# Efficient Annuitization:
# Optimal Strategies for Hedging Mortality Risk

Jason S. Scott, John G. Watson, and Wei-Yin Hu*

PRC WP2007-09
Pension Research Council Working Paper

Pension Research Council
The Wharton School, University of Pennsylvania
3620 Locust Walk, 3000 SH-DH
Philadelphia, PA 19104-6302
Tel: 215.898.7620 Fax: 215.573.3418
Email: prc@wharton.upenn.edu
http://www.pensionresearchcouncil.org

* The authors thank William Sharpe, Geert Bekaert, Steve Grenadier, and Jim Shearer for many excellent comments and suggestions. Any remaining errors or omissions are the authors' responsibility. The views expressed herein are those of the authors and not necessarily those of Financial Engines. JEL Classifications: D11, D91, E21, H55, J14, J26. Keywords: Annuities, annuitization, Social Security, pensions, longevity risk, insurance. Opinions and errors are solely those of the authors and not of the institutions with whom the authors are affiliated. 

Efficient Annuitization:
Optimal Strategies for Hedging Mortality Risk

Abstract

Two common explanations for the dearth of voluntary annuitization are bequest motives and liquidity demand, both of which create implicit costs for each annuitized dollar. Whenever costs prevent full annuitization, we demonstrate that efficient annuity allocations concentrate annuity-funded consumption late in life. This implies traditional immediate payout annuities are inefficient relative to recently introduced "delayed payout annuities" which have survival-contingent payments beginning years after purchase. For typical examples, a six percent delayed payout allocation has utility comparable to a thirty-nine percent immediate annuity allocation. Since retirees appear averse to large annuity purchases, delayed payout annuities could significantly improve retiree welfare.

**Jason S. Scott** (Corresponding Author)
Financial Engines, Inc.
1804 Embarcadero Rd.
Palo Alto, CA 94303
T: 650-565-4925 ● F:650-565-4905
E-mail: **jscott@financialengines.com**

**John G. Watson**
Financial Engines, Inc.
1804 Embarcadero Rd.
Palo Alto, CA 94303
T: 650-565-4922 ● F:650-565-4905
E-mail: **jwatson@financialengines.com**

**Wei-Yin Hu**
Financial Engines, Inc.
1804 Embarcadero Rd.
Palo Alto, CA 94313
T: 650-565-7771 ● F:650-565-4905
E-mail: **whu@financialengines.com**

# Efficient Annuitization: Optimal Strategies for Hedging Mortality Risk

## 1 Introduction

Yaari [1965] theorized that individuals with access to actuarially fair annuities and without a bequest motive would optimally annuitize all of their assets. Since this publication, it has been an open puzzle as to why individuals rarely voluntarily choose to annuitize much, if any, of their retirement assets. Recently, Davidoff, Brown and Diamond [2005] extended the puzzle by pointing out that full annuitization did not require all the axioms of expected utility maximization. Rather, since annuities facilitated consumption at lower prices, all that was essential to the result was a lack of a bequest motive and a preference for more over less.

Since the potential benefit for fully annuitizing is substantial[1], a significant amount of research over the ensuing decades has attempted to explain the lack of annuity demand. In broad terms, the research fell into two categories. First, some researchers focused on why the *benefits* to annuitization were smaller than originally envisioned. Other researchers hypothesized that unaccounted-for *costs* associated with annuitization dampened demand. (See Brown and Warshawsky [2004] for an excellent overview of this literature.)

Two explanations have been proposed that relate to reduced benefits from annuitization. First, some research has noted that many people already have a substantial fraction of their wealth already annuitized in the form of promised pension benefits such as Social Security or employer-based defined benefit plans. (See, for example, Brown

[1] Baseline annuity benefits have been estimated to increase effective available wealth by 50% or more. (See, for example, Mitchell, Poterba, Warshawsky and Brown [1999].)

and Poterba [2000], and Mitchell et al. [1999].) The benefits of additional annuitization are decreased if substantial existing wealth is already annuitized. A second factor reducing the benefits to annuitization is the ability of families to pool their mortality risk. For example, Kotlikoff and Spivak [1981] demonstrate that a couple can capture close to half the potential benefits associated with annuitization without actually purchasing any annuities. While both of these factors could substantially reduce the expected benefits to annuitization, they do not eliminate the benefit. Unless some costs associated with annuitization are introduced, these explanations merely reduce the available gains without altering the optimality of full annuitization.

Other authors tackle the annuity puzzle from the cost perspective. In Yaari's model, there are no downsides to annuitization. With only benefits and no costs, the optimal answer is full annuitization. First, some authors have noted the original Yaari work assumes actuarially fair annuity prices. If prices reflect insurance costs and the potential for adverse selection, then depending on the pricing structure, full annuitization may no longer be optimal (e.g., Friedman and Warshawsky [1988] and Mitchell, Poterba, Warshawksy and Brown [1999]). Another critical assumption made in Yaari's analysis was the absence of bequest motives. However, every dollar spent on annuities implies one less dollar available for a bequest. If individuals have a bequest motive, then full annuitization may again no longer be optimal. For example, Walliser [1999] calculates that, with reasonable assumptions for risk aversion and bequest motives, it is optimal to annuitize only 60% of wealth at age 65.

Demand for liquidity creates another implicit cost consideration. If future expenses are uncertain, then individuals would prefer to have a pool of liquid assets to insure against future needs. Since annuity purchases are largely irreversible, every dollar spent on an annuity is one less dollar available for unknown future expenses. As an example, some authors have pointed to expenses related to health uncertainty as a potential detriment to annuity demand (e.g., Brown and Warshawsky [2004] or Brugiavini [1993]).

Finally, a growing literature has focused on a market imperfection associated with many annuity contracts in the United States. These annuity contracts are not indexed to equity markets and thus preclude access to any available equity risk premium. Given this restriction, individuals may wish to reduce their annuity holdings or perhaps eliminate them altogether depending on their risk tolerance and the equity premium relative to annuities (e.g., Kapur and Orszag [1999]; Stabile [2003]; Milevsky and Young [2003]; Dushi and Webb [2004]; Dus, Maurer and Mitchell [2005]; Kingston and Thorp [2005]; Horneff, Maurer, Mitchell, and Dus [2006]; and Koijen, Nijman and Werker [2006]). Similar to bequests and liquidity, every incremental dollar annuitized carries an additional cost component. In this case, the cost stems from having one less dollar available to invest in the equity markets.

Once costs are introduced, the fundamental problem of optimal annuitization changes dramatically. First, since every dollar annuitized incurs additional costs, the optimal level of annuitization is generally a fraction of total wealth. Second, costs imply individuals are concerned with annuity efficiency. Annuity contracts that can deliver larger benefits per dollar invested (i.e., are more efficient) are always more desirable since they can provide equal benefits at lower costs compared to less efficient annuity contracts. Surprisingly, while the literature has explored costs associated with annuity purchases, no research has been done on annuity efficiency. The almost universal assumption has been that "buying an annuity" refers to purchasing an immediate annuity whose payments begin immediately and are guaranteed for as long as the individual lives.

This paper demonstrates that annuity contracts differ greatly in their efficiency. We first analyze optimal annuity purchases assuming access to flexible state-contingent annuity contracts. Given this framework, we demonstrate that the optimal allocation of annuity resources implies annuitized assets fund consumption late in retirement and non-annuitized assets fund consumption early in retirement. Since the optimal allocation of wealth involves a separation between non-annuity and annuity-funded years, immediate annuities are generally only optimal when costs are low enough to allow full annuitization. The size of the inefficiency of immediate annuities is surprising. For a typical example, we find that six percent of wealth efficiently allocated to state-contingent annuities can provide up to half of the available benefits from mortality risk sharing. To receive a similar level of welfare, immediate annuities would require an allocation of thirty-nine percent of wealth.

Unfortunately, state-contingent annuity contracts are a theoretical construct not currently available in the insurance market. However, we demonstrate that recently introduced delayed payout annuity contracts can provide all, or substantially all, of the benefits associated with full access to state-contingent annuity contracts. Because delayed payout annuities are much more efficient than immediate annuities, they should be highly desirable to any individual facing annuitization costs from bequests, liquidity concerns or asset allocation constraints. It will be interesting to see if these new annuity products are successful in the market. Theoretically, they should be strongly preferred to traditional immediate annuities. Since they have the potential to provide a majority of the annuity benefits with a modest five to ten percent annuity allocation, the usual cost-based explanations for low annuity demand (bequest motives, liquidity concerns, or market imperfections) are unlikely to apply to delayed payout annuities. A lack of market success for delayed payout annuities would thus create a new annuity puzzle.

The rest of this paper is organized as follows. Section 2 sets up the canonical problem of an individual entering retirement with a pool of wealth and an objective to maximize lifecycle utility through the purchase of either bonds or annuities. Section 3 explores optimal consumption assuming individuals are unconstrained in their allowed allocation of wealth to annuities and have access to flexible state-contingent annuities. Section 4 introduces the concept of annuity costs and considers the problem of optimal partial annuitization. Section 5 analyzes the efficiency of different annuity products at delivering the benefits of mortality risk sharing. In addition to a baseline analysis, the efficiency analysis considers two extensions. First, the situation where the annuity prices are actuarially unfair is analyzed. Second, the case where Social Security is a large fraction of total wealth is considered. Section 6 relates the current analysis to the existing literature, and Section 7 concludes by summarizing the key findings including normative suggestions for retirees struggling with longevity considerations.

## 2 Lifecycle Utility Problem Definition

Consider a retired individual that has an accumulated amount of wealth available to support retirement consumption. This individual has two flavors of investment options available. The first is zero coupon bonds. The prices for these zero coupon bonds are given by $B_t$ where:

$B_t$ = price today for \$1 payout in year $t$

Annuity products are also available. For this analysis, we consider a hypothetical set of simple contingent claim annuities that pay out \$1 $t$-periods into the future, provided the individual is alive at that time. The price for each of these annuities is given by $A_t$ where:

$A_t$ = price today for \$1 payout in year $t$, conditional on survival to period $t$

These types of annuities are sometimes referred to as "zero coupon" annuities due to their similarity to zero coupon bonds or as "Arrow" annuities due to their state-dependent payout structure. We will adopt the Arrow annuity terminology to refer to these contingent claim securities.

For simplicity, assume our individual has no bequest motive and wishes to allocate his wealth so as to maximize expected utility. We also assume at this point that there is no inflation, so dollar payouts at any point in time are equivalent to consumption units. Finally, assume that expected utility is additively separable in time, and is given by the expression below:

$$\text{(1a)} \qquad \Omega = \max\left\{\sum_{t=0}^{\infty} \Pi_t \cdot \Delta_t \cdot U(c_t)\right\}$$

In Eq.(1a), $\Pi_t$ is the probability that the individual is alive at period $t$ (as perceived by the individual planning at time 0), $\Delta_t$ is the individual's discount factor for utility at period $t$, $U$ is the undiscounted utility of consumption, assumed to have the same functional form for all periods, and $c_t$ is the consumption at period $t$. We assume that the discount factors

$\Delta_t$ and the probabilities $\Pi_t$ are less than or equal to one, with $\Delta_0 = \Pi_0 = 1$, and are strictly decreasing in $t$.

Let $x_t$ be the amount of consumption from bonds, and $y_t$ be the amount of consumption from annuities. The total consumption at time $t$, $c_t$, is given by:

$$(1b) \quad c_t = x_t + y_t .$$

Further, let $V_B$ equal the present value of all bonds purchased, and $V_A$ equal the present value of all annuities purchased. We then have:

$$(1c) \quad V_B = \sum_{t=0}^{\infty} B_t \cdot x_t ,$$
$$(1d) \quad V_A = \sum_{t=0}^{\infty} A_t \cdot y_t .$$

The present value of all purchases must equal the individual's current wealth, $W_0$. This is the budget constraint:

$$(1e) \quad V_B + V_A = W_0 ,$$

Equations (1) are a mathematical program, but the solutions are unbounded. The economics of the situation are apparent from the prices of the bonds and annuities. Whenever the cost of an annuity payout dollar in period $t$, $A_t$, does not equal the cost of a bond payout dollar in period $t$, $B_t$, then there is an opportunity for an individual to achieve infinite utility. For example, if $B_t > A_t$, then an individual could short sell $B_t$ and purchase $A_t$ in infinite quantities to maximize utility. To eliminate this possibility, we rule out short sales from the analysis: i.e.,

$$(1f) \quad x_t , y_t \geq 0.$$

These appear to be reasonable assumptions. First, there is no market where an individual can sell an arbitrary annuity, thus implying a non-negative $y_t$. Second, we would not expect a lender to allow borrowing at the bond rate using an annuity as collateral.[2]

**3 Unconstrained Optimal Consumption**

Solving the mathematical programming problem described by Eqs(1) yields the optimal consumption path $c_t$. We can reduce the optimization to an algebraic problem by writing down the Karush-Kuhn-Tucker conditions for optimality (Bazaraa and Shetty [1979]). Our results assume that the investor's utility function $U(c)$ is monotonically increasing and concave with marginal utility approaching infinity as $c$ approaches zero; i.e., our investor is risk-averse. This assumption leads to paths for which there is positive consumption in any period that the individual thinks they have a non-zero probability of being alive. We also assume that $U(c)$ is twice differentiable, and denote derivatives with respect to consumption by a prime. Since $U(c)$ is increasing and concave, we have $U'(c) > 0$ and $U''(c) < 0$ for all $c$. It then follows that the function $U'(c)$ has a unique inverse; we will put this inverse function to good use to calculate optimal consumption paths. Further, we will assume that annuity prices are such that they can be represented as a monotonically decreasing function, $P_t$, multiplied by the contemporaneous bond price. $P_t$ is the pricing discount received for agreeing to the annuity contract and can be viewed as the survivor function that prices the Arrow annuities. If this function equals the individual survivor function, $\Pi_t$, then the annuities are said to be 'actuarially fair'. Later

[2] Yaari [1965] and Davidoff, Brown, and Diamond [2005] make less restrictive assumptions about the ability to borrow. Loosening this constraint needlessly complicates our analysis without changing the fundamental implications when annuity purchases are constrained.

we will argue that monotonically decreasing $P_t$ is an extremely weak assumption on annuity prices.

To summarize, the initial assumptions for optimization are as follows:

| | |
|---|---|
| $U(c)$ | Twice differentiable |
| $U'(c) > 0$ | Monotonically increasing |
| $U'(c)\to\infty$ as $c\to 0$ | Infinite marginal utility at zero consumption |
| $U''(c) < 0$ | Concave |
| $A_t = P_t \cdot B_t$ | Annuity pricing structure |
| $P_t < 1$ | Annuity price always a discount relative to bonds |
| $P_t > P_{t+1}$ | Annuity discount function monotonically decreasing in time |

Optimal consumption at period t is given by the formula:

$$(2a)\quad \Pi_t \cdot \Delta_t \cdot U'(c_t) = \min\left(B_t, A_t\right) \cdot U'(c_0).$$

The intuition for this equation is straightforward. The right hand side of the equation corresponds to the utility cost of purchasing an incremental amount of consumption in period *t*. The left hand side of the equation represents the marginal increase in utility from extra consumption in period *t*. Optimality requires a consumption plan that equates these two quantities.

Given the assumed pricing structure for the Arrow annuities, $A_t < B_t$ for all $t \geq 0$.[3] We recover the standard result that all consumption is funded by annuity purchases ($x_t = 0$ and $c_t = y_t$). Eq. (2a) can be simplified to the following:

$$(2b) \quad \frac{U'(c_t)}{U'(c_0)} = \frac{P_t \cdot B_t}{\Pi_t \cdot \Delta_t}.$$

Since $U'(c)$ has a well-defined inverse, we can always solve Eq.(2b) for the consumption $c_t$, given the initial consumption $c_0$ and values for $\Delta$. The value of $c_0$ is then determined by imposing the budget constraint. Eq. (2b) also illustrates the well-known result that when annuity pricing is actuarially fair ($P_t = \Pi_t$) and the bond interest rate equals the future utility discount rate ($B_t = \Delta_t$), then consumption is constant across time.

## 4 Optimal Bond and Annuity Allocation with Limited Annuitization

The model described in the preceding section predicts full annuitization, because annuities are a strictly cheaper way to fund consumption in every time period. However, there are several opportunity costs associated with increasing levels of annuitization. These include fewer dollars available for bequests, fewer dollars available to meet unexpected liquidity demands, and more dollars in a restricted asset allocation required by the annuity contract. These factors can each lead to partial annuitization, as a retiree would trade off the marginal benefit of an extra annuity dollar against the foregone opportunities that a liquid (but more expensive) bond investment could satisfy.

[3] Assuming period zero corresponds to immediate consumption, then $A_0 = B_0 = 1$. However, this creates an indifference between bond and annuity based consumption at period 0. We adopt the convention that $A_0 = 1-\varepsilon$ to break ties. This assumption is immaterial to the overall results, but has the advantage of retaining optimality for 100% annuitization (as would be the case in continuous time models). In effect, period zero consumption occurs in the very near future with an infinitesimal chance of death over that time period.

Rather than pose a particular functional form for one of these sources of annuity opportunity costs, we analyze the problem of optimal partial annuitization by introducing the following constraint[4]:

$$(1g) \quad V_A \le \alpha W_0 .$$

This constrained model's first-order conditions involve a new tradeoff: the marginal utility of an extra dollar of annuity-funded consumption is balanced against the marginal cost imposed by the aggregate wealth constraint ($V_A + V_B = W_0$) plus the marginal cost of using one more dollar of the annuity budget constraint ($V_A \le \alpha W_0$). The annuity budget constraint's shadow price (its Lagrange multiplier) can be interpreted as representing the multiple sources of opportunity costs of annuitization discussed above. Whatever the exact source of these opportunity costs, it is necessary that any optimal solution for models that lead to partial annuitization use *efficient* allocations of annuities, where "efficient" in this context means that the limited annuity dollars provide the most expected utility.

We now turn to the problem of efficiently allocating annuity wealth to maximize expected utility given an available annuity budget. Constraining the amount of annuities that can be purchased forces the individual to split their wealth between bonds and annuities. The central question becomes what annuities are purchased when annuity dollars are scarce? To begin to answer that question, we first introduce the Bond - Annuity Separation Theorem.

**4.1 Bond - Annuity Separation Theorem:**

[4] Davidoff, Brown, and Diamond [2005] use a model with an equivalent constraint to describe the welfare gain from allowing annuities, moving from $\alpha$=0 to $\alpha$=1.

*Claim:* If the annuity discount function, $P_t$, decreases in time, then it is never optimal for annuitized payouts to precede non-annuitized portfolio payouts.

*Proof:* First, assume an optimal solution exists that consists of purchasing bond payouts for time $t+n$ ($x_{t+n} > 0$), $n$ periods after an annuity payout ($y_t > 0$). Next, construct an alternative solution derived by:

a) Reducing the consumption of period $t+n$ bonds by a small amount (ε)

b) Purchasing a similar amount of consumption in period $t+n$ via annuities

c) Reducing the consumption of period $t$ annuities by an amount that offsets the purchase of annuities in (b) and keeps the aggregate allocation to annuities constant.

d) Purchase enough consumption in period $t$ via bonds to offset the reduction in consumption from (c)

Denoting the alternate solution with an asterisk (*), the proposed alternate solution deviates from the presumed optimum by the following trades:

$$(3a) \quad x_{t+n}^{*} = x_{t+n} - \varepsilon$$

$$(3b) \quad y_{t+n}^{*} = y_{t+n} + \varepsilon$$

$$(3c) \quad y_{t}^{*} = y_{t} - \varepsilon \cdot \frac{A_{t+n}}{A_t}$$

$$(3d) \quad x_{t}^{*} = x_{t} + \varepsilon \cdot \frac{A_{t+n}}{A_t}$$

To assess whether the alternate strategy described by (3a) – (3d) is feasible, we must verify that none of the constraints are violated. First, the alternate complies with the no short sales constraints since the only reductions occur for $y_t$ and $x_{t+n}$, both of which are positive by assumption. Since ε can be arbitrarily small, there exists an ε small enough so the alternate strategy complies with the no short sales constraint. The next constraint to verify is the cap on aggregate wealth allocated to annuities. The aggregate allocation of wealth to annuities is unchanged from the assumed optimal solution and the proposed alternate solution (an amount of annuity wealth equal to $\varepsilon \cdot A_{t+n}$ has been transferred from period $t$ to period $t+n$).

Finally, we need to examine the budget constraint to determine if the alternate solution is feasible. Adding up the net costs of deviating from the presumed optimum yields the following impact on the budget constraint:

$$\text{Net Costs} = \varepsilon \cdot A_{t+n} / A_t \cdot (B_t - A_t) - \varepsilon \cdot (B_{t+n} - A_{t+n})$$

It suffices to show the net costs are negative to demonstrate that the alternate solution has higher expected utility than the assumed optimal solution. The alternate solution would then be one that has equal consumption as the assumed optimal, but with positive wealth available for additional consumption. The net costs are negative when:

(4a) $\quad \varepsilon \cdot A_{t+n} / A_t \cdot (B_t - A_t) - \varepsilon \cdot (B_{t+n} - A_{t+n}) < 0$

(4b) $\quad \Rightarrow A_{t+n} / B_{t+n} < A_t / B_t$

(4c) $\quad \Rightarrow P_{t+n} < P_t$

Condition (4c) is exactly the condition posited in the theorem, thus the claim is proven.

The separation theorem prohibits annuity-funded consumption from preceding bond-funded consumption in optimal solutions. This result is extremely useful in evaluating situations with constraints on the annuity allocation. All optimal plans have consumption first funded by the purchase of bonds and then later consumption funded via the purchase of annuities.[5]

The separation theorem relied upon the declining ratio of annuity to bond prices. But how reasonable is this assumption? First, consider this assumption in the case of actuarially fair annuity pricing. In this case, the ratio of annuity to bond prices is given by:

(5) $P_t = A_t \,/\, B_t = \Pi_t \cdot B_t \,/\, B_t = \Pi_t$

Since survival probabilities are strictly decreasing over time, the separation theorem holds given this pricing structure. A very weak condition for annuity prices would be the assumption that $A_t < B_t$ for all $t$. If this condition did not hold for a given $t$, then the bond in that time period would in essence dominate the annuity. For equal or less money the bond would provide the same consumption without requiring survival to collect. Rearranging the terms in Eq (4c) indicates that the condition for separation is similar. For arbitrary annuity pricing, the separation theorem requires:

(6) $A_{t+n} < A_t \cdot ( B_{t+n} \,/\, B_t )$ for all $n$

Eq (6) illustrates two different ways of purchasing consumption in period $t+n$. The left-hand side of Eq (6) purchases consumption in period $t+n$ using the $t+n$ Arrow annuity directly. The right-hand side of Eq (6) purchases $t+n$ consumption by first

[5] For the discrete model we analyze, there is the potential for bonds and annuities to fund consumption in a unique transition year. The separation theorem still holds since prior to the transition year, bonds fund consumption and afterward annuities fund consumption.

purchasing a period $t$ annuity and then investing the proceeds in the bond market between $t$ and $t+n$. The separation theorem requires that purchasing the annuity directly is the cheaper alternative. Similar to the argument that $A_t < B_t$ for all $t$, if Eq (6) does not hold then the period $t+n$ annuity is effectively dominated by a strategy that provides the same or more consumption with a weaker requirement on individual survival. Thus, it strikes the authors as very likely for Eq (6) to hold in any functioning annuity market.

We should note that while the prior discussion assumes that riskless bonds underlie the annuity contract, the separation theorem extends to annuity contracts with payouts contingent on stochastic market returns. As demonstrated in Appendix A, as long as the annuity discount rate relative to the price of the underlying security is declining, then it is never optimal to have annuity-based consumption precede the consumption from the underlying security alone. It is the pricing and not the stochastic nature of the returns that generates the result.

**4.2 Constrained Annuitization and Optimal Consumption**

We now return to the constrained optimization problem. The equation that governs optimal consumption in the constrained case requires the introduction of an additional Lagrange multiplier to account for the annuity constraint. In particular the equation is:

$$(2a') \quad \Pi_t \cdot \Delta_t \cdot U'(c_t) = \min\left(B_t, {}^{A_t}\!/_{\lambda}\right) \cdot U'(c_0).$$

The new Lagrange multiplier is λ. It takes a value of one for the unconstrained case. However, for the constrained case, the value of λ is bounded between zero and one. For a given value of λ, one can see how the separation theorem is maintained. Given $P_t$ is decreasing in time and given λ between zero and one, bonds are initially purchased until $P_t = \lambda$. After that point, annuities fund consumption.

In Section 5 we explore the efficiency with which various annuity bundles provide benefits compared to full access to Arrow annuities. A consideration in the efficiency analysis is how closely can the annuity bundle mimic the desired consumption path optimally selected with Arrow annuities? As we will see, increasing or constant annuity-based consumption paths will generally result in better performance for a given annuity product space. When Arrow annuities are available, the time path for consumption is defined by Eq. $(2a')$. Rearranging terms yields:

$$(2b') \quad U'(c_t) = \left(\frac{U'(c_0)}{\lambda}\right) \cdot \left(\frac{P_t}{\Pi_t}\right) \cdot \left(\frac{B_t}{\Delta_t}\right)$$

This equation is valid over the consumption region funded by annuity purchases.[6] Optimal allocations will equate the marginal utility of consumption in period t to the product of the three terms in Eq. $(2b')$. If the product of these terms is increasing over time, the consumption path will decrease due to the concavity of *U*. The first term is simply a positive scalar. It does not influence whether annuity-funded consumption is rising, falling or constant over time. The second term is the ratio of two discount factors. The first discount factor, $P_t$, measures the discount received, relative to bonds, for purchasing consumption via annuities. The second discount factor, $\Pi_t$, represents the individual discount factor for survival to period *t*. Actuarially fair annuity prices would equate these two discount factors, eliminating this term from influencing the time path of consumption. The third term is the ratio of the bond discount rate to the individual

[6] For the bond-funded region, the $P_t$ and λ terms drop from the above equation.

discount rate for future utility. To summarize, for time periods where annuities fund consumption,

$$\text{(7a) Annuity-funded consumption} \Rightarrow c_t \text{ is a decreasing function of } \left(\frac{P_t \cdot B_t}{\Pi_t \cdot \Delta_t}\right)$$

During the time frame where bonds are used to fund consumption (i.e. the early years), $\min(B_t, A_t / \lambda) = B_t$. During these years,

$$\text{(7b) Bond-funded consumption} \Rightarrow c_t \text{ is a decreasing function of } \left(\frac{B_t}{\Pi_t \cdot \Delta_t}\right)$$

We will refer back to these relationships when discussing whether a particular annuity product space is capable of replicating the desired time path of consumption.

## 5 Efficiency Loss: What if Arrow Annuities are not Available?

The previous analysis hypothesized the existence of Arrow annuities, which facilitated an arbitrary annuity-based consumption pattern. In reality, Arrow annuities are not available for purchase. Instead, the only available annuity contracts are those that involve bundles of Arrow annuities. The following analysis assesses the efficiency loss associated with annuity bundling.

### 5.1 Annuity Product Spaces

We consider five different annuity product spaces that allow increasingly flexible annuity-based consumption patterns. Throughout the analysis, we assume that annuity bundles do not carry discounts or premiums relative to the cost of purchasing the given pattern of annuity payouts via Arrow annuities.

**Annuity Product Space #0: No Annuities / Bonds Only**

This is the most restrictive product space. No annuities are available for purchase in any form. This optimization provides a helpful baseline to compare results from the other situations. We follow the convention in the literature of reporting an “Annuity Equivalent Wealth” or AEW. (See, for example, Brown and Warshawsky [2004], Brown and Poterba [2000] or Brown, Mitchell and Poterba [1999] for additional background on AEW.) Throughout this analysis, the AEW answers the question: “How much aggregate wealth is required for an individual with no access to annuities to be indifferent to a situation with $100 in wealth and the given (potentially constrained) access to annuities?”[7] To define AEW mathematically, let

$\Omega_i(w,\alpha) \equiv$ Max utility given wealth w, $\alpha$ percent annuities and product space i.

Then AEW is always defined by the following equation:

$$\Omega_0(AEW,0) = \Omega_i(\$100,\alpha) \qquad (8)$$

For example, suppose we are considering the case where a person has access to Arrow annuities and they are willing to allocate 10% of their assets to annuities. Suppose further that the reported AEW was $142. This would indicate that optimizing assuming $100 in wealth, access to Arrow annuities and a 10% constraint on annuity purchases provides the same expected utility as optimizing with $142 and no access to annuities.[8]

**Annuity Product Space #1: Immediate Annuities**

[7] In general, the AEW is a function of wealth. However, in all situations where we report AEW, we also assume CRRA utility. Given CRRA utility, the AEW ratio is wealth independent. We adopt a $100 wealth level for ease of exposition.

[8] The literature generally assumes the right hand side of Eq. (8) corresponds to full annuitization. Here, the definition of AEW is modified to be consistent with our focus on partial annuitization.

This is also a fairly limited annuity product space. There is only a single annuity available for purchase. This annuity is a bundle of all the Arrow annuities and it pays $1 in all periods where the annuitant is alive. An immediate annuity is assumed to have the following characteristics:

Payout structure = $1 payout in every year $t$, conditional on survival to period $t$

Price = $\sum_{t=0}^{\infty} A_t$

Given this annuity product space, the only annuity-funded consumption pattern available is constant consumption. To achieve any other consumption pattern requires the use of bonds to fill in the differences between desired and annuity-funded consumption.

**Annuity Product Space #2: Delayed Purchase Annuities**

Several authors have investigated the potential of delaying the purchase of an immediate annuity. (See, for example, Kapur and Orszag [1999], Dushi and Webb [2004], Milevsky and Young [2003], Dus, Maurer and Mitchell [2005] and Horneff, Maurer, Mitchell, and Dus [2006] who also provide a nice summary of this literature.) Our current framework imagines all decisions as occurring at time 0. To assess the potential benefit from delaying the immediate annuity purchase, we introduce $t$ different annuity products that replicate the cost and consumption pattern achieved by delaying the purchase of an annuity.

Before describing the payout and pricing structure for a delayed purchase annuity, let us first consider the assumed pricing for a delayed purchase Arrow annuity. Suppose our retiree decides to invest in a bond for $t$ periods and then purchases an Arrow annuity that pays out in $t+n$ periods. This strategy would net the retiree \$1 in period $t+n$ at a price equal to $B_t \cdot (A_{t+n} / A_t)$.[9] The period $t$ delayed purchase annuity would then have the following characteristics:

Payout structure = \$1 payout in every year $t+n$, given survival to $t+n$, $n \geq 0$.

$$\text{Price} = \left(\frac{B_t}{A_t}\right) \cdot \sum_{n=0}^{\infty} A_{t+n}$$

If the desired annuity-funded consumption pattern is increasing in time, then these annuities provide a more flexible product space to meet the demand.

Relative to immediate annuities, delayed purchase annuities could provide two additional benefits. First, delayed purchase annuities can provide a more flexible annuity-based consumption pattern. Second, they potentially provide a more efficient way to allocate scarce annuity dollars. This is in contrast to the literature where the equity risk premium generally plays the major role in driving the demand for delays. In that literature, annuitizing assets precludes capturing the equity risk premium, so generally investors delay some annuity purchases until the annuity premium provided by mortality risk sharing equals the equity risk premium. Our focus is not how the underlying assets are invested (e.g., bonds or stocks), but rather how the different annuity contract structures are able to deliver efficient longevity protection to retirees unwilling to fully annuitize.

---

[9] This assumes that an Arrow annuity purchased at period t for payout at period $t+n$ carries a price of $A_{t+n}$ / $A_t$. For example, actuarially fair annuity pricing and a flat term structure would conform to this assumption. More generally, any static term structure between period t and t+n is allowable.

### Annuity Product Space #3: Delayed Payout Annuities

Potentially one of the most important innovations in the annuity product space is the delayed payout annuity. In September 2004, MetLife introduced their Retirement Income Insurance℠ product that was described in a press release as:

> "an annuity that provides guaranteed lifetime income and is designed to generate income starting at a later age, for example an individual's 85th birthday, when other income sources may be running low." [MetLife, 2004]

This product is essentially a bundle of Arrow annuities that start at a future date instead of starting immediately.[10] These products generate the same payout stream as a delayed purchase annuity without the utility loss associated with bond investing prior to period *t*.

While Arrow annuities allow the option of arbitrary allocations to annuities over time, the separation theorem implies optimal annuity allocations are zero until a certain point in time and then support all consumption after that point in time. Given this optimal path of Arrow annuity demand, it is feasible that delayed payout annuities could provide all of the benefits of Arrow annuities even though they require bundled purchases. In fact, as long as the period-by-period growth in desired annuity-funded consumption exceeds the period-by-period growth in the payments from the delayed payout annuity, then delayed payout annuities provide all the benefits of Arrow annuities. We demonstrate this result more formally in Appendix B.

[10] In March 2006, The Hartford Financial Services Group introduced a similar delayed payout annuity product described as "a fixed payment annuity which guarantees lifetime income starting at a future point and continuing for the rest of one's life, no matter how long that might be."

This annuity product space allows a delayed payout annuity indexed to each time period $t$. The $t^{\text{th}}$ delayed payout annuity pays out starting $t$ years in the future and continuing until death. The period $t$ delayed payout annuity would then have the following characteristics:

Payout structure = \$1 payout in every year $t+n$, given survival to $t+n$, $n \geq 0$.

$$\text{Price} = \sum_{n=0}^{\infty} A_{t+n}$$

Notice how this product provides the same payouts as the delayed purchase annuity, but at a lower cost. Notice also, that if optimal consumption based on Arrow annuities is non-decreasing, then these delayed payout annuities can provide identical consumption at identical costs. It is in this sense that they can substitute for Arrow annuity availability.

**Annuity Product Space #4: Arrow Annuities**

This is the most flexible annuity product space considered. This annuity product space contains the annuity building blocks and thus can replicate or improve on any of the solutions based on the previous annuity product spaces.

The annuity product spaces are listed in order of increasing desirability. Each subsequent annuity product space either allows more flexibility or provides equal flexibility at a lower cost compared to the previous annuity product space. If $\text{AEW}_i$ refers to the AEW achieved given annuity product space $i$ is available, then for the above annuity product spaces the following holds:

$$\text{AEW}_i \geq \text{AEW}_j \text{ , } i > j$$

Appendix C details the additional optimization constraints required to solve the utility maximization given access to the various annuity product spaces.

### 5.2 Efficiency Analysis: Baseline, Actuarially Unfair Pricing and Social Security

To facilitate an efficiency analysis of the various annuity product spaces, it is necessary to make assumptions regarding the specific utility function and discounting parameters relevant to our retiree. The first assumption concerns the utility function, and is common across all of the cases investigated. In particular, we assume the utility function for our retiree is the following power utility function:

$$(9) \qquad U(c_t) = \frac{c_t^{1-\gamma} - 1}{1-\gamma}$$

Where gamma ($\gamma$) acts as the risk aversion parameter.

The second assumption we will make across all of the scenarios evaluated is that the discount rate for future utility, $\Delta_t$, is just offset by the bond interest rate, $B_t$. In addition, we assume that the annuity discount and bond interest rate is 3%. In other words, we assume:

$$(10) \quad \Delta_t = B_t = \left(\frac{1}{1.03}\right)^t \text{, for all } t$$

We now turn our attention to the evaluation of efficiency losses associated with various annuity product spaces in specific situations.

**Baseline Analysis:**

As a baseline analysis, we will assume annuity prices are actuarially fair. That is, we will assume that $P_t = \Pi_t$.[11] Given this assumption, we recover the standard result that optimal annuity-based consumption is constant over time. This can be readily seen in Eq (2b') where the marginal utility of consumption is constant over time. For the bond-funded years, optimal consumption is decreasing with time; decreasing survival probability leads one to purchase less bond-based consumption in later periods, even though the rate of time preference equals the bond rate of return. Evaluating (7b) with the baseline assumptions implies that the bond-funded consumption time path is a decreasing function of $1/\Pi_t$. Given a declining survival function, consumption during the bond-funded years is also declining. For these assumptions, optimal consumption decreases while it is funded by bonds and then stays constant once consumption is funded by annuities. Only annuity product spaces that can replicate this optimal path of declining bond-funded consumption followed by constant annuity-funded consumption can potentially provide equivalent benefits to Arrow annuities.

How does our retiree benefit from the introduction of various annuity product spaces? Figure 1 illustrates the AEW generated from allocating increasing amounts of annuity dollars utilizing the various annuity product spaces. For this example, the retiree is assumed to be male, age 65 with a risk aversion parameter ($\gamma$) equal to four. With these assumptions, the unconstrained AEW is \$154. Since the optimal unconstrained

[11] Generally, the GAM-94 mortality table is assumed to price the annuities. In contrast to SSA tables, GAM-94 caps mortality at 50% for the very aged. This cap creates implausible trends in the survival ratios for GAM vs. SSA populations. To eliminate this peculiarity from the analysis, we universally adopt the age 100+ mortality assumptions from the SSA with improvements from the SSA Actuarial Study #120. Given the extremely low probability of survival to these ages, this assumption is immaterial to the results.

consumption is constant over time, a retiree willing to invest all wealth in an annuity can achieve the maximum AEW with any of the annuity product spaces.

Figure 1 illustrates most of the key concepts in our analysis. First, since optimal annuity-based consumption is constant with Arrow annuities, delayed payout annuities provide equivalent benefits as Arrow annuities. The ability of delayed payout annuities to provide all or at least the vast majority of benefits achievable via Arrow annuities, the theoretically ideal annuity product set, is a pervasive result in our analysis. A second striking feature of Figure 1 is the inefficiency of immediate annuities. Delayed payout and Arrow annuities provide significant mortality risk sharing benefits from the first few dollars annuitized. Immediate annuities, in particular, and to a lesser extent delayed purchase annuities require substantially more wealth allocated to annuities to generate similar benefits.

Figure 1a explores the efficiency issue in more detail. Figure 1a reports essentially the same data as Figure 1; however, the y-axis is normalized to range from 0% ($100 AEW) to 100% ($154 AEW). The y-axis now corresponds to the percentage of potential annuity benefits delivered. Finally, Figure 1a only graphs the benefits associated with the first 50% of wealth allocated to annuities. The efficiency differences between the annuity product spaces are remarkable. Consider the amount of wealth required to achieve 50% of the total benefit available from annuitization, which in this case is about $27 in additional AEW. Arrow and delayed payout annuities require 6% of wealth allocated to annuities to deliver half the benefits of annuitization. With delayed purchase annuities, half the maximum benefit requires four times the allocation or 24% of wealth. Immediate annuities are even less efficient. They require our retiree to tie up

39% of wealth to achieve a similar level of benefits. A retiree who has a bequest motive or is simply uncertain about his future liquidity demands may find it very hard to commit almost 40% of their wealth to annuities. That same retiree may be willing to allocate 6% to a delayed payout annuity since it is such a small fraction of existing wealth and it generates a benefit equivalent to increasing consumption by 27%. If implicit costs such as bequest motives, liquidity concerns, or a limited annuity investment universe are core factors limiting annuity purchases, the advent of the delayed payout annuity could significantly impact annuity desirability and improve retiree welfare.

Table 1 reports information for scenarios that differ with respect to risk aversion and gender. In particular, data are reported for the following four scenarios:

Scenario #1: Male, 65, $\gamma = 4$, Figure 1 and Figure 1a Scenario

Scenario #2: Male, 65, $\gamma = 2$

Scenario #3: Male, 65, $\gamma = 1$

Scenario #4: Female, 65, $\gamma = 4$

For each scenario, results are reported assuming 5%, 10% and 20% of wealth are allocated to annuities. For each level of annuity allocation, a number of statistics are reported. First, the AEW for the various annuity product spaces is shown. Next, the age at which Arrow payouts begin is reported. For Arrow annuities, all consumption after this age is funded by annuities due to the separation theorem. The final two data elements reported correspond to the percentage of total benefit achieved by the Arrow annuity, and the level of immediate annuitization required to match the Arrow benefit.

The general patterns reported in Figure 1 and 1a are apparent in the other scenarios reported in Table 1. Arrow and delayed payout annuities are much more efficient compared to immediate or delayed purchase annuities. The maximum AEW benefit declines with decreasing levels of risk aversion. The maximum AEW assuming log utility ($\gamma$ equal to one) is \$134 compared to \$154 for $\gamma$ equal to four. Scenario #4 considers a 65 year-old female with a $\gamma$ parameter of four. This scenario highlights the impact of increasing life expectancy. Compared to a similarly risk-tolerant 65 year-old male, the female generates a lower AEW from access to annuity markets since lower mortality implies lower benefits from annuitization, *ceteris paribus*.[12] While the total benefit is reduced for individuals with lower mortality rates, the efficiency provided by Arrow annuities is increased. In this example, a five percent annuity allocation provides 47% of the total annuity benefit for the male, but 50% of the benefit for a female. In addition, the female would need to allocate 38% of wealth to immediate annuities to receive the same benefit, but the male would only need to allocate 36%.

**Private Mortality Analysis**

The analysis so far has assumed that the implied mortality table that the annuity provider uses to price the annuities is the same as the mortality table used by the individual to discount future utility, or in other words, annuity prices are actuarially fair. However, it is impossible for an insurance company to offer an actuarially fair annuity to everyone. Costs associated with offering annuities, capital requirements, and adverse selection issues all result in annuity prices that are less favorable compared to actuarially

[12] This may seem counter-intuitive: one might think that females have longer lives to insure, and thus should benefit more from the availability of annuities. One way to understand how longevity affects annuity demand is to consider an infinitely lived individual. Since survival is guaranteed, they would be indifferent between bonds and annuities. They receive no benefits from annuitization.

fair prices. This section analyzes the situation when the individual has private information about their mortality. In particular, the individual is assumed to have average life expectancy (represented by the Social Security population average mortality tables), but the annuity prices are assumed to be based on above-average life expectancy (represented by the Group Annuity Mortality or GAM tables). These two mortality tables have fairly significant differences. For example, the median life expectancy using the population average tables for a 65 year old male is a little above 17 years. Using the GAM 94 mortality tables, the median life expectancy increases to just over 20 years.

Figures 1, 1a and Table 1 report the AEW benefit for access to the various annuity product spaces assuming the same mortality table, GAM 94, is relevant for pricing and utility discounting. Figures 2, 2a and Table 2 report similar information assuming the individual discounts using population average mortality instead of GAM 94 mortality. Comparing Table 1 with Table 2 illustrates the impact of changing from actuarially fair pricing to private mortality inferior to pricing mortality. Before exploring the changes, first consider what has not changed. Annuities are still the cheapest method for attaining future consumption, thus full annuitization is still the optimal unconstrained solution. In fact, annuity prices have not changed at all, so the annuity prices still satisfy the requirements for the separation theorem.

However, several things have changed. First, Eq (2b) that governs optimal annuity-funded consumption is now:

$$(2b') \quad \frac{U'(c_t)}{U'(c_0)} = \frac{P_t \cdot B_t}{\Pi_t \cdot \Delta_t} = \frac{P_t}{\Pi_t}.$$

$P_t$ represents the pricing survival probability and $\Pi_t$ represents the private survival probability used to discount future consumption. Even with the assumption that the bond interest rate equals the individual discount rate, optimal annuity-funded consumption is no longer constant. The improvement in annuity prices is no longer sufficient to offset the discounting associated with survival. For the region funded by bonds, the consumption time path is still a decreasing function of $1/\Pi_t$. For the region funded by annuities, the time consumption time path is a decreasing function of $P_t / \Pi_t$. Given the mortality assumptions, $P_t / \Pi_t$ is always greater than one and is a decreasing function of $t$ until age 100. After 100, the mortality assumptions converge and the survival ratios stabilize. The concavity of $U$ thus implies optimal consumption is declining prior to age 100 and flat after age 100. Since optimal annuity-based consumption is decreasing given access to Arrow annuities, none of the annuity bundles are substitutes for full access to Arrow annuities even in the unconstrained case.

Turning to Figures 2, 2a and Table 2, the surprising result is just how little has changed especially over the region where relatively small amounts of wealth are annuitized. A reduction in life expectancy decreases the demand for future consumption. Since the benefit of annuities is cheap access to future consumption, the maximum annuity benefit has been reduced. While the maximum benefit has been reduced, the efficiency of delayed payout and Arrow annuities is relatively unchanged. The maximum benefit has been reduced from \$154 to \$149 for the baseline retiree. Figure 2 looks remarkably similar to Figure 1 in terms of the relative efficiency of the various annuity product spaces. As illustrated in Figure 2a, to get half the benefits available from mortality risk sharing, a retiree would need to allocate 6% to Arrow or delayed payout annuities, 24% to delayed purchase annuities, or 36% to immediate annuities. Recall that the analogous numbers for actuarially fair annuities were 6%, 24% and 39%, respectively. Similar to the retiree facing actuarially fair pricing, a retiree with private mortality rates markedly inferior to the pricing mortality can still garner a significant improvement in AEW ($124.5) from a fairly modest amount spent on annuities (6% of wealth). Essentially, the benefits from the first few dollars optimally spent on annuities are so significant that even relatively poor pricing does not substantially impact the desirability of annuitization.

A second striking feature of the private mortality analysis is the ability of delayed payout annuities to deliver benefits comparable to Arrow annuities. With these assumptions, optimal annuity-funded consumption declines until age 100 and then flattens. However, delayed payout annuities are not able to deliver declining consumption. Since delayed payout annuities can no longer replicate the optimal Arrow solution, the AEW for delayed payout annuities will be lower than the Arrow annuity AEW. While the AEW for delayed payout annuities is indeed lower, the degree of AEW loss is immaterial as illustrated in Figures 2, 2a and Table 2.

To better understand this result, Figure 2b compares the optimal consumption paths assuming a 20% allocation to either Arrow or delayed payout annuities. As expected, the optimal consumption path when Arrow annuities are available is declining while consumption is bond-funded, continues to decline after annuities start funding consumption and finally levels off after age 100. Over ages where Arrow annuities fund consumption, consumption declines from $5.81 to $5.03 representing about a 13.5% reduction in consumption. The consumption time path with delayed payout annuities is different. During the bond-funded region, the payout is virtually identical to the Arrow payout.[13] However, during the annuity funded region, the solutions diverge. The Arrow annuity solution continues to decline at a fairly similar rate to the bond-funded region. The delayed payout annuity has a substantial decline during the first year fully funded by the annuity, but is then required to be non-decreasing after that point. Examining the relative AEW for each case; however, clearly makes the point that the two consumption streams are very similar in terms of desirability. The AEW using Arrow annuities is $136.20 while the AEW using delayed payout annuities is $136.12. For these assumptions, access to cheap future consumption is much more important to delivering a high AEW than the ability to match a given consumption path.

**Social Security Analysis**

Table 1 reports the benefits of annuitization assuming all wealth currently available for investment is in either bonds or annuities. However, many individuals hold a significant portion of their retirement wealth in the form of Social Security benefits. Table 3 investigates the impact on the analysis of assuming 50% of wealth is already invested in an immediate annuity.[14] Having 50% of wealth pre-allocated to an immediate annuity will clearly reduce the benefit of additional annuitization. However, the basic result that Arrow and delayed payout annuities provide a significantly more efficient path to annuity benefits still holds.

---

[13] The delayed payout consumption is actually a fraction of a penny higher during the bond-funded years.

[14] While the fraction of wealth held by pre-retirees in the form of Social Security varies by wealth level, the median level is approximately 50% (Moore and Mitchell [2000]). Previous research focuses on a 50% allocation as indicative of the Social Security scenario (see, for example, Mitchell, Poterba, Warshawsky, and Brown [1999]).

For example, consider a 65 year-old male with a risk aversion parameter of four. Full annuitization has an AEW of $154. A 50% allocation of wealth to Social Security, which is modeled as an immediate annuity allocation, bestows a significant portion of the potential annuity benefit. The AEW for half Social Security wealth is $133. Even in this situation, the Arrow and delayed purchase annuities provide a very efficient method for improving AEW with relatively small additional allocations to annuities. If this person were to allocate an additional five percent of total wealth to Arrow annuities, the AEW improves from $133 to $143, a $10 increase in AEW. To get the same improvement in AEW from immediate annuities requires an incremental twenty percent of wealth invested in annuities.

Another interesting comparison between Arrow and immediate annuities is to consider the benefit from the first five percent allocated to annuities. From Table 1, the first five percent allocated to Arrow annuities yields a benefit to this individual of $25 if no wealth is previously allocated to immediate annuities. This amount drops to $10 if half of the wealth is allocated to immediate annuities via Social Security. For immediate annuities the comparable numbers are $4 and $3.5, respectively. Remarkably, allocating an additional five percent to Arrow or delayed purchase annuities with a large amount of Social Security wealth is still more than twice as efficient in generating AEW as immediate annuities without any Social Security wealth.

## 6 Literature Review and Implications

Our analysis suggests that delayed payout annuities could substantially improve retiree welfare. Importantly, virtually none of the annuity contracts analyzed in the literature or sold to individuals involve delayed payouts. This disparity provides an opportunity to extend the literature and hopefully improve retiree welfare. While delayed payout annuities are rare in the literature, we identify two papers that considered some form of delayed payout annuities. Both papers offer hints at the desirability of delayed payout annuities, but neither demonstrate or claim optimality for this form of annuity contract. We start our literature discussion with these papers, and then comment on implications of our analysis for related retirement literature.

Dus, Maurer, and Mitchell [2005] examine different phased withdrawal strategies including withdrawal strategies mixed with life annuities. They examine how a number of different withdrawal strategies perform on various non-utility based metrics such as expected present value (EPV) of a spending shortfall, EPV of benefit payouts, and EPV of bequests. Two of the annuitization strategies that are mixed with different withdrawal strategies included a delayed purchase strategy (called a "switching strategy") and a delayed payout strategy (called an "immediate purchase deferral strategy"). The authors note that "the deferred annuitization strategy is likely to be most attractive to those seeking to secure consumption while alive, without completely stripping their heirs of some unexpended funds." However, the reasons for the desirability of delayed payout are unclear. In fact, the authors indicate that "it is unclear what one might expect from these switching strategies, in terms of risk and rewards." Our analysis clarifies why the delayed payout annuity compares favorably to the delayed purchase annuity in all of the consumption and shortfall related measures.

Milevsky [2005] also considers an annuity product with delayed payout characteristics. Milevsky's paper "takes the approach that consumers will remain reluctant to annuitize a large lump sum at retirement." In terms of the annuity puzzle, Milevsky notes that "most people shun life annuities simply because they want to maintain control of their assets." The hypothesis for the main barrier to annuitization is that a "sudden irreversible transaction will never be popular." To help avoid the problem of a large annuitization event at retirement, Milevsky proposes that "slow annuitization over a very long period of time." To achieve this goal, a new insurance product is proposed that allows people to purchase future annuity payouts throughout their working years. There are two main advantages to this new annuity product. First, purchases are broken up into smaller transactions that occur over an individual's entire working career. Second, the price of securing a given annuity payout is extremely low. The low price stems from the sizeable discount for dollars forty or more years in the future and the potentially large mortality discount if payments begin at an age where survival is unlikely. Unfortunately, many barriers were identified to making this product a reality.[15]

We agree with Milevsky that asset control and irreversibility of annuity contracts likely play a large role in their lack of popularity. However, we would add one important caveat. Sudden, irreversible transactions are only a problem when they are a large fraction of wealth. If the transaction was a modest portion of total wealth then the significance of irreversibility is greatly reduced. It is still possible for existing retirees to attain a majority of annuity benefits with a modest allocation to a delayed payout annuity. While purchasing delayed payout annuities during the pre-retirement years may have some behavioral advantages, the benefit compared to a delayed payout annuity purchased at retirement is relatively small due to the high probability of survival until retirement. Spending a modest amount on a delayed payout annuity at retirement avoids the many

[15] Some of the barriers included: regulatory barriers to annuity payouts that start more than 30 years after purchase, lack of death benefit, and issues with inflation indexation for such a long horizon.

regulatory and administrative barriers Milevsky identified with his proposal while still maintaining the benefits of efficient annuitization.

In terms of the existing retirement literature, we identify three general classes of research that can either be extended or better understood in light of the results of our analysis. First, as identified in the introduction, there is a large and growing literature that considers annuity contracts in the context of portfolio construction and retirement consumption. This literature generally finds that delaying full annuitization is optimal. A market imperfection, namely the inability to access equity markets within an annuity contract, typically provides the fundamental incentive for partial annuitization.[16] Since this literature only considers immediate or delayed purchase annuities, expanding the analysis to included delayed payout annuities should significantly alter the results.

Another area of pension research explores the optimal point to initiate a given annuity payout. Our analysis also provides an intuition for the conclusions reached by this literature. For example, Coile, Diamond, Gruber and Jousten [2002] examine the question of the optimal delay in claiming the Social Security benefit. They find that "delaying benefit claim for a period of time after retirement is optimal in a wide variety of cases and that gains from delay may be significant." A critical feature of the analysis was the existence of non-annuitized wealth available to fund consumption prior to the onset of Social Security benefits. Viewed through the lens of our results, the optimality of delay becomes clearer. The Social Security benefit represents a fixed allocation to annuities. By delaying the onset of benefits, the individual uses annuity wealth to fund consumption later in life. Assuming non-annuity wealth is available to fund earlier

[16] Annuity pricing could also provide an incentive to delay. If the structure of annuity prices was such that costs exceed the mortality premium during the early periods, then delaying annuitization could be optimal.

consumption, the delay is utility-enhancing because the result is a more efficient allocation (i.e. more consistent with the Separation Theorem). Milevsky and Young [2003] also investigate a situation where individuals have an option for delay and come to similar conclusions with respect to the desirability of delaying the onset of payouts.

A third area of research closely related to the ‘optimal delay’ research is concerned with the public policy question of forced annuitization of retirement assets (for example, Milevsky and Young [2003] and Horneff, Maurer, Mitchell and Dus [2006]). If delaying annuitization is utility enhancing, then forcing annuitization too early results in utility loss. In essence, the authors are noting the utility loss associated with restricting retirees to immediate annuities when delayed purchase annuities were available (i.e., the loss associated with moving from annuity product space two to annuity product space one). Horneff, Maurer, Mitchell and Dus estimate that a female retiree with moderate risk aversion gains 11% of wealth with the option to delay annuitization compared to forced annuitization at age 65. While delayed purchase annuitization can be utility enhancing compared to immediate annuitization, additional benefits are available from delayed payout annuitization.

Our analysis suggests a policy change that would eliminate the identified utility loss associated with forced annuitization. Consider a policy that required assets of a retiree at a certain age to be used to purchase an annuity. However, this policy would allow the individual to select the date at which the annuity payouts begin. By providing individuals access to delayed payout annuities, the augmented policy would eliminate the utility loss identified by the prior research and even provide utility gains for many retirees relative to delaying the purchase of an immediate annuity.[17]

## 7 Conclusion

Since Yaari's [1965] seminal paper, other studies have identified several reasons (bequest motives, liquidity preferences, asset allocation constraints) why retirees may rationally choose to annuitize less than their full wealth. These barriers to full annuitization represent implicit costs to annuitization. Since these costs generally increase with the amount annuitized, optimizing agents have a preference for efficient annuities that deliver the maximum expected utility gain for a given annuity investment.

Using a standard lifecycle framework with flexible Arrow annuities as an analytical device, we determine several new insights. First, any optimal allocation of wealth between annuitized and non-annuitized assets involves a separation between non-annuity and annuity-funded consumption. Early consumption is based on non-annuitized assets, while annuities fund later consumption. The separation theorem in turn has implications for the relative efficiency of real-world annuity products. Because immediate annuities' constant payout streams violate the separation requirement, they are quite inefficient relative to Arrow annuities. We show that a recent innovation in annuity markets—delayed payout annuities—can effectively substitute for theoretical Arrow annuities in many settings, enabling much larger welfare gains than are achieved with similar allocations to immediate annuities. This ability of delayed payout annuities to virtually "complete the annuity market" in an Arrow-Debreu sense is due to the fact that delayed payout annuities obey the desired bond-annuity separation and allow typical

[17] This result assumes actuarially fair translation between payout start dates.

desired consumption streams to be matched. We further demonstrate that only small fractions (5-10%) of wealth are required for delayed payout annuities to achieve the majority of potential welfare gains from full annuitization. This stands in stark contrast to the need to invest large wealth fractions in immediate annuities in order to achieve similar welfare improvements.

Another potential obstacle to annuity markets is the impact of actuarially unfair pricing, relative to individual mortality. We analyze a model in which the retiree has a shorter life expectancy than implied by the pricing mortality table. These results show that delayed payout annuities are still strongly desirable; retirees facing actuarially unfair prices (relative to their own mortality probabilities) should not withdraw from the annuity market. In this setting, delayed payout annuities still effectively substitute for Arrow annuities, despite not being able to exactly replicate the desired consumption pattern.

Delayed payout annuities should realize higher demand than immediate annuities. If they do not, then the explanation is unlikely to be found in theoretical models that focus on implicit annuitization costs such as bequest motives, liquidity demands, or asset allocation constraints. Individuals unwilling to consider full annuitization should find delayed payout annuities a highly desirable way to get most of the annuity benefits without tying up a large fraction of their wealth in an irreversible transaction. Since the history of annuity markets has proven that most retirees are unwilling to annuitize large portions of wealth, these findings are important additions to the normative prescriptions of retirement economics.

## References


Bazaraa, Mokhtar, and C.M. Shetty. 1979. *Nonlinear Programming: Theory and Algorithms.* New York: John Wiley and Sons.

Brown, Jeffrey R., Olivia S. Mitchell, and James M. Poterba. 2001. "The Role of Real Annuities and Indexed Bonds in an Individual Accounts Retirement Program." In *Risk Aspects of Investment-Based Social Security Reform,* ed. John Y. Campbell and Martin Feldstein, 321-60. Chicago: University of Chicago Press.

Brown, Jeffrey R., and James M. Poterba. 2000. "Joint Life Annuities and Annuity Demand by Married Couples." *Journal of Risk and Insurance,* 67(4): 527-553.

Brown, Jeffrey R., and Mark J. Warshawsky. 2004. "Longevity-Insured Retirement Distributions From Pension Plans: Market and Regulatory Issues." In *Public Policies and Private Pensions*, ed. William G. Gale, John B. Shoven, and Mark J. Warshawsky, 332-69. Washington, DC: Brookings Institution Press.

Brugiavini, Agar. 1993. "Uncertainty Resolution and the Timing of Annuity Purchases." *Journal of Public Economics,* 50(1): 31-62.

Coile, Courtney, Peter Diamond, Jonathan Gruber, and Alain Jousten. 2002. "Delays in Claiming Social Security Benefits." *Journal of Public Economics,* 84(3): 357-85.

Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond. 2005. "Annuities and Individual Welfare." *American Economic Review,* 95(5): 1573-90.

Dus, Ivica, Raimond Maurer, and Olivia S. Mitchell. 2005. "Betting on Death and Capital Markets in Retirement: A Shortfall Risk Analysis of Life Annuities versus Phased Withdrawal Plans." *Financial Services Review,* 14: 169-96.

Dushi, Irena, and Anthony Webb. 2004. "Household Annuitization Decisions: Simulations and Empirical Analysis." *Journal of Pension Economics and Finance,* 3(2): 109-43.

Friedman, Benjamin M., and Mark J. Warshawsky. 1990. "The Cost of Annuities: Implications for Saving Behavior and Bequests." *Quarterly Journal of Economics,* 105(1): 135-54.

Horneff, Wolfram J., Raimond Maurer, Olivia S. Mitchell, and Ivica Dus. 2006. "Optimizing the Retirement Portfolio: Asset Allocation, Annuitization, and Risk Aversion." National Bureau of Economic Research Working Paper 12392.

Kapur, Sandeep, and J. Michael Orszag. 1999. "A Portfolio Approach to Investment and Annuitization During Retirement." http://www.ems.bbk.ac.uk/faculty/kapur/personal/annuity.pdf.

Kingston, Geoffrey, and Susan Thorp. 2005. "Annuitization and Asset Allocation with HARA Utility." *Journal of Pension Economics and Finance,* 4(3): 225-48.

Koijen, Ralph S.J., Theo E. Nijman, and Bas J.M. Werker. 2006. "Optimal Portfolio Choice with Annuitization." Center for Economic Research Discussion Paper No. 78.

Kotlikoff, Laurence J., and Avia Spivak. 1981. "The Family as an Incomplete Annuities Market." *Journal of Political Economy,* 89(2): 373-91.

MetLife 2004 Press Releases. 2006. MetLife, Inc. http://www.metlife.com/Applications/Corporate/WPS/CDA/PageGenerator/0,4132,P7328~S583,00.html?FILTERNAME=@URL\&FILTERVALUE=/WPS/ (accessed November 8, 2006).

Milevsky, Moshe A. 2005. "Real Longevity Insurance with a Deductible: Introduction to Advanced-Life Delayed Annuities (ALDA)." *North American Actuarial Journal,* 9(4): 109-22.

Milevsky, Moshe A., and Virginia R. Young. 2003. "Annuitization and Asset Allocation." http://www.ifid.ca/pdf_workingpapers/WP2003OCT15.pdf.

Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown. 1999. "New Evidence on the Money's Worth of Individual Annuities." *American Economic Review,* 89(5): 1299-318.

Moore, James F., and Olivia S. Mitchell. 2000. "Projected Retirement Wealth and Saving Adequacy." In *Forecasting Retirement Needs and Retirement Wealth*, ed. Olivia S. Mitchell, P. Brett Hammond, and Anna M. Rappaport, 68-94. Philadelphia: University of Pennsylvania Press.

Stabile, Gabriele. 2003. "Optimal Timing of the Annuity Purchases: A Combined Stochastic Control and Optimal Stopping Problem." http://amases2002.univr.it/amases/abstracts/Stabile.pdf.

Walliser, Jan. 1999. "Regulation of Withdrawals in Individual Account Systems." International Monetary Fund Working Papers 99/153.

Yaari, Menahem E. 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer." *Review of Economic Studies,* 32(2): 137-50.

**Figure 1**
**Annuity Equivalent Wealth by Annuity Product Space**
**Gamma = 4, Age = 65, Male**

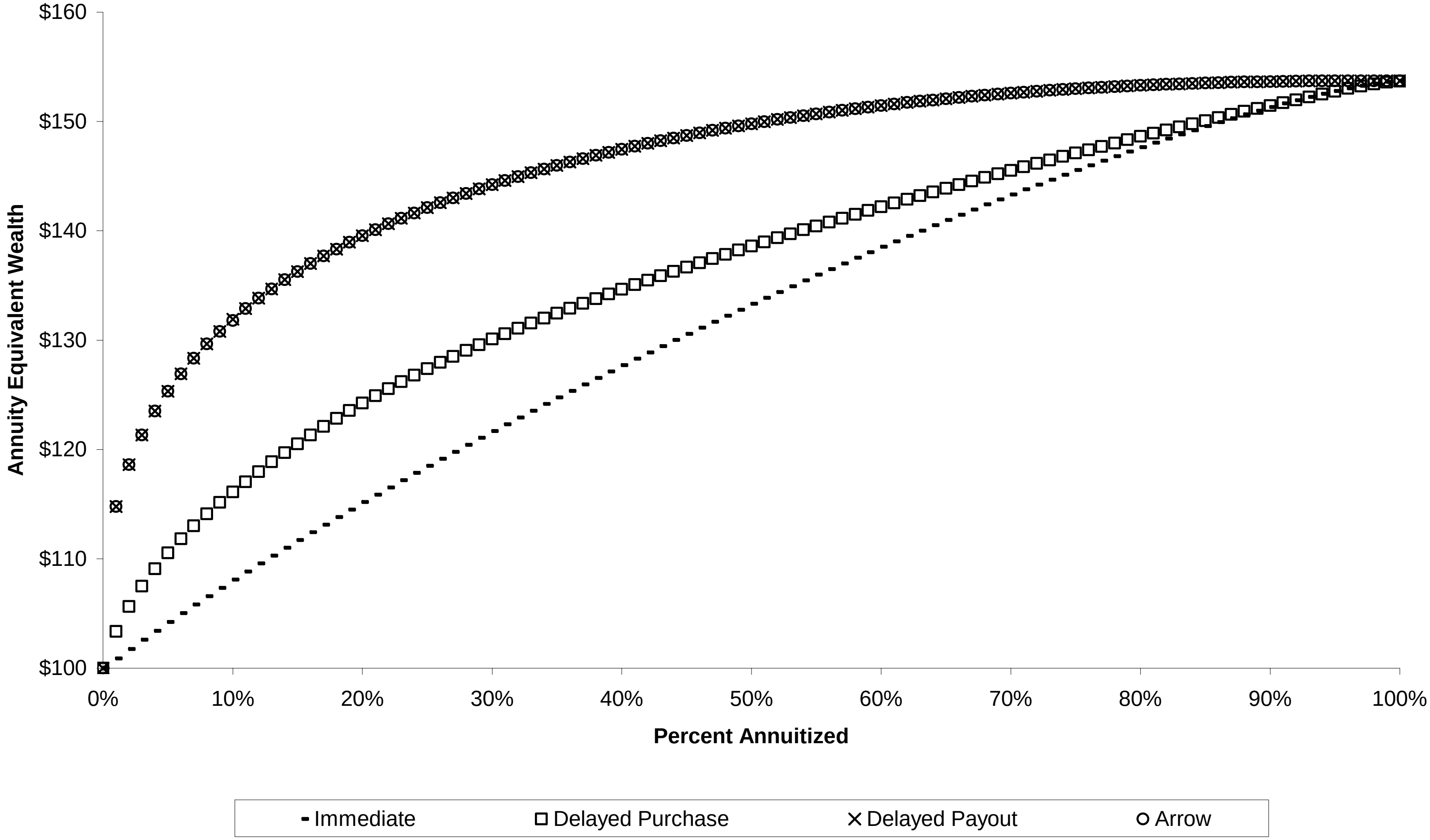


**Figure 1A**
**Efficiency Analysis of Annuity Product Spaces**
**Gamma = 4, Age = 65, Male**

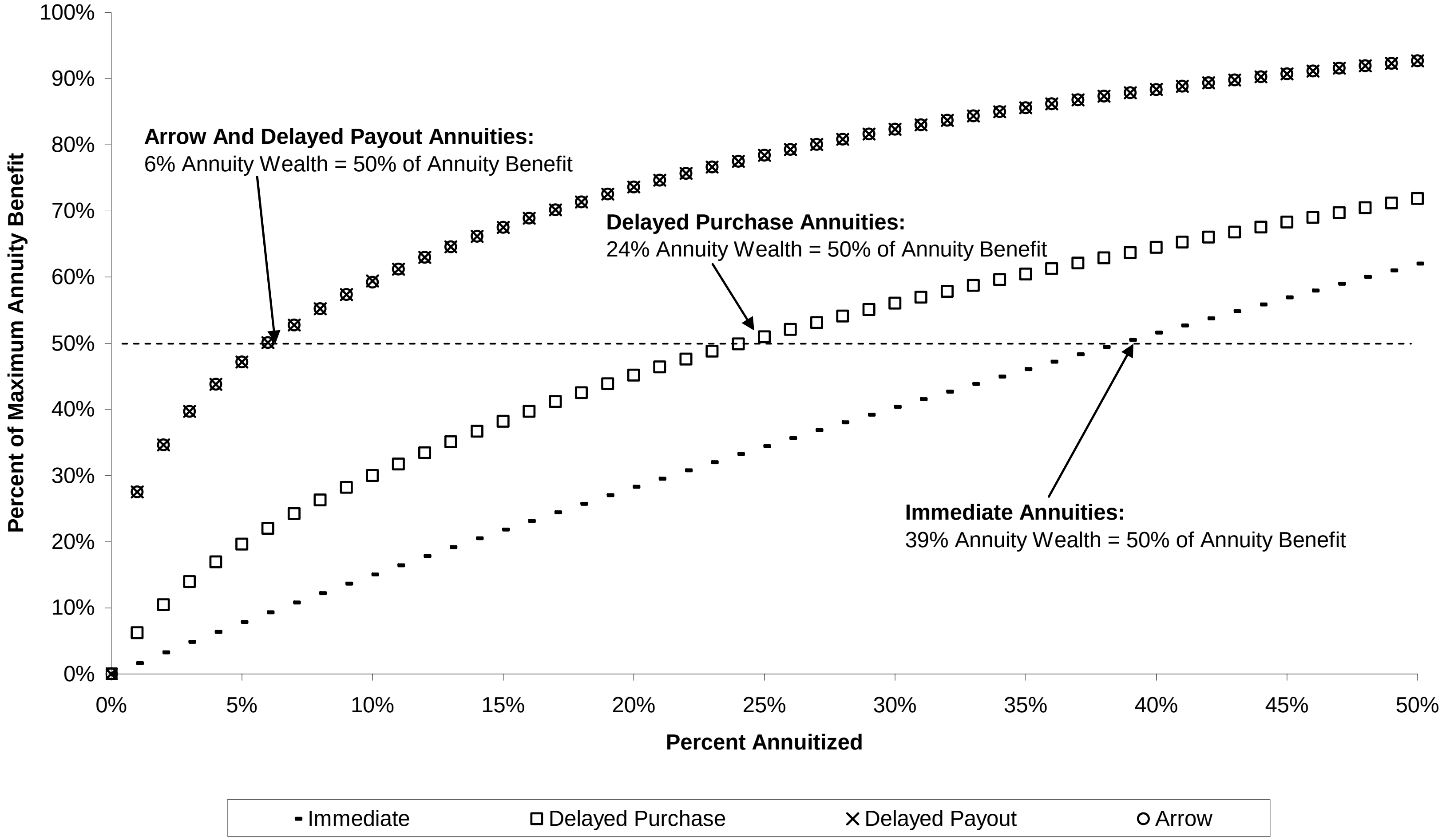


**Figure 2**
**Actuarially Unfair Annuity Pricing**
**Gamma = 4, Age = 65, Male**

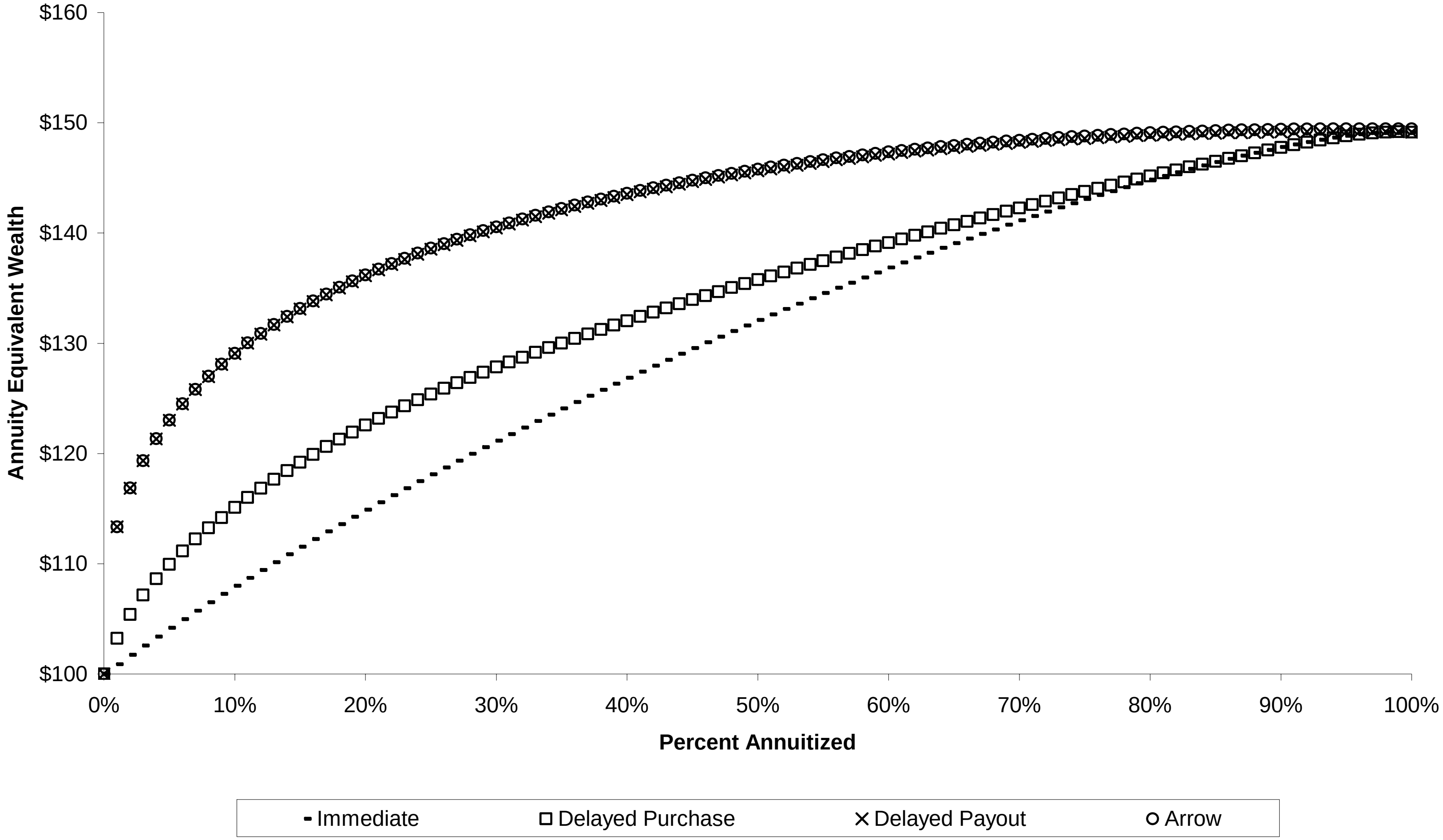


**Figure 2A**
**Efficiency Analysis with Actuarially Unfair Pricing**
**Gamma = 4, Age = 65, Male**

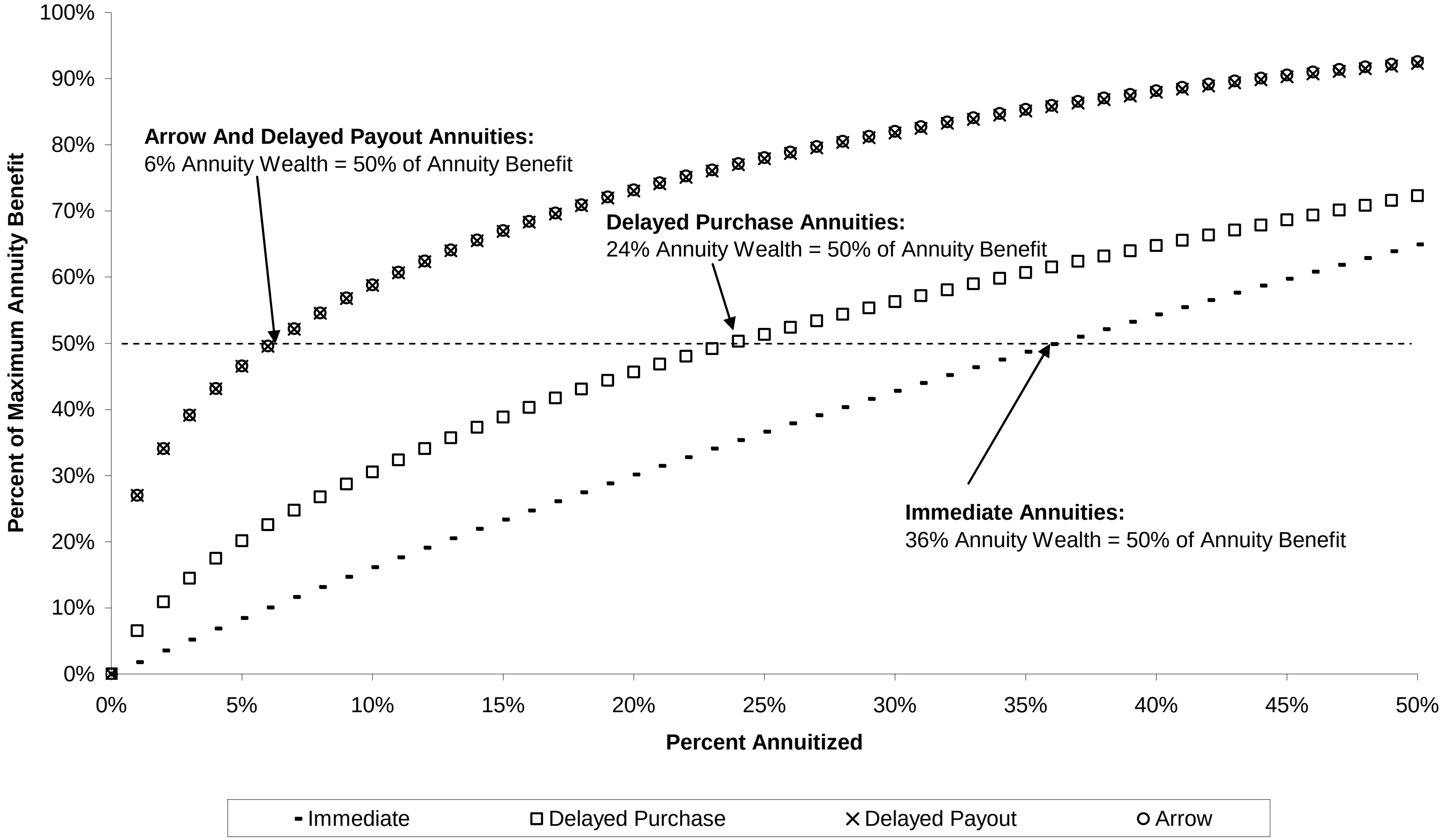
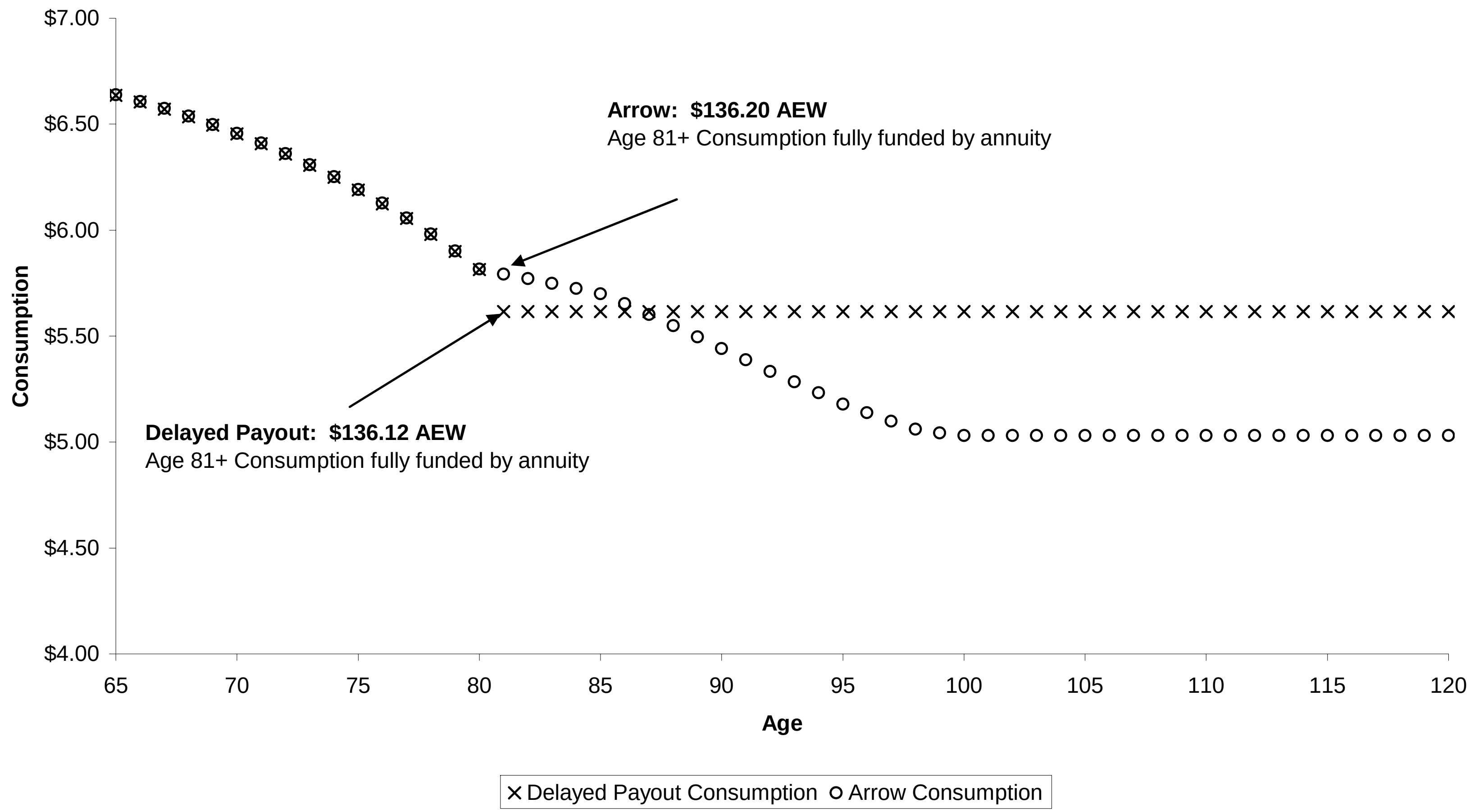


**Figure 2b**
**20% Annuity Allocation**
**Delayed Payout vs. Arrow Consumption Profile**
**Male, Age 65, Gamma = 4, Actuarially Unfair Pricing**



## Table 1
## Baseline Analysis: Actuarially Fair Prices

| | Scenario #1 | Scenario #2 | Scenario #3 | Scenario #4 |
|---|---|---|---|---|
| **Scenario Parameters**[1] | | | | |
| Age | 65 | 65 | 65 | 65 |
| Gender | Male | Male | Male | Female |
| Utility Gamma | 4 | 2 | 1 | 4 |
| Max AEW - Unconstrained Arrow Access[2] | $154 | $143 | $134 | $145 |
| **5% Annuity Allocation** | | | | |
| Immediate AEW | $104 | $104 | $103 | $103 |
| Delayed Purchase AEW | $111 | $107 | $105 | $110 |
| Delayed Payout AEW | $125 | $118 | $112 | $122 |
| Arrow AEW | $125 | $118 | $112 | $122 |
| Age Arrow Payouts Begin | 88 | 87 | 86 | 90 |
| Percent of Max Achieved with Arrow | 47% | 41% | 35% | 50% |
| Immediate Allocation to Equal Arrow AEW | 36% | 30% | 24% | 38% |
| **10% Annuity Allocation** | | | | |
| Immediate AEW | $108 | $107 | $106 | $107 |
| Delayed Purchase AEW | $116 | $111 | $108 | $115 |
| Delayed Payout AEW | $132 | $123 | $116 | $128 |
| Arrow AEW | $132 | $123 | $116 | $128 |
| Age Arrow Payouts Begin | 84 | 84 | 83 | 86 |
| Percent of Max Achieved with Arrow | 59% | 54% | 49% | 62% |
| Immediate Allocation to Equal Arrow AEW | 47% | 41% | 35% | 50% |
| **20% Annuity Allocation** | | | | |
| Immediate AEW | $115 | $113 | $110 | $113 |
| Delayed Purchase AEW | $124 | $118 | $113 | $122 |
| Delayed Payout AEW | $140 | $130 | $122 | $134 |
| Arrow AEW | $140 | $130 | $122 | $134 |
| Age Arrow Payouts Begin | 80 | 80 | 80 | 82 |
| Percent of Max Achieved with Arrow | 74% | 70% | 66% | 76% |
| Immediate Allocation to Equal Arrow AEW | 62% | 57% | 51% | 64% |

[1]*A GAM 94 mortality table with 100+ mortality set to population average is utilitized for all calculations*
[2]*Bond-only utility is normalized to $100 AEW*

# Table 2
## Actual Life Expectancy Below Annuity Pricing Life Expectancy[1]

| | Scenario #1 | Scenario #2 | Scenario #3 | Scenario #4 |
|---|---|---|---|---|
| **Scenario Parameters** | | | | |
| Age | 65 | 65 | 65 | 65 |
| Gender | Male | Male | Male | Female |
| Utility Gamma | 4 | 2 | 1 | 4 |
| Max AEW - Unconstrained Arrow Access[2] | $149 | $137 | $126 | $141 |
| **5% Annuity Allocation** | | | | |
| Immediate AEW | $104 | $104 | $103 | $103 |
| Delayed Purchase AEW | $110 | $106 | $104 | $109 |
| Delayed Payout AEW | $123 | $115 | $109 | $121 |
| Arrow AEW | $123 | $115 | $109 | $121 |
| Age Arrow Payouts Begin | 87 | 86 | 85 | 89 |
| Percent of Max Achieved with Arrow | 47% | 41% | 35% | 50% |
| Immediate Allocation to Equal Arrow AEW | 33% | 26% | 20% | 36% |
| **10% Annuity Allocation** | | | | |
| Immediate AEW | $108 | $107 | $105 | $107 |
| Delayed Purchase AEW | $115 | $110 | $107 | $114 |
| Delayed Payout AEW | $129 | $120 | $113 | $126 |
| Arrow AEW | $129 | $120 | $113 | $126 |
| Age Arrow Payouts Begin | 84 | 83 | 82 | 86 |
| Percent of Max Achieved with Arrow | 59% | 54% | 49% | 62% |
| Immediate Allocation to Equal Arrow AEW | 44% | 37% | 30% | 47% |
| **20% Annuity Allocation** | | | | |
| Immediate AEW | $115 | $112 | $109 | $112 |
| Delayed Purchase AEW | $123 | $116 | $110 | $121 |
| Delayed Payout AEW | $136 | $126 | $117 | $131 |
| Arrow AEW | $136 | $126 | $118 | $131 |
| Age Arrow Payouts Begin | 80 | 79 | 78 | 82 |
| Percent of Max Achieved with Arrow | 74% | 70% | 67% | 76% |
| Immediate Allocation to Equal Arrow AEW | 59% | 52% | 45% | 61% |

[1]*Prcing based on modified GAM 94 mortality table.  Individual mortality based on Social Security population average mortality.*
[2]*Bond-only utility is normalized to $100 AEW*

## Table 3
## Social Security Analysis: Baseline + 50% Preallocated to Immediate Annuity

| | Scenario #1 | Scenario #2 | Scenario #3 | Scenario #4 |
|---|---|---|---|---|
| **Scenario Parameters[1]** | | | | |
| Age | 65 | 65 | 65 | 65 |
| Gender | Male | Male | Male | Female |
| Utility Gamma | 4 | 2 | 1 | 4 |
| Base AEW - 50% Preallocated to Immediate | $133 | $127 | $122 | $128 |
| Max AEW - Unconstrained Arrow Access[2] | $154 | $143 | $134 | $145 |
| **5% Annuity Allocation** | | | | |
| Immediate AEW | $136 | $129 | $123 | $130 |
| Delayed Purchase AEW | $137 | $130 | $124 | $132 |
| Delayed Payout AEW | $143 | $134 | $126 | $136 |
| Arrow AEW | $143 | $134 | $126 | $136 |
| Age Arrow Payouts Begin | 84 | 83 | 82 | 86 |
| Percent of Max Achieved with Arrow | 49% | 42% | 35% | 51% |
| Immediate Allocation to Equal Arrow AEW | 20% | 16% | 13% | 21% |
| **10% Annuity Allocation** | | | | |
| Immediate AEW | $139 | $131 | $125 | $132 |
| Delayed Purchase AEW | $140 | $132 | $125 | $134 |
| Delayed Payout AEW | $147 | $137 | $128 | $139 |
| Arrow AEW | $147 | $137 | $128 | $139 |
| Age Arrow Payouts Begin | 80 | 80 | 79 | 82 |
| Percent of Max Achieved with Arrow | 66% | 61% | 55% | 69% |
| Immediate Allocation to Equal Arrow AEW | 28% | 25% | 22% | 29% |
| **20% Annuity Allocation** | | | | |
| Immediate AEW | $143 | $135 | $128 | $136 |
| Delayed Purchase AEW | $145 | $136 | $128 | $138 |
| Delayed Payout AEW | $151 | $140 | $131 | $142 |
| Arrow AEW | $151 | $140 | $131 | $142 |
| Age Arrow Payouts Begin | 75 | 75 | 74 | 76 |
| Percent of Max Achieved with Arrow | 85% | 82% | 79% | 87% |
| Immediate Allocation to Equal Arrow AEW | 38% | 36% | 34% | 39% |

[1]*A GAM 94 mortality table with 100+ mortality set to population average is utilitized for all calculations*
[2]*Bond-only utility is normalized to $100 AEW*

**Appendix A: Separation Theorem with Stochastic Returns**

The text derived a separation theorem assuming the availability of two types of assets, riskless zero coupon bonds, and Arrow annuities which were analogous annuitized assets. In this appendix, we extend the separation theorem to securities with stochastic payouts. The separation theorem does not require riskless securities. The main requirements are the existence of analogous securities that are available in either their normal form or in annuitized (i.e. survival contingent) form. As long as the pricing discount available for the annuitized asset becomes more favorable with horizon, the separation theorem will generally hold.

To demonstrate this result, first consider an investment, ***S***, that has a current $1 price, but has stochastic future values. Let $S_t$, refer to the potential values that portfolio ***S*** can take at time $t$. Now imagine a security that is analogous to our zero coupon bond, but instead delivers the payouts associated with investment ***S*** at time $t$. To retain the notation from the main text, we will use $B_t$ to refer to the price of this security. The only change is that now the payouts are stochastic. In particular, let

$B_t$ = Price today required to secure payouts $S_t$ at time $t$.

Notice, that all of these securities can be replicated through an investment in ***S***, so all of the prices must be $1. In other words,

$B_t$ = $1, for all $t$

In addition to the investment ***S***, our retiree has the ability to purchase Arrow annuities with payouts determined by the performance of portfolio ***S***. In particular:

$A_t$ = Price today of security which pays out $S_t$ given survival to $t$.

= Annuity discount × Price of security which pays out $S_t$

= $P_t \times \$1$

Similar to the problem described in the main text, our retiree can purchase either the payouts associated with portfolio ***S*** directly or they can purchase the annuitized version of those payouts. Let:

$x_t$ = Amount of dollars allocated to purchase payouts from $B_t$.

$y_t$ = Amount of dollars allocated to purchase payouts from $A_t$.

As long as the annuity discount function, $P_t$, decreases in $t$, then the separation theorem holds. The proof follows the same approach as the riskless security proof.

*Claim:* If the annuity discount function, $P_t$, decreases in time, then it is never optimal for annuitized payouts to precede non-annuitized portfolio payouts.

*Proof:* First, assume an optimal solution exists that consists of purchasing portfolio payouts at time $t+n$ ($x_{t+n} > 0$), $n$ periods after an annuity purchase ($y_t > 0$). Next, construct an alternative solution, denoted with an asterisk (*), that deviates from the presumed optimal by the following trades:

$$(3a') \quad x^*_{t+n} = x_{t+n} - \varepsilon$$

$$(3b') \quad y^*_{t+n} = y_{t+n} + \varepsilon$$

$$(3c') \quad y^*_t = y_t - \varepsilon \cdot \frac{A_{t+n}}{A_t}$$

$$(3d') \quad x^*_t = x_t + \varepsilon \cdot \frac{A_{t+n}}{A_t}$$

Similar arguments to those made in the main text imply the proposed alternate solution is feasible given the constraints. To complete the proof, we must assess the cost of implementing the alternate solution. For this problem we have

$$\text{Net Costs} = \varepsilon \cdot P_{t+n} / P_t \cdot (1 - P_t) - \varepsilon \cdot (1 - P_{t+n})$$

The net costs are negative when:

$$\varepsilon \cdot P_{t+n} / P_t \cdot (1 - P_t) - \varepsilon \cdot (1 - P_{t+n}) < 0$$

$$\Rightarrow P_{t+n} \cdot (1 - P_t) < (1 - P_{t+n}) \cdot P_t$$

$$\Rightarrow P_{t+n} < P_t$$

This inequality is satisfied by assumption, thus completing the proof. Independent of the stochastic nature of the payouts, if there are securities that can be purchased in annuitized or non-annuitized form, an efficient allocation will require purchasing the longer-horizon annuities first. The only general requirement for this result is that the annuity pricing discount relative to the non-annuitized asset improves with horizon.

## Appendix B: Conditions when Delayed Payout Substitutes for Arrow Annuities

Suppose the optimal Arrow annuity demand over time has the following form:

$$Y = \{0\ 0\ \ldots\ 0\ y_t\ y_{t+1}\ \ldots\ y_T\}$$

The $t^{\text{th}}$ element in the annuity demand vector, $Y$, corresponds to the amount of consumption in period $t$ funded by annuities. Note, the annuity demand vector is compliant with the separation theorem. Further, define the growth in annuity demand, $g_t$, as follows:

$$y_{t+1} = g_{t+1} \cdot y_t, \text{ for all } t \text{ such that } y_t > 0$$

The question now becomes under what conditions can delayed payout annuities (DPAs) exactly fulfill the desired annuity-supported consumption. To assess this, we first define what is meant by a DPA in terms of consumption payouts. A delayed payout annuity that begins payments at time $t$, $Z_t$, has the following payout structure:

$$Z_t = \{0\ 0\ \ldots\ 0\ z_t\ z_{t+1}\ \ldots\ z_T\}$$

We will let the first payment correspond to one unit of consumption without loss of generality. Thus, the payout vector looks like this:

$$Z_t = \{0\ 0\ \ldots\ 0\ 1\ z_{t+1}\ \ldots\ z_T\}$$

Analogous to annuity demand, we define a payout growth rate, $h_t$, that captures the growth in DPA payouts over time.

$$z_{t+1} = h_{t+1} \cdot z_t, \text{ for all } t \text{ such that } z_t > 0$$

We make two additional assumptions before deriving conditions under which DPAs substitute for Arrow annuities. First, we assume that a given DPA is priced consistent with the Arrow annuity prices. That is, a DPA that provides consumption starting at age 80 costs the same as funding that same pattern of consumption using Arrow annuities starting at age 80. Second, we assume that over a given period of time, the growth in DPA payouts is consistent across all DPAs independent of the start date of payouts. In other words, $h_t$ is the same for all DPAs with payment start dates prior to $t$.

With this collection of notation and assumptions, we can now evaluate whether DPAs can replicate the optimal Arrow annuity demand. To assess this, we examine the period-by-period annuity demand. Suppose annuity demand begins in period $t$, then we would need to purchase enough $Z_t$ to satisfy this annuity demand. Since the first payment from a DPA is one unit, we need to purchase $y_t$ units of the DPA with payments starting at period $t$. This DPA provides the following consumption stream:

$$y_t \text{ units of } Z_t = \{0\ 0\ \ldots\ 0\ y_t\ y_t z_{t+1}\ \ldots\ y_t z_T\}$$

Next we evaluate the annuity demand in period $t$+1. The aggregate annuity demand for period $t$+1 is $y_{t+1}$. Part of this annuity demand is fulfilled by the period $t$ DPA. The remainder must be satisfied by a purchase of a period $t$+1 DPA. More precisely:

$$\text{Demand for } Z_{t+1} = y_{t+1} - y_t z_{t+1} = g_{t+1} y_t - h_{t+1} y_t = (g_{t+1} - h_{t+1}) \cdot y_t$$

Assuming individuals cannot short DPAs, then replicating consumption in period $t$+1 requires the growth rate in annuity demand, $g_{t+1}$ to equal or exceed the growth rate in DPA payouts, $h_{t+1}$. Now that demand for period $t$ and $t$+1 DPAs are determined, we can investigate demand for period $t$+2 DPAs.

$$\text{Demand for } Z_{t+2} = y_{t+2} - h_{t+2} y_{t+1} = g_{t+2} y_{t+1} - h_{t+2} y_{t+1} = (g_{t+2} - h_{t+2}) \cdot y_{t+1}$$

The same condition holds for the $t$+2 growth rates that held for $t$+1 growth rates, namely annuity demand growth (i.e. consumption growth) must exceed DPA payment growth. A similar calculation for other time periods confirms the general result. A DPA market can perfectly substitute for an Arrow annuity market if:

(9) $g_t \geq h_t$ for all $t$ such that $y_{t-1} > 0$

In other words, as long as consumption growth exceeds DPA payout growth during the time period when annuities are demanded, then an investor can replicate the Arrow annuity consumption with DPA annuity consumption.

## Appendix C: Optimization with Alternative Annuity Product Spaces

In this appendix, we describe the mathematical problem for each of the annuity product spaces described in Section 5. In each case, the availability of specific annuity "bundles" requires further constraints on the problem. With some algebraic manipulation, we show that the optimization for each of the various annuity product spaces reduces to the baseline Arrow optimization problem with the addition of a single constraint. For ease of understanding, we start with the least constrained problem (Arrow annuities) and progress toward the most constrained problem (bonds and immediate annuities only). Each successive case can be seen as a more restrictive version of the preceding case (either having more restrictive constraints or subject to worse annuity pricing), so that the maximum obtainable utility must be non-increasing.

**Annuity Product Space #4: Arrow Annuities**

The objective function remains

$$max \sum_{t=0}^{\infty} \Pi_t \Delta_t U(c_t)$$

with the constraints

$$c_t = x_t + y_t$$

$$(1-\alpha)W_0 = \sum_{t=0}^{\infty} B_t x_t$$

$$\alpha W_0 = \sum_{t=0}^{\infty} A_t y_t$$

$$x_t \geq 0,\ y_t \geq 0$$

In the case where $\alpha$=1 (full annuitization is allowed), the optimal consumption profile for a CRRA utility function, $U(c) = c^{1-\gamma}/(1-\gamma)$, is given by

$$c_t = y_t = \frac{W_o(\Pi_t \Delta_t / A_t)^{1/\gamma}}{\sum_{s=0}^{\infty} A_s (\Pi_s \Delta_s / A_s)^{1/\gamma}}$$

In the case where $\alpha$=0, all money is invested in bonds and optimal consumption for the same CRRA utility function is given by

$$c_t = x_t = \frac{W_o(\Pi_t \Delta_t / B_t)^{1/\gamma}}{\sum_{s=0}^{\infty} B_s (\Pi_s \Delta_s / B_s)^{1/\gamma}}$$

**Annuity Product Space #3: Delayed Payout Annuities and Bonds**

Again we have the objective function

$$max \sum_{t=0}^{\infty} \Pi_t \Delta_t U(c_t)$$

but we modify the constraints to the following:

$$c_t = x_t + y_t$$

$$y_t = \sum_{s=0}^{t} z_s$$

$$(1-\alpha)W_0 = \sum_{t=0}^{\infty} B_t x_t$$

$$\alpha W_0 = \sum_{t=0}^{\infty} D_t z_t$$

$$x_t \geq 0,\ z_t \geq 0$$

Note that $z_t$ is the amount of income purchased from a delayed payout annuity that starts payments at time *t*. The total consumption from all annuities at time *t* is then the sum of income from all prior annuity purchases.

A delayed payout annuity that pays \$1 starting at time $t$ and every year thereafter, provided the individual is alive, has a cost $D_t$. Assuming that this annuity bundle is priced as the sum of its constituent Arrow annuities, we have

$$D_t = \sum_{s=t}^{\infty} A_s$$

One can show the following equality using summation by parts:

$$\alpha W_0 = \sum_{t=0}^{\infty} D_t z_t = \sum_{t=0}^{\infty} (D_t - D_{t+1}) y_t$$

Combining the above equations, we have

$$\alpha W_0 = \sum_{t=0}^{\infty} A_t y_t$$

which is the same annuity wealth constraint as for Arrow annuities.

Furthermore, we can enforce the conditions $z_t \geq 0$ by imposing the following constraints on $y_t$:

$$0 \leq y_o \leq y_1 \leq y_2 \leq \ldots \leq y_t \leq \ldots$$

Thus, the mathematical program for this annuity product space is the same as for Arrow annuities, but with the added condition that the annuity consumption path must be non-decreasing. Maximum utility for delayed payout annuities is bounded by maximum utility for Arrow annuities, with equality whenever optimal Arrow annuity-based consumption is non-decreasing through time.

**Annuity Product Space #2: Delayed Purchase Annuities and Bonds**

This problem has the same formulation as annuity product space #3, but with the annuity cost given by

$$D_t = \frac{B_t}{A_t}\sum_{s=t}^{\infty} A_s$$

Note that this pricing is always inferior to the pricing of delayed payout annuities because $B_t > A_t$ for $t>0$.

**Annuity Product Space #1: Immediate Annuities and Bonds**

With the choice of only immediate annuities beginning at time 0, we have the same objective function

$$max \sum_{t=0}^{\infty} \Pi_t \Delta_t U(c_t)$$

with the constraints:

$$c_t = x_t + y_t$$

$$y_t = z_0$$

$$(1-\alpha)W_0 = \sum_{t=0}^{\infty} B_t x_t$$

$$\alpha W_0 = D_0 z_0$$

$$x_t \geq 0,\ z_0 \geq 0$$

Again assuming that the cost of the immediate annuity is equal to the sum of the cost of the constituent Arrow annuities, we have

$$D_0 = \sum_{t=0}^{\infty} A_t$$

This annuity product space can be viewed as a subset of annuity product space #3 (delayed payout annuities), with the added restriction that

$$0 = z_1 = z_2 = \ldots = z_t = \ldots$$

Alternatively, one can view this problem as a special case of annuity product space #4 with the additional restriction that

$$y_0 = y_1 = y_2 = \ldots = y_t = \ldots$$

**Future Trading**

For all optimizations, it should be noted that decisions are assumed to occur at time zero and are not revisited as time passes. In many situations, future trading would not occur even if allowed. For example, consider the baseline problem with Arrow annuity availability (i.e. Annuity Product Space #4). At the two extremes for allowable annuity purchases, $\alpha = 0$ and $\alpha = 1$, allowing future trading is unnecessary since the original strategy is still optimal at any future point in time assuming the relative discount and interest rates between period t and t+n are unchanged as time advances.[18] When annuity purchases are allowed but constrained, $0 < \alpha < 1$, future trading is also unnecessary provided the implied period t allocation to annuities is consistent with the allowable allocation to annuities. However, a given strategy involving delayed payout and delayed purchase annuities may not in general be optimal when viewed from a later time period. As such, the potential exists for improved welfare stemming from future trading. Assessing the situations where large welfare improvements are available from future trading is the subject of future research.

[18] This can be verified by confirming that the original plan still satisfies the first order conditions at all future dates.