any profit on the sale of annuities, and reaches age ϵ with wealth of one unit, instead of slightly above 1 (as reported for W^{ϵ} in table 1). Thus its consumption when young, c^y, demand for annuities when old, ϕ, are depressed slightly, relative to subsequent cohorts. The lower demand of cohort 1 for conventional annuities also results in it obtaining a slightly more favourable pay-out ratio – 98.736% compared with 98.734% – which motivates it to invest a slightly higher proportion of its wealth at age 1 in conventional annuities, ϕ. In contrast, the proportion of wealth invested in annuities by cohort 2 at time 1 is reduced slightly. These considerations generate the variations that can be discerned from the table for the first four cohorts.

At the other extreme of the simulated period, the variation is driven by the fact that the final cohort, $T = 20$, cannot purchase conventional annuities when old. In response to the fact that it cannot insure against aggregate mortality risk, cohort 20 engages in precautionary saving, relative to the other cohorts considered for analysis. It therefore reduces its consumption rate when young, c^y, from 0.9065 to 0.9059. Table 1 indicates that cohort 20 also increases its supply of annuities, relative to older cohorts, so that the market-clearing pay-out ratio rises slightly. This increase in supply may be considered somewhat counter-intuitive: cohort 20 is motivated to increase its exposure to risk when young as a consequence of being exposed to more risk when old. Nevertheless, it is a direct consequence of adopting a preference relation that exhibits prudence.[15] This is

---

[14] Further simulations confirmed that with higher real interest rates the pay-out is even closer to 1.

[15] Preferences are described as exhibitting prudence if the sign on the third derivative of the util-

because prudence motivates precautionary saving, which increases the marginal utility of consumption when young, c^y, thereby increasing the perceived value of investment return in period 0. Of course, this is partly off-set by risk aversion of the preference relation, but the statistics reported in table 1 indicate that the former effect dominates in the case considered here.

The experience of the youngest cohort shown in Table 1 indicates how lack of access to an annuities market reduces lifetime utility; the results reported for V_0 exclude the (unspecified) constant term in the utility function. It is interesting, however, to note that the inter-cohort dynamics generated by the model mean that the highest expected lifetime utility is enjoyed by the second youngest cohort. This is because the second youngest cohort benefits from the higher supply of conventional annuities offered by the youngest cohort – a clear example of an inter-cohort transfer.

We follow this up with an exploration of compulsory annuitisation. Compared to the market solution, compulsory annuitisation results in a lower payout ratio on annuities; this is hardly surprising since compulsion almost doubles the demand for annuities and therefore results in favaourable terms for the suppliers. For the initial cohort welfare is markedly reduced; they face a combination of compulsion and a low payout ratio, with no offsetting financial gain through the sale of annuities. Subsequent cohorts, on average, make larger profits from the sale of annuities than when annuitisation is voluntary. The effects of this on expected welfare almost compensate for the extra costs arising from compulsion. Finally, the last cohort is appreciably better off with compulsion than without it. It makes larger capital gains as a supplier of annuities to its precursor, while its post-retirement choices are the same as those for the last cohort with voluntary annuitisation. Thus the dominant effect of compulsion is to transfer welfare from the first cohort to the last cohort.

---

ity function is positive. Leland (1968) proves that prudence is required for individuals to undertake precautionary saving.

t	ϕ'	π_0	V_0	w^ϵ	c^y	V_ϵ	ϕ	π_1
1	0.000000	0.000000	-0.506384					
1+ϵ				1.000000	0.902933	-0.501483		
2	0.204658	0.987358	-0.505505				0.536300	0.987358
2+ϵ				1.003975	0.906521	-0.495553		
3	0.205233	0.987341	-0.505499				0.535673	0.987341
3+ϵ				1.003990	0.906534	-0.495531		
4	0.205235	0.987341	-0.505499				0.535671	0.987341
4+ϵ				1.003990	0.906534	-0.495531		
5	0.205235	0.987341	-0.505499				0.535671	0.987341
5+ϵ				1.003990	0.906534	-0.495531		
6	0.205235	0.987341	-0.505499				0.535671	0.987341
6+ϵ				1.003990	0.906534	-0.495531		
7	0.205235	0.987341	-0.505499				0.535671	0.987341
7+ϵ				1.003990	0.906534	-0.495531		
8	0.205235	0.987341	-0.505499				0.535671	0.987341
8+ϵ				1.003990	0.906534	-0.495531		
9	0.205235	0.987341	-0.505499				0.535671	0.987341
9+ϵ				1.003990	0.906534	-0.495531		
10	0.205235	0.987341	-0.505499				0.535671	0.987341
10+ϵ				1.003990	0.906534	-0.495531		
11	0.205235	0.987341	-0.505499				0.535671	0.987341
11+ϵ				1.003990	0.906534	-0.495531		
12	0.205235	0.987341	-0.505499				0.535671	0.987341
12+ϵ				1.003990	0.906534	-0.495531		
13	0.205235	0.987341	-0.505499				0.535670	0.987341
13+ϵ				1.003990	0.906534	-0.495531		
14	0.205235	0.987341	-0.505499				0.535670	0.987341
14+ϵ				1.003990	0.906534	-0.495531		
15	0.205238	0.987341	-0.505499				0.535678	0.987341
15+ϵ				1.003990	0.906534	-0.495531		
16	0.205239	0.987341	-0.505498				0.535680	0.987341
16+ϵ				1.003990	0.906535	-0.495529		
17	0.205384	0.987351	-0.505502				0.536059	0.987351
17+ϵ				1.003991	0.906535	-0.495528		
18	0.205372	0.987350	-0.505503				0.536027	0.987350
18+ϵ				1.003991	0.906535	-0.495530		
19	0.205237	0.987341	-0.505494				0.535673	0.987341
19+ϵ				1.003990	0.906538	-0.495523		
20	0.205796	0.987380	-0.507015				0.537137	0.987380
20+ϵ				1.003993	0.905858	-0.497013		
21							0.000000	0.000000

Table 1: Market Equilibrium, Mortality-adjusted Annuities r=1% p.a., Relative Risk Aversion=4

	ϕ'	π_0	V_0	w^ε	c^y	V_ε	ϕ	π_1
1	0.000000	0.000000	-0.509541					
1+ε				1.000000	0.900599	-0.505284		
2	0.384482	0.982447	-0.506397				1.000000	0.982447
2+ε				1.006749	0.906665	-0.495073		
3	0.387093	0.982376	-0.506396				1.000000	0.982376
3+ε				1.006822	0.906730	-0.494962		
4	0.387121	0.982375	-0.506396				1.000000	0.982375
4+ε				1.006823	0.906731	-0.494962		
5	0.387121	0.982375	-0.506396				1.000000	0.982375
5+ε				1.006823	0.906731	-0.494962		
6	0.387121	0.982375	-0.506396				1.000000	0.982375
6+ε				1.006823	0.906731	-0.494962		
7	0.387121	0.982375	-0.506396				1.000000	0.982375
7+ε				1.006823	0.906731	-0.494962		
8	0.387121	0.982375	-0.506396				1.000000	0.982375
8+ε				1.006823	0.906731	-0.494962		
9	0.387121	0.982375	-0.506396				1.000000	0.982375
9+ε				1.006823	0.906731	-0.494962		
10	0.387121	0.982375	-0.506396				1.000000	0.982375
10+ε				1.006823	0.906731	-0.494962		
11	0.387121	0.982375	-0.506396				1.000000	0.982375
11+ε				1.006823	0.906731	-0.494962		
12	0.387121	0.982375	-0.506396				1.000000	0.982375
12+ε				1.006823	0.906731	-0.494962		
13	0.387121	0.982375	-0.506396				1.000000	0.982375
13+ε				1.006823	0.906731	-0.494962		
14	0.387121	0.982375	-0.506396				1.000000	0.982375
14+ε				1.006823	0.906731	-0.494962		
15	0.387121	0.982375	-0.506396				1.000000	0.982375
15+ε				1.006823	0.906731	-0.494962		
16	0.387121	0.982375	-0.506396				1.000000	0.982375
16+ε				1.006823	0.906731	-0.494962		
17	0.387121	0.982375	-0.506396				1.000000	0.982375
17+ε				1.006823	0.906731	-0.494962		
18	0.387121	0.982375	-0.506396				1.000000	0.982375
18+ε				1.006823	0.906731	-0.494962		
19	0.387121	0.982376	-0.506739				1.000000	0.982376
19+ε				1.006823	0.906738	-0.494946		
20	0.387112	0.982417	-0.504752				1.000000	0.982417
20+ε				1.006823	0.907452	-0.493410		
21							0.000000	0.000000

Table 2: Market Equilibrium: Compulsory Annuitisation $r = 1\% p.a.$, Relative Risk Aversion=4

5 Government Provision

An alternative to relying on the market to provide insurance against longevity risk is for the government to take on the role. The advantage of this is that the costs of unexpected developments to mortality rates can be shared with all future generations instead of simply being split among those currently alive. Given the private welfare function (5) of section 3.2 the government faces the maximisation problem

$$W_i\left[\rho_{i-1}, H_i\right] = \underset{D_i, \pi_i}{Max}\ V_{i,\varepsilon}\left(\pi_i, w_i^{\varepsilon}, \rho_{i-1}\right) + \frac{E\left(W_{i+1}\left[\rho_i, H_{i+1}\right]\right)}{R^G} \qquad i < T \tag{17}$$

Here R^G is the discount factor the government applies to the utility of successive cohorts. There is no need for $R^G = R$ although that provides a natural starting point. Since the government is fixing the pay-out instead of acting as a price-taker, we cannot identify any supply curve. However, we should expect the pay-out to be higher than that resulting from the market solution because the risks associated with the sale of annuities are more widely spread.

As with the private sector there is a wide range of possible evolutions of the mortality rates which one might explore. We once again limit ourselves to the case where the expected mortality rate of 2 is realised in each period. We might expect that in such circumstances the government will build up a precautionary balance of reserves which are not in fact called on. Since this process is gradual we explore a simulation of twenty cohorts. Figure 4 shows the government's stock of assets together with the dividend in each period. This shows precautionary balances being build up steadily from period 1 to period 16 with r=1% p.a. The dividend is negative (at -0.0026) in the first period but the combination of the interest earned on the resulting government balance and the profit from the sale of annuities means that from the third period onwards a positive dividend can be paid.

In figure 5 we show the annuity pay-out ratio and the welfare of the cohort which is aged 0 at the start of each period. The welfare figures take account of the uncertainty of the mortality rate between ρ_{i-2} (at age 0) and ρ_{i-1} (at age ε), with the required expectation calculated using the Gauss-Hermite quadrature. The annuity pay-out, 99.07% for cohort 1, falls gradually over time. It is higher than that delivered by the market solution for every cohort except the penultimate when it falls into line with the market solution. Cohort welfare is lower than with the market solution for the first cohort but higher for all subsequent cohorts when the real interest rate is 1% p.a. However, if we

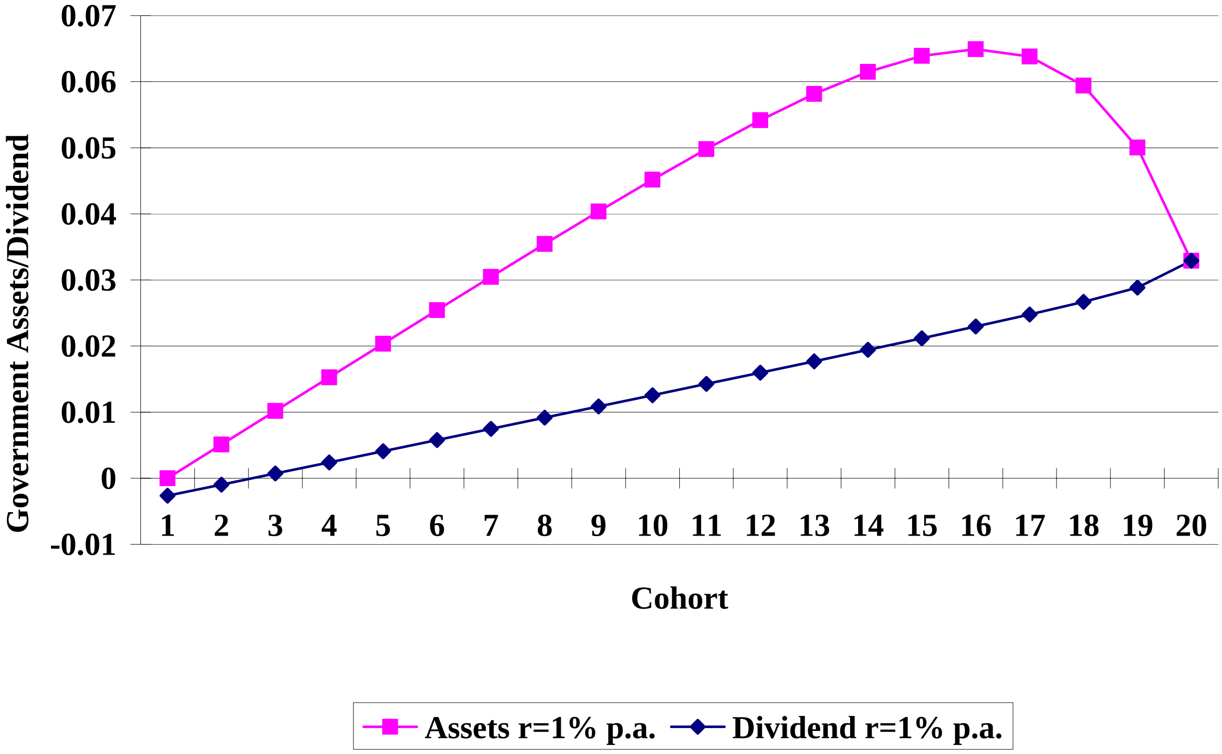


Figure 4: The Time-Profile of Government Assets and the Flat-rate Dividend, $r = 1\% p.a. \rho_i = \rho_i^e = 2$

evaluate the government policy function for the market outcomes, we find that the market solution delivers a value of -1.5511 while when the government provides annuities it is -1.5344. Thus, although government provision is not a Pareto-improvement, it does, as expected, result in higher welfare as measured by the government's policy criterion. This outcome should not be a surprise. If, in a much simpler framework, the resources available to each cohort were uncertain, and the government wanted to maximise the present discounted value of all future utility, it would undoubtedly undertake some precautionary saving at the start of the process. This means, that, if *ex post*, the resources available to the first cohort take their expected value, then the first cohort would be worse off than in the absence of such a policy. The effect of precautionary saving is not, however very marked. The welfare of each cohort rises by only about 0.7% per cohort, translating into a rate of 0.018% p.a.

The magnitude of precautionary saving can be reduced by raising the government discount factor R^G. With $r = 1\%$ p.a. and $R^G = R$ we have $R^G = e^{0.4}$ since each period

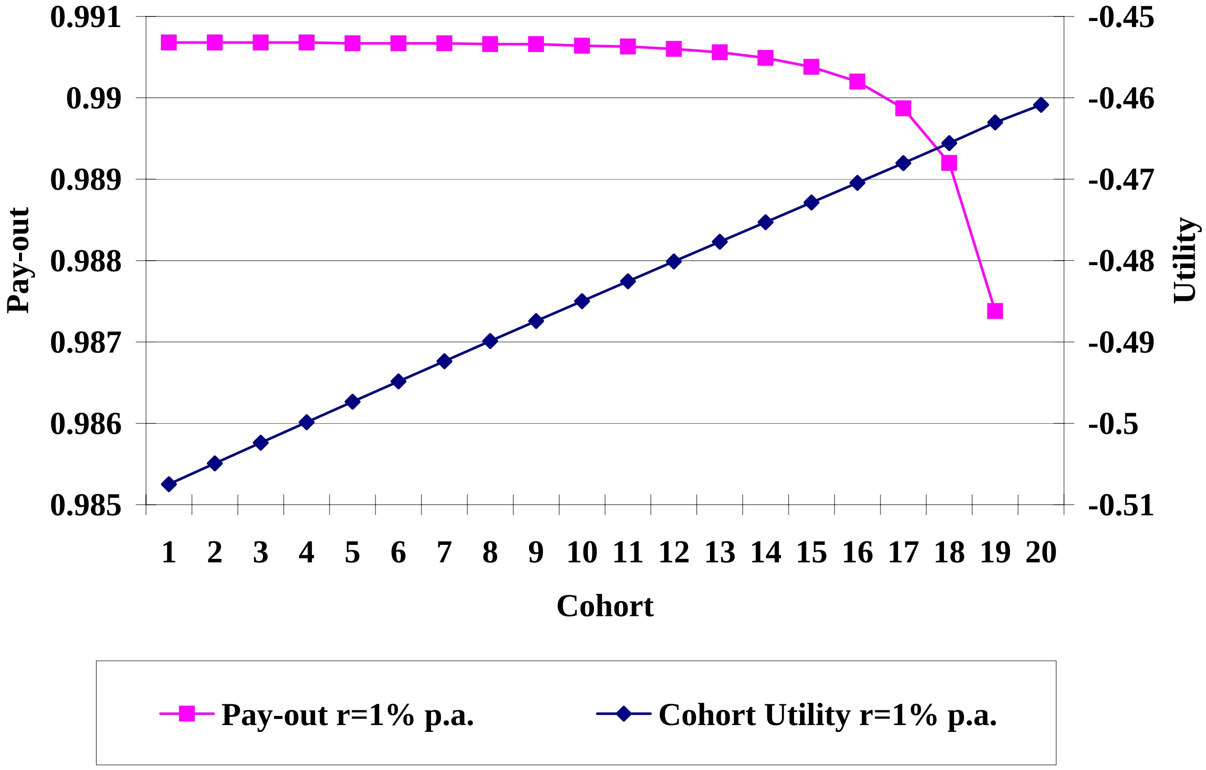


Figure 5: The Time-Profile of the Pay-out Ratio and Cohort Utility, $r = 1\% p.a. \rho_i = \rho_i^e = 2$

is assumed to last for forty years. If we set $R^G = e^{0.405}$ corresponding to a discount factor of 1.0125% p.a., then we find that the first period dividend is raised from -0.0026 to -0.0001. The first-period pay-out is reduced in its fifth decimal place. This results in cohort welfare rising slightly more slowly, but also delivers an initial value of -0.5036, which is higher than the welfare in the first period of the market outcome. This is shown in figure 6. This graph demonstrates, as Gordon & Varian (1988) show generally, that the opportunities for inter-generational risk-sharing made possible by fiscal policy can result in welfare increases for all cohorts.

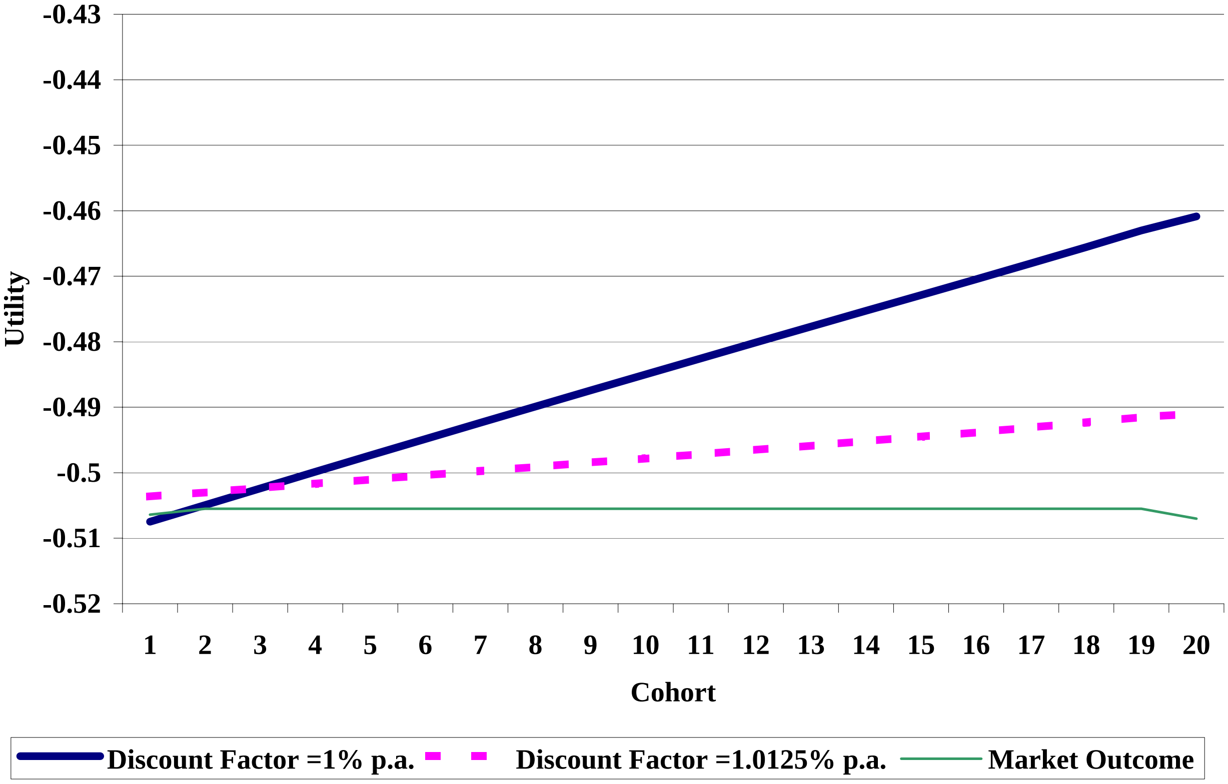


Figure 6: Cohort utility with different government discount factors compared to the market outcome. $\rho_i = \rho_i^e = 2$, $r = 1\%$ p.a.

6 Conclusions

This paper has set out a framework for exploring the issue of aggregate mortality risk by creating an overlapping generations structure in which young people carry the aggregate mortality risk of the old people. It has demonstrated a general point that annuity prices based on expected mortality rates are biased in favour of annuitants, although the bias is unlikely to exceed 1 1/2% even when the degree of uncertainty associated with the mortality rate of old people is high and the real rate of interest is low.

The structure of our model makes it natural to consider the annuity market in terms of demand and supply. We show that the demand curves generated by our model are close to linear curves which can be constructed by solving the model analytically at two points. These curves suggest that, if each cohort has the chance to carry aggregate mortality risk for itself, the demand for conventional annuities will fall to zero once the pay-out drops to 97% of the fair pay-out which would arise in a risk-neutral environment.

The construction of supply curves is slightly more complicated because the willingness of young people to sell annuities depends on the trading conditions that they expect when they reach old age. However we are able to derive a supply curve numerically and to find market-clearing pay-out ratios. We find, not surprisingly, that welfare is higher when old people are allowed to choose between buying annuities and carrying aggregate mortality risk for themselves than it they are obliged to invest all of their wealth in conventional annuities. Since demand for conventional annuities is also reduced by giving people a choice, we find that the pay-out on conventional annuities is higher when people have a choice than when conventional annuitisation is compulsory.

Alternatively mortality risk might be carried by the government. This allows it to be spread across all current and future cohorts, while the market only allows it to be spread across cohorts currently alive. We show that, if the government sets the pay-out on conventional annuities and also uses a (possibly negative) dividend to ensure intertemporal budget balance, it can deliver higher levels of inter-generational welfare than the market. Nevertheless if the government discount rate is the same as the market rate and is low (1% p.a.) the expected welfare of all cohorts may not be raised, as a result of precautionary saving undertaken when the policy is introduced. A small increase in the government discount rate (to 1.0125% p.a.) is adequate to ensure that the first cohort is also better off compared to the market solution.

There are three major conclusions from this paper. First of all, welfare can be improved by giving people a chance to carry some of the mortality risk and policy-makers should encourage the development of financial instruments which make this possible. Secondly, even with fairly high levels of risk aversion, low interest rates and substantial uncertainty about mortality rates of old people, market-clearing risk premium on annuities are unlikely to be very large. Thirdly, the ability of the government to spread risk across current and future cohorts means that the government has a substantial role to play in annuity markets provided that a policy of inter-generational risk-sharing can be sustained.

References

Banks, J. & Blundell, R. (2005), 'Private Pension Arrangements and Retirement in Britain', *Fiscal Studies* **26**, 35–54.

Blanchard, O. (1985), 'Debt, Deficits and Finite Horizons', *Journal of Political Economy* **93**, 223–247.

Bohn, H. (2005), Who bears what risk? an intergenertional perspective. www.econ.ucsb.edu/ bohn.

Constantinides, G., Donaldson, J. & Mehra, R. (2002), 'Junior can't Borrow', *Quarterly Journal of Economics* **117**, 269–296.

Davidoff, T., Brown, J. & Diamond, P. (2005), 'Annuities and Individual Welfare', *American Economic Review* **95**, 1573–1590.

Demange, G. (2002), 'On optimality of intergenerational risk-sharing', *Economic Theory* **20**, 1–27.

Diamond, P. & Koszegi, B. (2003), 'Quasi-hypoerbolic Discounting and Retirement', *Journal of Public Economics* **87**, 1839–1872.

Finkelstein, A. & Poterba, J. M. (2002), 'Selection effects in the United Kingdom individual annuities market', *Economic Journal* **112**, 28–50.

Friedberg, L. & Webb, A. (2006), Life is Cheap: using Mortality Bonds to Hegde Aggregate Mortality Risk. NBER Working Paper No 11984.

Gordon, R. & Varian, H. (1988), 'Intergenerational risk-sharing', *Journal of Public Economics* **37**, 185–202.

Hardy, M. (2005), A Matter of Life and Death. Financial Engineering News: www.fenews.com/fen46/topics act analysis/topcis-act-analysis.htm.

Lando, D. (1998), 'On Cox Processes and Credit Risky Secuirities', *Review of Derivatives Research* **2**, 99–120.

Lee, R. & Carter, L. (1992), 'Modelling and Forecasting US Mortality', *Journal of the American Statistical Association* **87**, 659–671.

Lee, R. & Miller, T. (2002), 'Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality', *Demography* **38**, 537–549.

Leland, H. (1968), 'Saving and Uncertainty: the Precautionary Demand for Saving', *Quarterly Journal of Economics* **82**, 443–464.

Oeppen, J. & Vaupel, J. (2002), ‘Broken Limits to Life Expectancy’, *Science* **296**, 1029–1031.

Pensions Commission (2005), Second report: Appendix e. uncertainties in life expectancy projections. www.pensionscommission.org.uk/publications/2005/annrep/appendices/appendix-e.pdf.

Schrager, D. (2006), ‘Affine Stochastic Mortality’, *Insurance: Mathematics and Economics* **38**, 81–97.

Willets, R. (2004), ‘The Cohort Effect: Insights and Explanations’, *British Actuarial Journal* **10**, 833–877.

Wilmoth, J. & Horiuchi, S. (1999), ‘Rectangularization Revisited: Variability at Age of Death within Human Populations’, *Demography* **36**, 475–494.

Yaari, M. (1965), ‘Uncertain Life-time, Life Insurance and the Theory of the Consumer’, *Review of Economic Studies* **32**, 137–150.