Pensions Institute

DISCUSSION PAPER PI-0211

Optimal Asset Allocation and The Real Option to Delay Annuitization: It's Not Now-or-Never

Moshe A. Milevsky and Virginia R. Young

September 2002

ISSN 1367-580X

The Pensions Institute
Cass Business School
City University
106 Bunhill Row London
EC1Y 8TZ
UNITED KINGDOM

http://www.pensions-institute.org/

Optimal Asset Allocation and
The Real Option to Delay Annuitization:
It's Not Now-or-Never

Moshe A. Milevsky[1]

and

Virginia R. Young[2]

Version: September 1, 2002

[1] Milevsky, the contact author, is an Associate Professor of Finance at the Schulich School of Business, York University, Toronto, Ontario, M3J 1P3, Canada, and the Director of the Individual Finance and Insurance Decisions (IFID) Centre at the Fields Institute. He can be reached at Tel: (416) 736-2100 ext 66014, Fax: (416) 763-5487, E-mail: milevsky@yorku.ca. This research is partially supported by a grant from the Social Sciences and Humanities Research Council of Canada and the Society of Actuaries. The author thanks N. Charupat, R. Chen, G. Daily, J. Green, M. Warshawsky, and participants at the York University finance seminar series, University of Cyprus economics department seminar series, and the annual meeting of the American Economics Association and the Western Finance Association.

[2] Young is an Associate Professor at the School of Business, University of Wisconsin-Madison, Madison, Wisconsin, 53706, USA. She can be reached at Tel: (608) 265-3494, Fax: (608) 263-3142, E-mail: vyoung@bus.wisc.edu. This research is partially supported by a grant from the Graduate School of the University of Wisconsin-Madison via a Vilas Associateship.

Optimal Asset Allocation and The Real Option to Delay Annuitization: It's Not Now-or-Never

Abstract: Asset allocation and consumption towards the end of the life cycle is complicated by the uncertainty associated with the length of life. Although this risk can be hedged with life annuities, empirical evidence suggests that voluntary annuitization amongst the public is not very common, nor is it well understood.

This paper develops a normative model of when, and if, one should purchase an *immediate life annuity*. This problem is particularly relevant given the increasing number of Defined Contribution pension plans in the U.S – for which participants must make this decision – and the corresponding trend away from Defined Benefit guarantees.

Specifically, our main qualitative argument is that there is a *real option* – akin to the corporate finance usage of the word – embedded in the decision to annuitize. A life annuity can be viewed as a project with a positive net present value. However, quite distinct from a fixed-income bond or period certain annuity, once purchased, a life annuity can never be sold, reversed, or exchanged. Its purchase is final because of the severe moral hazard involved in trying to terminate a life-contingent claim.

We use standard continuous-time technology to solve the optimal asset allocation and annuitization timing problem. We then define the value of the real option to defer annuitization (RODA) as the compensating utility loss from being unable to behave optimally.

By using reasonable capital market and actuarial parameters, we estimate that the real option to defer annuitization is quite valuable until the mid-70s or mid-80s. Of course, the precise values depend on one's gender, risk aversion, and subjective health assessment.

Finally, we show that low-cost variable immediate annuities, which are currently not widely available, greatly reduce the option value to wait and create substantial welfare gains. This might explain the large number of TIAA-CREF participants who rightfully choose to annuitize their DC pension plan, as a result of the availability of *both* fixed and variable payments in the payout stage.

JEL Classification: J26; G11

Keywords: Insurance; Life Annuities; Asset Allocation; Retirement.

"*Every project competes with itself delayed in time...*"

Steven Ross, FMA Keynote Lecture, 1995

1. Introduction, Motivation, and Objectives

Much of the economics literature has documented — and continues to puzzle over — the extremely low levels of voluntary annuitization exhibited among elderly retirees. Strictly speaking, this phenomenon is inconsistent with results of a standard Modigliani life-cycle model of savings and consumption, as described by Yaari (1965). In a life-cycle model with no bequest motives, Yaari (1965) demonstrated that all consumers hold life annuities as opposed to liquid assets. This implies that when given the chance, retirees should convert their liquid assets to life annuities because the latter provide protection against outliving one's money. The rationale behind Yaari's result is that returns from life annuities dominate all other assets in his model because the "living" inherit the assets and returns of the "dead." Moreover, at older ages, the higher probability of dying increases the relative return from life annuities, conditional on survival.

Nevertheless, despite the highly appealing arguments in favor of annuitization, there is little evidence that retirees are voluntarily embracing this arrangement. Modigliani (1986), Friedman and Warshawsky (1990), Mirer (1994), Poterba and Wise (1996), and Brown (1999, 2001), among others, have pointed out that very few people consciously choose to annuitize their marketable wealth.[3] Thus, in the face of poor empirical evidence, various theories have been proposed to salvage this aspect of the life-cycle hypothesis and to justify the low demand for longevity insurance. For example, in one of the earlier papers on this puzzle, Kotlikoff and Spivak (1981) argued that family risk pooling may be preferred to public annuity markets, especially given the presence of

[3] In the comprehensive Health and Retirement Survey (HRS) conducted in the U.S, only 1.57% of the HRS respondents reported annuity income. Likewise, only 8.0% of respondents with a defined contribution pension plan selected an annuity payout. The North American-based Society of Actuaries and LIMRA, as reported in Sondergeld (1997), conducted a study that shows only 0.3% of variable annuity contracts were annuitized during the 1992-1994 period. More recently (June 30, 2001), according to the National Association of Variable Annuities, of the $909 billion invested in variable annuities, only 2% were annuitized.

adverse selection and transaction costs. Indeed, a married couple function as a mini annuity market, as elaborated by Brown and Poterba (2000). Friedman and Warshawsky (1990) showed that average yields on individual life annuities during the late 1970's and early 1980's were lower than plausible alternative investments. The reduced yield was largely attributed to actuarial loads and profits, which have declined over time, according to recent work by Mitchell, Poterba, Warshawsky and Brown (MPWB, 1999). In a different vein, Kotlikoff and Summers (1981) argued that intergenerational transfers accounted for the vast majority of U.S. savings and therefore bequest motives *solve* the puzzle. Hurd (1989) and Bernheim (1991) echoed this view. In other words, individuals do not annuitize wealth simply because they want to bequeath assets.

Bernheim (1991) further argues that large pre-existing annuities in the form of Social Security and government pensions might serve as an additional deterrent to voluntary annuitization. In a distinct line of reasoning, Yagi and Nishigaki (1993) argue that the actual design of annuities impede full annuitization. One cannot obtain a life annuity that provides arbitrary payments contingent on survival, which is dictated by Yaari's (1965) model. They must be either fixed (in nominal or real terms) or variable (linked to an index). This constraint forces consumers to hold both marketable wealth and annuities. In related research, Milevsky (1998) computed the probability of success from mimicking the consumption from a fixed immediate annuity and investing the balance. His argument was that there is a "high probability" that equities will outperform fixed immediate annuities until late in life. However, he did not address the maximization of utility nor the option value embedded in the decision to annuitize.

In summary, many explanations exist for why people do *not* annuitize further wealth. Although these justifications have explanatory power, they fail to provide financial advice on optimal product design as well as normative strategies for the elderly. Furthermore, they cannot account for the casual observation that most people shun life annuities, simply because they want to maintain control of their assets.

In this paper, we pursue a different approach to the issue by calculating *when* an individual should annuitize her wealth. Motivated by Merton's (1971) consumption model and the recent Real Options paradigm, exposited by Trigeorgis (1993, 1996), we focus attention on the *real option* embedded in the decision to annuitize.[4] Heuristically, due to the irreversibility of annuitization, the decision to purchase a life annuity is akin to exercising an American-style mortality-contingent claim. It is only optimal to do so when the remaining time value of the option becomes worthless. Options derive their value from the volatility of the underlying state variables. Therefore, if one accounts for future mortality and investment uncertainty, the embedded option provides an incentive to delay annuitization until the option value has been eliminated. The option is *real* in the sense that it is not directly separable or tradable. The option is also *real* because it is personal; its value depends on the individual's subjective mortality rate and on her degree of risk aversion. We use the term "option to annuitize," in the same vein as Stock and Wise (1990), who describe the option value of retirement.

The tools we employ are standard in the finance and pension literature. By modifying Merton's model of consumption (1971), we maximize the expected discounted utility of consumption over the individual's random future lifetime. The decisions that the individual makes during this time are how much to consume, how much to invest in a risky asset, and when to annuitize. We derive the value function in this modified Merton framework and then locate the optimal stopping (annuitization) time. From a practical perspective, we estimate that the real option to delay annuitization (RODA) remains quite valuable until the mid-70s or mid-80s, depending on one's gender, level of risk aversion, and health status. Specifically, individuals with higher risk tolerance and greater health asymmetry are endowed with an even larger option value to wait. Also, women are more willing to defer annuitization because they expect to live longer; it is as if the mortality table has been shifted back. Finally, we show that the availability of (low-cost) variable immediate annuities reduces the value of the RODA and increases the relative appeal of annuitization.

We consider only the utility of consumption and do not include a utility of bequest. We do so to separate the "delay effects" of the desire to leave a bequest from the

---

[4] Strictly speaking, we use the term *real option* in the personal sense, as opposed to a financial option. Our real option cannot be traded, while financial options can be traded. Our real option exists because of the irreversibility of the decision to annuitize and because of the non-tradability and personal nature of the option. We refer the interested reader to other work by Berk (1999), Amram and Kulatilaka (1999), Ross

motivation of investing in a risky asset now, improving one's wealth, and annuitizing later. In other words, we *expect* the bequest motive to delay annuitization (Kotlikoff and Summers, 1981; Hurd, 1989; Bernheim, 1991), so we examine the delay induced by wanting to consume more.

As Yaari (1965) and many others have illustrated, the availability of a fairly priced life annuity relaxes the budget constraint that then induces greater utility of consumption. All else being equal, the consumer annuitizes wealth as soon as he or she is given the opportunity to do so. However, these classical arguments are predicated on the existence of only one financial asset from which the annuities are priced. This framework *de facto* assumes that the budget constraint will not improve over time. However, in practice, a risky asset is an alternative to the risk-free investment, and by taking a chance in the risky asset, the future budget constraint may improve. In other words, it might be worth waiting, since tomorrow's budget constraint may allow for a larger annuity flow and greater utility. In the meantime, of course, the individual is assumed to withdraw consumption from liquid wealth so as to maximize expected utility of consumption. In fact, when the volatility of the risky asset is set equal to zero, insurance loads are set to zero, and the mortality asymmetry is removed, the real option to delay has no value, which corresponds with Yaari's (1965) solution. Likewise, uncertainty about future interest rates, tastes for bequest, insurance loads, and product design all add value to the option to delay. Stated differently, our main argument is that retirees should refrain from annuitizing today because they may get an even better deal by self-annuitizing and consuming in the meantime, and then annuitizing tomorrow.

It is important to note that our argument is more than just a play on the equity risk premium, namely that sufficiently risk-tolerant consumers should invest in the risky asset. Rather, we argue that any multi-period framework that ignores the irreversibility of the life annuity purchase does not capture the option value in waiting. It is most likely that this is the reason why Richard's (1975) merging of Merton (1971) and Yaari (1965) also yields a full-annuitization result. In related research, Kapur and Orszag (1999) introduced immediate annuities into a Merton (1971) framework by assuming that the

---

(1995), Ingersoll and Ross (1992), and Hubbard (1994) for related literature on real options in the corporate finance sense.

risk-free rate is augmented by a mortality bonus that is proportional to the instantaneous hazard rate. However, the irreversibility is completely ignored in all these models since the annuity is treated as a tontine that can be renegotiated each period; a tontine is an arrangement whereby the survivors share the pool of money.

Likewise, Brugiavini (1993) examined the optimal time to annuitize and concludes that it should be early in the life cycle. However, her model is driven by adverse selection considerations and abstracts somewhat from the multiple sources of uncertainty that might induce people to wait. In a similar vein, Blake, Cairns, and Dowd (2000) conducted extensive computer simulations to determine the annuity and pension drawdown policy that provides the highest level of (exponential) utility. However, they did not examine the implications of annuitizing at different ages, as it pertains to the option value of waiting.

The main conceptual innovation in this paper is that we treat the right, but not the obligation, to annuitize as an option. We model this as an American option and locate the optimal exercise time. To value this option to wait, we proposed a methodology similar to the wealth-equivalent metric of MPWB (1999). We define the value of the *real option to defer annuitization* (RODA) as the (percentage) increase in wealth that would *substitute* for the ability to defer. We answer the question: How much would the consumer require in compensation for losing the opportunity to wait? This number is clearly preference-dependent, which means that the value of the RODA is *not* priced in the arbitrage-free framework of Black and Scholes (1973). Our valuation methodology is primarily due to the lack of secondary market for this real option.[5] Furthermore, the value of the RODA may actually be zero, in which case we argue that the consumer is better off annuitizing right now, since waiting can only destroy value.

---

[5] See Hodges and Neuberger (1989) for a similar approach to pricing financial options in the presence of transaction costs. Other researchers have considered problems related to retirement, but to our knowledge, none have followed our line of attack. Marcus (1985) values pension liabilities in a dynamic setting, although his method is different from ours because he assumes that the market for the liabilities is complete while our market for annuity transactions is incomplete. Similarly, Stanton (2000) values the "retirement" and "roll-over" options possessed by owners of 401(k) plans; he values these options in a complete market by ignoring the (somewhat) irreversibility of the option to retire and to remove one's money from the 401(k) as a lump-sum distribution. Sundaresan and Zapatero (1997) determine the optimal time to retire but use "no-arbitrage" arguments because they assume that the wage process is perfectly correlated with a tradable asset.

The remainder of this paper is organized as follows. In Section 2, we provide a simple 3-period model that illustrates our main argument. In Section 3, we do the same in multi-period continuous-time model and derive some numerical estimates for optimal time of annuitization and for the value of the RODA. We consider fixed annuities in Section 3.1, while in Sections 3.2 and 3.3, we extend our model to incorporate variable annuities and increasing annuities, respectively. Section 4 concludes the paper.

2. Simple 3-Period Model for Life Annuities

We first illustrate the option value of deferring annuitization with a simple, discrete 3-period example. In the next section, we generalize the model to the continuous-time setting. Our problem starts at time zero with a consumer aged (x) who has an initial endowment of \$1. All consumption takes place at the end of the period, and the probabilities of dying during the three periods are $q_x = 0.10$, $q_{x+1} = 0.25$, and $q_{x+2} = 0.60$.

If the individual is fortunate to survive to the end of third period, she consumes and immediately dies. For simplicity, we assume that both the consumer and the insurance company are aware of and agree on these probabilities of death. Also, for simplicity, we assume the consumer's subjective rate of time preference is set equal to the risk-free rate. We assume that the money is sitting in a qualified (registered) pension plan for which money withdrawn is taxed as ordinary income; therefore, taxes will not change the decision of when to annuitize tax-sheltered money. Later, we discuss the implications of relaxing some of these assumptions and the effect on the value of the RODA. We define c_i to be the consumption that takes place at the end of period i, i = 1, 2, or 3. The risk-free interest rate is 10% per period and is the interest rate with which annuities are priced.

If we assume no future source of uncertainty and no bequest, the consumer will annuitize her wealth of \$1 immediately, since her optimization problem is to maximize her expected discounted utility of consumption:

$$\max_{\{c_1, c_2, c_3\}} \frac{(1-0.1)u(c_1)}{1.1} + \frac{(1-0.1)(1-0.25)u(c_2)}{1.1^2} + \frac{(1-0.1)(1-0.25)(1-0.6)u(c_3)}{1.1^3}$$

such that

$$1=\frac{(1-0.1)c_1}{1.1}+\frac{(1-0.1)(1-0.25)c_2}{1.1^2}+\frac{(1-0.1)(1-0.25)(1-0.6)c_3}{1.1^3},$$

in which $u(c)$ is a twice-differentiable utility function that is increasing and strictly concave. Later, we assume that u exhibits constant relative risk aversion. We ignore the utility of bequest because this can only increase the value of the RODA. In other words, it is obvious that individuals with a strong utility of bequest will avoid (defer) annuitization. Our objective is to illustrate the conditions under which consumers with absolutely no utility of bequest will delay annuitization.

The annuity contract appears in the above equation by virtue of the (expected) mortality-adjusted discounting of consumption in the "budget" constraint. All else being equal, higher death rates increase the consumption attainable in the annuity market. The initial \$1 can be used to finance a higher consumption stream. Likewise, by setting all the death rates equal to zero, one tightens the budget constraint and reduces the feasible consumption set. This is akin to solving the problem without annuity markets.

The Lagrangian of this problem is

$$\begin{aligned}\mathsf{L}&=\frac{(1-0.1)u(c_1)}{1.1}+\frac{(1-0.1)(1-0.25)u(c_2)}{1.1^2}+\frac{(1-0.1)(1-0.25)(1-0.6)u(c_3)}{1.1^3}\\&\quad+\lambda\left(1-\frac{(1-0.1)c_1}{1.1}-\frac{(1-0.1)(1-0.25)c_2}{1.1^2}-\frac{(1-0.1)(1-0.25)(1-0.6)c_3}{1.1^3}\right),\end{aligned}$$

which after solving the first-order conditions, leads to the optimal consumption

$$c^*_i=c^*=\frac{1}{a_x},$$

where a_x is given by

$$a_x=\frac{(1-0.1)}{1.1}+\frac{(1-0.1)(1-0.25)}{1.1^2}+\frac{(1-0.1)(1-0.25)(1-0.6)}{1.1^3}=1.5789,$$

which is the time-zero price of a \$1 life annuity and is paid contingent on survival. The maximal utility equals $u(c^*)\,a_x$.

Our result is the classical annuity result, originally derived by Yaari (1965), which states that all liquid wealth is annuitized — held in the form of actuarial notes — and consumption is constant across all (living) periods. As mentioned earlier, in the absence of annuity markets, the budget constraint is tightened to equate present value of consumption and initial wealth, and the optimal consumption decreases as the probability of survival increases.

The constant consumption result is predicated on (*i*) the time preference being equal to the risk-free rate, and (*ii*) symmetric mortality beliefs. If these numbers are different, the optimal consumption stream might not be constant. In some cases, it might even induce holdings of non-annuitized assets. We refer the interested reader to Yagi and Nishigaki (1993) for a simple 3-period discrete-time example that generalizes the above result.

Numerically, in our case, if we set the coefficient of relative risk aversion (CRRA) equal to 1.5, then the utility function is $u(c) = -2/\sqrt{c}$. The individual's optimal consumption is $c^* = \$0.6334$, and her resulting expected discounted utility of consumption is -3.9679 utiles.

The above derivation is standard in the "annuity economics" literature, although possibly less well known to financial researchers. Our main point is that she can do better for herself than simply buying an annuity right now. Indeed, assuming an alternative, risky asset class, she can invest and consume now and defer annuitization until later. Suppose she consumes the same c^* = \$0.6334 at the end of the first period, which is what the annuity would have provided, and then reconsiders annuitization at that time. (In the continuous time models in Section 3, we determine the optimal consumption level knowing that she will annuitize later.) In the meantime, she invests her assets subject to a risky return. The risky return can fall into one of two states: There is a 0.70 probability of a good return equal to 45% during the first period and a 0.30 probability of a bad return equal to 0% during the period.

The utility of deferral captures the gains from taking a chance on next period's budget constraint. Specifically, the utility of deferral weighs next period's utility of consumption by the probability of each return-state. Indeed, in the event of a 45% return, the investor has \$1.45 units at the end of the first period, from which she consumes \$0.6334 to mimic the annuity. This leaves her with \$0.8166 for the 2nd-period budget constraint. Likewise, if the return from the risky asset is 0%, the investor is left with the original \$1 at the end of the first period, from which she consumes \$0.6334, leaving her with only \$0.3666 for the 2nd-period budget constraint. Clearly, the former is preferred to the latter, which will cause substantial regret. However, it may be worth taking this speculative risk.

Now, assuming she annuitizes at the end of the first period, her expected discounted utility from the decision to defer is greater than if she had annuitized immediately:

$$\frac{1-0.10}{1.1}\left(0.7u\left(\frac{0.8166}{a_{x+1}}\right)a_{x+1}+0.3u\left(\frac{0.3666}{a_{x+1}}\right)a_{x+1}+u(0.6334)\right)=-3.9193>-3.9679.$$

The above equation consists of two parts. The items within the bracket capture the expected utility from the stock market "gamble" plus the utility from the consumption at the end of the first period. This is then multiplied by the probability of actually surviving to the end of the period, and then discounted at the risk-free rate. Note that in this case, the individual prefers to consume then annuitize instead of annuitizing immediately.

Finally, if we were to compensate the individual with an additional \$0.0249 at time zero, she would be indifferent between annuitizing her wealth immediately and deferring for one period. In other words, had she started with \$1.0249, her expected utility if she annuitized immediately would be –3.9193 utiles, which is exactly the expected utility that she would have obtained from waiting. We conclude that the value of the RODA for one period is worth (at least) 2.49% of initial wealth.

Comments

A few technical insights emanate from this simple model:

- For the deferral to make financial sense, the stochastic return from the investable asset must exceed the mortality-adjusted risk-free rate in at least one state of nature. In our 3-period, 2-states-of-nature context, the return in the "good" state of nature (45%) must be greater than the one-plus risk-free rate (1.1) divided by the probability of survival (0.9). Otherwise, the expected utility from waiting will never exceed the initial utility.
- One does not require abnormally high investment returns in order to justify waiting. In fact, the entire analysis could have been conducted with a stochastic interest rate instead of a stochastic investment return. (Or both, for that matter.) The key insight is that waiting might change the budget constraint in the consumer's favor. The budget constraint might change on the left-hand side, which is an increase (or decrease) in

initial wealth, or on the right-hand side, with an increase (or decrease) in the interest rate with which the annuity is priced. As long as the risk-adjusted odds of a favorable change in the budget constraint are high enough, the option to wait has value. This insight is quite important since any possible change in the future price of the annuity provides an option value. This would include any changes in design, liquidity, or pricing that might improve tomorrow's budget constraint.

- When the coefficient of relative risk aversion (CRRA) is equal to one, which implies logarithmic utility, the value of the one-period option is (at least) 4.27%, which is higher than the case of CRRA = 1.5. As one would expect, the lower the CRRA, the higher the (utility-adjusted) incentive to take some financial risk and to defer the decision to annuitize. The same is true in the other direction, that is, a higher aversion to risk decreases the value of the RODA. For a high enough value – which in our case is CRAA = 2.1732 – the individual should not defer annuitization for one period since the risk is too high.
- We have not addressed the possibility of asymmetric mortality beliefs in our discrete-time analysis. We will model this explicitly in continuous time. We will show that if the consumer has a view of her own mortality that is different that the one used to price the annuity, the option to defer is even *more* valuable. Specifically, if the objective probability of death, q_x^O, which is used in the budget constraint to price the annuity, is different than the subjective probability of death, q_x^S, used in the objective function, then the value of the RODA increases. The individual is speculating on next period's budget constraint in the (risk-adjusted) hope it will improve. On the one hand, although annuities are a good deal to healthier people, they delay annuitizing longer in order to increases their wealth and thereby get an even better deal later. On the other hand, less healthy people want to delay annuitizing longer because the annuity's price is "too high."

We refer the interested reader to Hurd and McGarry (1995, 1997) for a discussion of experiments involving "subjective" versus "objective" assessments of survival probabilities. Also, more recently, Smith, Taylor, and Sloan (2001) claim that their "findings leave little doubt that subjective perceptions of mortality should be taken seriously." They state that individuals "longevity expectations are reasonably good predictions of future mortality." Asymmetry of beliefs might go a long way towards explaining why individuals who believe themselves to be less healthy than average are more likely to avoid annuities, despite having no declared bequest motive. In the classical Yaari (1965) framework, subjective survival rates do not play a role in the optimal policy. We will prove the surprising result that as long as the consumer disagrees with the insurance company's pricing basis regarding her subjective hazard rate – or personal health status – she will defer annuitization.

- In this section, the annuities are priced in a profitless environment in which loads and commissions are set to zero. Indeed, the above-mentioned study by MPWB (1999) finds values-per-premium dollar between 0.75 and 0.93, depending on the relevant mortality table, yield curve, gender, and age. In our context, this would imply another incentive to defer, since the return in the "good" state is more likely to exceed the mortality-adjusted risk-free rate. This would hold true as long as the proportional insurance loads do not increase as a function of age. Indeed, Table 3 on page 1308 of MPWB (1999) seems to indicate that loads decrease from age 55 to 75 when annuitant tables are benchmarked against the corporate yield curve. This, once again, provides an incentive to defer.
- Finally, one must be careful in referring to 2.49% as the value of the RODA. In theory, the individual might also defer for two periods, and then annuitize. Plus, the individual might find it better to consume something other than the amount that the annuity would have provided. To be absolutely precise, we should think of 2.49% as a lower bound on the value of the RODA, since the individual may defer for many periods or consume differently. We return to this issue in Section 3 where we develop a continuous-time model.

Having considered the basic intuition in a simple 3-period example, we now move to a continuous-time model in which some realistic estimates are developed for the value of the RODA.

3. Continuous-time Model of Option to Wait

In this section, we consider a continuous-time model of the option to defer annuitizing one's wealth. In Section 3.1 we consider the case in which the annuity has a fixed benefit. We extend this model in Section 3.2 by allowing the individual to invest a portion of wealth in a variable immediate annuity, while the remainder is "invested" in a fixed immediate annuity. In Section 3.3, we consider the case in which the annuity has an increasing benefit, and the individual chooses the rate of increase of the benefit.

3.1 Fixed Annuity Benefit

We assume that an individual can invest in a riskless asset whose price at time t, X_t, follows the process $dX_s = rX_s\, ds$, $X_t = x > 0$, for some fixed $r \geq 0$. Also, that individual can invest in a risky asset whose price at time s, S_s, follows geometric Brownian motion given by

$$\begin{cases} dS_s = \mu S_s ds + \sigma S_s dB_s, \\ S_t = S > 0, \end{cases}$$

in which $\mu > r$, $\sigma > 0$, and B_s is a standard Brownian motion with respect to a filtration $\{\mathcal{F}_s\}$ of the probability space $(\Omega, \mathcal{F}, \Pr)$. Let W_s be the wealth at time s of the individual, and let π_s be the amount that the decision maker invests in the risky asset at time s. Also, the decision maker consumes at a rate of c_s at time s. Then, the amount in the riskless asset is $W_s - \pi_s$, and wealth follows the process

$$\begin{cases} dW_s = d(W_s - \pi_s) + d\pi_s - c_s ds \\ \qquad = r(W_s - \pi_s)dt + \pi_s(\mu ds + \sigma dB_s) - c_s ds \\ \qquad = \left[rW_s + (\mu - r)\pi_s - c_s \right] ds + \sigma \pi_s dB_s, \\ W_t = w > 0. \end{cases} \tag{3.1}$$

We assume that the decision maker seeks to maximize (over admissible $\{c_s, \pi_s\}$ and over times of annuitizing his or her wealth, τ) the expected utility of discounted consumption. Admissible $\{c_s, \pi_s\}$ are those that are measurable with respect to the information available at time s, namely $\mathcal{F}_s$, that restrict consumption to be non-negative, and that result in (3.1) having a unique solution; see Karatzas and Shreve (1998). We also allow the individual to value expected utility via a subjective hazard rate (or force of mortality), while the annuity is priced by using an objective hazard rate.

Our financial economy is based on the (simpler) Geometric Brownian Motion (GBM) plus risk-free rate model originally pioneered by Merton (1971), as opposed to the more recent and richer models developed by Kim and Omberg (1996), Sorensen (1999) or Wachter (2002), for example. This is because we are primarily interested in the implications of introducing a mortality contingent claim into the asset allocation framework, as opposed to studying the impact of stochastic interest rates, or mean-reverting equity premiums *per se*. Thus, by avoiding – and not paying the computational price of – a more complex set-up, we are able to obtain closed-form solutions for option values, the optimal asset allocation, and timing strategies. While the simplicity is obviously a shortcoming of our model, we are confident that a richer financial economic model – which is on our agenda for future research – will not alter the qualitative existence of the option to wait. It will only impact the precise timing and value of the option. Along the same lines, we deliberately ignore the complex interaction between nominal and real consumption values. Thus, our model can either be phrased entirely in terms of nominal annuities, that are not inflation-protected, or entirely in terms of real annuities that provide CPI-linked income. However, we do not model the existence of both at the same time. Note, however, that real (after-inflation) annuities are extremely rare in the U.S., especially on the retail level.[6] In any event, we are more comfortable structuring the entire model in nominal terms, which is the more appropriate choice, for example, in a Defined Contribution pension plan. Indeed, our model in Section 3.3 will allow for annuities that increase (or escalate) at a fixed rate each year, and as such, do come close to what some practitioners consider a real annuity.

Apologies aside, we now more on to the insurance and actuarial assumptions. We let ${}_tp_x^S$ denote the subjective conditional probability that an individual aged (x) believes he or she will survive to age ($x + t$). It is defined via the subjective hazard function, λ_{x+s}^S, by the formula ${}_tp_x^S = \exp\left(-\int_0^t \lambda_{x+s}^S \, ds\right)$. We have a similar formula for the objective conditional probability of survival, ${}_tp_x^O$, in terms of the objective hazard function, λ_{x+s}^O. The actuarial present value of a life annuity that pays \$1 per year continuously to (x) is written $\bar{a}_x$. It is defined by $\bar{a}_x = \int_0^\infty e^{-rt} \, {}_tp_x dt$. If we use the subjective hazard rate to calculate the survival probabilities, then we write $\bar{a}_x^S$, while if we use the objective

---

[6] See Brown and Warshawsky (2001) where the authors claim that only one insurance company, ILOMA, offers this product in the U.S. market, and the company has had *zero sales* in the last few years.

(pricing) hazard rate to calculate the survival probabilities, then we write $\bar{a}_x^O$. Just to clarify, by objective $\bar{a}_x^O$, we mean the actual market prices of the annuity, whereas $\bar{a}_x^S$ denotes what the market price 'would have been' had the insurance company used the individual's personal and subjective assessment of her mortality.

The individual consumes at the rate of c_s between time t and the time of annuitization τ. If $\sigma = 0$, then we can show that the individual annuitizes *all* her wealth at time T for which $\mu = r + \lambda_{x+T}^O$. Furthermore, if $\lambda_{x+t}^S = \lambda_{x+t}^O$ for all $t > 0$, then the individual will consume (exactly) the annuity income after time T. As an approximation to the case for which $\sigma > 0$, we therefore assume that at some time τ, the individual annuitizes all her wealth W_τ and thereafter consumes at a rate of $\dfrac{W_\tau}{\bar{a}_{x+\tau}^O}$.

It follows that the associated value function of this problem is given by

$$\begin{aligned} U(w,t) &= \sup_{\{c_s,\pi_s,\tau\}} E\left[\int_t^\tau e^{-r(s-t)}\,{}_{s-t}p_{x+t}^S\,u(c_s)\,ds + \int_\tau^\infty e^{-r(s-t)}\,{}_{s-t}p_{x+t}^S\,u\left(\frac{W_\tau}{\bar{a}_{x+\tau}^O}\right)ds \,\middle|\, W_t = w\right] \\ &= \sup_{\{c_s,\pi_s,\tau\}} E\left[\int_t^\tau e^{-r(s-t)}\,{}_{s-t}p_{x+t}^S\,u(c_s)\,ds + e^{-r(\tau-t)}\,{}_{\tau-t}p_{x+t}^S\,u\left(\frac{W_\tau}{\bar{a}_{x+\tau}^O}\right)\bar{a}_{x+\tau}^S \,\middle|\, W_t = w\right], \end{aligned} \tag{3.2}$$

in which u is an increasing, concave utility function of consumption. Note that we assume that the individual discounts consumption at the riskless rate r. If we were to model with a subjective discount rate of say ρ, then this is equivalent to using r as in (3.2) and adding $\rho - r$ to the subjective hazard rate. Thus, there is no effective loss of generality in setting the subjective discount rate equal to the riskless rate r.

Note that we do not account for pre-existing annuities, say from Social Security in the U.S. We anticipate that such annuities will change the optimal time of annuitization, but we defer this problem to further research. Also, note that we take as given the annuity prices; we are not creating an equilibrium (positive) model of pricing as in the adverse selection literature (Akerlof, 1970; Rothschild and Stiglitz, 1976), but rather a normative model of how people should behave in the presence of these given market prices.

In this paper, we restrict our attention to the case in which the utility function exhibits constant relative risk aversion (CRRA), $-cu''(c)/u'(c)$. That is, u is given by

$$u(c) = \frac{1}{1-\gamma} c^{1-\gamma}, \ \gamma > 0, \ \gamma \neq 1.$$

For this utility function, the relative risk aversion equals γ, a constant. The utility function that corresponds to relative risk aversion 1 is logarithmic utility. In Appendix D, we include the formulas for logarithmic utility that parallel those for CRRA utility ($\gamma > 0$, $\gamma \neq 1$) that we develop in this section. In Appendix A, we show that for CRRA utility, we can assume first that the optimal stopping time is some fixed time in the future, say T. Based on that value of T, we then find the optimal consumption and investment policies. Finally, we find the optimal value of $T \geq 0$.

To this end, define the value function V by

$$V(w, t; T) = \sup_{\{c_s, \pi_s\}} E\left[\int_t^T e^{-r(s-t)} {}_{s-t}p_{x+t}^S \frac{1}{1-\gamma} c_s^{1-\gamma} \, ds + e^{-r(T-t)} {}_{T-t}p_{x+t}^S \frac{1}{1-\gamma} \left(\frac{W_T}{\bar{a}_{x+T}^O} \right)^{1-\gamma} \bar{a}_{x+T}^S \middle| W_t = w \right]. \tag{3.3}$$

V solves the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{cases} (r + \lambda_{x+t}^S) V = V_t + \max\limits_{\pi} \left[\frac{1}{2} \sigma^2 \pi^2 V_{ww} + (\mu - r) \pi V_w \right] + rwV_w + \max\limits_{c \geq 0} \left[-cV_w + \frac{1}{1-\gamma} c^{1-\gamma} \right], \\ V(w, \ T; T) = \frac{1}{1-\gamma} \left(\frac{w}{\bar{a}_{x+T}^O} \right)^{1-\gamma} \bar{a}_{x+T}^S. \end{cases} \tag{3.4}$$

To see how to get (3.4) from (3.3) formally, note that the effective discount rate, $r + \lambda_{x+t}^S$, is multiplied by V on the left-hand side of (3.4). On the right, the ds-term in the wealth process, namely $rw + (\mu - r)\pi - c$, is multiplied by V_w, ½ of the square of the dB_s-term in the wealth process is multiplied by V_{ww} (arising from Itô's lemma), and the utility of consumption that arises from the integral in (3.3) is also included. The boundary condition is given by setting $t = T$ in (3.3). See Björk (1998) for clear derivations of such HJB equations.

The optimal consumption and investment policies are given via the first-order conditions from (3.4) by

$$c^*(w, t) = \left(V_w(w, t) \right)^{-\frac{1}{\gamma}},$$

and

$$\pi^*(w,t) = -\frac{\mu - r}{\sigma^2}\frac{V_w(w,t)}{V_{ww}(w,t)},$$

respectively.

It is straightforward, but tedious, to show that for CRRA utility, *V* is given by

$$V(w,t;T) = \frac{1}{1-\gamma} w^{1-\gamma} \left[\left\{ \frac{\bar{a}^S_{x+T}}{\left(\bar{a}^O_{x+T}\right)^{1-\gamma}} \right\}^{\frac{1}{\gamma}} e^{-\frac{r-\delta(1-\gamma)}{\gamma}(T-t)} \left({}_{T-t}p^S_{x+t} \right)^{\frac{1}{\gamma}} + \int_t^T e^{-\frac{r-\delta(1-\gamma)}{\gamma}(s-t)} \left({}_{s-t}p^S_{x+t} \right)^{\frac{1}{\gamma}} ds \right]^{\gamma},$$

(3.5)

in which $\delta = r + \frac{(\mu - r)^2}{2\sigma^2\gamma}$. The optimal consumption and investment policies are given in feedback form by

$$C^*_t = c^*(W^*_t, t) = W^*_t\, k(t), \qquad (3.6)$$

and

$$\Pi^*_t = \pi^*(W^*_t, t) = \frac{\mu - r}{\sigma^2\gamma} W^*_t, \qquad (3.7)$$

respectively, in which W^*_t is the optimally controlled wealth before annuitization (time *T*). Here, the function *k* is defined by

$$k(t) = \left[\left\{ \frac{\bar{a}^S_{x+T}}{\left(\bar{a}^O_{x+T}\right)^{1-\gamma}} \right\}^{\frac{1}{\gamma}} e^{-\frac{r-\delta(1-\gamma)}{\gamma}(T-t)} \left({}_{T-t}p^S_{x+t} \right)^{\frac{1}{\gamma}} + \int_t^T e^{-\frac{r-\delta(1-\gamma)}{\gamma}(s-t)} \left({}_{s-t}p^S_{x+t} \right)^{\frac{1}{\gamma}} ds \right]^{-1}.$$

If we are in the case of logarithmic utility, as examined in Appendix D, then the optimal consumption rate is $C^*_t = \frac{W^*_t}{\bar{a}^S_{x+t}}$. It is interesting to note that if $r = 0$, in which case the denominator collapses to a (subjective) life expectancy, the consumption rate is precisely the minimum rate mandated by the IRS for annual consumption withdrawals from IRAs after age 71 in the U.S. Specifically, the proportion required to be withdrawn from one's annuity each year equals the start-of-year balance divided by the future expectation of life. Because $r > 0$ in reality, the minimum IRS-mandated consumption rate is less than what is optimal for individuals with logarithmic utility and with mortality equal to that in the IRS tables.

To find the optimal time of annuitization, differentiate V in (3.5) with respect to T. One can show that

$$\frac{\partial V}{\partial T} \propto \left[\frac{\gamma}{1-\gamma} \left(\frac{\bar{a}^S_{x+T}}{\bar{a}^O_{x+T}} \right)^{-\frac{1-\gamma}{\gamma}} - \frac{1}{1-\gamma} + \frac{\bar{a}^S_{x+T}}{\bar{a}^O_{x+T}} \right] + \bar{a}^S_{x+T} \left[\delta - (r + \lambda^O_{x+T}) \right]. \tag{3.8}$$

Thus, if the expression on the right-hand side of (3.8) is negative for all $T \geq 0$, then it is optimal to annuitize one's wealth immediately. However, if there exists a value $T^* > 0$ such that the right-hand side of (3.8) is positive for all $0 \leq T < T^*$ and is negative for all $T > T^*$, then it is optimal to annuitize one's wealth at time T^*. In all the examples we present below, one of these two conditions holds. The decision to annuitize is independent of one's wealth, an artifact of CRRA utility.

Note that if the subjective and objective forces of mortality are equal, then we have

$$\frac{\partial V}{\partial T} \propto \left[\delta - (r + \lambda_{x+T}) \right]. \tag{3.9}$$

If the hazard rate λ_x is increasing with respect to age x, then *either* $\delta \leq (r + \lambda_x)$, from which it follows that it is optimal to annuitize one's wealth immediately, *or* $\delta > (r + \lambda_x)$, from which it follows that there exists a time T in the future (possibly infinity) at which it is optimal to annuitize one's wealth. To interpret this result, recall that the rate of return (ignoring consumption) $r^a = 2\delta - r$; thus, the individual will annuitize his or her wealth when $r^a \leq r + 2\lambda_x$. It follows that annuitizing one's wealth is optimal as soon as the *excess* return, $r^a - r = \frac{(\mu - r)^2}{\sigma^2 \gamma}$, is exceeded by *twice* the hazard rate. Stated differently, we the optimal age to purchase a fixed immediate life annuity is as soon as the instantaneous force of mortality, λ_x, is greater than $\frac{(\mu - r)^2}{2\sigma^2 \gamma}$. Thus, one can think of the hazard rate as a form of excess return on the annuity due to the fact that the wealth reverts to the insurance company when the buyer of the annuity dies.

In all our examples, we observe that if the subjective force of mortality is *different* than the objective force of mortality, then the optimal time of annuitization increases from the T given by the zero of the right-hand side of (3.9). We can show mathematically that this is true if the subjective force of mortality varies from the objective force to the extent that $\bar{a}_x^S < 2\bar{a}_x^O$ for all x (see Appendix B), and we conjecture that it is true in general. Note that this inequality is automatically true for people who are less healthy because in this case, $\bar{a}_x^S < \bar{a}_x^O$ for all x. For an individual who is less healthy than the average person, the annuity will be too expensive, so that person wants to delay annuitizing her wealth. On the other hand, for an individual who is healthier than the average person, the annuity will be relatively cheap. However, such a healthy person will live longer on average and will be interested in receiving a larger annuity benefit by consuming less now and by waiting to buy the annuity later in life. Therefore, a healthy person is also willing to delay annuitizing her wealth in exchange for a larger annuity benefit (for a longer time).

Once again, we emphasize the counterintuitive nature of this result. Regardless of whether the individual believes she is healthier or less healthy, compared to the mortality assumption used by the insurance company, she will defer annuitization. In the meantime, she will consume at the optimal rate – which will depend on her subjective mortality assessment – and then convert her accumulated savings to a life annuity at time T.

Of course, by following the optimal policies of investment, consumption, and annuitizing one's wealth, an individual runs the risk of being able to consume less after annuitizing wealth than if she had annuitized wealth immediately at time $t = 0$. Naturally, there is the chance of the exact opposite, namely that the lifetime annuity stream will be higher. Therefore, to quantify this risk, we calculate the probability associated with various consumption outcomes. See Appendix C for the formula of this probability. We include calculations of it in Example 3.1 below.

For general subjective and objective forces of mortality, if it is optimal for the individual to annuitize his or her wealth at time T, then the value of the RODA, h, is defined to be the least amount of money that when added to current wealth makes the person indifferent between annuitizing now (with the extra wealth) and annuitizing at time T (without the extra wealth). Thus, h is given by

$$V(w, t; T) = V(w + h, t; 0), \tag{3.10}$$

in which T is the optimal time of annuitization and V is given by (3.5). In our examples, we express h as a percentage of wealth w. This is appropriate because V exhibits CRRA with respect to w.

3.1 Example Most mortality tables are discretized, but we require a continuous-time mortality law. We use a Gompertz force of mortality, which is common in the actuarial literature and was recently used for empirical annuity pricing by Frees, Carriere, and Valdez (1996). This model for mortality has also been employed in the economics literature for pricing insurance; see Johansson (1996), for example. The force of mortality is written $\lambda_x = \exp((x - m)/b)/b$, in which m is a modal value and b is a scale parameter. Note that the force of mortality increases exponentially with age. In this paper, we fit the parameters of the Gompertz, namely m and b, to the Annuity 2000 Mortality Table with projection scale G. For males, we fit parameters (m, b) = (88.18, 10.5); for females, (92.63, 8.78). Initially, we assume that the subjective and objective forces of mortality are equal. Throughout this paper, we assume that the seller of the annuity uses the female hazard rate to price annuities for women; similarly, for men. Figure 1 shows the graph of the probability density function (PDF) of the future-lifetime random variable under a Gompertz hazard rate that is fitted to the above-mentioned discrete mortality table.[7]

[7] We actually fit a Makeham hazard rate, or force of mortality, namely $\lambda + \exp((x - m)/b)/b$, in which $\lambda \geq 0$ is a constant that models an accident rate. However, the fitted value of λ was 0, so the effective form of the hazard rate is Gompertz (Bowers et al., 1997).

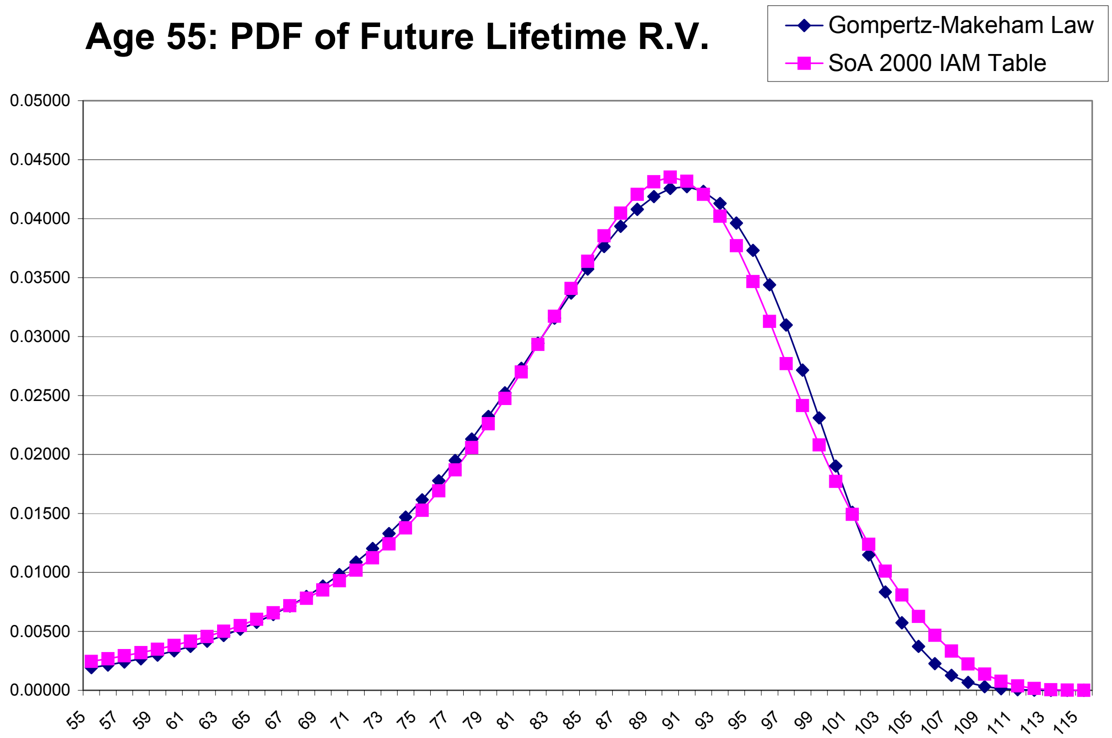

Figure 1

As for the capital market parameters, the risky stock is assumed to have drift $\mu =$ 0.12 and volatility $\sigma = 0.20$. This is roughly in line with Chicago's Ibbotson Associates (2000) numbers, which are widely used by practitioners when simulating long-term investment returns. We assume that the rate of return of the riskless bond is 0.06. We display values for the RODA, h, for two different levels of risk aversion, $\gamma = 1$ (logarithmic utility) and $\gamma = 2$. A variety of studies have estimated the value of γ to lie between 1 and 2. See, the paper by Friend and Blume (1975) that provides an empirical justification for constant relative risk aversion, as well as the more recent MPWB (1999) paper in which the CRRA value is taken between 1 and 2. In the context of estimating the present value of a variable annuity for Social Security, Feldstein and Ranguelova (2001) provide some qualitative arguments that the value of CRRA is less than 3 and probably even less than 2.

For a variety of ages, Table 1a provides the optimal age of annuitization, the value of the RODA as a percentage of initial wealth, and the probability of consuming less at the optimal time of annuitization than if one had annuitized one's wealth immediately. We refer to this as the *probability of deferral failure*.

Table 1a. The value of the real option to defer annuitization for males and females and for a coefficient of relative risk aversion of $\gamma = 1$ and 2. We assume the non-annuitized funds are invested in a risky asset with drift 0.12 and volatility 0.20. The risk-free rate is 0.06. The mortality is assumed to be Gompertz fit to the table IAM2000 with projection Scale G. For example, a 70-year-old female, with a coefficient of relative risk aversion of 2, will destroy a real option worth 5.2% of her wealth if she chooses to annuitize immediately. The optimal time for her to annuitize is at age 78.4.

	$\gamma = 1$, Female (Male)			$\gamma = 2$, Female (Male)		
Age	Optimal age of Annuitization	Value of RODA	Probability of Deferral Failure	Optimal age of Annuitization	Value of RODA	Probability of Deferral Failure
60	84.5 (80.3)	44.0 (32.0)%	0.311 (0.353)	78.4 (73.0)	15.3 (8.9)%	0.268 (0.321)
65	84.5 (80.3)	33.4 (21.9)	0.346 (0.391)	78.4 (73.0)	10.3 (4.3)	0.310 (0.372)
70	84.5 (80.3)	22.7 (12.3)	0.385 (0.431)	78.4 (73.0)	5.2 (0.8)	0.362 (0.435)
75	84.5 (80.3)	12.3 (4.2)	0.429 (0.470)	78.4 (Now)	1.2 (0.0)	0.428 (N/a)
80	84.5 (80.3)	3.7 (0.02)	0.473 (0.500)	Now (Now)	0.0 (0.0)	N/a (N/a)
85	Now (Now)	0.0 (0.0)	N/a (N/a)	Now (Now)	0.0 (0.0)	N/a (N/a)

Note that females wish to annuitize at older ages than do males because the mortality rate of females is lower than that of males. Also, note that more risk averse individuals wish to annuitize sooner, an intuitively pleasing result. Finally, the value of the RODA decreases as one gets closer to the optimal age of annuitization, as one expects.

The probability of deferral failure, although seemingly high, is balanced by the probability of ending up with more than, say, 20% of the original annuity amount. For example, for a 70-year-old female with $\gamma = 2$, the probability of consuming at least 20% more at the optimal age of annuitization than if she were to annuitize immediately is 0.474. Obviously, on a utility-adjusted basis this is a worthwhile trade-off as evidenced by the positive option value. See Table 1b for tabulations of the probability that the individual consumes at least 20% more at the optimal age of annuitization than if he or she were to annuitize immediately, for various ages and for $\gamma = 1$ and 2.

Table 1b. Assuming the individual self-annuitizes, and defers the purchase of a life annuity to the optimal age, this table indicates the probability of consuming at least 20% more at the time of annuitization, compared to if one annuitizes immediately. Thus, for example, a female (male) at age 65 with a coefficient of relative risk aversion of γ=1, has a 64.4% (59.6%) chance of creating a 20% larger annuity flow.		
	γ= 1, Female (Male)	γ= 2, Female (Male)
Age	Probability of consuming at least 20% more	Probability of consuming at least 20% more
60	0.644 (0.596)	0.631 (0.551)
65	0.602 (0.549)	0.565 (0.459)
70	0.552 (0.494)	0.474 (0.296)
75	0.493 (0.425)	0.316 (0.133)
80	0.414 (0.137)	N/a (N/a)
85	N/a (N/a)	N/a (N/a)

These "upside" probabilities decrease as the optimal age of annuitization approaches. Also, for a given age, they decrease as the CRRA increases; this makes sense because a less risk-averse person is less willing to face a distribution with a higher variance. □

3.2 Example Continue the assumptions in Example 3.1 as to the financial market. Suppose that we have a male, aged 60 with γ = 2, whose objective mortality follows that from Example 3.1; that is, annuity prices are determined based on the hazard rate given there. For this example, suppose that the subjective force of mortality is a multiple of the objective force of mortality; specifically, $\lambda_x^S = (1+f)\lambda_x^O$, in which f ranges from –1 (immortal) to infinity (at death's door). This transformation is called the *proportional hazards* transformation in actuarial science (Wang, 1996), and it is similar to the transformation examined by Johansson (1996) in the economic context of the value of increasing one's life expectancy.

In Table 2, we present the value of the RODA, the optimal age of annuitization, the optimal rate of consumption before annuitization (as a percentage of current wealth), and the rate of consumption after annuitization (also, as a percentage of current wealth). For comparison, if the male were to annuitize his wealth at age 60, the rate of consumption would be 8.34%. Also, the optimal proportion invested in the risky stock before annuitization is 75%.

Table 2. The value of the real option to defer annuitization (RODA) for a male, aged 60 with $\gamma = 2$. We assume the funds are invested in a risky asset with drift 0.12 and volatility 0.20. The risk-free rate is 0.06. The mortality is assumed to be Gompertz fit to the table IAM2000 with projection Scale G. Thus, for example, if the individual's subjective hazard rate is 20% higher (i.e. less healthy) than the mortality table used by the insurance company to price annuities, the optimal annuitization is at age 73.1, and the value of the RODA is 8.84% of the individual's wealth at age 60. While the 60-year-old male waits to annuitize, he consumes at the rate of 8.85% of assets, and once he purchases the fixed immediate annuity, his consumption rate – and standard of living – will increase to 11.26% of assets.

f	Optimal age of annuitization	Value of RODA	Consumption rate Before ann'n	Consumption rate After ann'n
-1.0	78.28	13.79%	7.55%	13.38%
-0.8	74.58	10.54	7.95	11.79
-0.6	73.71	9.68	8.18	11.47
-0.4	73.29	9.23	8.37	11.33
-0.2	73.09	8.99	8.54	11.26
0.0	73.03	8.87	8.70	11.24
0.2	73.08	8.84	8.85	11.26
0.5	73.31	8.93	9.06	11.33
1.0	74.04	9.34	9.38	11.59
1.5	75.21	10.00	9.68	12.03
2.0	76.96	10.89	9.98	12.76
2.5	79.71	12.01	10.26	14.12
3.0	85.38	13.38	10.55	18.01

Note that as the 60-year-old male's subjective mortality gets closer to the objective (pricing) mortality, then the optimal age of annuitization decreases. We hypothesize that the optimal age of annuitization will be a minimum when the subjective and objective forces of mortality equal, at least for increasing forces of mortality. We conjecture that this result is true in general, but we only have a proof of it when $\bar{a}_x^S < 2\bar{a}_x^O$; see Appendix B. Also, note that the consumption rate before annuitization increases as the person becomes less healthy, as expected.

Compare these rates of consumption with 8.34%, the rate of consumption if the male were to annuitize his wealth immediately. We see that if the male is healthy relative to the pricing force of mortality, then he is willing to forego current consumption in exchange for greater consumption when he annuitizes, at least up to f = -0.4. Past that point, the optimal rate of consumption before annuitization is greater than 8.34%. For a 60-year-old male with f = -0.2 (20% more healthy than average), see Figure 2 for a graph of his expected consumption rate as a percentage of *initial* wealth. We also graph the 25th and 75th percentiles of his random consumption. This individual expects to live to age 84.4. Note that he has roughly a 70% chance of consuming more throughout his life than he would if he were to annuitize at age 60. □

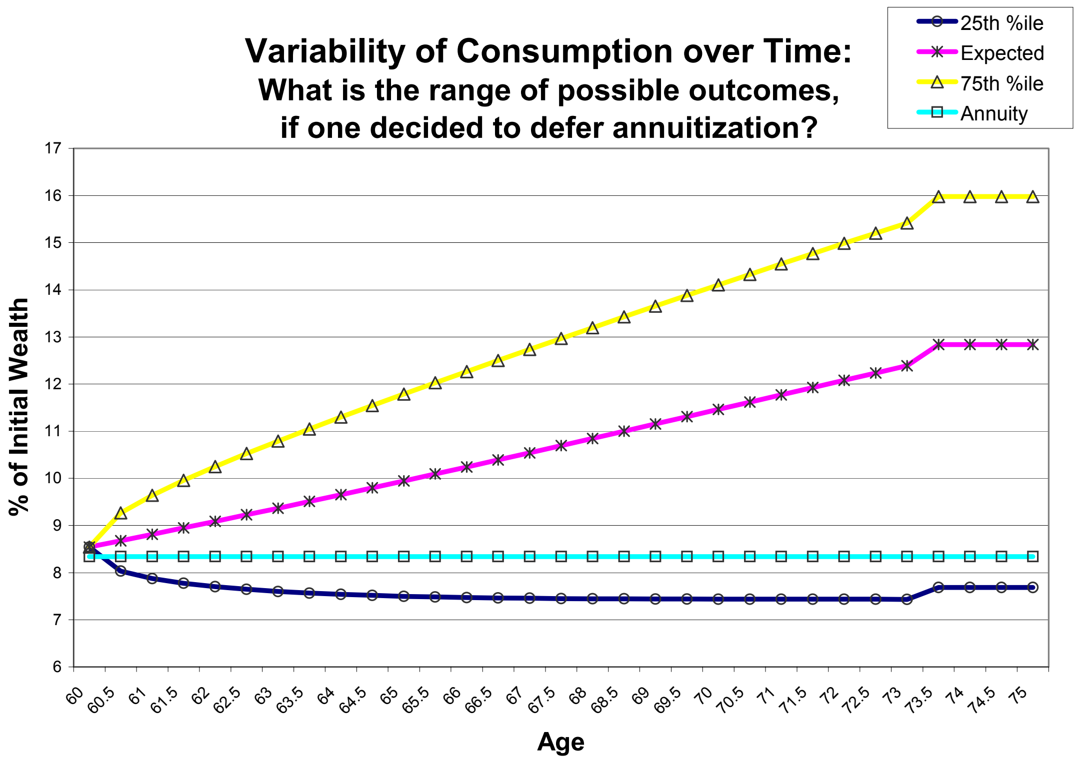

Figure 2. Expected consumption of a 60-year-old male who is 20% more healthy than average; 8.34% is the rate of consumption if he annuitizes his wealth at age 60. We also display the 25th and 75th percentiles of the distribution of consumption between ages 60 and 75.

3.2 Variable and Fixed Annuity Benefits

Up to this point, we have assumed that the only immediate annuities available upon annuitization are fixed. In other words, the annuities pay a fixed-dollar amount for the rest of the annuitant's life. Indeed, the bulk of the immediate annuity market is based on this concept. However, a recent development in the U.S. retirement market is the availability of variable, or fluctuating, immediate annuities. The full-fledged version of this developing product provides the consumer with a menu of risky and risk-free assets within an immediate annuity wrapper. These annuities, which were first pioneered by TIAA-CREF, entitle the annuitant to a fluctuating periodic payment that is based on the performance of the underlying assets supporting the account.

We stress that this innovation is not yet widely available[8] and is currently unavailable in many counties and pension jurisdictions.[9] For example, in Canada, one can only purchase fixed immediate annuities. Also, the few U.S.-based 401(k), 403(b), and 457 pension plans that *do* offer an immediate annuity option are reluctant to offer variable payout products, perhaps due to the complexity of the product and fiduciary concerns (Brown and Warshawsky, 2001). Unfortunately, most retail variable immediate annuities are plagued by high fees and hidden costs.

Nevertheless, this product is potentially available in the U.S., and some companies do allow a variety of asset allocations within the variable immediate annuity. The question we now address is: "How does the availability of a variable immediate annuity product affect the asset allocation decision and the optimal age at which to annuitize?" To this end, we introduce the symbol β to represent the proportion of wealth at the time of annuitization that is invested in the variable immediate annuity. It follows that $1 - \beta$ is the proportion invested in the fixed annuity. We assume that the mix between the variable and fixed annuities, namely β versus $1 - \beta$, stays fixed throughout the remaining life of the annuitant. Thus, we have a so-called *money mix* plan.[10]

---

[8] Fenton and Hecht (2001) recently reported that the total 1998 premiums for variable immediate annuities in the U.S. was only $300 million.

[9] Recently, the Diversified Services Group (DSG) released the results of its 2001 Immediate Variable Annuity Survey and Executive Forum (Wayne, PA, Oct. 29, PRNewswire). DSG concluded that the market opportunity for variable annuities is "substantial although still in its embryonic stage." According to Borden Ayers, DSG Principal and head of the DSG Retirement Income Management Practice, "the experience of participants in the Forum confirms our prior consumer research which indicates a strong market demand for the kinds of solutions that annuitization products are designed to provide. Immediate annuities, particularly immediate variable annuities, are not well understood by the distribution channels as a key building block in a holistic solution to address the problem of assuring retirement income protection in the face of substantially longer periods of retirement."

[10] We do not want to get side tracked by the minutia of the product features, but Metropolitan Life, for example, offers an immediate annuity in which this type of strategy can be followed. Other companies restrict the rebalancing frequency and may not allow movements from the fixed annuity to the variable annuity. Nevertheless, we abstract from these issues and take a portfolio approach to the problem.

Again, we consider CRRA utility and provide formulas for the power utility case. One can easily deal with the logarithmic case by letting γ approach 1 in the consumption, investment, and annuitization policies under power utility $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$, $\gamma > 1$, $\gamma \neq 1$.

To further capture the salient features of this product, we assume that the provider of the annuity has a nonzero insurance load on the fixed annuity such that the effective "return" on the fixed annuity is r', with $r' \leq r$. Similarly, there is a nonzero insurance load on the variable annuity such that the drift on the variable annuity is μ', with $\mu' \leq \mu$ and $\mu' - r' \leq \mu - r$. By insurance loads we simply mean that part of the gross earnings are used to pay for the insurance protection. For example, it is quite feasible that an investor can access a mutual fund whose expected return is the SP500 index minus 10 basis points (for example, Vanguard). However, within a variable immediate annuity, the insurance expense is on the order of 25 to 125 basis points. The extra fee is sometimes referred to as the Mortality and Expense (M&E) risk charge and is meant to compensate the insurance company for accepting the longevity risk.

For the mixture of a variable and a fixed annuity, define the value function V by

$$V(w,t;T) = \sup_{\{c_s, \pi_s, \beta\}} E\Bigg[\int_t^T e^{-r(s-t)} \,{}_{s-t}p_{x+t}^S \frac{1}{1-\gamma} c_s^{1-\gamma}\, ds$$

$$+ \int_T^\infty e^{-r(s-t)} \,{}_{s-t}p_{x+t}^S \frac{1}{1-\gamma} \left(\frac{W_T}{\bar{a}_{x+T}^{O,r'}} e^{\beta\left(\mu' - r' - \frac{\beta\sigma^2}{2}\right)(s-T) + \beta\sigma(B_s - B_T)} \right)^{1-\gamma} ds \,\Bigg|\, W_t = w \Bigg], \tag{3.11}$$

in which the second superscript on $\bar{a}_{x+T}^{O,r'}$, namely r', is the rate of discount used to calculate the actuarial present value of the annuity.

We can deal with the choice of β by noting that

$$E\left[W_T^{1-\gamma}e^{\beta(1-\gamma)\left(\mu'-r'-\frac{\beta\sigma^2}{2}\right)(s-T)+\beta(1-\gamma)\sigma(B_s-B_T)}\right]=E\left[E\left[W_T^{1-\gamma}e^{\beta(1-\gamma)\left(\mu'-r'-\frac{\beta\sigma^2}{2}\right)(s-T)+\beta(1-\gamma)\sigma(B_s-B_T)}\middle|\mathcal{F}_T\right]\right]$$
$$=E\left[W_T^{1-\gamma}E\left[e^{\beta(1-\gamma)\left(\mu'-r'-\frac{\beta\sigma^2}{2}\right)(s-T)+\beta(1-\gamma)\sigma(B_s-B_T)}\middle|\mathcal{F}_T\right]\right]$$
$$=E\left[W_T^{1-\gamma}\right]e^{\beta(1-\gamma)\left(\mu'-r'-\frac{\beta\sigma^2}{2}+\frac{\beta(1-\gamma)\sigma^2}{2}\right)(s-T)}$$
$$=E\left[W_T^{1-\gamma}\right]e^{\beta(1-\gamma)\left(\mu'-r'-\frac{\beta\gamma\sigma^2}{2}\right)(s-T)}.$$

Thus, the expectation is maximized if we maximize $\beta\left(\mu'-r'-\frac{\beta\gamma\sigma^2}{2}\right)$. It follows that the optimal value of β equals

$$\beta^*=\frac{\mu'-r'}{\sigma^2\gamma}.$$

Note that the optimal choice of β is independent of the optimal time T of annuitization. Naturally, if β^* is greater than one, the holdings are truncated by the seller of the annuity at 100% of the risky stock.

It follows that V solves the HJB equation

$$\begin{cases}(r+\lambda_{x+t}^S)V=V_t+\max_\pi\left[\frac{1}{2}\sigma^2\pi^2V_{ww}+(\mu-r)\pi V_w\right]+rwV_w+\max_{c\geq 0}\left[-cV_w+\frac{1}{1-\gamma}c^{1-\gamma}\right],\\ V(w,\ T;T)=\frac{1}{1-\gamma}\left(\frac{w}{\bar{a}_{x+T}^{O,r'}}\right)^{1-\gamma}\bar{a}_{x+T}^{S,\,r-\frac{(1-\gamma)(\mu'-r')^2}{2\sigma^2\gamma}}.\end{cases} \tag{3.12}$$

Note that this is the same HJB equation as in (3.4), except that the boundary condition reflects the mixture of the variable and fixed annuities. As in (3.11), the second superscript on the actuarial present value denotes the rate at which the annuity payments are discounted. It follows that V has the form as that given in equation (3.5), except that $\bar{a}_{x+T}^S=\bar{a}_{x+T}^{S,r}$ is replaced with $\bar{a}_{x+T}^{S,\,r-\frac{(1-\gamma)(\mu'-r')^2}{2\sigma^2\gamma}}$ and $\bar{a}_{x+T}^O=\bar{a}_{x+T}^{O,r}$ is replaced with $\bar{a}_{x+T}^{O,r'}$. Also, note that the optimal investment policy is similar in form to the one given in equation

(3.7). If there are no loads on the fixed and variable annuities, that is, if $r' = r$ and $\mu' = \mu,$ then the proportion of wealth invested in the risky asset from before annuitization equals the proportion after annuitization; however, in this case, the optimal strategy of the individual is to annuitize immediately. This latter result follows from the work of Yaari (1965). In the case of no loads, it is comforting that the individual will invest the same proportion of wealth whether or not annuities are available. This result, which is quite intuitive, is worth stressing. For a CRRA investor with no insurance loads on the annuities, the optimal mixture between risky and risk-free assets is invariant to whether the portfolio is annuitized or not.

The derivative of V with respect to T is proportional to

$$\frac{\partial V}{\partial T} \propto \left[\frac{\gamma}{1-\gamma} \left(\frac{\bar{a}_{x+T}^{S,\, r - \frac{(1-\gamma)(\mu'-r')^2}{2\sigma^2\gamma}}}{\bar{a}_{x+T}^{O,\, r'}} \right)^{-\frac{1-\gamma}{\gamma}} - \frac{1}{1-\gamma} + \frac{\bar{a}_{x+T}^{S,\, r - \frac{(1-\gamma)(\mu'-r')^2}{2\sigma^2\gamma}}}{\bar{a}_{x+T}^{O,\, r'}} \right] + \bar{a}_{x+T}^{S,\, r - \frac{(1-\gamma)(\mu'-r')^2}{2\sigma^2\gamma}} \left[\delta - \delta' - \lambda_{x+T}^{O}) \right], \tag{3.13}$$

in which $\delta' = r' + \dfrac{(\mu' - r')^2}{2\sigma^2 \gamma}$. We can use this equation to determine the optimal value of T in any given situation.

Example 3.3 In this example, assume that the financial market and mortality are as described in Example 3.1, except that for the variable annuity, the insurer has a 100-basis-point Mortality and Expense Risk Charge load on the return, so that the modified drift is $\mu' = 0.11,$ and for the fixed annuity, the insurer has a 50-basis-point spread on the return, so that the modified rate of return is $r' = 0.055.$ Assume that the individual has a CRRA of $\gamma = 2$, from which it follows that the individual will invest 75.0% in the risky stock before annuitization and 68.7% in the variable annuity after annuitization. In Table 3, we compare the optimal ages of annuitization and the values of the RODA when the individual can only buy a fixed annuity (compare with Table 1a) and when the individual can buy a money mix of variable and fixed annuities.

Table 3. The value of the real option to defer annuitization for males and females with a Coefficient of Relative Risk Aversion of $\gamma = 2$. We assume that the non-annuitized funds are invested in a risky asset with drift 0.12 and volatility 0.20. The risk-free rate is 0.06. The insurance loads on the variable and fixed annuities are 100 basis points and 50 basis points, respectively. The mortality is assumed to be Gompertz fit to the table IAM2000 with projection Scale G.

	Fixed Annuity, Female (Male)		Mixture of Variable (68.7%) and Fixed Annuities (31.3%), Female (Male)	
Age	Optimal age of annuitization	Value of RODA	Optimal age of Annuitization	Value of RODA
60	80.2 (75.2)	21.0 (13.4)%	70.8 (64.1)	3.4 (0.6)%
65	80.2 (75.2)	14.8 (7.5)	70.8 (Now)	1.3 (0.0)
70	80.2 (75.2)	8.5 (2.5)	70.8 (Now)	0.04 (0.0)
75	80.2 (75.2)	2.9 (0.003)	Now (Now)	0.0 (0.0)

Note that the option to defer is worth much less when the individual has the ability to invest in a money mix of variable and fixed annuities. This makes sense because the individual will be able to closely match the investing strategy before annuitization with the money mix of annuities. Indeed, when the insurance loads are zero, and the subjective hazard rate is equal to the objective hazard rate, the original Yaari (1965) result emerges once again. Also, note that the optimal time of annuitization and the value of the RODA are greater when there is an insurance load on the risk-free rate than when there is no load, as in Table 1a. ☐

3.3 Escalating Annuity Benefit

Suppose that the individual can buy an escalating annuity. An escalating annuity is one for which the payments increase at a (constant) rate *g*. These are known as COLA (Cost Of Living Adjustment) annuities and are available from all vendors that sell regular fixed annuities. These escalating annuities are popular as a hedge against (expected) inflation, since it is virtually impossible to obtain true inflation-linked annuities in the U.S.

It follows that the actuarial present value of an escalating annuity can be written $\bar{a}_x^{r-g}$; that is, the rate of discount r is reduced by the rate of increase of the payments g. As in the previous two sections, we consider CRRA utility and provide formulas for the power utility case. Thus, we can define the corresponding value function by

$$V(w,t;T) = \sup_{\{c_s,\pi_s,g\}} E\left[\int_t^T e^{-r(s-t)} \, {}_{s-t}p_{x+t}^S \frac{1}{1-\gamma} c_s^{1-\gamma} \, ds \right. \\ \left. + \int_T^\infty e^{-r(s-t)} \, {}_{s-t}p_{x+t}^S \frac{1}{1-\gamma} \left(\frac{W_T}{\bar{a}_{x+T}^{O,r-g}} e^{g(s-T)} \right)^{1-\gamma} ds \, \middle| W_t = w \right]. \tag{3.14}$$

This expression is maximized with respect to g if

$$\frac{1}{1-\gamma} \frac{\int_0^\infty e^{-(r-(1-\gamma)g)s} \, {}_s p_{x+T}^S ds}{\left[\int_0^\infty e^{-(r-g)s} \, {}_s p_{x+T}^O ds\right]^{1-\gamma}}$$

is maximized. The derivative of this expression with respect to g is proportional to

$$\frac{\int_0^\infty s e^{-(r-(1-\gamma)g)s} \, {}_s p_{x+T}^S ds}{\int_0^\infty e^{-(r-(1-\gamma)g)s} \, {}_s p_{x+T}^S ds} - \frac{\int_0^\infty s e^{-(r-g)s} \, {}_s p_{x+T}^O ds}{\int_0^\infty e^{-(r-g)s} \, {}_s p_{x+T}^O ds}.$$

Note that this is a difference of expectations of "*s*" with respect to two probability distributions. Also, note that if $\lambda_x^S = \lambda_x^O - c$ for all x and for some constant c, then the optimal value of g is $g^* = \frac{c}{\gamma}$. For example, if the individual is healthier to the extent that the subjective hazard rate is 0.01 less than the objective (pricing) hazard rate, then optimal rate of increase of the annuity payments (once the individual annuitizes his or her wealth) is $\frac{0.01}{\gamma}$. Note that in general, *healthier* people will want to buy escalating annuities with a *positive* g, while *sicker* people will want to buy escalating annuities with a *negative* g. This makes sense because healthier people anticipate living longer than normal, so they will be able to enjoy those larger annuity payments in the future. Finkelstein and Poterba (1999) show that this phenomenon occurs in the U.K. annuity market. On the other hand, sicker people will not live as long, so they demand higher payments now.

Example 3.4: In this example, assume that the financial market and mortality are as described in Example 3.2, except that the individual has the ability to choose the rate at which the annuity will increase. Consider an individual who is healthier than normal with f = -0.5; that is, the person has one-half the mortality rate of the average person. Suppose that the CRRA is γ = 1.5. In Table 4, we report the optimal age of annuitization and the value of the RODA. We compare these numbers with those when the individual can buy only a fixed annuity. It turns out that the optimal rate of escalation g = 0.02 (to two decimal places) for all ages and for both genders.

Table 4. The value of the RODA for males and females and for $\gamma = 1.5$. We assume the funds are invested in a risky asset with drift 0.12 and volatility 0.20. The risk-free rate is 0.06. The mortality is assumed to be Gompertz fit to the table IAM2000 with projection Scale G, while the individual has subjective mortality equal to one-half of the objective mortality.

	Fixed Annuity, Female (Male)		2% Escalating Annuity, Female (Male)	
Age	Optimal age of annuitization	Value of RODA	Optimal age of annuitization	Value of RODA
60	80.9 (76.1)	23.68 (15.59)%	78.5 (73.2)	17.41 (9.61)%
65	80.9 (76.1)	17.05 (9.24)	78.5 (73.2)	11.50 (4.88)
70	80.9 (76.1)	10.22 (3.57)	78.5 (73.2)	5.80 (0.96)
75	80.9 (76.1)	3.96 (0.15)	78.5 (Now)	1.29 (0.00)

Note that the individual is willing to annuitize sooner if there is a 2% escalating annuity available; however, there is still an advantage for the person to wait. □

4. Conclusion

This paper solves a particular type of portfolio asset allocation and consumption problem in a continuous-time financial economic setting. It is motivated by a pressing issue that is faced by almost all participants in Defined Contribution pension plans. The question is: "Should I annuitize and take the pension, or should I take the lump sum?" Indeed, as the number of DC and 401(k) plans[11] continues to increase at the expense of traditional DB plan, we believe that many Americans will face this exact dilemma. On the opposite side of the issue, trustees, and fiduciaries of these pension plans will have to decide which, if any, payout options should be offered to their participants. Indeed, we fully acknowledge the eventual superiority of immediate annuities and their implicit longevity insurance over all other non-mortality-contingent asset classes for those with no bequest motives. Yet, motivated by the *real options* literature, we hope to shift the question to: "Should I annuitize today or should I annuitize tomorrow?" We, therefore,

[11] According to the ERBI Databook on Employee Benefits, at year-end 2000, there were 325,000 401(k) plans with more than 42 million participants and over $1.8 trillion in assets.

believe that *ad hoc* rules that force annuitization at certain ages, such as 75 in the U.K. or possibly age 67 under certain Social Security proposals in the U.S., are not properly grounded in modern financial economic theory.

The main qualitative insight of this paper is that since this decision is completely irreversible, premature annuitization destroys the *real option* value contained in the ability to annuitize later. Specifically, by initially consuming (self-annuitizing) and then annuitizing later, the retiree stands to gain from the possibility of a relative improvement in the future budget constraint, as well as the resolution of uncertainty regarding one's future lifespan. The real option analogy can be taken a step further by arguing that a deterioration in health status, an increase in interest rates, better asset allocation and liquidity features, possible tax-law changes, or a reduction in actuarial loads will all serve to increase the future annuity payout if the retiree waits and is sufficiently risk tolerant.

From a practical point of view, we find that most individuals in their 60s and 70s should hold a substantial portion of their wealth in non-annuitized assets since the option value to wait is quite large. More importantly, the presence of asymmetric mortality beliefs – whether they are healthier or less healthy compared to the annuity population – and insurance loads that are absent from the non-annuitized assets create a real option to wait. We note that even when there is only a risk-free asset in the economy, a retiree with asymmetric mortality beliefs might be willing to give up the mortality credits embedded in the immediate annuity in exchange for the ability to self-annuitize and create an optimal consumption pattern.

However, we concur that the availability of low-cost variable (or fluctuating) immediate annuities reduces the value of the option to defer since the alternative asset class yields a similar pre-mortality return. Currently, though, very few DC plans – or countries – offer this payout option, even if a fixed immediate annuity is available.[12] Interestingly, the pension giant TIAA-CREF, which covers most of the academics in the U.S., does offer low-cost variable immediate annuities to their participants, where a

[12] In Canada, for example, one can only acquire a fixed immediate annuity in the retail market. Fluctuating, or variable, immediate annuities are not available. As such, there is no mechanism for obtaining mortality-credits together with equity-based asset returns. Participants in tax-sheltered RRSP plans can either manage the funds themselves – i.e., self annuitize subject to minimum withdrawal constraints – or purchase a fixed annuity.

limited number of equity-based and real-estate based investment returns are available within an annuity-structure. This might go some way in explaining the relatively high levels of annuitization selected by TIAA-CREF participants.[13]

In any event, we believe that our framework develops a parsimonious valuation methodology for computing the magnitude of the option value as well as the optimal time to exercise this irreversible option. In further research, we will examine the implications of introducing a bequest motive, more general utility preferences, pre-existing annuities, inflation uncertainty, more complex investment dynamics and stochastic interest rates into a model of asset allocation that includes realistic immediate annuities during the retirement years.[14] We also plan to examine the optimality of various gradual annuitization strategies as well as the possible benefits from purchasing deferred annuities, which are paid-for now but commence years later. Of course, analytic insights will be hard to obtain from such a general model, but we anticipate that the numerical results from solving the appropriate PDE will be revealing and rewarding.

In sum, analogous to the literature in the corporate finance arena, any irreversible *personal* financial decision should only be undertaken when the option value to wait — and do it tomorrow — has no value. Our results indicate that one should exercise caution when exercising the irreversible decision to annuitize one's wealth.

Appendix A

In this appendix, we verify the approach in Section 3 for CRRA utility, for which we assume first that the optimal stopping time is some fixed time in the future, say T. Based on that value of T, we then find the optimal consumption and investment policies. Finally, we find the optimal value of $T \geq 0$.

From Krylov (1980, p. 13), the HJB equation for U from (3.2) is

---

[13] See *Research Dialogues: The Retirement Patterns and Annuitization Decisions of a Cohort of TIAA-CREF participants* (August 1999) by John Ameriks. However, we caution that this particular immediate annuity is somewhat different from those modeled in this paper due to its participating insurance structure. Arguably, it is better described as a cross between an immediate life annuity and a medieval tontine given the fact that longevity risk is shared by the pool, as opposed to borne by the insurance company.

[14] In addition, outside of a qualified plan there are some tax benefits to holding immediate annuities versus alternative asset classes. See, for example, Brown et al. (1999) for details of the tax rules.

$$g - U + \sup_{c,\pi}\left[(rw + (\mu - r)\pi - c)U_w + \frac{1}{2}\sigma^2\pi^2 U_{ww} + U_t - (r + \lambda_{x+t}^S)U + f + U - g \right]_+ = 0,$$

in which

$$f = u(c) \text{ and } g = \bar{a}_{x+t}^S\, u\left(\frac{w}{\bar{a}_{x+t}^O} \right).$$

Thus,

$$(r + \lambda_{x+t}^S)U \ge U_t + rwU_w + \max_c \left[u(c) - cU_w \right] + \max_\pi \left[(\mu - r)\pi U_w + \frac{1}{2}\sigma^2\pi^2 U_{ww} \right],$$

with equality if

$$U(w,t) > \bar{a}_{x+t}^S\, u\left(\frac{w}{\bar{a}_{x+t}^O} \right).$$

Now, for $u(c) = \frac{1}{1-\gamma}c^{1-\gamma}$, $\gamma > 0$, $\gamma \neq 1$, we have $U(w,t) = \frac{1}{1-\gamma} w^{1-\gamma}\psi(t)$, in which

$$\frac{1}{1-\gamma}(r + \lambda_{x+t}^S)\psi \ge \frac{1}{1-\gamma}\psi' + \delta\psi + \frac{\gamma}{1-\gamma}\psi^{\frac{1-\gamma}{\gamma}},$$

with equality if

$$\frac{1}{1-\gamma}\psi(t) > \frac{1}{1-\gamma}\frac{\bar{a}_{x+t}^S}{(\bar{a}_{x+t}^O)^{1-\gamma}}.$$

The solution to this variational inequality is given by the bracketed term in equation (3.5) after one maximizes with respect to *T*, and we are done. □

Appendix B

In this appendix, we show that if the subjective force of mortality varies slightly from the objective force to the extent that $\bar{a}_x^S < 2\bar{a}_x^O$ for all *x*, then the optimal time of annuitization increases from the *T* given by the zero of the right-hand side of (3.9). In particular, if the individual is less healthy than normal ($\lambda_x^S > \lambda_x^O$ for all *x*), then $\bar{a}_x^S < \bar{a}_x^O$ for all *x*, from which it follows that the optimal time of annuitization is delayed. Also, if the individual is more healthy than normal but only to the extent that $\bar{a}_x^S < 2\bar{a}_x^O$ for all *x*, then the optimal time of annuitization is delayed.

Suppose $\overline{a}^S_{x+T} = \overline{a}^O_{x+T} + \varepsilon$, for some small ε, not necessarily positive. Then, equation (3.8) at the critical value T becomes

$$0 = \left[\frac{\gamma}{1-\gamma} \left(\frac{\overline{a}^O_{x+T} + \varepsilon}{\overline{a}^O_{x+T}} \right)^{-\frac{1-\gamma}{\gamma}} - \frac{1}{1-\gamma} + \frac{\overline{a}^O_{x+T} + \varepsilon}{\overline{a}^O_{x+T}} \right] + (\overline{a}^O_{x+T} + \varepsilon)\left[\delta - (r + \lambda^O_{x+T}) \right].$$

We can simplify this equation to

$$0 = (\overline{a}^O_{x+T} + \varepsilon)\left[\delta - (r + \lambda^O_{x+T}) \right] - \frac{1}{2}\left(-\frac{1-\gamma}{\gamma} - 1 \right)\left(\frac{\varepsilon}{\overline{a}^O_{x+T}} \right)^2 - \frac{1}{6}\left(-\frac{1-\gamma}{\gamma} - 1 \right)\left(-\frac{1-\gamma}{\gamma} - 2 \right)\left(\frac{\varepsilon}{\overline{a}^O_{x+T}} \right)^3 + \ldots,$$

if $\dfrac{\varepsilon}{\overline{a}^O_{x+T}}$ lies between –1 and 1. Thus, by the mean value theorem, there exists ε^* between 0 and $\dfrac{\varepsilon}{\overline{a}^O_{x+T}}$ such that

$$0 = (\overline{a}^O_{x+T} + \varepsilon)\left[\delta - (r + \lambda^O_{x+T}) \right] + \frac{1}{2\gamma}(\varepsilon^*)^2,$$

or equivalently,

$$0 = \left[\delta - (r + \lambda^O_{x+T}) \right] + \frac{(\varepsilon^*)^2}{2\gamma} \frac{1}{\overline{a}^O_{x+T}\left(1 + \dfrac{\varepsilon}{\overline{a}^O_{x+T}} \right)}.$$

The second term of the above equation is positive (and small) regardless of the sign of ε. Thus, T is determined by setting $\left[\delta - (r + \lambda^O_{x+T}) \right]$ equal to a negative number. It follows that, for λ^O_x increasing with respect to x, this value of T will be larger than the zero of (3.9).

Note that a sufficient condition for the above result is that $\dfrac{\varepsilon}{\overline{a}^O_{x+T}}$ lie between –1 and 1. Without difficulty, one can show that this requirement is equivalent to $\overline{a}^S_{x+T} < 2\overline{a}^O_{x+T}$. For less healthy people ($\lambda^S_x > \lambda^O_x$ for all x), we have $\overline{a}^S_x < \overline{a}^O_x$ for all x, so $\overline{a}^S_{x+T} < 2\overline{a}^O_{x+T}$ holds automatically. Also, there is some leeway in this inequality, so that even healthier people might have that $\overline{a}^S_{x+T} < 2\overline{a}^O_{x+T}$. Even when this inequality does not hold, we conjecture that we still have a delay in the time of annuitization beyond that given by the zero of the right-hand side of (3.9), as we see in Example 3.2. □

Appendix C

The probability that consumption (as a percentage of initial wealth) at optimal time of annuitization is *p*% less than the consumption if one annuitizes immediately equals

$$\Pr\left(\frac{W^*_T}{\bar{a}^O_{x+T}} < (1-0.01p)\frac{w}{\bar{a}^O_x}\middle| W_0 = w\right) = \Pr\left(w e^{\left(2\delta - r - \frac{(\mu-r)^2}{2\sigma^2\gamma^2}\right)T - \int_0^T k(s)ds + \frac{\mu-r}{\sigma\gamma}B_T} < (1-0.01p) w \frac{\bar{a}^O_{x+T}}{\bar{a}^O_x} \right)$$

$$= \Pr\left(e^{\frac{\mu-r}{\sigma\gamma}B_T} < \frac{(1-0.01p)\bar{a}^O_{x+T}}{\bar{a}^O_x} e^{-\left(2\delta - r - \frac{(\mu-r)^2}{2\sigma^2\gamma^2}\right)T + \int_0^T k(s)ds} \right)$$

$$= \Pr\left(B_T < \frac{\ln\left(\frac{(1-0.01p)\bar{a}^O_{x+T}}{\bar{a}^O_x}\right) - \left(2\delta - r - \frac{(\mu-r)^2}{2\sigma^2\gamma^2}\right)T + \int_0^T k(s)ds}{\frac{\mu-r}{\sigma\gamma}} \right)$$

$$= \Phi\left(\frac{\ln\left(\frac{(1-0.01p)\bar{a}^O_{x+T}}{\bar{a}^O_x}\right) - \left(2\delta - r - \frac{(\mu-r)^2}{2\sigma^2\gamma^2}\right)T + \int_0^T k(s)ds}{\frac{\mu-r}{\sigma\gamma}\sqrt{T}} \right).$$

Here, Φ denotes the cumulative distribution function of the standard normal. □

Appendix D

Let $u(c) = \ln(c)$. Then, the HJB equation for $V(w, t; T)$ is

$$\begin{cases} (r+\lambda^S_{x+t})V = V_t + \max_\pi\left[\frac{1}{2}\sigma^2\pi^2 V_{ww} + (\mu-r)\pi V_w\right] + rwV_w + \max_{c\geq 0}\left[-cV_w + \ln(c)\right], \\ V(w,\ T; T) = \bar{a}^S_{x+T}\ln\left(\frac{w}{\bar{a}^O_{x+T}}\right). \end{cases}$$

If we try $V(w, t) = \psi(t)\ln(w) + \phi(t)$, then ψ and ϕ satisfy the ordinary differential equations

$$\begin{cases} (r+\lambda^S_{x+t})\psi = \psi' + 1, \\ \psi(T) = \bar{a}^S_{x+T}, \end{cases}$$

and

$$\begin{cases}(r+\lambda_{x+t}^{S})\phi = \phi' + \delta\psi - \ln\psi - 1, \\ \phi(T) = -\overline{a}_{x+T}^{S} \ln(\overline{a}_{x+T}^{O}),\end{cases}$$

in which $\delta = r + \frac{(\mu - r)^2}{2\sigma^2}$. It follows that

$$\psi(t) = \overline{a}_{x+t}^{S},$$

and

$$\phi(t) = -\overline{a}_{x+T}^{S} \ln(\overline{a}_{x+T}^{O}) e^{-r(T-t)} {}_{T-t}p_{x+t}^{S} + \int_t^T e^{-r(s-t)} {}_{s-t}p_{x+t}^{S} (\delta \overline{a}_{x+s}^{S} - \ln(\overline{a}_{x+s}^{S}) - 1) ds.$$

The optimal consumption and investment policies are given by

$$C^*_t = \frac{W^*_t}{\overline{a}_{x+t}^{S}},$$

and

$$\Pi^*_t = \frac{\mu - r}{\sigma^2} W^*_t,$$

in which W^*_t is the optimally controlled wealth. These expressions are equal to the ones given by setting $\gamma = 1$ in equations (3.6) and (3.7), respectively. It is interesting that the optimal consumption rate is similar to one of the IRS-approved withdrawals plans for IRAs in the U.S.

If we calculate the derivative of V with respect to T, we obtain

$$\frac{\partial V}{\partial T} \propto \left[-\ln\left(\frac{\overline{a}_{x+T}^{S}}{\overline{a}_{x+T}^{O}} \right) - 1 + \frac{\overline{a}_{x+T}^{S}}{\overline{a}_{x+T}^{O}} \right] + \overline{a}_{x+T}^{S} \left[\delta - (r + \lambda_{x+T}^{O}) \right].$$

One can show that if one takes the limit as γ goes to 1 of the right-hand side of equation (3.8), then we get the above expression for $\frac{\partial V}{\partial T}$. □

References

Akerlof, G. A. (1970), The market for "lemons": Quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, 84: 488-500.

Berk, J. B. (1999), A simple approach for deciding when to invest, *American Economic Review*, 89 (December, 5): 1319-1326.

Bernheim, B. D. (1991), How strong are bequest motives? Evidence based on estimates of the demand for life insurance and annuities, *Journal of Political Economy*, 99 (5): 899-927.

Bowers, N. L., H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt (1997), *Actuarial Mathematics*, second edition, Society of Actuaries, Schaumburg, Illinois.

Björk, T. (1998), *Arbitrage Theory in Continuous Time*, Oxford University Press, New York, New York.

Black, F. and M. S. Scholes (1973), The pricing of options and corporate liabilities, *The Journal of Political Economy*, 81: 637-653.

Blake, D., A. J. Cairns, and K. Dowd (2000), PensionMetrics: Stochastic pension plan design during the distribution phase, *Pensions Institute Working Paper*.

Bodie, Z., A. Marcus, and R. Merton (1988), Defined benefit versus defined contribution plans, in Z. Bodie, J. Shove, and D. Wise (editors), *Pensions in the U.S. Economy*, NBER Research Project Report.

Brown, J., O. Mitchell, J. Poterba, and M. Warshawsky (1999), Taxing retirement income: Non-qualified annuities and distributions from qualified accounts, *National Tax Journal*, 14 (3): 219-234.

Brown, J. R. (1999), Are the elderly really over-annuitized? New evidence on life insurance and bequests, *NBER Working Paper* 7193.

Brown, J. R. (2001), Private pensions, mortality risk, and the decision to annuitize, *Journal of Public Economics*, 82 (1): 29-62.

Brown, J. R. and J. Poterba (2000), Joint life annuities and annuity demand by married couples, *Journal of Risk and Insurance*, 67(4): 527-554.

Brown, J. R. and M. J. Warshawsky (2001), Longevity-insured retirement distributions from pension plans: Market and regulatory issues, *NBER Working Paper* 8064.

Brugiavini, A. (1993), Uncertainty resolution and the timing of annuity purchases, *Journal of Public Economics*, 50: 31-62.

Feldstein, M. and E. Ranguelova (2001), Individual risk in an investment-based social security system, *American Economic Review*, 91 (4): 1116-1125.

Fenton, J. M. and J. Hecht (2001), *Making payout annuities pay*, Tower Perrin.

Finkelstein, A. and J. Poterba (1999), Selection effects in the market for individual annuities: New evidence from the United Kingdom, *NBER Working Paper* 7168.

Frees, E. W., J. Carriere, and E. Valdez (1996), Annuity valuation with dependent mortality, *Journal of Risk and Insurance*, 63 (2): 229-261.

Friedman, B. and M. Warshawsky (1990), The cost of annuities: Implications for saving behavior and bequests, *Quarterly Journal of Economics*, 105 (1): 135-154.

Friend, I. and M. E. Blume (1975), The demand for risky assets, *American Economic Review*, 65: 900-922.

Hodges, S. D. and A. Neuberger (1989), Optimal replication of contingent claims under transaction costs, *Review of Futures Markets*, 8: 222-239.

Hubbard, R. (1994), Investment under uncertainty: Keeping one's options open, *Journal of Economic Literature*, 32 (December, 4): 1816-1831.

Hurd, M. D. (1989), Mortality risk and bequest, *Econometrica*, 57 (July, 4): 779-813.

Hurd, M. D. and K. McGarry (1995), Evaluation of the subjective probabilities of survival in the health and retirement study, *Journal of Human Resources*, 30: 268-291.

Hurd, M. D. and K. McGarry (1997), The predictive validity of subjective probabilities of survival, *NBER Working Paper* 6193.

Ingersoll, J. E. and S. A. Ross (1992), Waiting to invest: Investment and uncertainty, *Journal of Business*, 65 (1): 1-29.

Johansson, P. O. (1996), The value of changes in life expectancy, *Journal of Health Economics*, 15: 105-113.

Kapur, S. and M. Orszag (1999), A portfolio approach to investment and annuitization during retirement, *Birkbeck College (University of London) Mimeo*, May 1999.

Karatzas, I. and S. Shreve (1998), *Methods of Mathematical Finance*, Springer-Verlag, New York.

Kim, T.S. and E. Omberg (1996), "Dynamic Nonmyopic Portfolio Behavior", Review of Financial Studies, Vol. 9. pg 141-161.

Kotlikoff, L. J. and A. Spivak (1981), The family as an incomplete annuities market, *Journal of Political Economy*, 89 (2): 373-391.

Kotlikoff, L. J. and L. H. Summers (1981), The role of intergenerational transfers in aggregate capital accumulation, *The Journal of Political Economy*, 89 (3): 706-732.

Krylov, N. V. (1980), *Controlled Diffusion Processes*, Springer-Verlag, Berlin.

Marcus, A. J. (1985), Spinoff/terminations and the value of pension insurance, *Journal of Finance*, 40 (3): 911-924.

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

Milevsky, M. A. (1998), Optimal asset allocation towards the end of the life cycle: To annuitize or not to annuitize?, *Journal of Risk and Insurance*, 65 (3): 401-426.

Mirer, T. W. (1994), The dissaving of annuity wealth and marketable wealth in retirement, *Review of Income and Wealth*, 40 (March, 1): 87-97.

Mitchell, O. S., J. M. Poterba, M. J. Warshawsky, and J. R. Brown (1999), New evidence on the money's worth of individual annuities, *American Economic Review*, December, 89 (5): 1299-1318.

Modigliani, F. (1986), Life cycle, individual thrift, and the wealth of nations, *American Economic Review*, 76 (3): 297-313.

Poterba, J. M. (1997), The history of annuities in the United States, *NBER Working Paper* 6004.

Poterba, J. M. and D. A. Wise (1996), Individual financial decisions in retirement savings plans and the provision of resources for retirement, *NBER Working Paper* 5762.

Richard, S. (1975), Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model, *The Journal of Financial Economics*, 2: 187-203.

Ross, S. A. (1995), Uses, abuses and alternatives to the net present value rule, *Financial Management*, 24 (3): 96-102.

Rothschild, M. and J. Stiglitz (1976), Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, *Quarterly Journal of Economics*, 90, 629-649.

Smith, V. K., D. H. Taylor, and F. A. Sloan (2001), Longevity expectations and death: Can people predict their own demise?, *American Economic Review*, 91 (4): 1126-1134.

Sondergeld, E. (1997), *Annuity Persistency Study*, LIMRA International and the Society of Actuaries.

Sorensen, C. (1999), Dynamic asset allocation and fixed income management, *Journal of Financial and Quantitative Analysis*, 34 (4): 513-532.

Stanton, R. (2000), From cradle to grave: How to loot a 401(k) plan, *Journal of Financial Economics*, 56: 485-516.

Stock, J. H. and D. A. Wise (1990), Pensions, the option value of work, and retirement, *Econometrica*, 58 (5): 1151-1180.

Sundaresan, S. and F. Zapatero (1997), Valuation, optimal asset allocation, and retirement incentives of pension plans, *Review of Financial Studies*, 10 (3): 631-660.

Trigeorgis, L. (1993), The nature of option interactions and the valuation of investments with multiple real options, *Journal of Financial and Quantitative Analysis*, 28 (1): 1-20.

Trigeorgis, L. (1996), *Real Options*, MIT Press, Cambridge, Massachusetts.

Wachter, J. A. (2002), Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets, *Journal of Financial and Quantitative Analysis*, 37 (1): 63-92.

Wang, S. S. (1996), Premium calculation by transforming the layer premium density, *ASTIN Bulletin*, 26 (1): 71-92.

Yaari, M. E. (1965), Uncertain lifetime, life insurance and the theory of the consumer, *Review of Economic Studies*, 32: 137-150.

Yagi, T. and Y. Nishigaki (1993), The inefficiency of private constant annuities, *The Journal of Risk and Insurance*, 60 (September, 3): 385-412.