Optimal DC Pension Fund Management and the Dangers of Longevity Risk

Colm Doyle *
Department of Banking & Finance
Smurfit School of Business
University College Dublin
Dublin, Ireland
Email: colm.doyle@ucdconnect.ie
Phone: +353 1 716 4350

John Cotter
Department of Banking & Finance
Smurfit School of Business
University College Dublin
Dublin, Ireland
Email: john.cotter@ucd.ie
Phone: +353 1 716 8900

Abstract

There is a pressing need to re-design the defined contribution pension scheme so that it can become an adequate replacement for the failing defined benefit scheme. This paper focuses on the importance of accounting for systematic longevity risk in light of this assertion. We discuss a proposed, plan member oriented, pension plan design that aims to deliver a desired standard of living in retirement through the use of the replacement ratio in the objective function of the optimal portfolio choice problem. We introduce an analytically tractable stochastic mortality model in order to facilitate working within a continuous-time dynamic programming framework. In order to gauge the impact of longevity risk we introduce, into the asset mix of the portfolio choice problem, a synthetic longevity-linked security. We determine the relative demand for this asset, and the value added through the introduction of this asset in terms of a utility gain. We find that for sufficiently risk averse plan members the value added by way of a reduction in conditional volatility is substantial, however these effects can be drowned out if the proportion of pension fund wealth allocated to risky assets is large.

EFM Classification: 370, 770

Keywords: defined contribution pension plan, longevity risk, optimal asset allocation, replacement ratio.

*Corresponding author and presenter.

1 Introduction

The high profile closure of many defined benefit (DB) pension schemes around the world has shed a spotlight on the numerable risks associated with pension funding. The closure of these schemes did nothing to mitigate these risks, but merely transferred them from the plan sponsor to the plan member, by accelerating the shift away from DB plans towards defined contribution (DC) plans. However, not only are the underlying risks essentially the same, the ultimate goal of both types of plan is, or at least should be, the same. This similarity is far from evident from the relative structure of the two types of plan. DB schemes were designed specifically for the purpose of providing a retirement income; and what's more, a retirement income that is directly related to the pre-retirement income level of the retiree so as to ensure that his or her standard of living is upheld. With a DC scheme it is largely left to the employee to figure out how much pension fund wealth he or she will need at retirement in order to provide a desired standard of living, and then principally determine the appropriate investment strategy required to deliver this fund.

Furthermore, with respect to the pension fund performance, with a DB plan the value of the pension fund assets is always expressed in terms of its liabilities. If the ratio of assets to liabilities exceeds one the fund is in surplus and if this ratio is less than one the fund is in deficit. Whether the fund is in surplus or deficit largely determines both the investment policy and contribution requirements. It appears that as DB schemes are dying out so too is this intuitive and reasoned approach to pension funding. By virtue of its reliance on the individual's own ability and resources the DC scheme does not facilitate an approach of this kind.

We propose that the DC plan should be designed so as to maintain the intuitive language of the DB scheme, without stipulating a transfer of funding risk back to the plan sponsor. The proposed design is based on an optimal dynamic asset allocation strategy so as to remove the burden of strategic decision making from the shoulders of the plan member. We measure the pension wealth in terms of the replacement ratio; that is, the ratio of retirement income to pre-retirement income, thereby giving each individual plan member a clearer understanding of the type of pension he or she is likely to receive. The asset allocation strategy is determined as that which maximises the expected value of this ratio. The expected replacement ratio acts as an implicit funding ratio, with the desired retirement income the implicit fund liability.

If we are to measure DC pension fund assets in terms of the fund's implicit liability, it is imperative that the uncertainty associated with this implicit liability is properly understood and accounted for. It was the failure to accurately account for the uncertainty associated with scheme liabilities that led to the current crisis in DB funding. There are three principal risk factors driving this uncertainty — interest rate risk, longevity risk and labour income risk. We focus in this paper on longevity risk, which, until recently, has been largely ignored by those considering optimal pension fund asset allocation, both in practice, and in the academic literature.

By longevity risk we are referring to the risk that a population of interest will live longer on average than expected. In this discussion the relevant population is the pool of annuitants on the annuity book of the life company underwriting the annuity contract used to define the replacement ratio. This type of longevity risk is what is referred to as systematic longevity risk, as opposed to idiosyncratic longevity risk, which is the risk that a given individual will live longer than he or she is expected to live given current expectations of future survival probabilities. Systematic longevity risk is essentially the risk that these expectations will change over time. Increasing life expectancies drive up the cost of annuities and hence pension fund liabilities, be they explicit (DB) or implicit (DC).

It is necessary, whenever attempting to account for longevity risk, to first identify a model that captures the time-varying dynamics of mortality rates. We derive a continuous time, affine mortality model, based on a mortality improvements following a Cox-Ingersoll-Ross process. By specifying the initial mortality intensity curve as a Gompertz law it is possible to derive an explicit functional form for the survival probabilities. The mathematical tractability of this model greatly facilitates working within the stochastic dynamic programming framework of the asset allocation model. The simulation of the processes governing the dynamics of the value function, and asset allocation strategies, is, as a result, much more efficient.

There are currently no efficient means by which to hedge longevity risk, although there is a market for longevity linked securities beginning to emerge, and an increased focus on the design and valuation of such securities. In order to gauge the impact of longevity risk we introduce, into the asset mix of the portfolio choice problem, a synthetic longevity-linked security, based on the structure of a longevity bond i.e. a bond with a principal repayment proportional to the mortality rate. We determine the relative demand for this asset, and the value added through the introduction of this asset in terms of a utility gain. We find that for sufficiently risk averse plan members the value added by way of a reduction in conditional volatility is substantial, however these effects can be drowned out if the proportion of pension fund wealth allocated to risky assets is large. We also express the value of the longevity hedge, in nominal terms, as the increase in contribution rate required to ensure that the probability of falling below some minimal permissible replacement ratio is the same in both the complete and incomplete markets.

The remainder of this paper is organised as follows: Section 2 discusses the need to rethink the current design of DC pension plans in light of the failures of the traditional DB plan. Section 3 focuses on the issues associated with developing the type of robust asset allocation strategy required for the proposed DC plan design. Section 4 introduces the continuous-time stochastic mortality model and provides the closed form solution for the survival probability (the derivation of which is presented in the appendix). Section 5 derives the optimal asset allocation strategies in the presence of longevity risk, governed by stochastic mortality model introduced in section 4. It also includes a simulation analysis that highlights the effect of the longevity risk through examining the demand for the longevity-linked asset, and corresponding utility gains and contribution rate effects. Section 6 concludes.

2 Redesigning the DC Plan

2.1 Problems with the traditional DC design

Broadly speaking there are two factors that determine the size of an individual's accumulated pension fund at the time of retirement — the amount of fund contributions, and the investment return. The defining feature of a DC plan is that the contribution rate is known in advance. The name 'defined contribution' however is principally derived from the fact that it is the employer's contribution that is known in advance. This in contrast to a DB plan in which the employer often needs to increase his/her/its contribution rate to counteract low investment returns or increases in plan liabilities. There are generally few restrictions DC plan members increasing their contributions, up to a point; such restrictions would unnecessarily discourage retirement saving.

There is a danger that DC plans could be mis-sold to the employees of a company on the basis that they "cost" less than a corresponding DB plan. This danger is made more apparent once we consider that individuals tend to exhibit a present bias in their time preferences (Benzoin et al., 1989, Laibson, 1997); that is, individuals tend to disproportionately value present consumption over future consumption. It is unlikely that a given plan member would willingly increase his or her contribution rate beyond a level which they perceive as absolutely necessary to obtain their desired standard of living in retirement. In light of this we argue that the primary objective of any DC plan design should be to frame the retirement saving problem in terms that make it explicit to the plan member what precisely is required of him or her in order to best obtain a particular standard of living in retirement.

The DC plan as currently designed implicitly requires that the plan member determine the investment strategy that will best deliver the desired standard of living. This is a highly complex optimisation problem that must account for an immense degree of uncertainty. The typical DC plan member could not hope to identify an optimal investment strategy in this environment. Huberman and Jiang (2006), in fact, identified a tendency for DC plan members to adopt an investment strategy that allocates an equal proportion of assets to each of the funds in which they are invested. Although this observation does not necessarily imply that plan members are irrational in their choices it is highly unlikely that such a naive approach would be investor optimal. The observation is made all the more revealing if one considers that the plan members studied are those that are actively involved in the management of their pension fund.

Cronquist and Thaler (2004), in studying the design of the partially privatised Swedish state pension system; highlighted that (after the initial advertising campaign to encourage citizens to take an active approach had subsided) in excess of 90% of new entries were entirely invested in the default fund. These results are similar in principle to those of Madrian and Shea (2001) who, in analysing the effect of auto-enrollment on participation rates in 401(k) plans, showed that the increase in the number of employees participating in the 401(k) plan as a result of auto-enrollment, brings with it a stark increase in the number of employees using the default investment option. The implications of this observation are growing increasingly important as auto-enrollment is becoming more and more popular with employers and governments alike, as a means of increasing participation rates in private pension arrangements.

2.2 A plan member oriented design

Merton (2007) proposed that DC plans should be designed "based on questions that most people find reasonable, such as the following: What standard of living do you desire to have? What standard of living are you willing to accept? What contribution or savings rate are you willing or able to make?" The principle behind this proposition is that the pension plan is designed with an intrinsic optimal dynamic asset allocation strategy that is sensitive to the answers to these questions. Given the answers to any two of these three questions the third would be determined by the asset allocation model. Or similarly, given the answer to any one of the three, the choices available for the other two would be restricted to within some upper and lower bounds, again determined by the asset allocation model. The plan member could vary the three determinants so as to find the combination that is most attractive to him or her. The dynamic asset allocation model is therefore the key component of this proposed plan design. Designing the plan based on an optimal dynamic asset allocation model would go a long way towards addressing the problems of disinterest and investor naivety among plan members discussed in the previous section.

The investment strategy in such a design would not be the concern of the plan member; it would be determined by the asset allocation model which would be developed and monitored by industry experts. In addition, the default investment option could be set so as to ensure an efficient accumulation of wealth that satisfies the retirement goals of the typical plan member with high probability. Blake et al. (2009) discuss designing a DC pension plan in this vein; that is with all technical complexity regarding investment policy and risk management hidden from the plan member. Their emphasis is on the importance of designing the plan with the express purpose of delivering a retirement income as opposed to a lump sum nominal amount.

The concept of a replacement ratio is important when developing a dynamic asset allocation model with this aim. The replacement ratio is the ratio of the level of retirement income to the level of pre-retirement income. The replacement ratio is widely considered to be the most appropriate proxy measure for a standard of living in retirement, if cost of living increases in retirement are factored in to the annuity payments, i.e. it is assumed that the plan member's income prior to retirement is sufficient to support the plan member's standard of living at that time. As the replacement ratio is measured in terms of an income per year (or month, week etc.) not simply a nominal amount, it is implicitly assumed that one can convert the accumulated wealth of their pension fund into a stream of income payments. This of course can be accomplished through the purchase of an annuity; or in the case of a DB plan, the purchase of an annuity can perfectly hedge the pension liability.

With a robust optimal dynamic asset allocation model in place the implications of the uncertainty associated with the state variables could be measured quantitatively in terms of the effect that possible unexpected changes might have on the realised replacement ratio. By generating a probability distribution for the replacement ratio, this information could then be communicated to the plan member using readily understood downside risk measures such as the fund's value-at-risk (VaR) or expected shortfall. If the plan member feels that the fund is overly risky he or she could choose to re-evaluate the desired replacement ratio and/or increase the contribution rate so as to facilitate a more risk-averse investment approach. Revisions to this balance could then be made at time intervals that suit the plan member so as that he or she could take advantage of positive movements, or correct for negative movements, in the underlying state variables. Considering the time preference bias characterised by most it would, perhaps, be beneficial to stipulate that the contribution rate cannot be reduced in the event of a positive movement and the risk-exposure cannot be increased in the event of a negative movement in the underlying variables. This would enable a general targeting of the specified desired replacement ratio.

3 Dynamic Asset Allocation and Longevity Risk

3.1 Asset allocation and the replacement ratio

The DC plan design discussed in the previous section rests on the development of a robust optimal asset allocation model. It was proposed that in developing such a model, pension fund wealth should be measured in terms of the replacement ratio, so as to give a more intuitive understanding of the type retirement the fund will support. Cairns et al. (2006a) present a model based on this principle — the *stochastic lifestyling* model, and this model forms the basis for the asset allocation model presented in section 5.

The optimal asset allocation strategy is derived within the dynamic programming framework, pioneered by Merton (1969). The objective function used is the expected value of the utility derived from the replacement ratio; that is, the fund wealth at retirement as measured as a percentage of the fund wealth required to deliver a retirement income equivalent to the pension plan member's pre-retirement income. The use of the replacement ratio in the objective function, in contrast to the nominal amount of the pension fund assets, introduces added uncertainty. The pre-retirement salary is unknown, and so too is the price of the annuity at the retirement date; if the asset allocation model were to be truly robust it is essential that this uncertainty be taken into account.

The incorporation of labour-income into the asset allocation problem is essential regardless of the consideration of the replacement ratio since the pension plan member's labour-income is what drives the contributions to the fund. We consider the case where labour markets are complete so that the associated risks can be fully hedged. This is so that the model remains analytically tractable within the continuous time dynamic programming framework. There are a number of papers concerned with the life-cycle portfolio choice problem with incomplete labour-income markets (see for example, Polkovnichenko, 2007, Viceira, 2001), but in a discrete time environment so as to facilitate numerical simulation and estimation. Munk and Sorenson (2010) explore the effect of stochastic labour-income in continuous time. They derive a closed form solution for the optimal asset allocation when labour-income is instantaneously correlated with the interest rate and asset prices. They also consider, however, the case where labour-income is non-hedgeable, solving the Hamilton-Jocobi-Bellman equations numerically using a backward induction technique. Our focus, however, is with the uncertainty associated with the other factor determining the replacement ratio — the annuity rate at retirement.

The annuity rate at retirement depends on both the term structure of interest rates (which determines the cost of each future payment) and the term structure of mortality rates (which determines the expected number of future payments) at that time. Both term structures are decidedly uncertain over long time horizons. The risk that mortality rates will change over time is, as mentioned, commonly referred to as systematic longevity risk. The plan member's exposure to longevity risk therefore enters the problem through the annuity rate i.e. the plan member is exposed to the risk that mortality rates will decrease unexpectedly over the life of plan driving up the cost of annuities, thereby driving down the realised replacement ratio from its expected value. Cairns et al. (2006a) do not account for systematic longevity in the development of the stochastic lifestyling model. The survival probabilities used in their valuation of the annuity are assumed to be deterministic, therefore the implicit mortality rates are also assumed to be deterministic.

3.2 The annuity decision

Implicit in the assumption that the replacement ratio is the perfect measure for pension fund wealth is the assumption that the plan member wishes to fully annuitise his or her wealth upon retirement. Yarri (1965) in his seminal work on the subject concluded that, under given assumptions (including no bequest motive and the availability of actuarially fair annuities), this action is indeed optimal. However, it is well observed that there are much fewer people participating in annuity markets than would be expected given the theoretical justifications for doing so. This observation has become known as the annuity market participation puzzle, and has been discussed at length in the literature (see for example, Davidoff et al., 2005, Inkmann et al., 2011).

Much of this literature attempts to explain the annuity market puzzle by considering an environment that is less restrictive than that proposed by Yarri (1965), e.g. the agent possesses a bequest motive, has access to equity markets in retirement, has greater information with regard to his or her own life expectancy than the annuity provider etc. This type of analysis is concerned with optimal consumption over the life-cycle, in the traditions of economic theory. We, however, omit the consumption variable entirely from our pension plan design problem; a decision that is largely attributable to behavioural considerations.

Under our design, the contribution rate is determined exogenously to the portfolio choice problem. It is unlikely that a pension plan that is designed based on a life-cycle portfolio choice and consumption model would be popular by virtue of the fact that the contribution rate would vary randomly over time. Regardless of any theoretical justification that the model is ultimately to their favour, individuals are likely to yield to their behavioural biases.

The particular behavioural bias presenting itself here is known in the literature as *mental accounting* (see Thaler, R. H., 1990, Choi et al., 2009). Broadly speaking, this is the tendency for individuals to disconnect financial decisions from one another; in this case, the tendency for individuals to disconnect their retirement saving and investment decisions from their consumption decisions. The assumption that individuals wish to fully annuitise their pension fund wealth does not preclude a bequest motive for example, but simply that the bequest is funded through means aside from the pension fund contributions, perhaps even partly through the annuity payments themselves, which of course could not be considered optimal within the traditional life-cycle problem.

3.3 Hedging longevity risk over the life-cycle

Historically research in the area of longevity risk and its effects on life-cycle saving, investment and consumption has focused primarily on idiosyncratic longevity risk; that is, the risk that the individual in question will outlive his or her savings. This line of research dates back to the seminal work of Yaari (1965) discussed in the previous section; and later, to that of Richard (1975), who did not consider the annuity decision but focused on the portfolio choice problem in the presence of a stochastic time of death. More recently, many studies have focused on the combined problem of life-cycle asset allocation and the hedging of idiosyncratic longevity risk.

The general approach taken by these studies is to incorporate illiquid annuities [1] as an asset class; be they traditional fixed annuities (e.g. Milevsky et al., 2007, Horneff et al., 2008) or investment linked variable annuities (e.g. Horneff et al., 2009, Koijen et al., 2011). The results from these studies are largely consistent and suggest a gradual hedging of longevity risk beginning many years prior to retirement and continuing into retirement. We assume, however, that a full annuitisation decision is embedded in the design of the DC pension plan, thereby fully hedging all exposure to idiosyncratic longevity risk.

Systematic longevity risk is relatively underexplored in the context of life-cycle asset portfolio choice problems. However, as systematic longevity risk has become more of a focus in the broader pension and insurance literature, and indeed, industry [2], we have begun to see its consideration in the context of life-cycle portfolio choice emerge as a subject of interest. Horneff et al. (2010) consider this issue in what is, in principle, an expansion on the earlier work by Horneff et al. (2008). They allow the survival probabilities to vary stochastically over time, and consider deferred annuities as opposed to immediate annuities, which provide for a more efficient tool for hedging systematic longevity risk.

Cocco and Gomes (2012) consider the life-cycle portfolio choice and consumption problem in the presence of systematic longevity risk, but allow for the hedging of that longevity risk through endogenous saving and retirement decisions. In addition, they consider an asset mix that includes a synthetic liquid mortality-linked security designed for the specific purpose of providing a longevity risk hedge. This is precisely the approach we adopt in section 5. In the absence of the mortality-linked asset the authors show that individuals would indeed seek to hedge against the longevity risk by saving more and retiring later, but this comes at a significant cost to utility.

The introduction of the mortality-linked asset enables a far more efficient hedge and thus reduces this cost tremendously. A paper with a similar focus to that of Cocco and Gomes (2012), and one which takes an analytical approach more in line with our analysis in section 5, is the earlier work of Menoncin (2008). He too introduced a synthetic mortality-linked security, namely a longevity bond, into the asset mix of a portfolio choice model, but in the absence of a consumption variable. The stochastic mortality model considered, however, was limited in it's capacity to capture true mortality dynamics, and therefore too, in the potential for longevity risk analysis. Our analysis is based on a stochastic mortality model far more in keeping with true mortality dynamics.

[1] Illiquid annuities in this context means that the annuity purchase is an irreversible transaction.

[2] The Life & Longevity Markets Association (LLMA), established to promote a liquid traded market in longevity risk in 2010, now has as members: AVIVA, AXA, Deutsche Bank, J.P. Morgan, Legal & General, Morgan Stanley, Munich Re, Pension Corporation, Prudential PLC, RBS and Swiss Re.

4 The Stochastic Mortality Model

4.1 Mortality modelling

Large unexpected increases in life expectancy over the last 50 years or so contributed greatly to the current crisis in DB funding levels. Since these increases were not factored into the liability calculations liabilities were drastically underestimated. Plan sponsors were, as a result, required to take action in order to decrease the mounting deficits. Cocco and Gomes (2012) note that members of the Calpers benefits and programme administration committee have cited changes to post-retirement mortality assumptions as the most influential force driving increases in employer contributions. It is reasonable to assume that these changes also, in part, drove the large increases in the equity exposure observed over the last two decades, given that one cannot increase contribution rates retrospectively.

These increases in the allocations to equities had dramatic effect in the wake of the stock market crashes of 2000 and 2008. As discussed in section 2, the proposed design of DC schemes closely relates to the design of traditional DB schemes; in that, the fund is managed with the expressed aim of delivering a replacement ratio. It is imperative that we recognise the faults with the traditional DB design so that we can correct for them in any new DC design; the treatment of longevity risk clearly warrants attention in this regard.

In order to give an accurate portrayal of these dangers it is essential that we identify a mortality model that captures the characteristics of true mortality dynamics. There is a vast literature focused on the forecasting of future mortality rates but, as studies such as Currie et al. (2004) show, it is very difficult to do so with any accuracy. There are many popular discrete time mortality models that can be shown to fit historical data well and which are straightforward to simulate (e.g. Lee and Carter, 1992, Cairns, Blake and Dowd, 2006b) but they do not lead to analytical formula for spot survival probabilities.

In the continuous time dynamic programming framework we require a continuous time model that is analytically tractable. There is a class of mortality models that have come to prominence recently for this very reason; they are the class of *affine* mortality models (see for example, Schrager, 2006). This class of models is characterised by assumption that the logarithm of the survival probability can be represented as an affine function of the stochastic force of mortality. Dahl and Møller (2006) propose a particularly flexible affine mortality model that allows the stochastic force of mortality to follow a time-inhomogeneous Cox-Ingersoll-Ross (CIR) model. This model provides the framework for our stochastic mortality model.

4.2 The model

We take as a starting point an initial curve for the mortality intensity, $\lambda_0(x)$, for an individual aged x at time 0. Future mortality is then viewed as a stochastic process $\lambda(x,t)_{t\epsilon[0,T]}$, where $\lambda(x,0) = \lambda_0(x)$.[3] We model changes in the mortality intensity via a stochastic mortality improvement process $\zeta(x,t)$, with $\zeta(x,0) = 1$ for all x . The mortality process is determined as the product of the deterministic component of the mortality

[3] In the asset allocation problem that is the focus of this paper, time 0 would be the time at which the pension plan member joined the plan, and time T the time of retirement.

intensity, determined by the initial mortality curve, and the mortality improvement process, i.e. $\lambda(x,t) = \lambda_0(x+t)\zeta(x,t)$.

If the mortality improvement process is allowed to follow a CIR process,

$$d\zeta(x,t) = \big(\theta - \delta\zeta(x,t)\big)dt + \sigma_\zeta\sqrt{\zeta(x,t)}dZ_\zeta(t) \tag{1}$$

(assuming that $2\theta(x,t) \geq \sigma_\zeta(x,t)^2$ so that the process remains strictly positive), it is straightforward to show (by application of Itō's lemma) that the mortality process follows:

$$d\lambda(x,t) = (\theta_\lambda(x,t) - \delta_\lambda(x,t)\lambda(x,t))dt + \sigma_\lambda(x,t)\sqrt{\lambda(x,t)}dZ_\zeta(t) \tag{2}$$

where,

$$\theta_\lambda(x,t) = \theta\lambda_0(x+t) \;\; ; \;\; \delta_\lambda(x,t) = \delta - \frac{\frac{d}{dt}\lambda_0(x+t)}{\lambda_0(x+t)} \;\; ; \;\; \sigma_\lambda(x,t) = \sigma_\zeta\sqrt{\lambda_0(x+t)} \tag{3}$$

which is a time-inhomogeneous CIR process. Now, following Dahl and Møller (2006), we define the survival probability $F(x,t,T,\zeta(t))$; that is the probability at time t that an individual aged x at time 0 will live to time T, by

$$F(x,t,T,\zeta(t)) = \mathrm{E}\Big[e^{-\int_t^T \lambda(x,\tau)d\tau} \mid \mathcal{I}(t)\Big] \tag{4}$$

where $\mathcal{I}(t)$ is the natural filtration of the process $\zeta(x,t)$. We also define the corresponding martingale

$$M(t,T,\zeta(t)) = \mathrm{E}\left[e^{-\int_0^T \lambda(\tau)d\tau}\middle|\mathcal{I}(t)\right] = e^{-\int_0^t \lambda_0(\tau)\zeta(\tau)d\tau}F(t,T,\zeta(t)) \tag{5}$$

Applying Itō's lemma to $M(t,T,\zeta(t))$ we find that $F(x,t,T,\zeta(t))$ satisfies a Black-Scholes type PDE of the form

$$\begin{aligned}\frac{\partial}{\partial t}F(x,t,T,&\zeta(t)) + (\theta_\lambda(x,t) - \delta_\lambda(x,t)\lambda)\frac{\partial}{\partial\lambda}F(x,t,T,\zeta(t)) \\ &+ \frac{1}{2}(\sigma_\lambda(x,t))^2\frac{\partial^2}{\partial\lambda^2}F(x,t,T,\zeta(t)) - \lambda F(x,t,T,\zeta(t)) = 0\end{aligned} \tag{6}$$

Since $\lambda(x,t)$ follows a time-inhomogeneous CIR process the solution to this PDE has an exponential affine form; that is,

$$F(x,t,T,\zeta(t)) = e^{\alpha(x,t,T)-\beta(x,t,T)\lambda(x,t)} \tag{7}$$

It is straightforward to show that $\alpha(x,t,T)$ and $\beta(x,t,T)$ must therefore satisfy

$$\frac{\partial}{\partial t}\beta(x,t,T) = \delta_\lambda(x,t)\beta(x,t,T) + \frac{1}{2}(\sigma_\lambda(x,t))^2\beta(x,t,T)^2 - 1 \tag{8}$$

$$\frac{\partial}{\partial t}\alpha(x,t,T) = \theta_\lambda(x,t)\beta(x,t,T) \tag{9}$$

with $\qquad \beta(x,T,T) = 0 \quad$ and $\quad \alpha(x,T,T) = 0$

The choice of initial mortality curve $\lambda_0(x)$ is crucial, both with regard to ensuring that the model accurately represents the real world mortality term structure, and in determining whether analytic solutions to equations (8) and (9) can be found. A Gompertz law

$$\lambda_0(x) = \frac{1}{b(x)} e^{\frac{1}{b(x)}(x-m(x))} \tag{10}$$

for, cohort dependent, constant parameters $m(x)$ and $b(x)$, as in Milevsky (2001), succeeds on both counts.

Proposition 1. If we define the initial mortality λ_0 as in (10), and allow the mortality improvement process to follow a CIR process as in (1), then

$$\beta(x,t,T) = \frac{2b}{z'(t)}\left[\frac{I_{-\nu+1}(z'(T))I_{\nu-1}(z'(t)) - I_{\nu-1}(z'(T))I_{-\nu+1}(z'(t))}{I_{\nu-1}(z'(T))I_{-\nu}(z'(t)) - I_{-\nu+1}(z'(T))I_{\nu}(z'(t))}\right] \tag{11}$$

$$\alpha(x,t,T) = -\frac{\theta\delta}{\sigma_\zeta^2}(T-t) + \ln\left(\frac{I_{-\nu+1}(z'(T))I_{\nu}(z'(t)) - I_{\nu-1}(z'(T))I_{-\nu}(z'(t))}{I_{-\nu+1}(z'(T))I_{\nu}(z'(T)) - I_{\nu-1}(z'(T))I_{-\nu}(z'(T))}\right)^{\frac{\theta\delta}{\sigma_\zeta^2}} \tag{12}$$

where

$z'(t) = 2\sigma_\zeta b\sqrt{\lambda_0(x+t)}$, and $I_\nu(x)$ is the modified Bessel function of the first kind, with $\nu = \delta b$.

Proof. See Appendix A.

This closed form representation for the survival function, $F(t,T,\zeta(t))$, allows for the efficient estimation of survival probabilities and life expectancies, which in turn can be used to value annuities and other mortality contingent claims. For the remainder of the paper we will suppress the explicit x dependence as our portfolio analysis focuses on one particular age cohort, with all parameter values estimated to correspond to that particular cohort. What remains is to determine these appropriate parameter estimates.

4.3 Mortality dynamics under the model

4.3.1 Calibration

The calibration of any stochastic mortality model is always a contentious issue. The past century has seen life-expectancies increase at an unprecedented rate [4] due to medical and technological advances. Although there is little or no data available dating back beyond the last century, considering that humans have been on the earth for tens of thousands of years it is easy to argue that such a rate of increase is uncharacteristic of mortality dynamics in general. However, the undoubted causal relationship between technological advancement and life-expectancy presents a counter point to this argument, if one accepts

[4] Life expectancy at birth for males in the US increased from 58 to 76 in the years from 1933 to 2009 (figures from the Human Mortality Database).

that the rate of technological advancement is set to continue. Then again, "technological advancement" is such a broad term, and it could be argued that the observed relationship with life-expectancy was driven by a relatively small number of monumental advances, advances of the kind that are unlikely to be seen again. This type of back and forth is typical of a broad debate focused on the forecasting of future mortality rates, and the relative importance of recent mortality trends to this endeavour.

Despite this controversy however, there is some what of a consensus (albeit a tentative one) in the stochastic mortality modeling literature when it comes to model calibration (see for example Lee and Carter, 1992, Cairns, Blake and Dowd, 2006b); that is, to assume that the mortality dynamics of the last half century or so are indicative of mortality dynamics going forward. Since we are concerned with the effect that changing mortality rates have on the value of an implicit pension fund liability, whether the mortality model can predict the true mortality rates is secondary to whether the model can reflect the expectations of the market for annuities (while being flexible enough to accommodate any realised mortality improvements). The consensus approach is therefore the most appropriate approach for our purposes.

For the calibration of the model introduced in the previous section, we use the death rates of US males over the 40 year period (chosen to match the life of the pension fund we consider in section 5) from 1969 to 2009 as a proxy for the force of mortality over this time[5]. The model structure calls first for the calibration of the base mortality curve $\lambda_o(t)$, which was specified by the Gompertz law (10). It is well observed that the Gompertz law fails to accurately model the true mortality rates for both very young and very old individuals. Therefore in order for the law to hold on average, these rates must be excluded from the calibration.

The calibration of the mortality improvement process is not as straight forward as that of the base mortality curve. The principle reason for this, aside from the fact that the structure itself is more complex, is that in modelling parallel shifts to the mortality curve we are implicitly assuming that all shocks to mortality affect all age groups equally. This does not reflect what has been observed to be the case in reality; that is, that mortality rates decline faster in response to shocks for younger age groups (see for example, Lee and Carter, 1992). As a result we can neither follow the time-varying dynamics of the mortality rate for one specific age group (as this would disproportionately favour the specific age group), nor can we follow that of the mortality rate for an individual over his life-time (as this would induce additional volatility through the variation across the age groups).

To circumvent this problem we take the change in area under the mortality curve (or more accurately the estimated Gompertz curve) as a proxy measure for the parallel shift to the curve.

Insert figure 1 here

Figure 1 shows the mortality improvements based on the individual age groups (blue lines), an individual aged 25 at time 0 (red line), and the proxy measure just described (green line). It is clear that the variance across the age groups is too severe to consider any one group as representative of the whole, and the jaggedness of the line corresponding to the individual, presents a degree of volatility not in keeping with the typical age

[5]Data obtained from the Human Mortality Database.

grouping, as we would expect. The proxy measure however finds a balance between these two extremes.

A further structural drawback of the model that should be taken into account before calibration is that the CIR process governing the mortality improvement dynamics is mean reverting. This implies that if mortality improvements over a given time interval were faster than expected, the probability of further mortality improvements over subsequent time intervals would be reduced; we would not typically expect this type of behaviour. However, if we set $\theta = \frac{1}{2}\sigma_\zeta^2$, i.e. the minimum permissible value, we can ensure that the mean reversion effect is minimal.

Insert table 1 here

Table 1 shows the maximum likelihood parameter estimates for the CIR process subject to this constraint, along with the least-squares parameter estimates for the base mortality curve.

4.3.2 Comparison with the Lee-Carter model

In order to show that our model is comparable in character to the types of models used in practice we present here a direct comparison with the Lee-Carter model (Lee and Carter, 1992). The Lee-Carter model assumes that the mortality rates follow:

$$\ln\big(\lambda(x,t)\big) = \phi(x) + \eta(x)k(t) + \epsilon_{x,t} \tag{13}$$

where $k(t)$ determines the general change to mortality over time, $\eta(x)$ determines to what relative degree the mortality rate for the age group x responds to this general mortality change, $\phi(x)$ is average over time of the log of the mortality rates for the age group x, and $\epsilon_{x,t}$ is an error term. Ordinary regression methods do not apply to the Lee-Carter model as there are no given regressors. Singular Value Decomposition (SVD) is instead the method of choice for parameter estimation. Lee and Carter (1992) outline a close approximation to SVD, which suffices for our purposes here.

Typically, when forecasting using the Lee-Carter model, the time varying component $k(t)$ is set to follow a random-walk with drift. We therefore let $k(t)$ be determined by

$$k(t) = \mu_k + k(t-1) + \sigma_k \varepsilon_t \tag{14}$$

where μ_k is the drift term and ε_t is a standard normal random variable. Upon calibration we find parameter estimates for μ_k and σ_k of -1.4305 and 2.899 respectively.

Insert figure 2 here

To show that our model is similar in character to the Lee-Carter model we have estimated survivor and longevity fan charts using both models. These charts can be seen in figure 2, with those on the left estimated using our model, and those on the right estimated using the Lee-Carter model. Survivor and longevity fan charts are two alternative ways for representing graphically the uncertainty associated with future mortality rates.

The survivor fan chart in figure 2 (top) shows the probability that an man aged 65 will survive to any given age, for 1000 different simulated paths. The longevity fan chart (bottom) shows how the life expectancy of men aged 65 evolves over a 40 year period, again for 1000 simulated paths. There is an important difference between these two charts and that is that the survivor fan chart relates to just one age cohort (men aged 65 at time 0), whereas the longevity fan chart considers multiple cohorts (40 to be exact; men aged 65 at time 0, men aged 65 at time 1, etc.).

Our model implicitly assumes that mortality improvements are the same for all age cohorts, as does the Lee-Carter model. It is known, though, that mortality dynamics differ across different age cohorts and most recent mortality models do in fact include an age cohort parameter (e.g. Cairns, Blake and Dowd, 2006b); however since we are concerned primarily with defined contribution pension schemes which tailor to individuals, the age cohort of the individual plan member is all that is relevant.

We include the longevity fan chart in figure 2 because it gives a more intuitive understanding of mortality uncertainty, giving a clear indication that the our model allows for considerable variation in future mortality rates. If we compare the graphs on the left of figure 2 with those on the right, we see that our model represents mortality dynamics broadly similar to those represented by the Lee-Carter model. The observed differences at advanced ages are a result of estimation difficulties innate to mortality rate estimation. The Lee-Carter model attempts to fit the sparse data subset of advanced age death rates whereas our model effectively extrapolates from the data set excluding these rates. There is currently no consensus as to what constitutes best practice in this regard. It is reasonable to posit, therefore, that we have identified a continuous-time stochastic mortality model that is both analytically tractable and representative of true mortality dynamics, in so far as they can be modelled at present. This will enable us to model longevity risk realistically in the pension fund asset allocation problem addressed in the next section.

5 The Impact of Longevity Risk on the Replacement Ratio

In this section we look to analyse the impact of systematic longevity risk on the replacement ratio delivered by an optimal dynamic pension fund asset allocation strategy. We will consider the pension funding problem of an individual DC plan member who joins the plan at age 25, retires at age 65 and uses the total accumulated pension fund wealth at that time to purchase an annuity. The plan member is assumed to contribute a fixed proportion of his salary to the pension fund throughout his working life.[6] We will not consider the plan member's time of death since the purchase of the annuity renders the consideration financially immaterial.

As discussed in section 2, when dealing with pension plans, utility is derived from the replacement ratio as opposed to a nominal cash amount; therefore, when looking to determine the optimal asset allocation underlying the design of the proposed plan member oriented pension plan design, it is necessary to account for this. We therefore look to identify the asset allocation strategy which maximises

$$U\left(\frac{W(T)}{Y(T)a(T,\zeta(T))}\right) \tag{15}$$

[6] We will consider the plan member to be male so as to be consistent with the mortality estimates of section 4

where T is the time of retirement; $U(\cdot)$ is the plan member's utility function incorporating his individual level of risk aversion; $W(T)$ is the terminal nominal wealth of the pension fund; $Y(T)$ is the plan member's salary immediately prior to retirement; and $a(T, \zeta(T))$ is the price of an annuity at retirement that pays one unit of wealth per year for the remainder of the plan member's life. We notice here that the price of the annuity is dependent on the evolution of the mortality improvement state variable $\zeta(t)$.

The price of the annuity at retirement is given by

$$a(T, \zeta(T)) = \sum_{\tau=T}^{\infty} e^{-r(\tau-T)} F(T, \tau, \zeta(T)) = \sum_{\tau=T}^{\infty} e^{-r(\tau-T)} e^{\alpha(T,\tau)-\beta(T,\tau)\lambda_0(T)\zeta(T)}$$

We will let the annuity rate prior to retirement, $t < T$, be given by

$$a(t, \zeta(t)) = e^{-r(T-t)} \sum_{\tau=T}^{\infty} e^{-r(\tau-T)} F(t, \tau, \zeta(t)) \tag{16}$$

This is the fair price of the corresponding deferred annuity; that is, the price at which the pension plan member could exchange the accumulated pension fund wealth at time $t < T$ for a stream of annual payments commencing at retirement. This price factors in the possibility that the plan member could die before retirement, which might seem at first to contradict our assumption that the plan member's time of death is irrelevant. However this assumption relates to the plan member only, which is reasonable since the entire notion of retirement planning is built on the premise that there will actually be a retirement period. A life company calculating the fair value of a deferred annuity on the other hand, would factor in the possibility of premature death; (16) merely reflects this consideration. Applying Itō's lemma we get

$$\mathrm{d}a(t, \zeta) = a(t, \zeta)\Big[\big(r + \lambda_0(t)\zeta\big)\mathrm{d}t + \sigma_\zeta\sqrt{\zeta}\psi(t, \zeta)\mathrm{d}Z_\zeta(t)\Big] \tag{17}$$

where we have defined $\psi(t, \zeta) = \dfrac{1}{a(t, \zeta)}\dfrac{\partial}{\partial \zeta}a(t, \zeta)$ as the semielasticity of the annuity price with respect to the mortality improvement factor.

A significant advantage gained from working with the stochastic mortality model derived in the previous section is that we have a analytical expression for the transition densities of the mortality improvement process (by virtue of it following a CIR process). Using this analytical expression we can derive a corresponding analytical expression for the conditional expectation of the annuity price at retirement (see appendix B).

$$\begin{aligned}\mathrm{E}\Big[a\big(T, \zeta(T)\big) \mid \mathcal{I}(t)\Big] = \sum_{\tau=T}^{\infty} & \frac{e^{-r(\tau-T)+\alpha(T,\tau)}}{1 + \frac{\sigma_\zeta^2}{2\delta}(1 - e^{-\delta(T-t)})\lambda_0(T)\beta(T, \tau)} \\ & \cdot \exp\Big\{\frac{-\zeta(t)e^{-\delta(T-t)}\lambda_0(T)\beta(T, \tau)}{1 + \frac{\sigma_\zeta^2}{2\delta}(1 - e^{-\delta(T-t)})\lambda_0(T)\beta(T, \tau)}\Big\}\end{aligned} \tag{18}$$

We will now consider the optimal asset allocation strategy. We will first consider the incomplete market case; that is, the case where the longevity risk cannot be hedged. We will then compare this to the complete market case so as to gain an understanding of the value of the longevity hedge.

5.1 The incomplete market case

We allow the pension fund to invest in two assets, a risk-free asset, or bond, B, and a risky asset, or stock, S; and the plan member seeks to maximise his expected terminal utility (utility at retirement) by varying the proportions of pension fund wealth invested in each according to the optimal asset allocation policy. The growth of the risk free asset is assumed to be deterministic i.e. $B(t) = B(0)\exp(rt)$ and the price of the risky asset is assumed to satisfy

$$\mathrm{d}S(t) = S(t)\Big[(r + \xi\sigma_S)\mathrm{d}t + \sigma_S \mathrm{d}Z_S(t)\Big] \tag{19}$$

where $Z_S(t)$ is a standard Brownian motion independent of $Z_\zeta(t)$, and $\xi\sigma_S$ is the risk premium on the asset i.e. ξ is the market price of risk.

Now, in order to specify the dynamics of the replacement ratio we must define the salary dynamics, and determine the wealth dynamics. We assume that all uncertainty associated with plan member's future salary is driven by $Z_S(t)$, and therefore can be hedged through an appropriate investment in the risky asset. The plan member's salary, $Y(t)$ is governed by

$$\mathrm{d}Y(t) = Y(t)\Big[(r + \mu)\mathrm{d}t + \sigma_Y \mathrm{d}Z_S(t)\Big] \tag{20}$$

where σ_Y allows for the possible correlation between salary and equity returns. If we denote the proportion of pension fund wealth invested in the generic risky asset at time t by $p(t)$ the wealth process is given by

$$\mathrm{d}W(t) = W(t)\Big[\big(r + p(t)\xi\sigma_S\big)\mathrm{d}t + p(t)\sigma_\zeta S\mathrm{d}Z_S(t)\Big] + \pi Y(t)\mathrm{d}t$$

We can now define the replacement ratio, $X(t)$:

$$X(t, \zeta) = \frac{W(t)}{Y(t)a(t, \zeta(t))} \tag{21}$$

We find, again using Itō's lemma (multiple times),

$$\begin{aligned}\mathrm{d}X(t, \zeta) = \frac{\pi}{a(t, \zeta)}\mathrm{d}t + X(t)\Big[\Big(&-\big(r + \lambda_0(t)\zeta + \mu\big) + \sigma_Y^2 + p(t)\sigma_S(\xi - \sigma_Y) \\ &+ \psi^2(t, \zeta)\sigma_\zeta^2\zeta\Big)\mathrm{d}t - \psi(t, \zeta)\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(t) + (p(t)\sigma_S - \sigma_Y)\mathrm{d}Z_S(t)\Big]\end{aligned} \tag{22}$$

We define the value function as:

$$V(t,X,\zeta) = \sup_{p\in\mathcal{Q}} \mathrm{E}_P\Big[U\big(X(T,\zeta(T))\big)\Big|\mathcal{G}(t)\Big] \tag{23}$$

where E_P is the expectation with respect to the real world probability measure P, $\mathcal{G}(t)$ is the filtration containing all information up to time t and $\mathcal{Q}$ is the set of all admissible asset allocation strategies.

The Hamilton-Jacobi-Bellman (HJB) equation of optimality is therefore

$$\begin{aligned}\sup_q \Bigg[V_t + &\Bigg(\frac{\pi}{a(t,\zeta)} + x\Big(-\big(r+\lambda_0(t)\zeta+\mu\big) + \sigma_Y^2 + p(t)\sigma_S(\xi-\sigma_Y) + \psi^2(t,\zeta)\sigma_\zeta^2\zeta\Big)\Bigg)V_x \\ &+ \big(\theta-\delta\zeta\big)V_\zeta + \frac{1}{2}x^2\Big((p(t)\sigma_S-\sigma_Y)^2 - \psi^2(t,\zeta)\sigma_\zeta^2\lambda_0(t)\zeta\Big)V_{xx} \\ &- \psi(t,\zeta)\sigma_\zeta^2\lambda_0(t)\zeta V_{x\zeta} + \frac{1}{2}\sigma_\zeta^2\lambda_0^2(t)\zeta V_{\zeta\zeta}\Bigg] = 0\end{aligned} \tag{24}$$

subject to the boundary condition $V(T,X(T),\zeta) = U\big(X(T)\big)$.[7]

If we solve the first order condition for $p(t)$, and if we denote by $p^*(t)$ the resulting optimal asset allocation strategy, we find that

$$p^*(t) = \frac{\sigma_Y}{\sigma_S} - \frac{(\xi-\sigma_Y)}{\sigma_S}\frac{V_x}{xV_{xx}} \tag{25}$$

Inserting this expression into the HJB equation yields the following PDE:

$$\begin{aligned}V_t + V_x\Big(\frac{\pi}{a(t,\zeta)} + x\big(-(r+\lambda_0(t)\zeta+\mu) + \xi\sigma_Y + \psi^2(t,\zeta)\sigma_\zeta^2\zeta\big)\Big) + V_\zeta\big(\theta-\delta\zeta\big) \\ -\frac{1}{2}\frac{V_x^2}{V_{xx}}(\xi-\sigma_Y)^2 + \frac{1}{2}x^2V_{xx}\psi^2(t,\zeta)\sigma_\zeta^2\zeta - xV_{x\zeta}\psi(t,\zeta)\sigma_\zeta^2\zeta + \frac{1}{2}V_{\zeta\zeta}\sigma_\zeta^2\zeta = 0\end{aligned} \tag{26}$$

Following Cairns et al. (2006a), we first consider the case where $\pi = 0$, i.e. the case where there is a single contribution at time zero. Furthermore, we will restrict ourselves to the case of power utility, i.e.

$$U\big(X(t)\big) = \begin{cases} \frac{1}{\gamma}\big(X(t)\big)^\gamma & \gamma<1, \gamma\neq 0 \\ \log\big(X(t)\big) & \gamma = 0\end{cases}$$

[7]Here $V_t = \frac{\partial}{\partial t}V(t,x,\zeta)$; $V_x = \frac{\partial}{\partial x}V(t,x,\zeta)$; $V_{xx} = \frac{\partial^2}{\partial x^2}V(t,x,\zeta)$; $V_\zeta = \frac{\partial}{\partial\zeta}V(t,x,\zeta)$; $V_{\zeta\zeta} = \frac{\partial^2}{\partial\zeta^2}V(t,x,\zeta)$; $V_{x\zeta} = \frac{\partial}{\partial\zeta}\Big(\frac{\partial}{\partial x}V(t,x,\zeta)\Big)$

(where $1-\gamma$ is the coefficient of relative risk aversion).

In choosing to work with power utility we have that the value function is homogeneous in the replacement ratio and therefore takes the form

$$V(t,x,\zeta) = \frac{1}{\gamma}g(t,\zeta)x^{\gamma} \tag{27}$$

with $g(T,\lambda) = 1$ for all λ. This leads us to the following proposition.

Proposition 2. The value function takes the form

$$V(t,x,\zeta) = \frac{1}{\gamma}x^{\gamma}\exp\Big\{\gamma\Big(-\mu+\xi\sigma_Y+\frac{1}{2}\frac{1}{(1-\gamma)}(\xi-\sigma_Y)^2\Big)\big(T-t\big)\Big\} \cdot a\big(t,\zeta\big)^{\gamma}\mathrm{E}_P\Big[a\big(t,\zeta(T)\big)^{-\gamma}\Big|\mathcal{G}(t)\Big] \tag{28}$$

Proof. See Appendix C.

Now, by extension of theorem 3.4.1 of Cairns et al. (2006a) we find that, for the case where $\pi > 0$,

$$V(t,x,\zeta) = \frac{1}{\gamma}\Big(x+\frac{\pi f(t)}{a\big(t,\zeta\big)}\Big)^{\gamma}\exp\Big\{\gamma\Big(-\mu+\xi\sigma_Y+\frac{1}{2}\frac{1}{(1-\gamma)}(\xi-\sigma_Y)^2\Big)\big(T-t\big)\Big\} \cdot a\big(t,\zeta\big)^{\gamma}\mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\gamma}\Big|\mathcal{G}(t)\Big] \tag{29}$$

where

$$\pi f(t) = \frac{1}{Y(t)}\mathrm{E}_Q\left[\int_t^T e^{-r(s-t)}\pi Y(s)\mathrm{d}s\Big|\mathcal{F}(t)\right] = \pi\frac{\exp\Big[(\mu_Y-\xi\sigma_{Y_1})(T-t)\Big]-1}{\mu_Y-\xi\sigma_Y} \tag{30}$$

$\mathcal{F}(t)$ here is the filtration generated by $Z_S(\tau)$ up to time t and Q is the unique risk neutral measure. $\pi Y(t)f(t)$ is the market price at time t for the premiums payable between t and T.

From (25) we find that

$$p^*(t) = \frac{\sigma_Y}{\sigma_S}+\frac{(\xi-\sigma_Y)}{\sigma_S}\frac{1}{1-\gamma}\Big(1+\frac{\pi Y(t)f(t)}{W(t)}\Big) \tag{31}$$

This is precisely the same optimal asset allocation we would find in the case where wealth is measured solely in terms of salary. Thus, background longevity risk does not impact the asset allocation in our framework when the market is incomplete. It does however impact the realised replacement ratio as we would expect.

5.2 The complete market case

5.2.1 The mortality-linked security

In order to gauge the impact of the longevity risk we introduce a synthetic mortality-linked asset, $L(t,\zeta)$, that permits a perfect longevity hedge, and look to determine the relative demand for this asset. We define $L(t,\zeta)$ as follows:

$$L(t,\zeta(t)) = \sum_{\tau=t}^{\infty} e^{-r(\tau-t)} M(t,\tau,\zeta(t)) \tag{32}$$

where $M(t,\tau,\zeta(t))$ is as defined in (5), with

$$\mathrm{d}M(t,T,\zeta(t)) = -\sigma_\zeta \lambda_0(t)\sqrt{\zeta(t)}\beta(t,T)M(t,T,\zeta(t))\mathrm{d}Z_\zeta(t) \tag{33}$$

This asset can be viewed as a portfolio of 'longevity bonds' with maturities matching the expected annuity payments. A longevity bond is generally defined as a bond whose payout is proportional to the number of individuals alive at maturity in some underlying population of individuals. We assume here that the payout from one of the longevity bonds making up the portfolio $L(t,\lambda)$, is equal to the "realised" survival probability. This is analogous to the number of individuals surviving from a given population, in that the survival probability dynamics are driven by the number of people surviving from one period to the next relative to the number predicted; however the population driving the mortality rate dynamics must necessarily be much larger than that underlying any given longevity bond, covering all age cohorts. We have omitted a longevity risk premium due to the computational advantages gained by doing so. These computational advantages outweigh any potential informational value the inclusion of a risk premium would add considering the comparative nature of our analysis, and the role of the longevity bond as solely an efficient hedging tool.

Applying Itō's lemma to (32), we get

$$\mathrm{d}L(t,\zeta) = L(t,\zeta)\big[r\mathrm{d}t + \sigma_\zeta\sqrt{\zeta}\psi(t,\zeta(t))\mathrm{d}Z_\zeta(t)\big] \tag{34}$$

It is clear that an investment in the asset $L(t,\lambda)$ is sufficient to provide a complete hedge against the longevity risk exposure induced by the use of the replacement ratio as the measure of pension fund wealth.

5.2.2 The optimal longevity hedging strategy

If we denote again the proportion of pension fund wealth invested in the risky asset $S(t)$ at time t by $p(t)$, and that invested in the asset $L(t,\lambda)$ by $q(t)$ (and therefore the proportion invested in the risk-free asset by $1-p(t)-q(t)$), the wealth process is given by

$$\mathrm{d}W(t) = W(t)\Big[\big(r + p(t)\xi\sigma_S\big)\mathrm{d}t + q(t)\psi(t,\zeta(t)\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(t) + p(t)\sigma_\zeta S\mathrm{d}Z_S(t)\Big] + \pi Y(t)\mathrm{d}t$$

The replacement ratio, (21), therefore follows:

$$\mathrm{d}X(t,\zeta) = \frac{\pi}{a(t,\zeta)}\mathrm{d}t + X(t)\Big[\Big(- \big(r + \lambda_0(t)\zeta + \mu\big) + \sigma_Y^2 + p(t)\sigma_S(\xi - \sigma_Y) - \big(q(t)-1\big)\psi^2(t,\zeta)\sigma_\zeta^2\zeta\Big)\mathrm{d}t + \big(q(t)-1\big)\psi(t,\zeta)\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(t) + (p(t)\sigma_S - \sigma_Y)\mathrm{d}Z_S(t)\Big] \tag{35}$$

and the HJB equation of optimality is

$$\begin{aligned}\sup_q\Bigg[V_t + \Bigg(\frac{\pi}{a(t,\zeta)} + x\Big(- \big(r + \lambda_0(t)\zeta + \mu\big) + \sigma_Y^2 + p(t)\sigma_S(\xi - \sigma_Y)\\ - \big(q(t)-1\big)\psi^2(t,\zeta)\sigma_\zeta^2\zeta\Big)\Bigg)V_x + \big(\theta - \delta\zeta\big)V_\zeta + \frac{1}{2}x^2\Big((p(t)\sigma_S - \sigma_Y)^2 \qquad (36)\\ + \big(q(t)-1\big)\psi^2(t,\zeta)\sigma_\zeta^2\zeta\Big)V_{xx} + \Big(\big(q(t)-1\big)\psi(t,\zeta)\sigma_\zeta^2\lambda_0(t)\zeta\Big)V_{x\zeta} + \frac{1}{2}\sigma_\zeta^2\lambda_0^2(t)\zeta V_{\zeta\zeta}\Bigg] = 0\end{aligned}$$

If we now solve the first order conditions for $p(t)$ and $q(t)$, and if we denote by $p^*(t)$ and $q^*(t)$ the resulting optimal strategies, we find that

$$p^*(t) = \frac{\sigma_Y}{\sigma_S} - \frac{(\xi - \sigma_Y)}{\sigma_S}\frac{V_x}{xV_{xx}} \tag{37}$$

$$q^*(t,\zeta) = 1 - \frac{V_x}{xV_{xx}} - \frac{1}{\psi(t,\zeta)}\frac{V_{x\zeta}}{xV_{xx}} \tag{38}$$

Inserting these two expressions into the HJB equation yields the following PDE:

$$\begin{aligned}V_t + V_x\Big(\frac{\pi}{a(t,\zeta)} + x\big(- (r + \lambda_0(t)\zeta + \mu) + \xi\sigma_Y\big)\Big) + V_\zeta\big(\theta - \delta\zeta\big) - \frac{1}{2}\frac{V_x^2}{V_{xx}}\Big((\xi - \sigma_Y)^2\\ + \psi^2(t,\zeta)\sigma_\zeta^2\zeta\Big) + \frac{V_xV_{x\zeta}}{V_{xx}}\psi(t,\zeta)\sigma_\zeta^2\zeta - \frac{1}{2}\frac{V_{x\zeta}^2}{V_{xx}}\sigma_\zeta^2\zeta + \frac{1}{2}V_{\zeta\zeta}\sigma_\zeta^2\zeta = 0\end{aligned}$$

If we consider, again, first the case where $\pi = 0$, and to the case of power utility, we can express the value function as:

$$V(t,x,\zeta) = \frac{1}{\gamma}x^\gamma g(t,\zeta)^{1-\gamma} \tag{39}$$

with $g(T,\zeta) = 1$ for all ζ. Now, following the approach presented for the incomplete market case, in the proof of proposition 2 (appendix C), it can be shown that the value function, for the general case ($\pi \geq 0$), takes the form

$$\begin{aligned}V(t,x,\zeta) = \frac{1}{\gamma}\Big(x + \frac{\pi f(t)}{a(t,\zeta)}\Big)^\gamma \exp\Big\{\gamma\Big(- \mu + \xi\sigma_Y + \frac{1}{2}\frac{1}{(1-\gamma)}(\xi - \sigma_Y)^2\Big)\big(T - t\big)\Big\}\\ \cdot a\big(t,\zeta\big)^\gamma \mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\frac{\gamma}{1-\gamma}}\Big|\mathcal{G}(t)\Big]^{1-\gamma} \qquad (40)\end{aligned}$$

From (37) and (38) we find that

$$p^*(t) = \frac{\sigma_Y}{\sigma_S} + \frac{(\xi - \sigma_Y)}{\sigma_S}\frac{1}{1-\gamma}\Big(1 + \frac{\pi Y(t) f(t)}{W(t)}\Big) \tag{41}$$

$$q^*(t,\zeta) = \frac{1 + \dfrac{\pi Y(t) f(t)}{W(t)}}{\psi(t,\zeta)} \frac{\dfrac{\partial}{\partial \zeta}\mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\frac{\gamma}{1-\gamma}}\Big|\mathcal{G}(t)\Big]}{\mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\frac{\gamma}{1-\gamma}}\Big|\mathcal{G}(t)\Big]} \tag{42}$$

Since we know the probability density function for $\zeta(T)$ (conditional on information at time $t < T$) it is possible to calculate (using numerical integration techniques) $q^*(t,\zeta)$ for given values of $t < T$ and ζ. For $t = T$ we have that $q^*(T,\zeta) = -\dfrac{\gamma}{1-\gamma}$ for all ζ.

5.2.3 The value of the longevity hedge

The optimal asset allocation strategies $p^*(t)$ and $q^*(t)$ in (41) and (42) are given as a proportion of the total nominal pension fund wealth $W(t)$. They therefore require a short position in the risk-free asset constituting a loan against future contributions to the fund.

Insert figure 3 here

Figure 3 shows the evolution of $p^*(t)$ (red) and $q^*(t)$ (blue) over the 40 year life of the DC fund considered for 1000 simulated paths, conditional on the risky asset following a single, randomly selected, path. The parameter values used here, and throughout the remainder of the paper (except where stated otherwise) were chosen as $\pi = 0.1$, $\xi = 0.2$, $\mu = 0$, $\sigma_S = 0.2$ and $\sigma_Y = 0.05$, in line with typical values used throughout the literature. We can see that initially both proportions are significantly above one, indicating the large initial short position in the risk-free asset. The size of this short position reduces quickly over time as contributions are received and used to pay back the implicit initial borrowings, while the nominal value of the fund assets increases.

We also see that there is a similar degree of variation across the paths for both the $p^*(t)$ and $q^*(t)$ strategies. This is, perhaps, counterintuitive, on the basis that it is solely the mortality rate that differs across the paths. We might expect a much greater degree of variation across the $q^*(t)$ paths, since $q^*(t)$ is the optimal longevity risk hedging strategy. However, both strategies are dependent on the ratio of the present value of future contributions to the nominal pension fund wealth, and it is the variation due to the effect of the varying mortality rate on this ratio that is driving the observed behaviour.

It is difficult to gauge the effect of the mortality rate variation on the demand for the longevity risk hedge from figure 3 due to the overwhelming influence of the value of future contributions over the early stages of the plan. If we, however define the augmented pension fund wealth as

$$\tilde{W}(t) = W(t) + Y(t)\pi f(t) \tag{43}$$

and express the asset allocation strategies as a proportion of $\tilde{W}(t)$, i.e. we treat the future contributions as part of the current pension fund wealth, we can observe far more distinct behaviour.

Insert figure 4 here

Figure 4 shows the two strategies expressed in this way for 3 different degrees of relative risk aversion: 3 (green), 6 (red) and 9 (blue). We can see that the risky asset allocations (dashed lines) are constant across both time and path; that is, the proportion of augmented wealth allocated to the risky asset does not change deterministically over time or in response to shocks to mortality. The allocation to the longevity bond portfolio is both time varying and responds to mortality shocks, as we would expect. The upward sloping time variation is attributable to the asymmetry between the individual longevity bond prices, which implicitly factor in the possibility that the plan member could die prior to retirement, and the annuity payments, which are conditional on survival to retirement.

As we would expect, the more risk averse investor invests more, on average, in the longevity bond portfolio, and less in the risky asset. The significant variation across the different paths shows that there is at least some benefit to adopting a dynamic hedging strategy, however the extent of this benefit cannot be made clear by merely observing the asset allocation strategies themselves. In order to gauge this extent we need to consider the value derived from the hedge; that is, the impact the addition of the hedging asset has on the value function of the problem, i.e. the plan member's expected utility at retirement.

Insert figure 5 here

Figure 5 shows 1000 simulated value function paths (again conditional on single, randomly selected, risky asset path) for both the incomplete market case (blue) and the complete market case (red), for 4 different levels of relative risk aversion: 1 (top left)[8], 3 (top right), 6 (bottom left) and 9 (bottom right).

For an individual with a level of relative risk aversion equal to one the cases are very similar; that is, as the plan member is near risk neutral he does not place much value on the longevity risk hedge and therefore follows much the same strategy as he would if the hedge were unavailable. As the level of risk aversion increases, however, the relative risk associated with the incomplete market case increases, becoming quite stark for the higher levels. There is no observable difference between the expected terminal utility at time zero, at any level of risk aversion. This stems from the mortality linked asset's role as purely a hedging instrument.

As we have discussed in section 3, the proposed pension plan design calls for the contribution rate to be determined as an exogenous variable in order to accommodate some of the behavioural biases typical of pension plan members, thereby making the plan more attractive. The responsibility of the pension plan, with regard to the contribution rate decision, is to provide reliable information concerning the replacement ratio the plan member is likely to receive given a particular contribution rate. How the plan member uses this information to determine the contribution rate most appropriate to them is left open. It is clear that if the plan member were to choose the contribution rate resulting in a desired expected terminal utility, the availability of the longevity hedge would not affect this decision. There is, however, a vast set of criteria by which the plan member could make this decision, and on which the longevity risk hedge may have an effect.

[8]For ease of calculation we in fact set the coefficient of relative risk aversion here to be .99

When communicating the pension fund wealth information to the plan member it could be confusing to speak in terms of a theoretical utility derived from this wealth. In the context of the proposed pension plan design, the utility function is a convenient quantitative tool, used as a means of accounting for varying risk preferences in the dynamic asset allocation model; it should be of little concern to the plan member. We will therefore, from now on, speak solely in terms of the replacement ratio, which offers an intuitive understanding of pension fund wealth, thereby facilitating the decision making process.

Insert figure 6 here

The probability distribution of the replacement ratio, if known, would present most, if not all, of the information relevant to the plan member; principally, the probability of meeting, failing to meet, or surpassing some desired, or permissible, replacement ratio. Figure 6 presents probability densities for a selection of key parameter values, estimated using Monte Carlo simulation over 30,000 paths. The parameter values varied are the market price of risk coefficient, ξ, and the coefficient of relative risk aversion, γ; two parameters that significantly affect the proportion of (augmented) pension wealth allocated to the risky asset, S. The purpose of this was to highlight the relative effects of longevity risk and market risk.

For both a relatively low level of risk aversion and relatively high market price of risk (top left), we do not observe any benefit (by way of a reduction in risk) from the longevity hedge i.e. the histogram representing the density estimate of complete market case (blue) eclipses that representing that of the incomplete market case. This is due primarily to the fact that optimal allocation to the risky asset is high for these parameter values, and therefore the proportion of total portfolio variance attributable to the risky asset far outweighs that attributable to the mortality linked security.

If, however, we consider a relatively high level of risk aversion and a relatively low market price of risk (bottom right), we observe a considerable benefit. Here, with a low allocation to the risky asset, the principal driver behind the replacement ratio variance is the annuity rate, and therefore the effect of the hedge is more pronounced. The hedge effectiveness can be seen to be increasing from the density at the top left to that at the bottom right in line with a decreasing risky asset allocation.

The value derived from the longevity risk hedge evidently depends on the individual plan member's attitude towards risk. It is likely that an individual classified as risk averse (by whatever means) for the purpose of determining the asset allocation strategy would likewise exhibit risk averse behaviour when deciding on an appropriate contribution rate. We will, for illustrative purposes, assume that the plan member chooses the contribution rate that ensures there is a certain probability the replacement rate will be above some minimal permissible value; or, in other words, the contribution rate that gives a desired value-at-risk (VaR). We can then define the value of the longevity hedge (to the individual plan member) as the increase in contribution rate required to ensure that this VaR is same in both the complete and incomplete markets.

Insert table 2 here

Table 2 presents these required contributions for 3 different confidence levels (90%, 95% and 99%), based on estimated probability densities analogous to those in figure 6, for a selection of different relative risk aversion and market price of risk parameter values. The required increases are modest across all scenarios, ranging from no increase at the 90% confidence level for the scenario corresponding to the largest risky asset allocation, to an increase of 5.22% (i.e. from 10% to 10.522%) at the 99% confidence level for the scenario corresponding to the least risky asset allocation. The increases may be modest, but over a 40 year period even modest increases in the contribution rate can have a significant impact on the nominal fund wealth, i.e. the cost of trying to make up the difference at retirement could be steep; a harsh lesson learned by many DB plans in recent times.

6 Conclusions

Over the past decade or so the global pensions markets have experienced a shift from a defined benefit (DB) plan dominated environment to one in which defined contribution (DC) plans are playing an ever-increasing role. This prevailing trend has brought with it a demand for a change in the way in which DC plans are designed. We have discussed a design proposal that looks to, not only retain the plan member oriented principles of the DB plan but correct for many of the plan shortcomings that saw so many DB plans fail so dramatically.

This design hinges on the development of a robust dynamic asset allocation strategy that can react optimally to changes in key underlying state variables. It was the failure to anticipate, and react appropriately to, changing market conditions, slowing economies and age demographic shifts that resulted in the DB plan failures. We have argued that these risks are not specific to DB plans, and must be accounted for in any DC design going froward. We have focused primarily on the risk that age demographics will change, or particularly, the risk that populations will age due to an average increase in life expectancy.

We refer to the risk that average life expectancy will increase as systematic longevity risk. If we re-design defined contribution pensions so that they address the problem of funding a desired level of retirement income, as would be desirable by most individuals in the market for pensions, there is no escaping the problem of systematic longevity risk. We have seen that longevity risk manifests itself in the proposed pension plan design problem through the annuity rate; that is, longevity risk in the pension plan design problem is the risk that the expected annuity rate at retirement will decrease over time as a result of corresponding unexpected decreases in mortality rates.

We worked within a continuous-time dynamic programming framework to identify the optimal dynamic asset allocation strategy. We took as the objective function of the optimization problem the expected utility derived from the replacement ratio at retirement. We identified an affine stochastic mortality model that remained analytically tractable within the continuous time stochastic dynamic programming framework. The model combined the renowned Gompertz law for mortality with the widely used CIR term structure model, which we applied to the mortality improvements. This combination ensured that the model parameters could be easily understood in terms of real world phenomena. The use of the Gompertz law enabled us to find closed form solutions to the resulting Ricatti equations, giving us a functional form for the survival probabilities. This allows for the efficient estimation of future survival probabilities and life expectancies, which in turn can be used to value annuities and other mortality contingent claims. We provided a straightforward scheme by which to calibrate the model to historical death rates and showed that, when calibrated to US data, the model captures real world mortality dynamics, as compared with the most popular discrete time model used in practice.

In order to gauge the impact of longevity risk we introduced, into the asset mix of the portfolio choice problem, a synthetic longevity-linked security, of a similar type to what has already been proposed in the literature. We determined the relative demand for this asset, and the value added through the introduction of this asset in terms of a utility gain. We found that for sufficiently risk averse plan members the value added by way of a reduction in conditional volatility is substantial, however these effects can be drowned out if the proportion of pension fund wealth allocated to risky assets is large. We showed that, if the individual plan member considered was to choose his contribution rate such that the probability of the fund falling below some minimal permissible value was kept to a given level, the benefit in terms of saved contributions derived from the availability of an efficient longevity hedge could be significant.

References

Blake, D., Cairns, A.J.G. and Dowd, K. (2009) *Designing a Defined-Contribution Plan: What to Learn from Aircraft Designers.* Financial Analysts Journal, 65, 1, 37-42.

Benzion, U., Rapoport, A. and Yagil, J. (1989) *Discount Rates Inferred from Decisions: An Experimental Study.* Management Science, 35, 3, 270-284.

Cairns, A.J.G., Blake, D. and Dowd, K. (2006a) *Stochastic lifestyling: Optimal dynamic asset allocation for defined contribution pension plans.* Journal of Economic Dynamics & Control, 30, 843-877.

Cairns, A.J.G., Blake, D. and Dowd, K. (2006b) *A two-factor model for stochastic mortality with parameter uncertainty.* The Journal of Risk and Insurance, 73, 4, 687-718.

Choi, J.J., Laibson, D. and Madrian, B.C. (2009) *Mental Accounting in Portfolio Choice: Evidence from a Flypaper Effect.* The American Economic Review, 99, 5, 2085-2095.

Cocco, J.F. and Gomes,F.J. (2012) *Longevity Risk, retirement savings, and financial innovation.* Journal of Financial Economics, 103, 507-529.

Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985) *A Theory of the Term Structure of Interest Rates.* Econometrica, 53, 2, 385-407.

Cronquist, H. and Thaler, R.H. (2004) *Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience.* The American Economic Review, Papers and Proceedings, 94, 2, 424-428.

Currie, I.D., Durban, M. And Eilers, P.H.C. (2004) *Smoothing and forecasting mortality rates.* Statistical Modelling, 4, 279-298.

Dahl, M. And Moller, T. (2006) *Valuation and hedging of life insurance risks with systematic mortality risk.* Insurance: Mathematics and Economics, 39, 193-217.

Davidoff, T., Brown, J.R. and Diamond, P.A. (2005) *Annuities and Individual Welfare.* The American Economic Review, 95, 5, 1573-1590.

Horneff , W.J., Maurer, R.H., Mitchell, O.S. and Stamos, M.Z. (2009) *Asset allocation and location over the life cycle with investment-linked survival-contingent payouts.* Journal of Banking & Finance, 33, 16881699.

Horneff, W., Maurer, R. and Rogalla, R. (2010) *Dynamic portfolio choice with deferred annuities.* Journal of Banking & Finance, 34, 26522664.

Horneff , W.J., Maurer, R.H. and Stamos, M.Z. (2008) *Life-cycle asset allocation with annuity markets.* Journal of Economic Dynamics & Control, 32, 3590 3612.

Huberman, G. and Jiang, W. (2006) *Offering versus Choice in 401(k) Plans: Equity Exposure and Number of Funds.* The Journal of Finance, 61, 2, 763-801.

Inkmann, J., Lopes, P. and Michaelides, A. (2011) *How Deep is the Annuity Market Participation Puzzle?* The Review of Financial Studies, 24, 1, 279-319.

Koijen, R.S.J., Nijman, T.E., and Werker, B.J.M. (2011) *Optimal Annuity Risk Management.* Review of Finance, 15, 799-833.

Laibson, D. (1997) *Golden Eggs and Hyperbolic Discounting.* Quarterly Journal of Economics, 112, 2, 443-477.

Lee, R. D. and Carter, L. R. (1992) *Modelling and forecasting US mortality.* Journal of the American Statistical Association 87, 659-675.

Madrian, B.C. and Shea, D.F. (2001) *The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior.* The Quarterly Journal of Economics, 116, 4, 1149-1187.

Menoncin, F. (2008) *The role of longevity bonds in optimal portfolios.* Insurance: Mathematics and Economics, 42, 343-358.

Merton, R.C. (1969) *Lifetime portfolio selection under uncertainty: the continuous-time case.* Review of Economics and Statistics 51, 247257.

Merton, R.C. (2007) *The future of Retirement Planning.* In The Future of life-cycle saving and investment, The Research Foundation of CFA Institutue.

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

Milevsky, M.A. and Young,V.R. (2007) *annuitisation and asset allocation.* Journal of Economic Dynamics & Control, 31, 3138-3177.

Munk, C. and Sorensen, C. (2010) *Dynamic asset allocation with stochastic income and* *interest rates.* Journal of Financial Economics, 96, 433-462.

Polkovnichenko, V. (2007) *Life-Cycle Portfolio Choice with Additive Habit Formation Preferences and Uninsurable Labor Income Risk.* The Review of Financial Studies, 20, 1, 83-124.

Polyanin, A.D. and Zaitsev, V.F. (2003) *Handbook of Exact Solutions for Ordinary Differential Equations.* 2nd Edition, Chapman & Hall/CRC, Boca Raton.

Richard, S. (1975) Optimal consumption, *portfolio and life insurance rules for an uncertain lived individual in a continuous time model.* Journal of Financial Economics 2, 187203.

Schrager, D.F. (2006) *Affine Stochastic Mortality.* Insurance: Mathematics and Economics, 38, 81-97.

Thaler, R. H. (1990) *Anomalies: Saving, Fungibility, and Mental Accounts.* Journal of Economic Perspectives, 4, 193-205.

Viceira, L.M. (2001) *Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income.* The Journal of Finance, 56, 2, 433-470.

Yaari, Y.E. (1965) *Uncertain Lifetime, Life Insurance, and the Theory of the Consumer.* The Review of Economic Studies, 32, 2, 137-150.

Appendix A

We first find that

$$\theta_\lambda(x,t) = \frac{\theta}{b}e^{\frac{1}{b}(x+t-m)} \qquad \delta_\lambda(x,t) = \delta - \frac{1}{b} \qquad \sigma_\lambda(x,t) = \sigma_\zeta\sqrt{\frac{1}{b}e^{\frac{1}{b}(x+t-m)}} \tag{44}$$

Letting $D = \frac{1}{b}e^{\frac{1}{b}(x-m)}$, we have (suppressing the x dependence)

$$\frac{\partial}{\partial t}\beta(t,T) = (\delta - \frac{1}{b})\beta(t,T) + \frac{1}{2}\sigma_\zeta^2 De^{\frac{1}{b}t}\beta(t,T)^2 - 1 \tag{45}$$

$$\frac{\partial}{\partial t}\alpha(t,T) = \theta De^{\frac{1}{b}t}\beta(t,T) \tag{46}$$

We can solve equation this coupled ODE if we first make the following transformation:

$$\beta(t,T) = -\frac{1}{\frac{\sigma_\zeta^2}{2}De^{\frac{1}{b}t}y(t,T)}\frac{\partial}{\partial t}y(t,T) \tag{47}$$

This gives us

$$\frac{\partial^2}{\partial t^2}y(t,T) - \delta\frac{\partial}{\partial t}y(t,T) - \frac{\sigma_\zeta^2}{2}De^{\frac{1}{b}t}y(t,T) = 0 \tag{48}$$

It can be shown (see Polyanin and Zaitsev, 2003) that this ODE has the general solution:

$$y(t,T) = e^{\frac{\delta}{2}t}C_1(T)J_\nu\Big(2b\sqrt{\frac{-\sigma_\zeta^2}{2}D}e^{\frac{1}{2b}t}\Big) + e^{\frac{\delta}{2}t}C_2(T)Y_\nu\Big(2b\sqrt{\frac{-\sigma_\zeta^2}{2}D}e^{\frac{1}{2b}t}\Big) \tag{49}$$

where J_ν and Y_ν are the Bessel functions of first and second kind respectively, $\nu = \delta b$, and $C_1(T)$ and $C_2(T)$ are arbitrary functions of T. Now

$$\frac{\partial}{\partial t}y(t,T) = \frac{\delta}{2}y(t,T) + C_1(T)e^{\frac{\delta}{2}t}\frac{\mathrm{d}}{\mathrm{d}t}J_\nu(z(t)) + C_2(T)e^{\frac{\delta}{2}t}\frac{\mathrm{d}}{\mathrm{d}t}Y_\nu(z(t)) \tag{50}$$

where,

$$\frac{\mathrm{d}}{\mathrm{d}t}J_\nu(z(t)) = \frac{\mathrm{d}}{\mathrm{d}z(t)}J_\nu(z(t))\frac{\mathrm{d}}{\mathrm{d}t}z(t) \quad ; \quad \frac{\mathrm{d}}{\mathrm{d}t}Y_\nu(z(t)) = \frac{\mathrm{d}}{\mathrm{d}z(t)}Y_\nu(z(t))\frac{\mathrm{d}}{\mathrm{d}t}z(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}z(t)}J_\nu(z(t)) = J_{\nu-1}(z(t)) - \frac{\nu}{z(t)}J_\nu(z(t)) \quad ; \quad \frac{\mathrm{d}}{\mathrm{d}z(t)}Y_\nu(z(t)) = Y_{\nu-1}(z(t)) - \frac{\nu}{z(t)}Y_\nu(z(t))$$

So

$$\begin{aligned}
\frac{\partial}{\partial t}y(t,T) =& \frac{\delta}{2}y(t,T) + C_1(T)e^{\frac{\delta}{2}t}\Big[\frac{1}{2b}z(t)\big(J_{\nu-1}(z(t)) - \frac{\nu}{z(t)}J_\nu(z(t))\big)\Big] \\
&+ C_2(T)e^{\frac{\delta}{2}t}\Big[\frac{1}{2b}z(t)\big(Y_{\nu-1}(z(t)) - \frac{\nu}{z(t)}Y_\nu(z(t))\big)\Big] \\
=& \frac{\delta}{2}y(t,T) - \frac{\nu}{2b}y(t,T) + \frac{1}{2b}z(t)e^{\frac{\delta}{2}t}\Big[C_1(T)J_{\nu-1}(z(t)) + C_2(T)Y_{\nu-1}(z(t))\Big] \\
=& \frac{1}{2b}z(t)e^{\frac{\delta}{2}t}\Big[C_1(T)J_{\nu-1}(z(t)) + C_2(T)Y_{\nu-1}(z(t))\Big]
\end{aligned} \tag{51}$$

We therefore have

$$-\frac{1}{y(t,T)}\frac{\partial}{\partial t}y(t,T) = -\frac{1}{2b}z(t)\frac{C_1(T)J_{\nu-1}(z(t)) + C_2(T)Y_{\nu-1}(z(t))}{C_1(T)J_{\nu}(z(t)) + C_2(T)Y_{\nu}(z(t))} \tag{52}$$

Since $\frac{\sigma_\zeta^2}{2}De^{\frac{1}{b}t} = -\frac{1}{4b^2}(z(t))^2$,

$$\begin{aligned} -\frac{1}{\frac{\sigma_\zeta^2}{2}De^{\frac{1}{b}t}y(t,T)}\frac{\partial}{\partial t}y(t,T) &= \frac{2b}{z(t)}\frac{C_1(T)J_{\nu-1}(z(t)) + C_2(T)Y_{\nu-1}(z(t))}{C_1(T)J_{\nu}(z(t)) + C_2(T)Y_{\nu}(z(t))} \\ &= \beta(t,T) \end{aligned} \tag{53}$$

In order to find the particular solution we must find the functions $C_1(T)$ and $C_2(T)$ that satisfy $\beta(T,T) = 0$ and $\alpha(T,T) = 0$.

$$\beta(T,T) = 0 \Rightarrow C_1(T)J_{\nu-1}(z(t)) + C_2(T)Y_{\nu-1}(z(t)) = 0 \Rightarrow \frac{C_1(T)}{C_2(T)} = -\frac{Y_{\nu-1}(z(T))}{J_{\nu-1}(z(T))} \tag{54}$$

We know that

$$Y_{\nu-1}(z(t)) = \cot(\pi(\nu-1))J_{\nu-1}(z(t)) - \csc(\pi(\nu-1))J_{-\nu+1}(z(t)) \tag{55}$$

So

$$\begin{aligned} \beta(t,T) &= \frac{2b}{z(t)}\frac{C_1(T)J_{\nu-1}(z(t)) + C_2(T)\big(\cot(\pi(\nu-1))J_{\nu-1}(z(t)) - \csc(\pi(\nu-1))J_{-\nu+1}(z(t))\big)}{C_1(T)J_{\nu}(z(t)) + C_2(T)Y_{\nu}(z(t))} \\ &= \frac{2b}{z(t)}\frac{\big(C_1(T) + C_2(T)\cot(\pi(\nu-1))\big)J_{\nu-1}(z(t)) - C_2(T)\csc(\pi(\nu-1))J_{-\nu+1}(z(t))}{C_1(T)J_{\nu}(z(t)) + C_2(T)Y_{\nu}(z(t))} \end{aligned}$$

We can let $C_2(T) = -1$ and $C_1(T) = \cot(\pi(\nu-1)) - \frac{J_{-\nu+1}(z(T))}{J_{\nu-1}(z(T))}\csc(\pi(\nu-1))$, therefore

$$\begin{aligned} \beta(t,T) &= \frac{2b}{z(t)}\frac{-\frac{J_{-\nu+1}(z(T))}{J_{\nu-1}(z(T))}\csc(\pi(\nu-1))J_{\nu-1}(z(t)) + \csc(\pi(\nu-1))J_{-\nu+1}(z(t))}{C_1(T)J_{\nu}(z(t)) + C_2(T)Y_{\nu}(z(t))} \\ &= \frac{2b}{z(t)}\frac{-\csc(\pi(\nu-1))\big(J_{-\nu+1}(z(T))J_{\nu-1}(z(t)) - J_{\nu-1}(z(T))J_{-\nu+1}(z(t))\big)}{-\frac{J_{-\nu+1}(z(T))}{J_{\nu-1}(z(T))}\csc(\pi(\nu-1))J_{\nu}(z(t)) + \csc(\pi(\nu))J_{-\nu}(z(t))} \\ &= \frac{2b}{z(t)}\left[\frac{J_{-\nu+1}(z(T))J_{\nu-1}(z(t)) - J_{\nu-1}(z(T))J_{-\nu+1}(z(t))}{J_{-\nu+1}(z(T))J_{\nu}(z(t)) - J_{\nu-1}(z(T))J_{-\nu}(z(t))}\right] \end{aligned} \tag{56}$$

Now,

$$J_{\nu-1}(z(t)) = \frac{\partial}{\partial z(t)}J_{\nu}(z(t)) + \frac{\nu}{z(t)}J_{\nu}(z(t)) \tag{57}$$

$$J_{-\nu+1}(z(t)) = -\frac{\partial}{\partial z(t)}J_{-\nu}(z(t)) - \frac{\nu}{z(t)}J_{-\nu}(z(t)) \tag{58}$$

For ease of notation, if we let $\dfrac{J_{-\nu+1}(z(T))}{J_{\nu-1}(z(T))} = C(T)$, then

$$\beta(t,T) = \frac{2b}{z(t)}\left(\frac{\nu}{z(t)} + \frac{\frac{\partial}{\partial z(t)}\big(C(T)J_\nu(z(t)) + J_{-\nu}(z(t))\big)}{\big(C(T)J_\nu(z(t)) + J_{-\nu}(z(t))\big)}\right) \tag{59}$$

We have from equation our initial ODE for $\alpha(t,T)$,

$$\frac{\partial}{\partial t}\alpha(t,T) = \theta D e^{\frac{1}{b}t}\beta(t,T) = -\frac{\theta}{2b^2\sigma_\zeta^2}(z(t))^2\beta(t,T) \qquad \text{with} \qquad \alpha(T,T) = 0$$

So

$$\begin{aligned}
\frac{\partial}{\partial t}\alpha(t,T) &= -\frac{\theta\delta}{\sigma_\zeta^2} - \frac{\theta\delta}{\sigma_\zeta^2}\frac{1}{2b}z(t)\frac{\frac{\partial}{\partial z(t)}\big(C(T)J_\nu(z(t)) + J_{-\nu}(z(t))\big)}{\big(C(T)J_\nu(z(t)) + J_{-\nu}(z(t))\big)} \\
\Rightarrow \alpha(t,T) &= -\frac{\theta\delta}{\sigma_\zeta^2}\int_t^T \mathrm{d}\tau - \frac{\theta\delta}{\sigma_\zeta^2}\frac{1}{2b}\int_t^T z(\tau)\frac{\frac{\partial}{\partial z(\tau)}\big(C(T)J_\nu(z(\tau)) + J_{-\nu}(z(\tau))\big)}{\big(C(T)J_\nu(z(\tau)) + J_{-\nu}(z(\tau))\big)}\mathrm{d}\tau \\
&= -\frac{\theta\delta}{\sigma_\zeta^2}(T-t) - \frac{\theta\delta}{\sigma_\zeta^2}\int_{z(t)}^{z(T)} \frac{\frac{\partial}{\partial z(\tau)}\big(C(T)J_\nu(z(\tau)) + J_{-\nu}(z(\tau))\big)}{\big(C(T)J_\nu(z(\tau)) + J_{-\nu}(z(\tau))\big)}\mathrm{d}z(\tau) \\
&= -\frac{\theta\delta}{\sigma_\zeta^2}(T-t) - \frac{\theta\delta}{\sigma_\zeta^2}\left[\ln\left(\frac{C(T)J_\nu(z(T)) + J_{-\nu}(z(T))}{C(T)J_\nu(z(t)) + J_{-\nu}(z(t))}\right)\right] \qquad (60)
\end{aligned}$$

Since $z(t)$ is complex for positive b values it is more appropriate to write $\alpha(t,T)$ and $\beta(t,T)$ in terms of the modified Bessel function $I_\nu(z)$, where

$$I_\nu(z) = \frac{z^\nu}{(iz)^\nu}J_\nu(iz) \tag{61}$$

If we let $z'(t) = 2b\sqrt{\dfrac{\sigma_\zeta^2 D}{2}}e^{\frac{1}{2b}t}$, i.e. $z(t) = z'(t)i$, then we have that $J_\nu(z(t)) = i^\nu I_\nu(z'(t))$. It is straightforward to show that

$$\beta(t,T) = \frac{2b}{z'(t)}\left[\frac{I_{-\nu+1}(z'(T))I_{\nu-1}(z'(t)) - I_{\nu-1}(z'(T))I_{-\nu+1}(z'(t))}{I_{\nu-1}(z'(T))I_{-\nu}(z'(t)) - I_{-\nu+1}(z'(T))I_\nu(z'(t))}\right] \tag{62}$$

$$\alpha(t,T) = -\frac{\theta\delta}{\sigma_\zeta^2}(T-t) + \ln\left(\frac{I_{-\nu+1}(z'(T))I_\nu(z'(t)) - I_{\nu-1}(z'(T))I_{-\nu}(z'(t))}{I_{-\nu+1}(z'(T))I_\nu(z'(T)) - I_{\nu-1}(z'(T))I_{-\nu}(z'(T))}\right)^{\frac{\theta\delta}{\sigma_\zeta^2}} \tag{63}$$

Appendix B

$\zeta(T)$, conditional on information at time $t < T$, has the following probability density function (see Cox et al.,1985):

$$f(\zeta(T),T;\zeta(t),t) = \frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\exp\Big\{-\frac{2\delta(\zeta(T)+\zeta(t)e^{-\delta(T-t)})}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\Big\}$$
$$\cdot\Big(\frac{\zeta(T)}{\zeta(t)e^{-\delta(T-t)}}\Big)^{\frac{\gamma}{\sigma_\zeta^2}-\frac{1}{2}} I_{\frac{2\gamma}{\sigma_\zeta^2}-1}\Big(\frac{4\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\sqrt{\zeta(T)\zeta(t)e^{-\delta(T-t)}}\Big)$$

where $I_\nu(z)$ here is again the modified Bessel function of the first kind. Recall, in section 4, in order to minimise the rate of mean reversion we set $\theta = \sigma_\zeta^2/2$. If we do the same here we see the density simplifies to

$$f(\zeta(T),T;\zeta(t),t) = \frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\exp\Big\{-\frac{2\delta(\zeta(T)+\zeta(t)e^{-\delta(T-t)})}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\Big\}$$
$$\cdot I_0\Big(\frac{4\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\sqrt{\zeta(T)\zeta(t)e^{-\delta(T-t)}}\Big)$$

Now,

$$\mathrm{E}\Big[a\big(T,\zeta(T)\big)\mid\mathcal{I}(t)\Big] = \int_0^\infty\sum_{\tau=T}^\infty \exp\big\{-r(\tau-T)+\alpha(T,\tau)-\lambda_0(T)\beta(T,\tau)\zeta(T)\big\}$$
$$\cdot\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\exp\Big\{-\frac{2\delta(\zeta(T)+\zeta(t)e^{-\delta(T-t)})}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\Big\}$$
$$\cdot I_0\Big(\frac{4\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\sqrt{\zeta(T)\zeta(t)e^{-\delta(T-t)}}\Big)\mathrm{d}\zeta(T)$$

$$= \int_0^\infty\sum_{\tau=T}^\infty \exp\big\{-r(\tau-T)+\alpha(T,\tau)-\lambda_0(T)\beta(T,\tau)\zeta(T)\big\}$$
$$\cdot\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\exp\Big\{-\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}(\zeta(T)+\zeta(t)e^{-\delta(T-t)})\Big\}$$
$$\cdot\sum_{k=0}^\infty\frac{1}{(k!)^2}\bigg(\Big(\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\Big)^2\zeta(T)\zeta(t)e^{-\delta(T-t)}\bigg)^k\mathrm{d}\zeta(T)$$

$$= \sum_{\tau=T}^\infty\sum_{k=0}^\infty\frac{1}{(k!)^2}\Big(\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\Big)^{2k+1}\big(\zeta(t)e^{-\delta(T-t)}\big)^k$$
$$\cdot\exp\big\{-r(\tau-T)+\alpha(T,\tau)\big\}\exp\Big\{-\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}\zeta(t)e^{-\delta(T-t)}\Big\}$$
$$\cdot\int_0^\infty\zeta(T)^k\exp\Big\{-\Big(\frac{2\delta}{\sigma_\zeta^2(1-e^{-\delta(T-t)})}+\lambda_0(T)\beta(T,\tau)\Big)\zeta(T)\Big\}\mathrm{d}\zeta(T)$$

$$
= \sum_{\tau=T}^{\infty} \sum_{k=0}^{\infty} \frac{1}{(k!)^2} \Big(\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \Big)^{2k+1} \big(\zeta(t) e^{-\delta(T-t)} \big)^k
$$

$$
\cdot \exp\Big\{ - r(\tau - T) + \alpha(T,\tau) \Big\} \exp\Big\{ - \frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \zeta(t) e^{-\delta(T-t)} \Big\}
$$

$$
\cdot \frac{k!}{\left(\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} + \lambda_0(T)\beta(T,\tau) \right)^{k+1}}
$$

$$
= \sum_{\tau=T}^{\infty} \frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \exp\Big\{ - r(\tau - T) + \alpha(T,\tau) \Big\}
$$

$$
\cdot \exp\Big\{ - \frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \zeta(t) e^{-\delta(T-t)} \Big\} \frac{1}{\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} + \lambda_0(T)\beta(T,\tau)}
$$

$$
\cdot \sum_{k=0}^{\infty} \frac{1}{k!} \left[\frac{\left(\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \right)^2 \zeta(t) e^{-\delta(T-t)}}{\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} + \lambda_0(T)\beta(T,\tau)} \right]^k
$$

$$
= \sum_{\tau=T}^{\infty} \frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \exp\Big\{ - r(\tau - T) + \alpha(T,\tau) \Big\}
$$

$$
\cdot \frac{\exp\Big\{ - \frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \zeta(t) e^{-\delta(T-t)} \Big\}}{\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} + \lambda_0(T)\beta(T,\tau)} \exp\left\{ \frac{\left(\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} \right)^2 \zeta(t) e^{-\delta(T-t)}}{\frac{2\delta}{\sigma_\zeta^2 (1 - e^{-\delta(T-t)})} + \lambda_0(T)\beta(T,\tau)} \right\}
$$

$$
= \sum_{\tau=T}^{\infty} \frac{e^{-r(\tau - T) + \alpha(T,\tau)}}{1 + \frac{\sigma_\zeta^2}{2\delta}(1 - e^{-\delta(T-t)}) \lambda_0(T)\beta(T,\tau)} \exp\Big\{ \frac{-\zeta(t) e^{-\delta(T-t)} \lambda_0(T)\beta(T,\tau)}{1 + \frac{\sigma_\zeta^2}{2\delta}(1 - e^{-\delta(T-t)}) \lambda_0(T)\beta(T,\tau)} \Big\}
$$

Appendix C

Substituting $V(t,x,\zeta) = \dfrac{1}{\gamma} g(t,\zeta)x^{\gamma}$ into (16) we find that (following the notation used for $V(t,x,\lambda)$) the PDE reduces to

$$g_t + \Big(\theta - \delta\zeta - \gamma\psi\sigma_\zeta^2\zeta\Big)g_\zeta + \frac{1}{2}\sigma_\zeta^2\zeta g_{\zeta\zeta} \tag{64}$$
$$+ \gamma\Big(\big(- (r + \lambda_0(t)\zeta + \mu\big) + \xi\sigma_Y + \frac{1}{2}\frac{1}{1-\gamma}\big((\xi - \sigma_Y)^2\big) + \frac{1}{2}(\gamma+1)\psi^2\sigma_\zeta^2\zeta\Big)g = 0$$

The solution to this PDE has the following Feynman-Kac representation

$$g(t,\zeta) = \mathrm{E}_{\hat{P}}\Big[\exp\Big\{\gamma\int_t^T \psi\big(\tau,\zeta(\tau)\big)\mathrm{d}\tau\Big\}\Big|\mathcal{G}(t)\Big] \tag{65}$$

where

$$\psi\big(t,\zeta(t)\big) = \big(- (r + \lambda_0(t)\zeta + \mu\big) + \xi\sigma_Y + \frac{1}{2}\frac{1}{1-\gamma}\big((\xi - \sigma_Y)^2\big) + \frac{1}{2}(\gamma+1)\psi^2\sigma_\zeta^2\zeta \tag{66}$$

with

$$\mathrm{d}\zeta(t) = \Big(\theta(t) - \delta\zeta(t) - \gamma\psi\sigma_\zeta^2\zeta(t)\Big)\mathrm{d}t + \sigma_\zeta\sqrt{\zeta(t)}\mathrm{d}\hat{Z}_1(t) \tag{67}$$

and where the measure is changed to $\hat{P}$, defined by the density

$$\frac{\mathrm{d}\hat{P}}{\mathrm{d}P}\Big|_t = \exp\Big\{ - \gamma\int_0^t \psi\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(\tau)$$
$$- \frac{1}{2}\gamma^2\int_0^t \psi^2\sigma_\zeta^2\zeta\mathrm{d}\tau\Big\} \tag{68}$$

The process $\hat{Z}_\lambda(t)$ is a brownian motion under $\hat{P}$ and is determined by

$$\hat{Z}_\lambda(t) = Z(t) + \gamma\int_0^t \psi\sigma_\zeta\sqrt{\zeta}\mathrm{d}\tau \tag{69}$$

$\dfrac{\mathrm{d}\hat{P}}{\mathrm{d}P}\Big|_t$ is the Radon-Nikodým derivative of $\hat{P}$ with respect to P, so

$$\mathrm{E}_{\hat{P}}\Big[X(T)\Big|\mathcal{G}(t)\Big] = \frac{1}{\frac{\mathrm{d}\hat{P}}{\mathrm{d}P}\Big|_t}\mathrm{E}_P\Big[X(T)\frac{\mathrm{d}\hat{P}}{\mathrm{d}P}\Big|_T\Big|\mathcal{G}(t)\Big] \tag{70}$$

We also have that

$$
\begin{aligned}
a\big(t,\zeta(t)\big) &= a\big(0,\zeta(0)\big)\exp\Big\{\int_0^t \psi\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(\tau) + \int_0^t \big(r+\lambda_0\zeta-\frac{1}{2}\psi^2\sigma_\zeta^2\zeta\big)\mathrm{d}\tau\Big\} \\
&\Rightarrow \int_0^t \psi\sigma_\zeta\sqrt{\zeta}\mathrm{d}Z_\zeta(\tau) = \ln\Big(\frac{a\big(t,\zeta(t)\big)}{a\big(0,\zeta(0)\big)}\Big) - \int_0^t \big(r+\lambda_0\zeta-\frac{1}{2}\psi^2\sigma_\zeta^2\zeta\big)\mathrm{d}\tau
\end{aligned} \tag{71}
$$

Therefore,

$$
\begin{aligned}
\frac{\mathrm{d}\hat{P}}{\mathrm{d}P}\Big|_t &= \exp\Big\{-\gamma\ln\Big(\frac{a\big(t,\zeta(t)\big)}{a\big(0,\zeta(0)\big)}\Big) + \gamma\int_0^t \big(r+\lambda_0\zeta-\frac{1}{2}\psi^2\sigma_\zeta^2\zeta\big)\mathrm{d}\tau \\
&\qquad -\frac{1}{2}\gamma^2\int_0^t \psi^2\sigma_\zeta^2\zeta\mathrm{d}\tau\Big\} \\
&= \Big(\frac{a\big(t,\zeta(t)\big)}{a\big(0,\zeta(0)\big)}\Big)^{-\gamma}\exp\Big\{\gamma\int_0^t \big(r+\lambda_0\zeta\big)\mathrm{d}\tau - \frac{1}{2}\gamma(1+\gamma)\int_0^t \psi^2\sigma_\zeta^2\zeta\mathrm{d}\tau\Big\}
\end{aligned} \tag{72}
$$

and

$$
\begin{aligned}
g(t,\zeta) &= \Big(\frac{a\big(t,\zeta(t)\big)}{a\big(0,\zeta(0)\big)}\Big)^{\gamma}\exp\Big\{-\gamma\int_0^t \big(r+\lambda_0\zeta\big)\mathrm{d}\tau + \frac{1}{2}\gamma(1+\gamma)\int_0^t \psi^2\sigma_\zeta^2\zeta\mathrm{d}\tau\Big\} \\
&\cdot \mathrm{E}_P\Big[\exp\Big\{\gamma\int_t^T \psi\big(\tau,\zeta(\tau)\big)\mathrm{d}\tau\Big\}\Big(\frac{a\big(T,\zeta(T)\big)}{a\big(0,\zeta(0)\big)}\Big)^{-\gamma}\exp\Big\{\gamma\int_0^T \big(r+\lambda_0\zeta\big)\mathrm{d}\tau \\
&\qquad -\frac{1}{2}\gamma(1+\gamma)\int_0^T \psi^2\sigma_\zeta^2\zeta\mathrm{d}\tau\Big\}\Big|\mathcal{G}(t)\Big]
\end{aligned} \tag{73}
$$

Cancelling terms and taking out what is known we get

$$
\begin{aligned}
g(t,\zeta) &= \exp\Big\{\gamma\Big(-\mu+\xi\sigma_Y+\frac{1}{2}\frac{1}{(1-\gamma)}(\xi-\sigma_Y)^2\Big)\big(T-t\big)\Big\} \\
&\qquad \cdot a\big(t,\zeta\big)^{\gamma}\mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\gamma}\Big|\mathcal{G}(t)\Big]
\end{aligned} \tag{74}
$$

Therefore

$$
\begin{aligned}
V(t,x,\zeta) &= \frac{1}{\gamma}x^{\gamma}\exp\Big\{\gamma\Big(-\mu+\xi\sigma_Y+\frac{1}{2}\frac{1}{(1-\gamma)}(\xi-\sigma_Y)^2\Big)\big(T-t\big)\Big\} \\
&\qquad \cdot a\big(t,\zeta\big)^{\gamma}\mathrm{E}_P\Big[a\big(T,\zeta(T)\big)^{-\gamma}\Big|\mathcal{G}(t)\Big]
\end{aligned} \tag{75}
$$

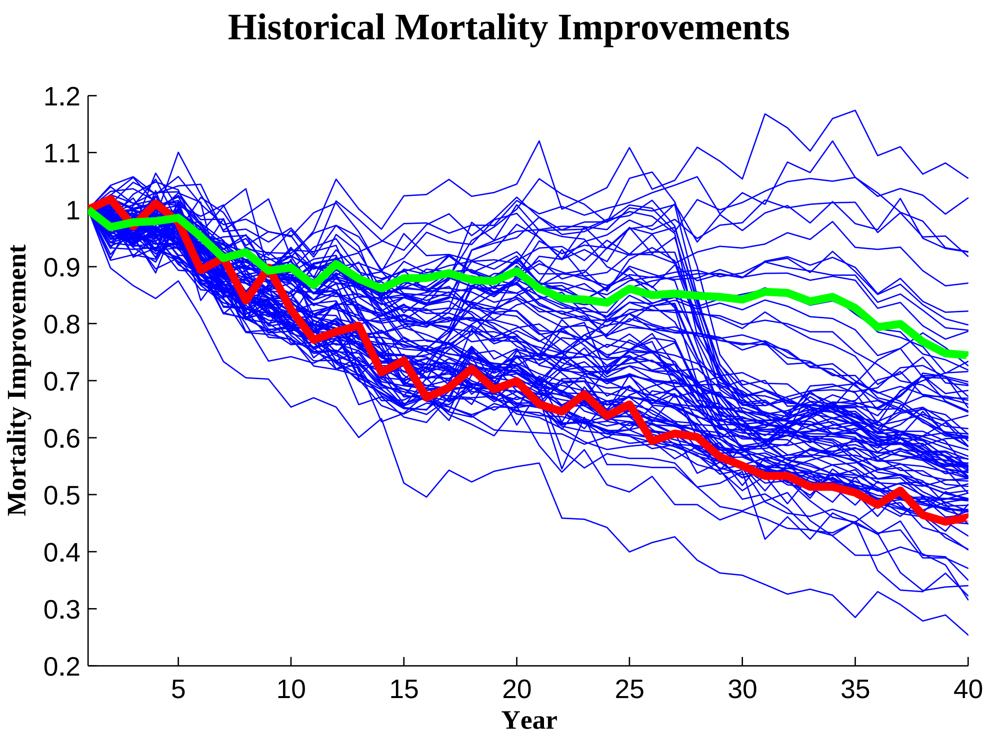

Figure 1: This figure presents the historical mortality improvements based on a number of different measures: the individual age groups (blue lines), an individual aged 25 at time 0, 26 at time 1, etc. (red line), and our proxy measure - the area under the base mortality curve (green line). It is clear that the variance across the age groups is too severe to consider any one group as representative of the whole, and the jaggedness of the line corresponding to the individual, presents a degree of volatility not in keeping with the typical age grouping, as we would expect. The proxy measure however finds a balance between these two extremes.

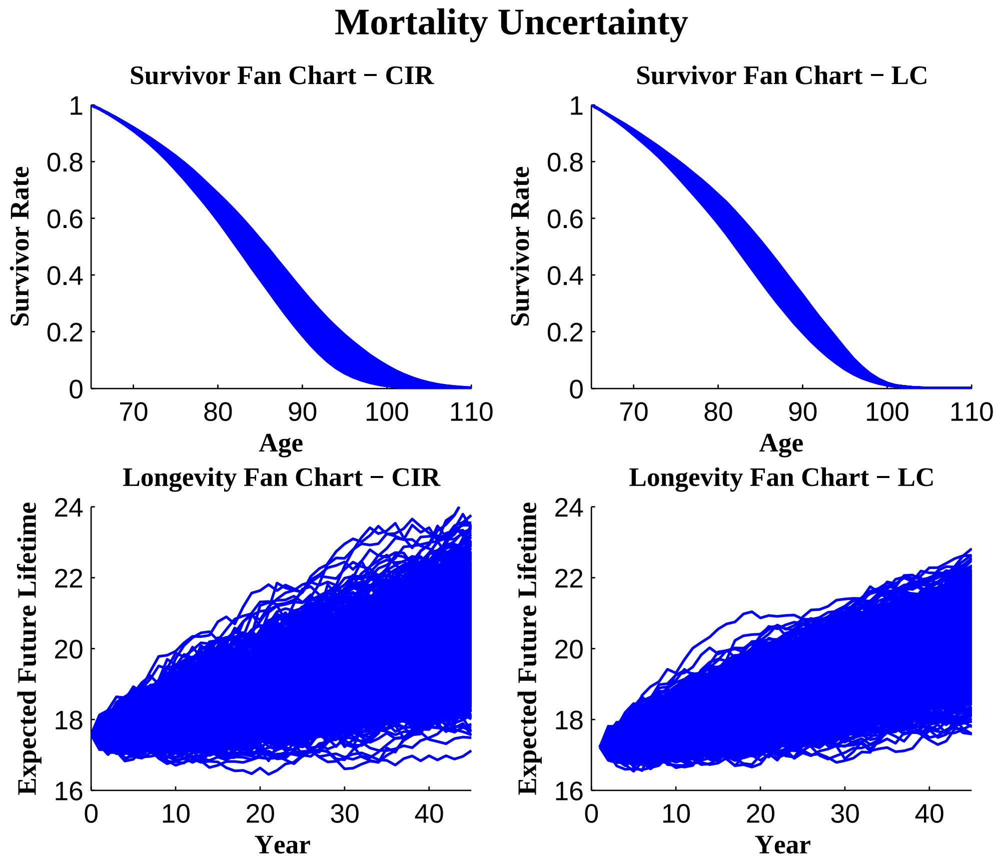

Figure 2: This figure presents survivor and longevity fan charts estimated using the 2 models considered. Those on the left are estimated using our model, based on the CIR mortality improvements, and those on the right are estimated using the Lee-Carter model. Survivor and longevity fan charts are two alternative ways for representing graphically the uncertainty associated with future mortality rates. The survivor fan chart (top) shows the probability that an man aged 65 will survive to any given age, for 1000 different simulated paths of the underlying mortality process. The longevity fan chart (bottom) shows how the life expectancy of men aged 65 evolves over a 40 year period, again for 1000 simulated paths of the mortality process.

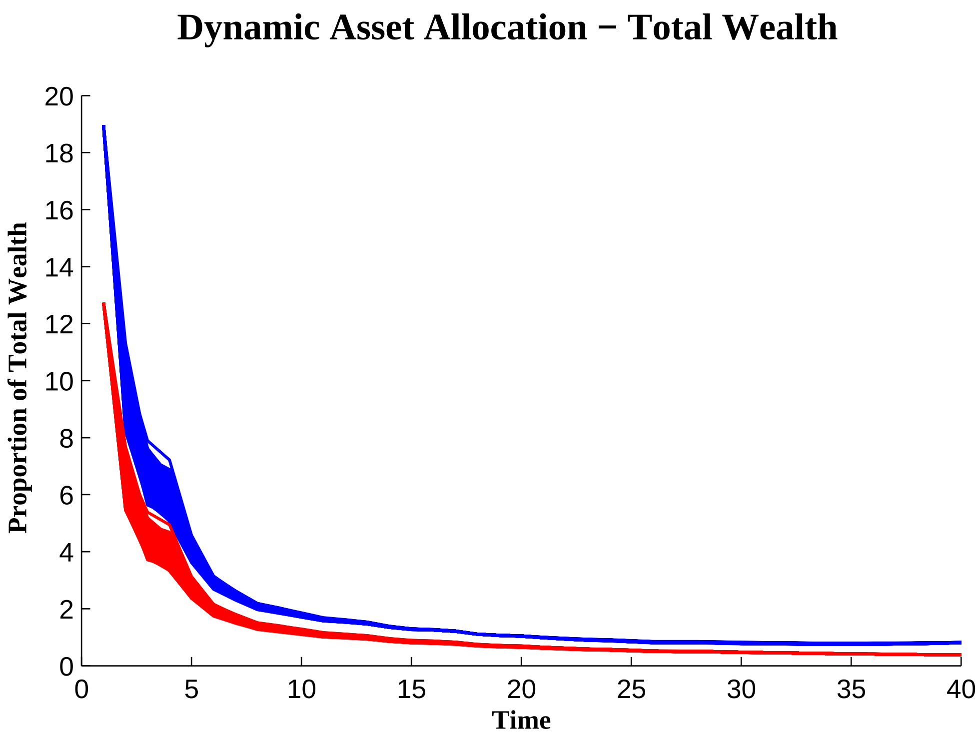

Figure 3: This figure shows the evolution of the optimal dynamic asset allocation for the risky asset (red) and mortality-linked security (blue), over the 40 year life of the DC fund considered, for 1000 simulated paths conditional on the risky asset following a single, randomly selected, path. We can see that initially both proportions are significantly above one, indicating the large initial short position in the risk-free asset. This size of this short position reduces quickly over time as contributions are received and used to pay back the initial borrowings, while the nominal value of the fund assets increases.

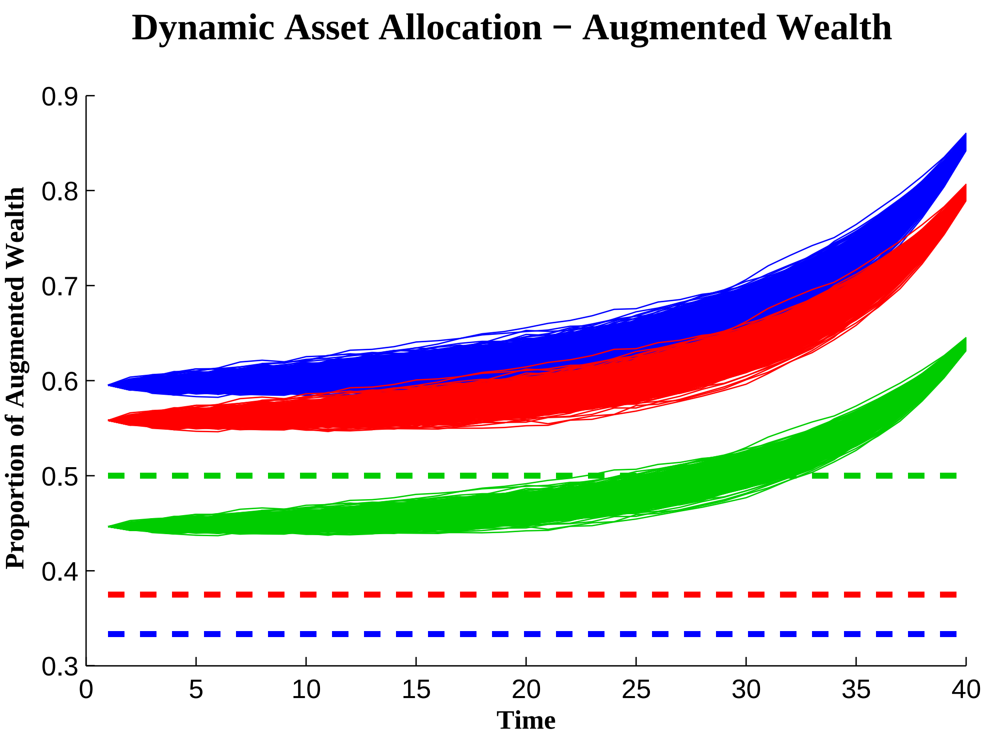

Figure 4: This figure shows the two optimal asset allocation strategies expressed in terms of the augmented pension fund wealth for 3 different degrees of relative risk aversion: 3 (green), 6 (red) and 9 (blue). We can see that the risky asset allocations (dashed lines) are constant across both time and path; that is, the proportion of augmented wealth allocated to the risky asset does not change deterministically over time or in response to shocks to mortality. The allocation to the longevity bond portfolio is both time varying and responds to mortality shocks, as we would expect. The upward sloping time variation is attributable to the asymmetry between the individual longevity bond prices, which implicitly factor in the possibility that the plan member could die prior to retirement, and the annuity payments, which are conditional on survival to retirement.

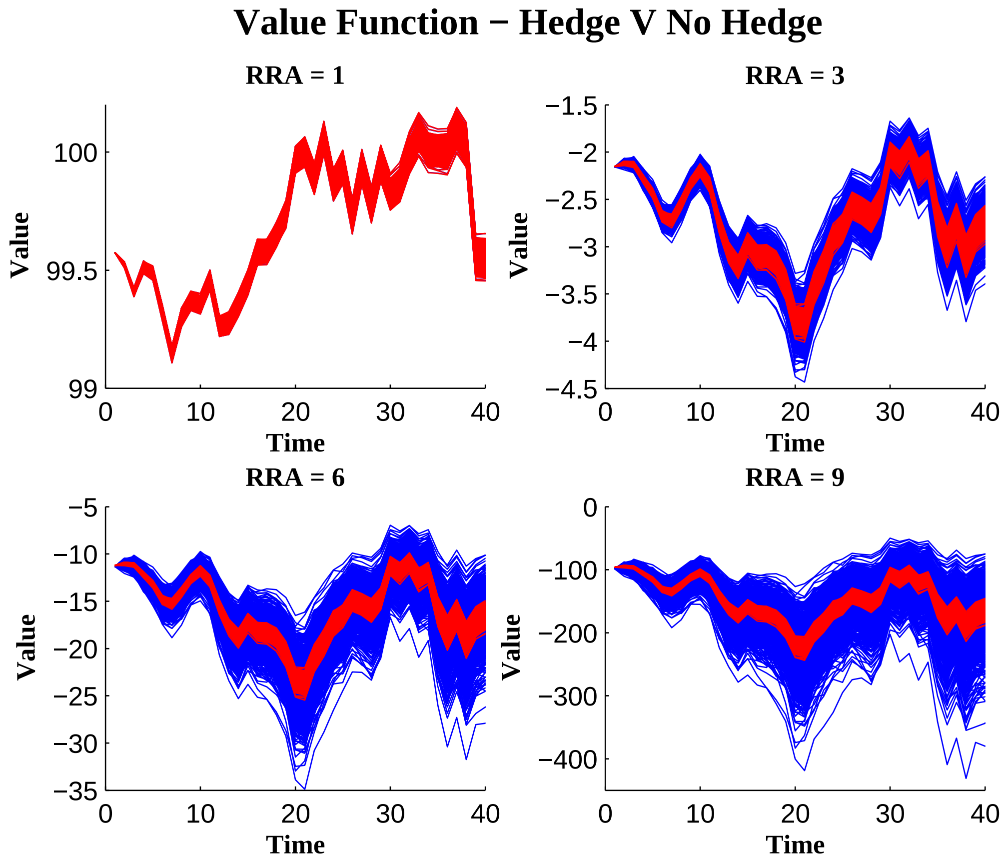

Figure 5: This figure presents 1000 simulated value function paths for both the incomplete market case (blue) and the complete market case (red),for 4 different levels of relative risk aversion: 1 (top left), 3 (top right), 6 (bottom left) and 9 (bottom right). For an individual with a coefficient of relative risk aversion equal to 1 the cases are very similar; that is, as the plan member is near risk neutral he does not place much value on the longevity risk hedge and therefore follows much the same strategy as he would if the hedge were unavailable. As the level of risk aversion increases, however, the relative risk associated with the incomplete market case increases, becoming quite stark for the higher levels.

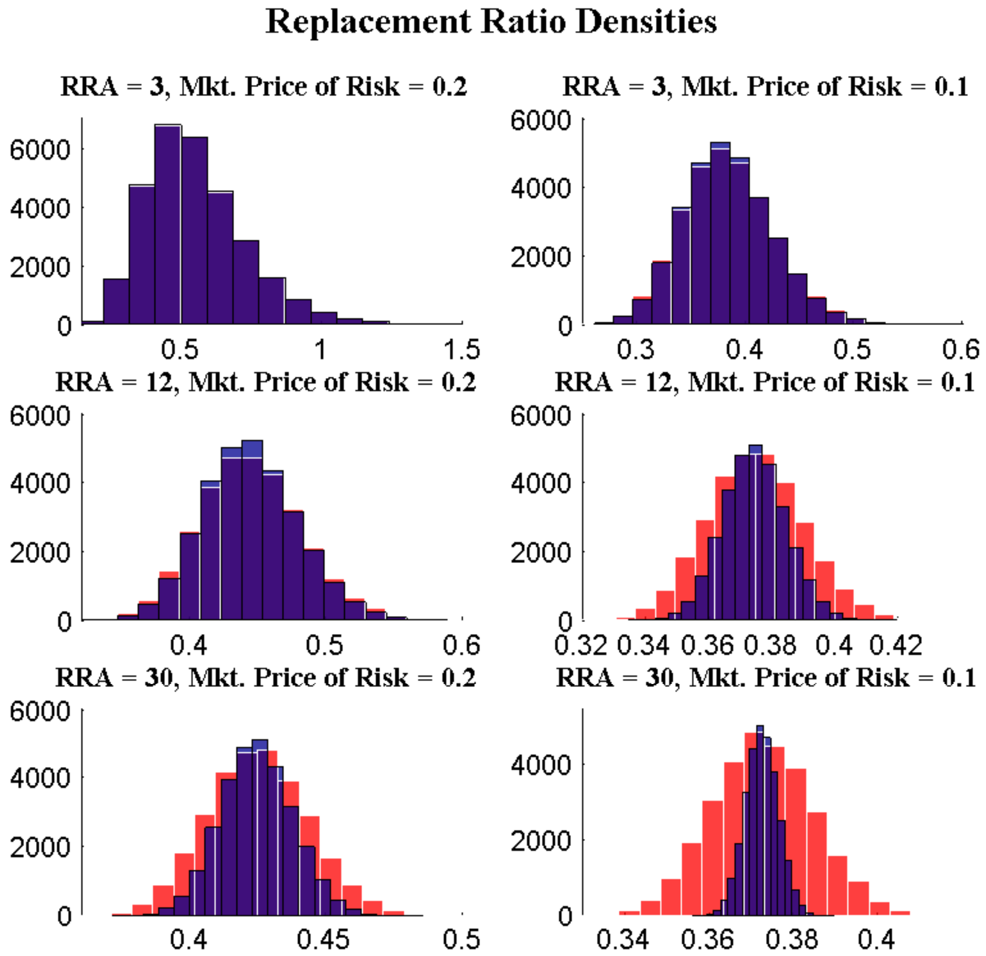

Figure 6: This figure shows the estimated replacement ratio probability densities for a selection of key parameter values, estimated using Monte Carlo simulation over 30,000 paths. The parameter values varied are the market price of risk coefficient, ξ, and the coefficient of relative risk aversion, γ; two parameters that significantly affect the proportion of (augmented) pension wealth allocated to the risky asset, S. For both a relatively low level of risk aversion and relatively high market price of risk (top left), we do not observe any benefit (by way of a reduction in risk) from the longevity hedge i.e. the histogram representing the density estimate of complete market case (blue) eclipses that representing that of the incomplete market case. This is due primarily to the fact that optimal allocation to the risky asset is high for these parameter values, and therefore the proportion of total portfolio variance attributable to the risky asset far outweighs that attributable to the mortality linked security. If, however, we consider a relatively high level of risk aversion and a relatively low market price of risk (bottom right), we observe a considerable benefit. Here, with a low allocation to the risky asset, the principal driver behind the replacement ratio variance is the annuity rate, and therefore the effect of the hedge is more pronounced. The hedge effectiveness can be seen to be increasing from the density at the top left to that at the bottom right in line with a decreasing risky asset allocation.

Table 1: Parameter estimates

	$\lambda_o(x,t)$		$\zeta(x,t)$		
Parameter	b	m	θ	δ	σ_ζ
Estimate	10.05559	84.5957	0.000194	0.008367	0.019674

This table shows the maximum likelihood parameter estimates for the CIR mortality improvement process, along with the least-squares parameter estimates for the base mortality curve.

Table 2: Increase in contributions required to offset longevity risk

RRA	Mkt. Price of Risk	Asset Allocation		Rep. Ratio StDev		Cont. Increase		
		S	L	Hedge	No Hedge	90%	95%	99%
	0.1	0.333		0.041	0.042	0.58%	0.62%	0.80%
3	0.15	0.417	0.667	0.095	0.096	0.12%	0.41%	0.52%
	0.2	0.500		0.177	0.178	0.00%	0.26%	0.76%
	0.1	0.292		0.020	0.023	1.09%	1.35%	1.83%
6	0.15	0.333	0.833	0.044	0.046	0.61%	0.67%	0.86%
	0.2	0.375		0.075	0.077	0.42%	0.49%	0.68%
	0.1	0.271		0.010	0.015	1.88%	2.33%	3.22%
12	0.15	0.292	0.917	0.021	0.025	1.10%	1.38%	1.81%
	0.2	0.313		0.035	0.038	0.83%	0.91%	1.19%
	0.1	0.262		0.006	0.013	2.49%	3.19%	4.51%
21	0.15	0.274	0.952	0.012	0.017	1.73%	2.16%	2.93%
	0.2	0.286		0.019	0.024	1.30%	1.61%	2.07%
	0.1	0.258		0.004	0.012	2.85%	3.66%	5.22%
30	0.15	0.267	0.967	0.008	0.015	2.14%	2.70%	3.77%
	0.2	0.275		0.013	0.019	1.67%	2.09%	2.81%

This table shows the required contributions for 3 different confidence levels (90%, 95% and 99%), based on estimated probability densities analogous to those in figure 6, for a selection of different relative risk aversion and market price of risk parameter values. The required increases are modest across all scenarios, ranging from no increase at the 90% confidence level for the scenario corresponding to the largest risky asset allocation, to an increase of 5.22% (i.e. from 10% to 10.522%) at the 99% confidence level for the scenario corresponding to the least risky asset allocation. The proportion of pension fund wealth invested in the risky asset, S, and the mortality-linked asset, L, immediately prior to retirement, are shown in the third and fourth columns respectively so as to give an idea of the relative increases and decreases across the different scenarios.