

Yaari's LifeCycle Model in the 21st Century: Consumption Under a Stochastic Force of Mortality

H. Huang, M.A. Milevsky¹ and T.S. Salisbury

Version: 6 January 2011

¹Huaxiong Huang is Professor of Mathematics and Statistics at York University. Moshe A. Milevsky is Associate Professor of Finance, York University, and Executive Director of the IFID Centre. Thomas S. Salisbury is Professor of Mathematics and Statistics at York University, all in Toronto, Canada. The contact author (Milevsky) can be reached via email at: milevsky@yorku.ca. The authors acknowledge the helpful comments of seminar participants at the Lifecycle, Insurance, Finance and Economics (LIFE) conference at the Fields Institute, Georgia State University and the University of Technology, Sydney. This research is supported in part by grants from NSERC, MITACS and The IFID Centre.

Abstract

Yaari's LifeCycle Model in the 21st Century: Consumption Under a Stochastic Force of Mortality

We extend the lifecycle model (LCM) of consumption over a random-length lifecycle (a.k.a. the Yaari model) to a world in which (i.) the force of mortality obeys a diffusion process as opposed to being deterministic, and (ii.) a consumer can adapt their consumption strategy to new information about their mortality rate (a.k.a. health status) as it becomes available. We compare and contrast the optimal consumption paths to investigate the impact of mortality *rate* uncertainty vs. simple *lifetime* uncertainty, but assuming the actuarial survival curves are the same. In addition to deriving the optimal consumption function, our main result is that when utility preferences are logarithmic the initial consumption rates are identical to the Yaari model. But, in a CRRA framework in which the coefficient of relative risk aversion is greater (smaller) than one the consumption rate is higher (lower) and a stochastic force of mortality changes optimal behavior even when faced with the same probability of survival. Numerical examples are provided to illustrate the magnitude of this effect. Our results should be relevant to researchers interested in calibrating the lifecycle model as well as those who provide normative guidance (a.k.a. financial advice) to retirees based on the LCM.

1 Introduction and Motivation

Assume that two hypothetical retirees – i.e. consumers who are not expecting any future labour income – approach a financial economist for guidance on how they should spend their accumulated financial capital over their remaining lifetime; a time horizon they both acknowledge is stochastic. Assume both retirees have time-separable and rational preferences and seek to maximize discounted utility of lifetime consumption with the same elasticity of intertemporal substitution ($1/\gamma$), the same subjective discount rate (ρ) and the same initial financial capital constraint (F_0). They have no declared bequest motives and – for whatever reason – neither are willing (or able) to invest in anything other than a risk-free asset with instantaneous return (r); which means they are *not* looking for guidance on asset allocation. All they want is an optimal consumption plan ($c^*(t); t \geq 0$) guiding them from time zero (retirement) to the last possible time date of death ($t \leq D$). Most importantly, both retirees agree-on and share the same probability-of-survival curve denoted by $p(s)$. In other words they currently live in the same “health state” and the same effective biological age. For example, they both agree on a $p(35) = 5\%$ probability that either of them survive for 35 years and a $p(20) = 50\%$ probability that either of them survive for 20 years, etc.

Over 45 years ago, Menahem Yaari (1965) in a classical and highly-cited paper showed exactly how to solve such a problem. He derived the Euler-Lagrange equation for the optimal trajectory of wealth and the related consumption function. He was the first to show how to work with lifetime uncertainty in a lifecycle model (LCM) and amongst other results provided a rigorous foundation for Irving Fisher’s claim that lifetime uncertainty effectively increases consumption impatience and is akin to behavior under higher subjective discount rates.

In Yaari’s (20th century) model both of the above-mentioned retirees would be told to follow identical consumption paths until their random date of death. In fact, they would both be guided to optimally consume $c(t)^* = F(t)/a(t)$, where $a(t)$ is a *function of time only* and is related to an actuarial annuity factor. We will explain this factor in more detail, later in the paper.

But here is the impetus for this paper. Although both retirees appear to have the same longevity risk assessment and agree-on the survival probability curve $p(s)$, they have *differing views about the volatility of their health as proxied by a mortality rate volatility*. In the language of 21st century actuarial science, the first retiree (#1) believes that his instantaneous force of mortality (denoted by $\lambda^{\text{DfM}}(t)$) will grow at a deterministic rate until he eventually dies, while the second retiree (#2) believes that her force of mortality (denoted by $\lambda^{\text{SfM}}(t)$) will grow at stochastic (but measurable) rate until a random date of death. As such, the remaining lifetime random variable is doubly stochastic. While this distinction might sound farfetched and artificial, a growing number of researchers in the actuarial literature are using this model for the biology of the model, and not the simplistic (old) model used

by economists. The actuaries motivation in advocating for a stochastic force of mortality, is to generate more robust pricing and reserving for mortality-contingent claims. These studies – although not central to this paper and reviewed in Cairns, Blake and Dowd (2006) for example – have all argued that SfM models better reflect the uncertainty inherent in demographic projections *vis a vis* the inability of insurance companies to diversify mortality risk entirely. We ask: *how do the recent actuarial model impact the economics of the problem?*

When one thinks about it, real-life mortality rates are indeed stochastic, capturing (unexpected) improvements in medical treatment, or (unexpected) epidemics, or even (unexpected) changes to the health status of an individual. Rational consumers choosing to make saving and consumption decisions using models based on deterministic mortality rates would likely agree to re-evaluate those decisions if their views about the values of those mortality rates change dramatically. Our thesis is that economic decision-making can only be improved if mortality models reflect the realistic nature of mortality rates.

We will carefully explain the mathematical distinction between deterministic and stochastic forces of mortality (SfM) in section #2 of this paper, but just to make clear here, at time zero both our hypothetical retirees agree on the survival probability curve $p(s)$. However, at any future time their survival probability curves will deviate from each other depending on the realization of the mortality rate between now and then.

So, motivated by these (21st century) models of mortality, in this paper we derive the optimal consumption function for both retirees; one who believes in – and operates under a – stochastic mortality and one who doesn't. Stated differently, we will solve the Yaari (1965) model where the optimal consumption plan is given as a function of wealth, time *and the evolving mortality rate as a state variable*. Indeed, with thousands of references to the Yaari (1965) paper in the economic literature, and the growing interest in stochastic mortality models in the actuarial community, we believe these results will be of interest to both communities of researchers.

Recall that in the Yaari model conditioning on the mortality rate was redundant or unnecessary since its evolution over time was deterministic. All one needed was the value of wealth $F(t)$ and time t . But, in a stochastic mortality model, the mortality rate itself becomes a state variable. In this paper we show how the uncertainty of mortality interacts with longevity risk aversion (γ) – which is the reciprocal of the intertemporal elasticity of substitution – to yield an optimal consumption plan.

To briefly preview our results, we describe the conditions under which retiree #1 (deterministic mortality) will start-off consuming more than retiree #2 (stochastic mortality), as well the conditions under which retiree #1 consumes less than retiree #2, and the (surprising) conditions under which they both consume exactly the same. We provide numerical examples under a variety of specific mortality models and examine the magnitude of this effect.

The remainder of this paper is organized as follows. In section #2 we explain in more detail exactly how a stochastic model of mortality differs from the more traditional (and widely used in economics) deterministic force of mortality. In section #3 we take the opportunity to review Yaari (1965) and set our notation and benchmark for the stochastic model. In section #4 we characterize the optimal consumption function in the stochastic mortality model under the most general assumptions, and prove a theorem regarding the relationship between consumption in this model vs. Yaari (1965). In section #5 we make some specific assumptions regarding the stochastic mortality rate and illustrate the magnitude of this effect, and section #6 summarizes our main results and concludes the paper. The appendix contains mathematical details and algorithms that are not central to our main economic contributions.

First, we explain exactly the difference between deterministic and stochastic force of mortality.

2 Understanding The Force of Mortality

In most of the relevant papers in the LCM literature over the last 45 years – starting with Yaari (1965), Richard (1975), Davies (1985), Leung (1990) – the force of mortality from time zero to the last possible date of death is known with certainty. Ergo, the conditional survival probabilities over the entire retirement horizon are known and predictable at time zero. So, if a 65-year-old retiree is told (by his doctor) that he faces a 5% chance of surviving to age 100 and a 37% chance of surviving to age 90, then by definition there is a 13.5% = $(0.05/0.37)$ probability of surviving to age 100, if he is still alive at age 90. In other words, he makes consumption decisions today that trade-off utility in different states of nature, knowing that if-and-when he reaches the age of 90, there will only be a 13.5% chance he will survive to age 100. In the language of actuarial science, the table of individual $\{q_{x+i}; i = 0, \dots, N\}$ mortality rates are known in advance. This is the essence of a deterministic force of mortality and textbook life contingencies. If q_{65} is the retiree’s probability of dying between age 65 and 66, while q_{66} is the probability of the same retiree dying between age 66 and 67, then the probability of surviving from age 65 to age 67 is $(1 - q_{65})(1 - q_{66})$.

In stark contrast, under a stochastic force of mortality the above predictability – or multiplicative relationship – breaks down. While a 65-year-old might currently face a 5% estimated probability of surviving to age 100 and a 37% chance of reaching age 90, there is absolutely no guarantee that the conditional survival probability from any future age, to age 100 (given the observed mortality rates), will satisfy the ratio. At time zero there is an expectation of what the probability will be at age 90. But, the probability itself is random. This way of thinking – which might be new to economists – is the essence of a stochastic force of mortality and is the impetus for our paper.

Here it is formally. Let $\lambda(t)$ denote the mortality rate of a cohort of a population, which may be stochastic or deterministic. Let $\mathcal{F}_t = \sigma\{\lambda(q) \mid q \leq t\}$ be the filtration determined by λ . Then individuals in the population have lifetimes of length ζ satisfying

$$P(\zeta > s \mid \zeta > t, \mathcal{F}_\infty) = e^{-\int_t^s \lambda(q) dq}. \quad (1)$$

Assume further that $\lambda(t)$ is a Markov process, and define the survival function $p(t, s, \lambda)$ by

$$p(t, s, \lambda) = E \left[e^{-\int_t^s \lambda(q) dq} \mid \lambda(t) = \lambda \right]. \quad (2)$$

This gives the conditional probability of surviving from time t to time s , given knowledge of the mortality rate at time t . Therefore

$$P(\zeta > s \mid \zeta > t, \mathcal{F}_t) = E \left[e^{-\int_t^s \lambda(q) dq} \mid \mathcal{F}_t \right] = p(t, s, \lambda(t)). \quad (3)$$

If $t = 0$ then we write $p(s, \lambda)$ for $p(0, s, \lambda)$.

Our basic problem in this paper will be to compare optimal consumption under two models that share a common initial value λ_0 of the mortality rate, as well as a common survival function $p(t, \lambda_0)$. Typically one will be deterministic and one stochastic. When we do actual computations, we will either choose a specific deterministic model and calibrate a stochastic model to it. Or conversely, we will choose a stochastic model and calibrate the deterministic model to it. Both possibilities are discussed below. It should be clear from the context which model we are discussing. But when it is necessary to make this distinction explicitly, we will write $\lambda^{\text{DfM}}(t)$ and $\lambda^{\text{SfM}}(t)$.

2.1 Deterministic force of Mortality (DfM)

Let $\lambda_0 = \lambda(0)$ be the initial value of the mortality rate. In the deterministic case,

$$p(t, \lambda_0) = e^{-\int_0^t \lambda(q) dq}, \quad (4)$$

and we can recover $\lambda(t)$ as $-p_t(t, \lambda_0)/p(t, \lambda_0)$, where the t -subscript denotes the time derivative. In other words, if we start with a concrete stochastic model, and obtain the survival curve $p(t, \lambda_0)$ from it, the above formula determines the calibration of the deterministic force of mortality model to it. This approach is computationally simpler, but has the disadvantage that neither the stochastic nor deterministic model is in a simple form, familiar to practitioners. In other words, a “simple” model for the stochastic force of mortality rates leads to a “complicated” model for the deterministic force of mortality, and vice versa.

When doing actual calculations we will start by assuming that $\lambda(t)$ follows a standard Gompertz model. In other words, that

$$d\lambda(t) = \eta\lambda(t) dt \quad (5)$$

so $\lambda(t) = \lambda_0 e^{\eta t}$. The usual form for Gompertz is $\lambda(t) = b^{-1} e^{(x+t-m)/b}$, so here we are using $\eta = 1/b$ and $\lambda_0 = b^{-1} e^{(x-m)/b}$. This model is simple, and takes advantage of long experience calibrating the Gompertz model to real populations.

Note that in the deterministic setting,

$$p(t, s, \lambda(t)) = e^{-\int_t^s \lambda(q) dq} = e^{-\int_0^s \lambda(q) dq} / e^{-\int_0^t \lambda(q) dq} = p(s, \lambda_0) / p(t, \lambda_0). \quad (6)$$

This will typically NOT be true in the stochastic setting. As long as we keep in mind that we are calibrating at time 0 (i.e. to $p(t, \lambda_0)$ only) that should not cause problems.

Table #1: Conditional Survival Probability: Deterministic Mortality								
	$x = 65$	$x = 70$	$x = 75$	$x = 80$	$x = 85$	$x = 90$	$x = 95$	$x = 100$
To Age 65	1.000							
To Age 70	0.9479	1.000						
To Age 75	0.8659	0.9135	1.000					
To Age 80	0.7429	0.7837	0.8580	1.000				
To Age 85	0.5733	0.6047	0.6620	0.7716	1.000			
To Age 90	0.3696	0.3899	0.4268	0.4975	0.6447	1.000		
To Age 95	0.1758	0.1855	0.2031	0.2367	0.3067	0.4757	1.000	
To Age 100	0.0500	0.0527	0.0577	0.0673	0.0872	0.1353	0.2844	1.000
λ_x	0.0081	0.0137	0.0232	0.0394	0.0667	0.1129	0.1911	0.3234

Table #1 displays a typical deterministic mortality survival probability “matrix” of values together with the corresponding mortality rate at each age x , on the bottom row. Note that these numbers were generated using a (deterministic) Gompertz model in which $m = 89.335$ and $b = 9.5$. Indeed, given the initial probability of survival from age 65 to any age $y > 65$ (which is the first column in Table #1) one can solve for the conditional survival probability from age y to any age $z > y$, by dividing the two probability values. This is the essence of equation (6). Alas, when mortality rates are stochastic all numbers $p(t, s, \lambda(t))$ beyond the first column in Table #1, are unknown at time zero.

2.2 Stochastic Force of Mortality (SfM)

There are many possible stochastic models to choose from. Starting from the models of Lee and Carter (1992) to a review by Cairns, Blake and Dowd (2006) as well as Wills and Sherris (2010), actuaries have employed a variety of specifications for the stochastic $\lambda(t)$, subsequently used to price mortality and longevity risk. In what follows in the numerical examples, we adopt a lognormal mortality rate, which is often called the Dothan model for

interest rates in the derivative pricing literature. Although it might seem natural to have constant drift and diffusion coefficients, in order to calibrate to a given deterministic model, we allow a time-dependent growth coefficient. For most of the numerical examples provided later-on we take:

$$d\lambda(t) = \mu(t)\lambda(t) dt + \sigma\lambda(t) dB(t) \quad (7)$$

where $B(t)$ is a Brownian motion. This is obviously the source of randomness in the stochastic force of mortality.

There are many ways to select (or calibrate) a stochastic force of mortality to a particular survival curve. The details on how to actually compute this are provided in the second part of the appendix.

With the probability background out of the way, we now review the Yaari (1965) model which is based on a deterministic force of mortality.

3 Review of the Yaari (1965) Model

The lifecycle model (LCM) with a random date of death and assuming no bequest motive, can be written as follows:

$$J = \max_c E \left[\int_0^D e^{-\rho t} u(c(t)) 1_{\{t \leq \zeta\}} dt \right], \quad (8)$$

where $\zeta \leq D$ is the remaining lifetime satisfying $\Pr[\zeta > t] = p(t, \lambda_0)$, defined above in Section #2. When the mortality rate is deterministic one can obviously assume independence between the optimal consumption $c^*(t)$ and the lifetime indicator variable $1_{\{t \leq \zeta\}}$, so that by Fubini's theorem we can re-write the value function as:

$$\begin{aligned} J &= \max_c \int_0^D e^{-\rho t} u(c(t)) E[1_{\{t \leq \zeta\}}] dt \\ &= \max_c \int_0^D e^{-\rho t} u(c(t)) p(t, \lambda_0) dt. \end{aligned} \quad (9)$$

From this perspective, there really isn't any more randomness in the model. This is a problem within the calculus of variations subject to some constraints on the function $c(t)$. The wealth (budget) constraint can be written as:

$$F_t(t) = v(t, F(t))F(t) + \pi_0 - c(t), \quad (10)$$

with boundary conditions $F(0) = W > 0$ and $F(D) = 0$. The parameter π_0 denotes a constant income rate which we include in this section for comparison with Yaari's model, but which in subsequent sections will be taken to equal zero; $c(t)$ is the consumption rate and the control variable in our problem; $v = v(t, F)$ is the interest rate function defined by:

$$v(t, F) = \begin{cases} r + \xi\lambda(t), & F \geq 0, \\ R + \lambda(t), & F < 0, \end{cases} \quad (11)$$

where $r \leq R \leq \infty$. The (new) parameter ξ indicates and denotes the availability of “actuarial notes” as introduced and described by Yaari (1965). In words, $v := v(t, F)$ is the investment return r when wealth is positive so that $F \geq 0$. It is equal to the borrowing rate $R + \lambda(t)$ when wealth is negative, ie. $F < 0$. Note the composite structure which includes the mortality rate $\lambda(t)$ for an individual at time t to reflect the reality that unsecured loans are available as long as they are purchased with life insurance protection for the creditor.

When $\xi = 1$ and $R = r$, the function $v(t, F)$ collapses to $r + \lambda(t)$ and we are in Case B of Yaari (1965). When $\xi < 1$ the investor earns less than their mortality rate when holding actuarial notes, due to adverse selection considerations. Finally, when $R = \infty$ we are effectively imposing a (no) borrowing constraint and when $\xi = 0$ as well we do not allow for holding actuarial notes (or purchasing additional annuities.) For the purposes of this paper and from this point onwards we assume that $R = \infty$ and $\xi = 0$. In a follow-up paper. In that sense, this paper only brings a part of the Yaari model to the 21st century. we plan to examine the impact and availability of actuarial notes, i.e. the case when $\xi > 0$.

Although this wasn’t explicitly imposed in the Yaari (1965) model, in this paper we operate under a constant relative risk aversion (CRRA) formulation for the utility function. In principle this should mean using $\bar{u}(c)$, where:

$$\bar{u}(c) = \frac{c^{1-\gamma} - 1}{1 - \gamma} \quad (12)$$

for $\gamma > 0$ and $\gamma \neq 1$, with the understanding that when $\gamma = 1$ we define $\bar{u}(c) = \ln[c]$. This family of utilities varies continuously with γ .

Of course, it makes no difference to our optimization problems if we shift \bar{u} by an arbitrary additive constant. So to make scaling relationships easier to express, actual calculations will be carried out using the equivalent utilities

$$u(c) = \frac{c^{1-\gamma}}{1 - \gamma} \quad (13)$$

for $\gamma > 0$ and $\gamma \neq 1$. When $\gamma = 1$ we take $u(c) = \bar{u}(c) = \ln[c]$.

The marginal utility of consumption is the derivative of utility with respect to c , which is simply

$$u_c = c^{-\gamma} > 0. \quad (14)$$

To avoid the distractions of inflation models and assumptions, throughout this paper we assume that the interest rate r is expressed in real (after-inflation) terms and therefore consumption $c(t)$ is expressed in real terms as well.

Finally, when $v(t) = r$ during the entire interval $(0, D)$ then as a consequence of the Euler-Lagrange Theorem, the optimal financial capital trajectory $F(t)$ must satisfy the following linear second-order non-homogenous differential equation over the values for which $F(t) \neq 0$.

$$F_{tt}(t) - \left(\frac{r - \rho - \lambda(t)}{\gamma} + v \right) F_t(t) + v \left(\frac{r - \rho - \lambda(t)}{\gamma} \right) F(t) = - \left(\frac{r - \rho - \lambda(t)}{\gamma} \right) \pi_0. \quad (15)$$

When the pension income rate $\pi_0 = 0$ the differential equation collapses to the homogenous case.

3.1 Explicit Solution: Gompertz Mortality

When the (deterministic) mortality rate function obeys the (pure) Gompertz law of mortality

$$\lambda(t) = \frac{1}{b} \exp\left(\frac{x+t-m}{b}\right), \quad (16)$$

the survival probability is

$$p(t, \lambda_0) = \exp\left\{-\int_0^t \lambda(q) dq\right\} = \exp\{b\lambda_0(1 - e^{t/b})\}. \quad (17)$$

Here x denotes the age at time 0, m is called the modal value and b is the dispersion coefficient for the Gompertz model. To simplify notation let

$$k(t) = \frac{r - \rho - \lambda(t)}{\gamma}, \quad (18)$$

and recall from the budget constraint that:

$$c(t) = vF(t) - F_t(t) + \pi_0, \quad (19)$$

$$c_t(t) = vF_t(t) - F_{tt}(t). \quad (20)$$

Equation (15) can be rearranged as

$$F_{tt}(t) - vF_t(t) + k(t)(vF(t) - F_t(t)) = -k(t)\pi_0, \quad (21)$$

which then leads to

$$k(t)c^*(t) - c_t^*(t) = 0 \quad (22)$$

The solution to this basic equation is

$$\begin{aligned} c^*(t) &= c^*(0)e^{\int_0^t k(s)ds} = c^*(0)e^{\int_0^t \left(\frac{r-\rho-\lambda(s)}{\gamma}\right)ds} = c^*(0)e^{\left(\frac{r-\rho}{\gamma}\right)t} e^{-\frac{1}{\gamma} \int_0^t \lambda(s) ds} \\ &= c^*(0)e^{\left(\frac{r-\rho}{\gamma}\right)t} (p(t, \lambda_0))^{1/\gamma}, \end{aligned} \quad (23)$$

where $c^*(0)$ is the optimal initial consumption rate, to be determined, which is the one free constant resulting from equation (22). Note that when the interest rate r is equal to the subjective discount rate ρ , and $\gamma = 1$ (i.e. log utility), the optimal consumption rate at any age $x+t$ is the probability of survival to that age times the initial consumption $c^*(0)$. However, when $\gamma > 1$, which implies higher levels of risk aversion, the optimal consumption rate will decline at a slower rate as the retiree ages. Longevity risk aversion induces people to behave as if they were going to live longer than determined by the actuarial mortality

rates. We will explore the impact of γ on the optimal consumption path in a stochastic force of mortality model, later in Section #4, which is why it's important to focus on this here.

Mathematically one can see that $(p(t, \lambda_0))^{1/(\gamma+\varepsilon)}$ is greater than $(p(t, \lambda_0))^{1/\gamma}$ for any $\varepsilon > 0$ since $p(t, \lambda_0) < 1$ for all t . Finally, note that in the Gompertz mortality model evaluating $(p(t, \lambda_0))^{1/\gamma}$ for a given (x, m, b) triplet is equivalent to evaluating $p(t, \lambda_0)$ under the same x, b values, but assuming that $m^* = m + b \ln \gamma$. This then implies that one can tilt/define a new deterministic mortality rate $\hat{\lambda}_0 = \gamma \lambda_0$ and derive the optimal consumption as if the individual was risk neutral. This will be used later in the explicit expression for $F(t)$ and $c^*(t)$

Moving on to a solution for $F(t)$, we now substitute the optimal consumption solution (23) into equation (19) to arrive at yet another first-order ODE, but this time for $F(t)$:

$$F_t(t) - vF(t) - \pi_0 + c^*(0)e^{(\frac{r-\rho}{\gamma})t}(p(t, \lambda_0))^{1/\gamma} = 0. \quad (24)$$

Writing down the canonical solution to this equation leads to:

$$F(t) = \left(\pi_0 \int_0^t e^{-rs} ds - c^*(0) \int_0^t e^{(\frac{r-\rho}{\gamma})s} (p(s, \lambda_0))^{1/\gamma} e^{-rs} ds + F(0) \right) e^{rt}, \quad (25)$$

where F_0 denotes the free initial condition from the ODE for $F(t)$ in equation (24). Recall that we still haven't specified $c^*(0)$, the initial consumption). We will do so (eventually) by using the terminal condition $F(D) = 0$.

To represent the wealth trajectory explicitly define the following (new) Gompertz Present Value (GPV) function

$$\begin{aligned} a_x^T(r, m, b) &= \int_0^T p(s, \lambda_0) e^{-rs} ds = \int_0^T e^{-\int_0^s (r+\lambda(t)) dt} ds \\ &= \int_0^T e^{-\int_0^s \left(r + \frac{1}{b} e^{\left(\frac{x-m+t}{b} \right)} \right) dt} ds \\ &= \frac{b\Gamma(-rb, \exp\{\frac{x-m}{b}\}) - b\Gamma(-rb, \exp\{\frac{x-m+T}{b}\})}{\exp\{(m-x)r - \exp\{\frac{x-m}{b}\}\}}. \end{aligned} \quad (26)$$

The function $a_x^T(r, m, b) = a(t)$ is the age- x cost of a life-contingent annuity that pays \$1 per year continuously provided the annuitant is still alive, but only until time $t = T$, which corresponds to age $x + T$. If the individual survives beyond age $(x + T)$ the payout stops. Naturally, when $T = \infty$ the expression collapses to a conventional single premium income annuity (SPIA).

Note that $\Gamma(A, B)$ is the incomplete Gamma function. In other words, equation (26) is analytic and in closed-form.

The reason for introducing the GPV is that combining equation (25) with equation (26) leads to the (very tame looking) expression

$$F(t) = \left(F(0) + \frac{\pi}{r} \right) e^{rt} - a_x^t(r - k, m^*, b) c^*(0) e^{rt} - \frac{\pi_0}{r}, \quad (27)$$

where recall that $m^* = m + b \ln[\gamma]$. Then, using the boundary condition $F_\tau = 0$, where τ is the wealth depletion time, we obtain an explicit expression for the initial consumption

$$c^*(0) = \frac{(F(0) + \pi_0/r) e^{r\tau} - \pi_0/r}{a_x^\tau(r - k, m^*, b) e^{r\tau}}. \quad (28)$$

3.2 Consumption Under DfM: Numerical Examples

In our numerical examples we assume an 86.6% probability that a 65-year-old will survive to the age of 75, a 57.3% probability of reaching 85, a 36.9% probability of reaching 90, a 17.6% probability of reaching age 95 and a 5% probability of reaching 100. These are the values generated by the Gompertz law with $m = 89.335$ and $b = 9.5$. To complete the parameter specifications required for our model, we assume the subjective discount rate (ρ) is equal to the risk-free rate $r = 2.5\%$. Within the context of a lifecycle model, this implies that the optimal consumption rates would be constant over time in the absence of longevity and mortality uncertainty.

We are now ready for some results. Assume a 65-year-old with a (standardized) \$100 nest egg. Initially we allow for no pension annuity income ($\pi_0 = 0$) and therefore all consumption must be sourced to the investment portfolio which is earning a deterministic interest rate $r = 2.5\%$. The financial capital $F(t)$ must be depleted at the very end of the lifecycle, which is time $D = (120 - 65) = 55$ and there are no bequest motives. So, according to equation (28), the optimal consumption rate at retirement age 65 is \$4.605 when the risk aversion parameter is $\gamma = 4$ and the optimal consumption rate is (higher) \$4.121 when the risk aversion parameter is set to (higher) $\gamma = 8$.

As the retiree ages ($t > 0$) he/she rationally consume less each year – in proportion to the survival probability adjusted for γ . For example, in our baseline $\gamma = 4$ level of risk aversion, the optimal consumption rate drops from \$4.605 at age 65, to \$4.544 at age 70 (which is $t = 5$), then \$4.442 at age 75 (which is $t = 10$), then \$3.591 at age 90 (which is $t = 25$) and \$2.177 at age 100 (which is $t = 35$), assuming the retiree is still alive. A lower real interest rate (r) leads to a reduced optimal consumption/spending rate. All of this can be sourced to equation (23).

Thus, one of the important insights from the Yaari (1965) model is that a fully rational consumer will actually spend less as they progress through retirement. The optimizer spends more at earlier ages and reduces spending with age, even if his/her subjective discount rate (SDR) is equal to (or less than) the real interest rate in the economy.

Intuitively the individual deals with longevity risk by planning to reduce consumption – if that risk materializes – in proportion to the survival probability, linked to their risk aversion. The Yaari (1965) model provides a rigorous foundation to Irving Fisher (1930) statement in his book *Theory of Interest* (page 85): “... *The shortness of life thus tends powerfully to increase the degree of impatience or rate of time preference beyond what it*

would otherwise be. . .” and (page 90) “Everyone at some time in his life doubtless changes his degree of impatience for income. . . When he gets a little older, . . . he expects to die and he thinks: instead of piling up for the remote future, why shouldn’t I enjoy myself during the few years that remain?”¹

3.3 Time-zero Consumption Ratio = Initial Withdrawal Rate

Finally, in the very specific case when $\pi_0 = 0$ (which implies that the wealth depletion time is $\tau = D$) and the subjective discount rate $\rho = r$, the retiree must rely exhaustively on his/her initial wealth F_0 . We get

$$\frac{c^*(0)}{F(0)} = \frac{1}{a_x^D(r - \lambda_0/\gamma, m^*, b)} \quad (29)$$

We now have all the ingredients to compare with a stochastic model. This ratio is often called the Initial Withdrawal Rate (IWR) amongst financial practitioners and in the retirement spending literature.

4 Optimal Consumption: General Results

In this section we obtain the *most general* optimal consumption strategy for a retiree maximizing expected discounted utility of consumption, which will include the Yaari (1965) model as a special case. Since our main focus now is on the mortality model, at this stage we make the additional assumption $\rho = r$, that is, that the subjective discount rate equals the interest rate in the economy. Also, in contrast to the discussion in the previous section, we assume no exogenous pension income, so that $\pi_0 = 0$, which then precludes any borrowing. Once again we assume a fixed terminal horizon D , which denotes the last possible date of death. The mathematical formulation is to find

$$J = \max_{c(s) \text{ adapted}} E \left[\int_0^T e^{-\int_0^s (r + \lambda(q)) dq} u(c(s)) ds \mid \lambda(0) = \lambda, F(0) = F \right]. \quad (30)$$

Whereas in section #3 of this paper we used basic calculus of variations to derive the optimal trajectory of wealth and the consumption function in the Yaari (1965) model, given the inclusion of mortality as a state variable we must resort to dynamic programming techniques to obtain the optimality conditions. Regardless of the different techniques, we will show how the optimal consumption function collapses to the Yaari (1965) model when the volatility of mortality is zero.

¹For additional (case specific) examples of the Yaari (1965) model in action during the non-labour income retirement phase, we refer the interested reader to Milevsky and Huang (2010) or a recent paper by Lachance (2010).

Define:

$$J(t, \lambda, F) = \max_{c(s) \text{ adapted}} E \left[\int_t^T e^{-\int_t^s (r+\lambda(q)) dq} u(c(s)) ds \middle| \lambda(t) = \lambda, F(t) = F \right]. \quad (31)$$

As in the deterministic mortality model, the wealth process (which we shall soon see is stochastic) satisfies $dF(t) = (rF(t) - c(t))dt$. Assume that there is an optimal control. Then for that control,

$$\begin{aligned} E \left[\int_0^T e^{-\int_0^s (r+\lambda(q)) dq} u(c(s)) ds \middle| \mathcal{F}_t \right] \\ = e^{-\int_0^t (r+\lambda(q)) dq} J(t, \lambda(t), F(t)) + \int_0^t e^{-\int_0^s (r+\lambda(q)) dq} u(c(s)) ds \end{aligned} \quad (32)$$

is a martingale. This will likewise give a supermartingale under a general choice of c . Applying Itô's lemma, we obtain the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\sup_c \{u(c) - cJ_F\} + J_t - (r + \lambda)J + rFJ_F + \mu(t)\lambda J_\lambda + \frac{\sigma^2 \lambda^2}{2} J_{\lambda\lambda} = 0. \quad (33)$$

If there is any possibility of confusion, we will denote this value function $J^{\text{SM}}(t, \lambda, F)$.

For deterministic mortality, HJB can be obtained by sending $\sigma \rightarrow 0$ with $\mu(t) = \eta$, which was equal to $1/b$ in the Yaari (1965) model derived in Section #3, as

$$\sup_c \{u(c) - cJ_F\} + J_t - (r + \lambda)J + rFJ_F + \eta\lambda J_\lambda = 0. \quad (34)$$

Moving on to the optimal consumption plan, we solve the HJB equation under CRRA utility as follows: let

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \quad J = \frac{F^{1-\gamma}}{1-\gamma} a(t, \lambda) \quad (35)$$

and apply the 1st order condition $c^* = J_F^{-\frac{1}{\gamma}}$. We obtain $c^* = Fa^{-\frac{1}{\gamma}}$ and get the following equation for $a(t, \lambda)$:

$$a_t - (r\gamma + \lambda)a + \gamma a^{1-\frac{1}{\gamma}} + \mu(t)\lambda a_\lambda + \frac{\sigma^2 \lambda^2}{2} a_{\lambda\lambda} = 0 \quad (36)$$

with boundary condition $a(T, \lambda) = 0$.

We now solve the PDE for $a(t, \lambda)$, which we re-write as:

$$\beta_t + 1 - v\beta + \mu(t)\lambda\beta_\lambda + \frac{\gamma-1}{2\beta} \sigma^2 \lambda^2 \beta_\lambda^2 + \frac{1}{2} \sigma^2 \lambda^2 \beta_{\lambda\lambda} = 0. \quad (37)$$

for $\beta = \beta(t, \lambda) = a(t, \lambda)^{1/\gamma}$. Here also the new variable $v = r + \frac{\lambda}{\gamma}$. The boundary conditions are $\beta(T, \lambda) = 0$, $\beta_\lambda(t, \infty) = 0$ and at $\lambda = 0$ we solve $\beta_t + 1 - v\beta = 0$. Note that the optimal consumption rate is $c = F/\beta$, using shorthand notation.

We are now ready for our main theorem which implicitly answers the question posed in the introduction and motivation to this paper.

4.1 Stochastic Force of Mortality: Main Theorem

Denote by $c^{\text{SfM}}(t, \lambda, F)$ the optimal consumption at time t , given $\lambda(t) = \lambda$ and $F(t) = F$, under a stochastic force of mortality (SfM) model. Denote by $c^{\text{DfM}}(t, F)$ the optimal consumption at time t , when $F(t) = F$, under a deterministic force of mortality (DfM) model.

THEOREM: *Assume that the survival functions for the two models agree: $p^{\text{SfM}}(t, \lambda_0) = p^{\text{DfM}}(t, \lambda_0)$ for every $t \geq 0$, and that utility is CRRA(γ).*

- (a) $\gamma > 1 \implies c^{\text{SfM}}(0, \lambda_0, F) \geq c^{\text{DfM}}(0, F)$.
- (b) $\gamma = 1 \implies c^{\text{SfM}}(0, \lambda_0, F) = c^{\text{DfM}}(0, F)$.
- (c) $0 < \gamma < 1 \implies c^{\text{SfM}}(0, \lambda_0, F) \leq c^{\text{DfM}}(0, F)$.

PROOF: To see this, we change point of view, and work exclusively with the stochastic model. So we drop the SfM superscript, and write $p = p^{\text{SfM}}$, $J = J^{\text{SfM}}$, $c^* = c^{\text{SfM}}$, $\lambda = \lambda^{\text{SfM}}$, etc. Within that model, we pose two different optimization problems, depending on the level of information available about $\lambda(t)$. The value function $J(t, \lambda, F)$ solves the problem given before in (31), where $c(t)$ can be any suitable process adapted to \mathcal{F}_t . But we define a new value function $J^1(t, F)$ in which we impose an additional constraint on $c(t)$, namely that it be deterministic. More precisely,

$$\begin{aligned} J(0, \lambda_0, F_0) &= \max_{c(s) \text{ adapted}} E \left[\int_0^T e^{-\int_0^s (r+\lambda(q)) dq} u(c(s)) ds \right] \\ J^1(0, F_0) &= \max_{c(s) \text{ deterministic}} E \left[\int_0^T e^{-\int_0^s (r+\lambda(q)) dq} u(c(s)) ds \right] \\ &= \max_{c(s) \text{ deterministic}} \int_0^T e^{-rs} p(s, \lambda_0) u(c(s)) ds. \end{aligned} \quad (38)$$

We let c^* denote the optimal control for J , and c^1 denote the optimal control for J^1 .

Since every deterministic control $c(t)$ is also adapted, we have the basic relationship

$$J(0, \lambda_0, F_0) \geq J^1(0, F_0). \quad (39)$$

On the other hand, the above expression is exactly what the old deterministic model would have given. That is,

$$J^1(0, F_0) = J^{\text{DfM}}(0, F_0) \quad (40)$$

and $c^1 = c^{\text{DfM}}$.

We know from our earlier scaling arguments that $J(t, \lambda, F) = a(t, \lambda)F^{1-\gamma}/(1-\gamma)$ and $c^*(t, \lambda, F) = a(t, \lambda)^{-1/\gamma}F$ for some function $a \geq 0$. Likewise $J^{\text{DfM}}(t, F) = a_1(t)F^{1-\gamma}/(1-\gamma)$ and $c^1 = a_1^{-1/\gamma}F$ for some $a_1 \geq 0$. If $\gamma > 1$ then $1-\gamma < 0$, so $a(0, \lambda_0) \leq a_1(0)$, so $c^* \geq c^1$ at $t = 0$. This shows (a). Likewise if $0 < \gamma < 1$ then $a(0, \lambda_0) \geq a_1(0)$, so $c^* \leq c^1$ at $t = 0$. This shows (c).

Recall that when $\gamma = 1$, we have $u(c) = \ln[c]$. Earlier, when $\gamma \neq 1$, we had $u(c) = c^{1-\gamma}/(1-\gamma)$ and could make use of a scaling relation. In other words, if c is optimal for F , then kc is optimal for kF , and that leads to the expression $J(t, \lambda, kF) = k^{1-\gamma}J(t, \lambda, F)$. Or in other words,

$$J(t, \lambda, F) = F^{1-\gamma}J(t, \lambda, 1). \quad (41)$$

With logarithmic utility, the corresponding expression is that

$J(t, \lambda, kF) = J(t, \lambda, F) + (\ln k) \int_t^T e^{-r(s-t)} p(t, s, \lambda) ds$. Or in other words,

$$J(t, \lambda, F) = J(t, \lambda, 1) + (\ln F) \int_t^T e^{-r(s-t)} p(t, s, \lambda) ds. \quad (42)$$

Likewise,

$$J^{\text{DfM}}(t, F) = J^{\text{DfM}}(t, 1) + (\ln F) \int_t^T e^{-r(s-t)} \frac{p(s, \lambda_0)}{p(t, \lambda_0)} ds. \quad (43)$$

The first order conditions in the optimization problem then imply that

$$c^* = F / \int_t^T e^{-r(s-t)} p(t, s, \lambda) ds, \quad c^{\text{DfM}} = F / \int_t^T e^{-r(s-t)} \frac{p(s, \lambda_0)}{p(t, \lambda_0)} ds. \quad (44)$$

These agree when we send $t \rightarrow 0$, showing (b). **QED**

The theorem certainly proves that $\gamma = 1$ is a point of indifference. The invariance of mortality volatility when utility is logarithmic is reminiscent of similar results in consumption theory where income negates substitution effects. More on this later.

Note that we only use $J^1(0, F)$ above, not $J^1(t, F)$. If we had, we would have had to be careful. The correct definition is that

$$J^1(t, F) = J^{\text{DfM}}(t, F) = \max_{c(s)} \int_t^T e^{-r(s-t)} \frac{p(s, \lambda_0)}{p(t, \lambda_0)} u(c(s)) ds \quad (45)$$

rather than

$$\max_{c(s)} E \left[\int_t^T e^{-\int_t^s (r+\lambda(q)) dq} u(c(s)) ds \right] = \max_{c(s)} \int_t^T e^{-r(s-t)} E[p(t, s, \lambda(t))] u(c(s)) ds. \quad (46)$$

These quantities have connections to annuities, as suggested by the fact that the optimal consumption rates given above are, as a fraction of wealth, inverse annuity prices. In particular, $\int_t^T e^{-rs} p(s, \lambda_0) ds$ is the (actuarial) price of a deferred annuity, purchased at time 0 with payments starting at time t . While $\int_t^T e^{-rs} E[p(t, s, \lambda(t))] ds$ is a ‘‘forward’’ annuity price. That is, if at time 0 an insurance company guarantees (a retiree) the right to buy an annuity at time t at a price determined at time 0, then this is that price (computed actuarially, i.e. by discounting mean cash flows).

5 Optimal Consumption: Numerical Examples

We started with a particular survival probability at time zero, namely the Gompertz mortality curve with parameters $m = 89.335$ and $b = 9.5$. The age $x = 65$ survival probabilities to any age $y > x$ are given in Table #1. Both hypothetical retirees agree on these numbers, which means that their initial mortality rate is $\lambda_0 = (1/9.5) \exp\{(65 - 89.335)/9.5\} = 0.008125$.

Over time retiree #1 believes his mortality rate will grow at a rate $\eta = (1/9.5) = 0.10526316$ per year, while retiree #2 believes it will evolve stochastically with a time-dependent growth rate of $\mu(t)$ and a volatility σ . The actual curve $\mu(t)$ depends on the selected parameter for volatility, since $\mu(t)$ is constrained to match $p(0, \lambda_0)$. The actual process for extracting $\mu(t)$ for any given value of σ is rather complicated (although it is not central to our analysis) and is placed in the appendix of this paper. With these numbers in hand – and specifically the function $\mu(t)$ for the drift of the mortality rate – we can proceed to solve the PDEs given in equation (36) and (37), which then lead to the desired optimal consumption function and the initial portfolio withdrawal rate at age 65.

Table #2: Optimal Retirement Portfolio Withdrawal Rates $c^*(0)/F_0$						
Mortality Volatility	$\gamma = 0.5$	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$
$\sigma = 0$	7.59%	6.12%	5.58%	5.02%	4.78%	4.61%
$\sigma = 15\%$	7.52%	6.12%	5.60%	5.04%	4.80%	4.62%
$\sigma = 25\%$	7.44%	6.12%	5.62%	5.06%	4.82%	4.63%
Notes: Retirement age 65, interest rate $r = 2\%$, mortality $\lambda_0 = 0.0081$						

Table #2 provides a variety of numerical examples across different values of (mortality volatility) σ and (risk aversion) γ , once again assuming that the retirees are both at age $x = 65$ with observable mortality rate $\lambda_0 = 0.0081$. As we proved in section #3, and discussed above, the consumption rate is the same across all levels of mortality volatility when $\gamma = 1$. It increases relative to DfM when $\gamma > 1$ and decreases relative to DfM when $\gamma < 1$. Notice that impact of stochastic mortality on optimal withdrawal rates is reduce as the value of risk aversion increases. Notice how at a coefficient of relative risk aversion $\gamma = 10$, the portfolio withdrawal rates are (approximately) 4.6%

Note that the σ values provided are rather *ad hoc* and have not been estimated from any particular demographic dataset. We refer the interested reader to recent actuarial papers – such Bauer, et. al. (2008) – for an empirical discussion around the estimation of the volatility of mortality. Our objective here is to explore whether or not volatility has a (noticeable) impact on behavior.

6 Discussion and Conclusion:

In this article we extended the classical lifecycle model of consumption over a random-length lifecycle, to a model in which individuals can adapt behavior to new information about mortality rates. Yaari (1965) was the first to include lifetime uncertainty in a Ramsey-Modigliani lifecycle model and amongst other results, he provided a rigorous foundation for Irving Fisher’s claim that lifetime uncertainty increases consumption impatience and is akin to higher subjective discount rates. When the mortality rate itself is stochastic, this analogy is no longer meaningful and – to our knowledge – his work not been extended into the realm of 21st century models of mortality and longevity risk.

We built this extension by assuming that (i) the instantaneous force of mortality is stochastic and obeys a diffusion process as opposed to being deterministic, and (ii) that a utility-maximizing consumer can adapt their consumption strategy to new information about their mortality rate (a.k.a. current health status) as it becomes available. Our diffusion model for the stochastic force of mortality was quite general and the models were inspired by (a.k.a. borrowed from) the recent literature in the actuarial science arena which has argued for stochasticity in mortality rates.

We focused our attention on the “retirement income” stage of the lifecycle model (LCM) where health considerations are likely to be more prevalent and to avoid complications induced by wages, labor and human capital consideration. Thus, we refer to our general consumer as the retiree and the retirement consumption rate $c(t)$ scaled by wealth W , as the initial withdrawal rate.

In most² of the papers in the continuous-time LCM literature the force of mortality from time zero to the last possible date of death is known with certainty at time zero. Ergo, the conditional survival probabilities over the entire retirement horizon are known and predictable. So, if a 65-year-old retiree is told (by his doctor) that he faces a 5% chance of surviving to age 100 and a 37% chance of surviving to age 90, then by definition there is a $13.5\% = (0.05/0.37)$ probability of surviving to age 100, if he is still alive at age 90. In other words, he makes consumption decisions today that trade-off utility in different states of nature, knowing that if-and-when he reaches the age of 90, there will only be a 13.5% chance he will survive to age 100. In the language of actuarial science, the table of q_x mortality factors are known in advance. This is the essence of a deterministic force of mortality – which is textbook life contingencies – and is at the core of most economic lifecycle models such as Richard (1975), Davies (1981), Leung (1990) or Lachance (2010) that model consumption and portfolio choice problems with lifetime uncertainty.

²An exception would be the paper by Cocco and Gomes (2009) who recognize the stochasticity of mortality and use a Lee-Carter type model to numerically calibrate a lifecycle model. While our paper is motivated by the same consideration, our objective here is to extend the Yaari (1965) model, while retaining the essence of his other non-mortality assumptions.

In stark contrast, under a stochastic force of mortality the above predictability simply breaks down. While a 65-year-old might currently face a 5% chance of surviving to age 100 and a 37% chance of reaching age 90, there is absolutely no guarantee that the conditional survival probability from any future age, to age 100 (given the observed mortality rates), will satisfy the ratio. At time zero there is an expectation of what the probability will be at age 90. But, the probability itself is random. This way of thinking – which might be new to economists – is the essence of a stochastic force of mortality. The actuarial science literature has virtually exploded in this area, but so far it has had little impact on the economics (consumption theory) literature.

So, to phrase our question differently: how does an increase in the information set and the ability to act on this information change the optimal consumption function?

To compare results with previous literature and set notation, in the first part of this paper we re-derived the optimal consumption function – under a deterministic force of mortality (DfM) using techniques from the calculus of variations. We provided a closed-form expression for the entire consumption rate function under a Gompertz mortality assumption. With those benchmark results in place, we derived the optimal consumption strategy under a stochastic force of mortality (SfM), by expressing and solving the relevant Hamilton-Jacobi-Bellman (HJB) equation. In addition to the time variable, two state variables in the resulting PDE are current wealth and the current mortality rate. The methodology and mechanics were explained carefully and in great detail in section #4.

This basic set-up enabled us to extract a number of insights regarding the impact of stochastic mortality rates on optimal consumption in general and retirement spending rates in particular. Indeed, framing our model within the context of a retiree is especially relevant in the 21st century. A greater fraction of the population is retiring without access to defined benefit (DB) pension income, is faced with increased mortality (rate volatility) and health uncertainty and must personally decide how to spend their accumulated wealth over their remaining lifecycle.

Although a variety of different authors – most recently Cocco and Gomes (2009) – have explored similar problems, here is a summary of results and insights we believe are new to the literature.

1. As one might expect, the ability to adapt consumption to information about health status and unexpected changes in mortality rates is welfare enhancing. The value function – i.e. discounted utility of lifetime consumption – is uniformly higher at time zero, even though the trajectory of consumption in a SfM model is non-smooth. Recall that a by-product of Yaari (1965) is a smooth consumption path.
2. Retirees with (i) no bequest motives, (ii) constant relative risk aversion (CRRA) preferences, and (iii) subjective discount rates equal to the interest rate are expected to consume less as they age since they prefer to allocate consumption into states of nature

where they are most likely to be alive. This is the conventional diminishing marginal utility argument. In our model, while this is not true in every state of nature, it is true on average. In particular, a positive shock to the mortality rate in the form of pleasant health news (perhaps a cure for cancer) will reduce consumption instantaneously and further than expected at time zero, and a negative shock to the mortality rate (for example, being diagnosed with terminal cancer) will increase consumption beyond what was expected.

3. Our particular representation enables researchers to easily compare the consumption strategy of retirees who can react to changes in mortality rates and health status, with retirees who must set their plans in stone at time zero and cannot adapt to (or might even be ignorant of) new information. With proper calibration, this allows us to run a stochastic vs. deterministic horse race.
4. When the coefficient of CRRA (denoted by γ) is equal to one, and the retiree has logarithmic utility preferences, the optimal consumption rate at time zero is identical in both models. In other words, a retiree who cannot adjust their consumption plan as mortality rates evolve starts-off with the exact same consumption rate as the (more knowledgeable) consumer who can adapt to changes in mortality rates and health status. Although the path of their respective consumption will diverge over time – depending on the evolution of mortality rates – initially they are the same. We found this result to be most interesting.
5. In contrast, when the coefficient of CRRA is greater than one and the retiree is more risk-averse compared to a logarithmic utility maximizer, the initial consumption rate is higher in the stochastic model vs. the deterministic model. In other words, as one might expect the ability to adapt to changes in health status and new information about mortality rates allows the retiree to be more generous at time zero – relative to their “information ignorant” neighbor.
6. When the coefficient of the CRRA is between one and zero, which is at the razor’s edge of longevity risk neutrality, the result is reversed. The canonical retiree in a stochastic mortality model will consume less compared to their neighbor who is operating under deterministic mortality assumptions. Once again, the switch-over point is $\text{CRRA} = 1$, logarithmic utility. Notwithstanding the above results, the absolute consumption rate at time zero is uniformly higher the lower the coefficient of relative risk aversion, which is identical to the Yaari (1965). This is a manifestation of longevity risk aversion. The retiree is concerned about living a long time, and therefore consumes less today to protect themselves and self-insure consumption in old age.

7. The calibration of our economic model leads to some interesting by-products in actuarial finance. In particular, in order to construct a stochastic force of mortality that matches or fits a pre-determined Gompertz survival curve – the most popular and frequently used analytic law in this literature – one requires a lognormal diffusion process in which the drift itself grows even faster than exponentially over time. This can be tricky to navigate and caution is required when solving the optimal consumption PDE derived and presented in Section #4. The numerical technique requires us to move backwards thru a space-time grid in which one of the main parameter values in (explosively) large. Likewise, if one assumes a constant coefficient lognormal model, which is a geometric Brownian motion (GBM) of mortality, the resulting survival curve exhibits tails which are quite thicker than the Gompertz model, for the same life expectancy value. Either way, selecting diffusion and drift rate parameters that match a mortality table in expectation lead to large biases in the tail probabilities. We provide an alternative formulation in the appendix.
8. Depending on the actual model selected for the time-zero survival probability, for example Gompertz, or exponentially distributed, or logistic, the changes in the optimal consumption rates can range from as little as 1% to as large as 10% depending on initial age, volatility of mortality, and the actual value of the coefficient of relative risk aversion. In section #5 we offered a number of case-specific insights relating the magnitude of the change in consumption in a stochastic vs. deterministic model. It would therefore be interesting to see if actual consumer behavior as measured in some of the longitudinal databases, can be better described by a stochastic mortality representation.

Moving forward, there are a number of avenues to pursue in future research. A natural extension to this kind of thinking would be to explore the impact of stochastic investment returns as well as mortality rates and thus include a strategic asset allocation dimension. Another item on the research agenda – and one that is occupying us currently – is to dig further into this framework and explore the role of health and mortality-contingent claims in stochastic mortality model. Recall that one of Yaari’s (1965) noted results is that lifecycle consumers with no bequest motives should hold all of their wealth in actuarial notes. This was extended and his assumptions were relaxed by Davidoff, Brown and Diamond (2005). But these are essentially pension annuities in a deterministic mortality model. However, in the presence of a stochastic mortality, it is no longer clear how an insurance company would price pension annuities, given the systematic risk involved. In such a model, a retiree would have to choose between investing wealth in a tontine pool, with corresponding stochastic returns or purchasing a pension annuity with a deterministic consumption flow, but possibly paying a risk-premium for the privilege. In fact, the optimal portfolio allocation might be a mixture between tontines and annuities. Alas, we leave this work for a subsequent paper.

References

- [1] Babbel, D.F. and C.B. Merrill (2006), Rational Decumulation, *working paper* Wharton Financial Institutions Centre.
- [2] Bauer, D., M. Borger, J. Rus, H.J. Zwiesler (2008), The volatility of mortality, *Asia-Pacific Journal of Risk and Insurance*, Vol. 3. pg. 184-211.
- [3] Brillinger, D.B. (1961) A Justification of some Common Laws of Mortality, *Transactions of the Society of Actuaries*, Vol. XIII, pp.116-119.
- [4] Bodie, Z., J.B. Detemple, S. Ortuba and S. Walter (2004), Optimal consumption-portfolio choice and retirement planning, *Journal of Economic Dynamics and Control*, Vol. 28, pg. 1115-1148.
- [5] Butler, M. (2001), Neoclassical life-cycle consumption: a textbook example, *Economic Theory*, Vol. 17, pg. 209-221.
- [6] Carrière, J.F. (1994), An investigation of the Gompertz law of mortality, *Actuarial Research Clearing House*, Vol. 2, pg. 1-34
- [7] Cairns, A.J.G., D. Blake and K. Dowd (2006), A two-factor model for stochastic mortality with parameter uncertainty, *Journal of Risk and Insurance*, Vol. 73(4).
- [8] Cocco, J.F. and F.J Gomes (2009), Longevity Risk and Retirement Savings, *Working Paper, London Business School*
- [9] Davidoff, T, J.R. Brown and P. Diamond (2005), *Annuities and Individual Welfare*, American Economic Review, Vol 95(5)
- [10] Dahl, M. (2004), Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts, *Insurance: Mathematics and Economics*, Vol. 35, pg. 113-136
- [11] Davies, J.B. (1981), Uncertain lifetime, consumption and dissaving in retirement, *Journal of Political Economy*, Vol. 89, pg. 561-577.
- [12] Deaton, A. (1991), Saving and liquidity constraints, *Econometrica*, Vol. 59(5), pg. 1221-124
- [13] Dothan, L.U. (1978), On the Term Structure of Interest Rates, *The Journal of Financial Economics*, Vol. 6, pp. 59-69.
- [14] Fisher, I. (1930), *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*, The Macmillan company, New York.

- [15] Gutterman, S. and I.T. Vanderhoof (1998), Forecasting Changes in Mortality: A Search for a Law of Causes and Effects, *The North American Actuarial Journal*, Vol. 2, No. 4, pp. 135-138.
- [16] Kingston, G. and S. Thorp (2005), Annuitization and asset allocation with HARA utility, *Journal of Pension Economics and Finance*, Vol. 4(3), pg. 225-248.
- [17] Kotlikoff, L. (2008), Economics' Approach to Financial Planning, Boston University working paper and *Journal of Financial Planning*, March.
- [18] Lachance, M. (2010), A realistic and full calibration of life-cycle model, forthcoming, *Journal of Pension Economics and Finance*.
- [19] Liu, Hong (2005), Lifetime consumption and investment: Retirement and Constrained Borrowing, *working paper*, John M. Olin School of Business.
- [20] Leung, S. F. (1994), Uncertain Lifetime, The Theory of the Consumer and the Life Cycle Hypothesis, *Econometrica*, Vol. 62(5), pg. 1233-1239
- [21] Leung, S. F (2007), The existence, uniqueness and optimality of the terminal wealth depletion time in life-cycle models of saving under certain lifetime and borrowing constraint, *Journal of Economic Theory*, Vol. 134, pg. 470-493.
- [22] Martin-Jimenez, S. and A. R. Sanchez Martin (2007), An evaluation of the life cycle effects of minimum pensions on retirement behavior, *Journal of Applied Econometrics*, Vol. 22, pg. 923-950.
- [23] Milevsky, M.A. and S.D. Promislow (2001), Mortality derivatives and the option to annuitize, *Insurance: Mathematics and Economics*, Vol 29, pg. 299-318
- [24] Milevsky, M.A. and H. Huang (2010), Spending Retirement on Planet Vulcan: The Impact of Longevity Risk Aversion on Optimal Withdrawal Rates, *Working Paper, York University*
- [25] Modigliani, F. and R. Brumberg (1954), Utility analysis and the consumption function: An interpretation of cross-section data, published in *Post Keynesian Economics*, New Brunswick, Rutgers University Press.
- [26] Modigliani, F. (1986), Lifecycle, individual thrift and the wealth of nations, *American Economic Review* Vol. 76(3), pg. 297-313.
- [27] Norberg, R (2010), Forward Mortality and other vital signs – Are they the way forward?, *Insurance: Mathematics and Economics*, Vol. 47(2), pg. 105-112.

- [28] Pitacco, E., Denuit, M., Haberman, S., and Olivieri, A. (2008), *Modelling Longevity Dynamics for Pensions and Annuity Business*, Oxford University Press, UK
- [29] Phelps, E.S. (1962), The accumulation of risk capital: a sequential utility analysis, *Econometrica*, Vol. 30(4), pg. 729-743.
- [30] Ramsey, F.P. (1928), A mathematical theory of saving, *The Economic Journal*, Vol. 38(152), pg.543-559
- [31] Richard, S.F. (1975), Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model, *Journal of Financial Economics*, Vol. 2, pg. 187-203.
- [32] Sheshinski, E. (2008), *The Economic Theory of Annuities*, Princeton University Press, New Jersey
- [33] Wills, S. and M. Sherris (2010), Securitization, structuring and pricing of longevity risk, *Insurance: Mathematics and Economics*, Vol. 46(1), pg. 173-185.
- [34] Yaari, M.E. (1965), Uncertain Lifetime, Life Insurance and the Theory of the Consumer, *The Review of Economic Studies*, Vol. 32(2), pg. 137-150

7 Appendix: Matching Time-Zero Survival Curves

Given a deterministic model (Gompertz in our numerical examples), we can compute the time-zero survival function $p(t, \lambda_0)$. We plan to match this using a stochastic model, by a suitable choice of parameters. This means that at time 0 the two models deliver identical survival probabilities. Recall that at times other than $t = 0$ the comparison will no longer be meaningful, even controlling for the current observed mortality rate, because the mismatch between conditional survival probabilities means that the two models give different views of lifetimes going forward.

Let $\Lambda(t) = e^{-\int_0^t \lambda(q) dq}$ and define a pseudo-density $q(t, \lambda)$ by the formula

$$E[\Lambda(t)\phi(\lambda(t))] = \int_0^\infty \phi(\lambda)q(t, \lambda) d\lambda. \quad (47)$$

Then $p(t, \lambda_0) = \int_0^\infty q(t, \lambda) d\lambda$. By Itô's lemma,

$$\begin{aligned} \phi(\lambda(t))\Lambda(t) &= \phi(\lambda_0) + \int_0^t \Lambda(s) \left[\mu(s)\lambda(s)\phi'(\lambda(s)) + \frac{1}{2}\sigma^2\lambda(s)^2\phi''(\lambda(s)) - \lambda(s)\phi(\lambda(s)) \right] ds \\ &\quad + \int_0^t \Lambda(s)\mu(s)\lambda(s)\phi'(\lambda(s)) dB(s). \end{aligned} \quad (48)$$

Take expectations and differentiate with respect to t . We get

$$\int_0^\infty \phi(\lambda)q_t(t, \lambda) d\lambda = \int_0^\infty \left[\mu(t)\lambda\phi'(\lambda) + \frac{1}{2}\sigma^2\lambda^2\phi''(\lambda) - \lambda\phi(\lambda) \right] q(t, \lambda) d\lambda \quad (49)$$

with initial condition $q(0, \cdot) = \delta_{\lambda_0}$. Using integration by parts (for ϕ vanishing fast at 0 and ∞), we have

$$q_t(t, \lambda) = -\mu(t)\frac{\partial}{\partial \lambda} [\lambda q(t, \lambda)] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial \lambda^2} [\lambda^2 q(t, \lambda)] - \lambda q(t, \lambda). \quad (50)$$

So if $\mu(t)$ is known for $0 \leq t \leq t_1$, then all expectations $\int_0^\infty q(t_1, \lambda)\phi(\lambda) d\lambda$ can be found by solving the forward equation for q and then integrating against ϕ .

Let $\lambda_{(1)} = \int_0^\infty \lambda q d\lambda$ and $\lambda_{(2)} = \int_0^\infty \lambda^2 q d\lambda$ be the first two moments of $q(t, \lambda)$. Note that the zeroth moment is the survival probability, so we can integrate (by parts) the forward PDE for q and the product of λ and the forward PDE and obtain the following relationships

$$\begin{aligned} \lambda_{(1)} &= -\frac{dp}{dt}, \\ \lambda_{(2)} &= \mu(t)\lambda_{(1)} - \frac{d\lambda_{(1)}}{dt}. \end{aligned} \quad (51)$$

Combined the two expressions, we have

$$\mu(t) = \frac{\frac{d\lambda_{(1)}}{dt} + \lambda_{(2)}}{\lambda_{(1)}}. \quad (52)$$

Replacing $\mu(t)$ in the forward PDE for q and obtain an integro-differential equation

$$q_t(t, \lambda) = -\frac{\frac{d\lambda_{(1)}}{dt} + \lambda_{(2)}}{\lambda_{(1)}} \frac{\partial}{\partial \lambda} [\lambda q(t, \lambda)] + \frac{\sigma^2}{2} \frac{\partial^2}{\partial \lambda^2} [\lambda^2 q(t, \lambda)] - \lambda q(t, \lambda), \quad (53)$$

or

$$q_t(t, \lambda) = \frac{-\frac{\partial^2 p}{\partial t^2} + \int_0^\infty \lambda^2 q(t, \lambda) d\lambda}{\frac{\partial p}{\partial t}} \frac{\partial}{\partial \lambda} [\lambda q(t, \lambda)] + \frac{\sigma^2}{2} \frac{\partial^2}{\partial \lambda^2} [\lambda^2 q(t, \lambda)] - \lambda q(t, \lambda), \quad (54)$$

which we can solve numerically with the initial condition $q(0, \lambda) = \delta(\lambda - \lambda_0)$.

We solve the integro-differential equation for q numerically first, obtain the value of $\mu(t)$. We then solve the HJB equation for optimal consumption as before, with the constant μ now replaced by the function $\mu(t)$ at $\lambda = 0$.

Finally, we should record a couple of remarks about the form of $\mu(t)$. First of all,

$$\mu(0) = \eta. \quad (55)$$

To see this, observe that $p_t(0, \lambda_0) = -E[\lambda_0 \Lambda_0] = -\lambda_0$. So $\lambda_0^2 = E[\lambda_0^2 \Lambda_0^2] = p_{tt}(0, \lambda_0) + \lambda_0 \mu(0)$. But $p_{tt}(0, \lambda_0)$ can be computed explicitly, since it is Gompertz, to give $\lambda_0^2 - \lambda_0 \eta$. This implies that $\mu(0) = \eta$.

Second, note that $\mu(t)$ should be increasing in σ . The mean $E[e^{-\int_0^t \lambda(q) dq}]$ doesn't change with σ , so by convexity of the exponential, the median of this quantity must decrease as we increase the variance. In other words, $\mu(t)$ must rise. Put another way, this expectation is driven by the possibility of relatively larger values of the exponent, ie of abnormally low values of λ . As σ rises, the impact of longevity risk gets more pronounced, and to compensate for that the growth rate $\mu(t)$ must also rise.