

Too Risk Averse to Purchase Insurance?

A Theoretical Glance at the Annuity Puzzle*

Antoine Bommier François Le Grand[†]

March 2014

Abstract

This paper suggests a new explanation for the low level of annuitization, which is valid even if one assumes perfect markets. We show that, as soon there is a positive bequest motive, sufficiently risk averse individuals should not purchase annuities. A model calibration accounting for lifetime risk aversion generates a willingness-to-pay for annuities, which is significantly smaller than the one generated by a standard Yaari (1965) model. Moreover, the calibration predicts that riskless savings finance one third of consumption, in line with empirical findings.

Keywords: annuity puzzle, insurance demand, bequest, intergenerational transfers, risk aversion, multiplicative preferences.

JEL codes: D11, D81, D91.

1 Introduction

Among the greatest risks in life is that associated with life duration. A recently retired American man aged 65 has a life expectancy of about 17.5 years. However, there is more than a 22% chance that he will die within the first 10 years and more than a 20% chance that he will live for longer than 25 years. Savings required to sustain 10 or 25 years of retirement vary considerably, and one would expect a strong demand for annuities, which are financial securities designed to deal with lifetime uncertainty. A number of papers have stressed the utility gains that would be generated by the

*We are grateful to Edmund Cannon, Alexis Direr, Glenn Harrison, Lee Lockwood, Thomas Post, James Poterba, Ray Rees, Harris Schlesinger (the editor), an anonymous referee and seminar participants at ETH Zurich, University of Paris I, University of Zurich, 2011 Summer Meetings of the Econometric Society and CEAR/MRIC Behavioral Insurance Workshop 2012 (LMU, Munich) for their comments.

[†]Antoine Bommier is Professor at ETH Zurich; abommier@ethz.ch François Le Grand is Assistant Professor at EMLyon Business School and Associate Researcher at ETH Zurich; legrand@em-lyon.com. Both authors gratefully acknowledge financial support from Swiss-Re.

annuitization of wealth at retirement. It is generally estimated that individuals would be willing to give up to as much as 25% of their wealth at retirement to gain access to a perfect annuity market (see Mitchell, Poterba, Warshawsky and Brown (1999) among others). According to standard theoretical predictions, even when individuals have a bequest motive, they should fully annuitize the expected value of their future consumption. However, empirical evidence quite consistently shows that the market for private annuities is on aggregate very small, in sharp contradiction with the theoretical predictions. For example, Johnson, Burman, and Kobes (2004) report that in the US, private annuities finance less than 1% of household income for people over 65.¹ Similarly, they also observe that only 5% of people over 65 purchase private annuities. James and Song (2001) find similar results for other countries, such as Canada, the United Kingdom, Switzerland, Australia, Israel, Chile and Singapore.²

A number of explanations to this puzzle have been suggested, relying on market imperfections or rationality biases.³ For example, due to imperfect health insurance, individuals would need to store a substantial amount of liquidities, unfair annuity pricing would make them unattractive assets, and framing effects would play an important role in agents' decisions to annuitize.

In this paper we emphasize that, even if the annuity market were perfect, a low (or even zero) level of annuitization may be fully rational. Our explanation relies on the role of risk aversion. We show that a high level of risk aversion together with a positive bequest motive is sufficient to predict a negative demand for annuities. Even though the role of risk aversion has not been studied in isolation, the intuition that annuities are perceived as a risky gamble was first evoked by Brown (2007) and Brown, Kling, Mullainathan and Wrobel (2008), who emphasized the perceived riskiness of annuities.

The reason why the effect of risk aversion has remained unexplored is that the literature has mainly focused on time additively separable preferences, or on Epstein and Zin (1989) specification, while neither model is suited to studying the role of risk aversion (See Bommier, Chassagnon, LeGrand (2012), henceforth BCL).⁴ In the current paper, the role of risk aversion is investigated in the expected utility framework, through the concavification of the lifetime utility function as introduced by Kihlstrom and Mirman (1974). We prove that the demand for annuities decreases with risk aversion and eventually vanishes when risk aversion is large enough.

The fact that annuity demand decreases – and does not increase – with risk aversion might seem

¹Roughly one half of income stems from public pensions, 17% from company sponsored pension payments and one third is financed from savings.

²Some studies, such as Benartzi, Previtro, and Thaler (2011), argue, however, that the observed levels of annuitization are not as puzzling as usually claimed. They provide evidence to show that most individuals actually choose to annuitize their wealth when possible and point out that most retirement plans do not offer this possibility.

³See Brown (2007), as well as the following section for a literature review.

⁴Davidoff, Brown and Diamond (2005) consider non-additive separable preferences featuring habit formation and show that it helps explain the annuity puzzle. However, they do not investigate the role of risk aversion.

counter-intuitive. Insurance demand is generally found to increase with risk aversion. However, this correlation does not hold any longer when irreplaceable commodities, such as life, are at risk. As was explained by Cook and Graham (1977), rational insurance decisions aim at equalizing marginal utilities of wealth across states of nature. With irreplaceable commodities, this may generate risk taking behavior. Whenever this is the case, risk aversion should limit these risk taking behaviors and reduce the demand for insurance.⁵

Annuities provide an example where purchasing insurance is risk increasing. Lifetime is uncertain, but living a long time is generally considered to be a good outcome, while dying early is seen to be a bad outcome. For a given amount of savings, purchasing annuities, rather than bonds for example, involves reducing bequest in the case of early death (i.e., a bad outcome), while increasing consumption in the event of survival (i.e., a good outcome). Thus, for a given level of savings, annuities transfer resources from bad to good states of the world and are, as such, risk increasing. If first period consumption were exogenous and inter-vivos transfers were ruled out, simple dominance arguments as in BCL would directly imply that the demand for annuities decreases with risk aversion. In the current paper, the result is obtained with endogenous consumption smoothing and the introduction of inter-vivos transfers. Moreover, we prove that when risk aversion is large enough, annuity demand eventually vanishes.

In order to evaluate the contribution of risk aversion in solving the annuity puzzle, we calibrate a life-cycle model in which agents can invest in bonds and annuities. Calibrating risk aversion and bequest motives to plausible levels generates considerably lower willingnesses-to-pay for annuities than those obtained with the standard model of Yaari (1965), indicating that risk aversion may be an important factor to explain the low levels of annuitization. Our calibration implies that one third of the agents' consumption is financed by riskless savings, which is in line with the empirical findings of Johnson, Burman, and Kobes (2004). This contrasts with the standard Yaari model in which riskless savings do not contribute at all to consumption financing, even if agents have bequest motives.

The remainder of the paper is structured as follows. Section 2 discusses the related literature. We then present a two-period model and derive our theoretical predictions in Section 3. In Section 4, the model is extended to an N -period setting and calibrated. Numerical simulations then derive the optimal life-cycle strategy of agents facing realistic mortality rates. Section 5 concludes.

⁵This was also noticed by Drèze and Rustichini (2004), who provided an example where insurance demand may decrease with risk aversion (see their Proposition 9.1).

2 Related literature

The literature on annuities was initiated by Yaari's (1965) seminal contribution. Agents who do not care for bequest but value consumption should invest all their wealth in annuities.⁶ Full annuitization is no longer optimal when bequest motives are introduced. However, Davidoff, Brown and Diamond (2005), as well as Lockwood (2012a), prove that the optimal behavior consists in annuitizing the discounted value of all future consumptions. The low level of observed annuitization was then identified as a puzzle, for which different explanations were suggested.

A first explanation is related to unfair pricing of annuities, as reported by Mitchell, Poterba, Warshawsky and Brown (1999), Finkelstein and Poterba (2002) and (2004). Lockwood (2012a) demonstrates that this aspect, together with bequest motives of a reasonable magnitude, may be sufficient to explain the low level of annuitization.

Another possible explanation is that inadequate insurance products such as health or long term care insurance may encourage people to save a large amount of liquid assets. As a result of adverse selection issues, annuities are not very liquid and difficult to sell back. The optimal strategy while facing uninsurable risks may then involve investing wealth in buffer assets, such as bonds or stocks rather than in annuities. Sinclair and Smetters (2004), Yogo (2009), Pang and Warshawsky (2010), among others, emphasize this explanation. Pashchenko (2013) shows that the illiquidity of housing wealth together with public annuities, minimum annuity purchase requirement, and bequest motives can quantitatively contribute to explain the low demand for annuities.

A related channel is the fact that annuities diminish individuals' investment opportunity sets by preventing savings in high return and high risk assets. Milevsky and Young (2007) and Horneff, Maurer, Mitchell and Stamos (2010) argue that the annuity puzzle stems from the lack of annuities backed by high-risk and high-return assets.

Finally, behavioral economics provides a whole range of explanations. For example Brown, Kling, Mullainathan and Wrobel (2008) emphasize that framing effects could be at the origin of the low demand for annuities.⁷ Brown (2007) reviews other behavioral hypotheses, such as regret aversion, financial illiteracy and the illusion of control or loss aversion. Hu and Scott (2007) also point out the role of loss aversion.

Interestingly enough, papers discussing these behavioral aspects also underline the role of annuity riskiness. In particular Brown, Kling, Mullainathan and Wrobel (2008, p.305) explain that "annuities appear riskier than the bond", since purchasing annuities generates a substantial loss in the event of early death. Similarly, Brown (2007) explains that agents seem to be willing to

⁶See corollary 1 in Davidoff, Brown and Diamond (2005).

⁷Framing effects describe the fact that individuals' choices may depend on the formulation of alternatives and especially if they are focused on gains or losses.

purchase insurance that pays off well in the case of bad events, while annuities pay in the case of good events (i.e., survival). Agnew, Anderson, Gerlach and Szykman (2008) confirm through lab experiments the importance of annuity riskiness perception.⁸ The role of framing is also highlighted by Benartzi, Previtro, and Thaler (2011, p. 156), who state that “while economists tend naturally to think about annuitization as a risk-reducing strategy like the purchase of insurance, many consumers may not share this point of view”. It seems that agents are extremely sensitive to the riskiness of annuities, and that risk aversion may therefore play a significant role.

The role of risk aversion has not hitherto been formalized. The reason is that most papers use Yaari’s approach, based on an assumption of additive separability of preferences, which imposes risk aversion to be equal to the inverse of the intertemporal elasticity of substitution. As underlined in several papers (e.g., Epstein and Zin (1989)), the additive framework is ill-suited for the analysis of the role of risk aversion, since it cannot disentangle aspects of preferences over certain outcomes from the ones related to risky gambles.

A few papers on annuities focus on Epstein and Zin’s (1989) approach to disentangling risk aversion from the elasticity of substitution.⁹ However, as shown in BCL, Epstein-Zin utility functions are not well ordered in terms of risk aversion. This generates surprising results when studying the relation between risk aversion and savings choices. For example, in a simple two-period model, simple dominance arguments developed in BCL indicate that precautionary savings rise with risk aversion.¹⁰ The same conclusion is drawn when considering well ordered specifications based on expected utility or on rank dependent expected utility (see Drèze and Modigliani (1972), Yaari (1987), or Bleichrodt and Eeckhoudt (2005) among others). On the contrary, Kimball and Weil (2009) prove that this relation is ambiguous for Epstein-Zin preferences.

A simple and robust way of studying risk aversion involves remaining within the expected utility framework and increasing the concavity of the lifetime – and not instantaneous – utility function, as initially suggested by Kihlstrom and Mirman (1974). This approach has notably been followed by van der Ploeg (1993), Eden (2008), and Van den Heuvel (2008). In the case of choice with lifetime uncertainty, this approach was first used in Bommier (2006) and leads to novel predictions on a number of topics, including the relation between time discounting and risk aversion, the impact of mortality change and the value of life. In particular, as highlighted in BCL, these preferences are well ordered in terms of risk aversion and deliver meaningful results when studying intertemporal choice problems. They were shown to generate realistic lifecycle consumption profiles (Bommier (2013)). In the present paper we consider such an approach in a framework accounting for bequests and inter-vivos transfers.

⁸Gazzale and Walker (2011) reach a similar conclusion using neutral-context laboratory experiments.

⁹See e.g. Ponzetto (2003), Inkmann, Lopes and Michaelides (2011) and Horneff, Maurer and Stamos (2010).

¹⁰Precautionary savings can be defined as the optimal savings due to the uncertainty of the second-period income.

3 The model

3.1 Description

The economy is populated by a single agent, who cares for someone else. This heir is not modeled and his single attribute is to accept inter-vivos transfers or bequests. The economy is affected by a mortality risk. The agent may live for one period with probability $1 - p$ or for two periods with probability $p \in (0, 1)$.

We assume that the agent can transfer consumption from the first period to the second period, either through annuity or bond savings. The annuity market is supposed to be perfectly fair and the bond market pays off an exogenous riskless gross rate of return $1 + R$. Investing one unit of consumption in riskless savings in period 1 returns $1 + R$ consumption units in the second period, while the same investment in annuities produces $\frac{1+R}{p}$ second period consumption units.

The agent is endowed with an initial constant wealth W_0 and has no other source of income. In the first period, the agent consumes c_1 out of his wealth. He is left with wealth $W_0 - c_1$ that he allocates either to annuities a , or savings s . In the second period, the agent faces two alternatives. First, with probability $1 - p$, the agent dies and his capitalized savings $(1 + R)s$ are left to his heir, while his annuities are completely lost, for both the agent and his heir. Second, with probability p , the agent survives and in the second period, he enjoys the benefits from his riskless saving and his annuity payment, the total of which is equal to $(1 + R)s + \frac{1+R}{p}a$. Out of this sum, the agent consumes c_2 and hands down the remaining money to his heir through an inter-vivos transfer.

3.2 Preferences

Given our previous description, the economy is ex post described by only three variables: the first period consumption c_1 , the second period agent's status x_2 (i.e., dead or alive, and if he is alive how much he consumes), and the amount of money τ left to the heirs, through either bequests or inter-vivos transfers. Modeling agents' behavior involves comparing lotteries whose consequences are the previous triplet $(c_1, x_2, \tau) \in \mathbb{R}^+ \times (\mathbb{R}^+ \cup \{d\}) \times \mathbb{R}^+$ where d denotes the death state. We constrain consumption, as well as savings and intergenerational transfers to being non-negative. The idea is that an agent cannot force his heir to give him money, or accept a negative bequest.

The agent enjoys felicity $u_1(c_1)$ from the first period consumption, felicity $u_2(c_2)$ from his second period status and felicity $v(\tau)$ from the transfer to his heir. The agent is assumed to be an expected utility maximizer with the following utility index defined over the set of consequences $\mathbb{R}^+ \times (\mathbb{R}^+ \cup \{d\}) \times \mathbb{R}^+$:

$$U(c_1, x_2, \tau) = \phi(u_1(c_1) + u_2(x_2) + v(\tau)).$$

The function ϕ , which makes the link between lifetime felicity and utility, governs risk aversion. This transformation does not modify ordinal preferences and consequently has no impact in deterministic environments. As shown by Kihlstrom and Mirman (1974), augmenting the concavity of the function ϕ provides the standard (and only) way to discuss the role of risk aversion while remaining in the expected utility framework. Such an approach has received little attention because it was thought to lead to time inconsistencies, or to history dependent preferences. Bommier (2013) showed however that the framework of Kihlstrom and Mirman (1974) is not incompatible with the assumption of preference stationarity, provided that one uses an exponential functional form for ϕ , as we will do later on.

Most of the applied literature on intertemporal choice has focused on the special case of a linear transformation ϕ and has associated the words “risk aversion” to measures of the curvature of the functions u_1 , u_2 and v . This is a rather unfortunate terminology as agents with different functions u_1 and u_2 and v cannot be compared in terms of risk aversion, since they do not have the same preferences over certain outcomes.¹¹ A sounder terminology would consist in using the words “elasticity of substitution” when commenting on the curvature of the functions u_1 , u_2 and v , and keeping the expression risk aversion to discuss properties of the function ϕ . We will adhere to this terminology. However, to insist on the difference between our terminology and the usual (but inappropriate) one, we introduce the term “lifetime” before any mention of the words risk aversion. In short, what we call “lifetime risk aversion” is what should have been called “risk aversion” and is exclusively related to the curvature of the function ϕ .¹² In order to further help the reader take some distance with the usual additive model, we moreover use the terms “felicity” when mentioning the functions u_1 , u_2 and v , and keep the term “utility” for the function $U(c_1, x_2, \tau_1) = \phi(u_1(c_1) + u_2(x_2) + v(\tau))$. The usual approach assumes that (lifetime) utility is additive in felicity (i.e., that the transformation ϕ is linear), which involves assuming risk neutrality with respect to lifetime felicity, or equivalently, intertemporal correlation neutrality. The current paper explores the case of a concave function ϕ . We thus focus on the preferences exhibiting a positive lifetime risk aversion, which is consistent with the results of the recent experimental study of Andersen, Harrison, Lau and Rutstrom (2011).

Without loss of generality, we normalize felicity functions as follows. First, the second period felicity when dead is normalized to 0: $u_2(d) = 0$. Second, leaving nothing to his heir also provides $v(0) = 0$. Finally, the function ϕ is normalized with $\phi'(0) = 1$. We also assume that all functions

¹¹As discussed in Kihlstrom and Mirman (1974, Section 2 and in particular 2.1, p.365 and following) and in Epstein and Zin (1989, Section 4, p.950), studying the role of risk aversion requires leaving ordinal preferences unchanged. Individuals who would rank deterministic outcomes differently cannot be compared in terms of risk aversion.

¹²Lifetime risk aversion is sometimes also called *multivariate risk aversion* (see for example Richard (1975)), or *correlation aversion* (see for example Epstein and Tanny (1980), Bommier (2007) or Dorfmeister and Krapp (2007)) since it refers to risk aversion with many commodities.

are regular and more precisely: (i) u_1 , the restriction of u_2 to \mathbb{R}^+ , and v are twice continuously differentiable, increasing and strictly concave and (ii) ϕ is twice continuously differentiable and increasing. Moreover, in order to always obtain strictly positive consumption levels, we assume that marginal utilities of consumption tend to infinity when consumption tends to zero: $\lim_{c \rightarrow 0^+} u_1'(c) = \lim_{c \rightarrow 0^+} u_2'(c) = +\infty$.

Regarding the second period felicity u_2 , we also assume that there exist second period consumption levels such that $u_2(c_2) > 0 = u_2(d)$. This means that for some levels of second period consumption, the agent prefers life to death. We denote c_2^* the minimum level of second period consumption that makes life preferable to death. Formally:

$$c_2^* = \inf\{c_2 > 0 | u_2(c_2) > 0\}. \quad (1)$$

With some specifications, we have $c_2^* = 0$, which means that life is preferable to death no matter the level of consumption. But with other specifications (e.g., when assuming a constant elasticity of substitution smaller than one), this minimal level c_2^* is strictly positive. In this case, if the agent does not enjoy a sufficient second period consumption, he would prefer to die rather than remain alive.

The function v measures to what extent transfers to heirs and bequests are valued by the agent. This is a shortcut to take into account the agent's altruism, and measure how the agent cares for his heir. Such a modeling choice for bequests has already been made in the literature, for example by Hurd and Smith (2002), De Nardi (2004), Kopczuk and Lupton (2007), De Nardi, French and Jones (2010), Ameriks, Caplin, Laufer, and Van Nieuwerburgh (2011), Lockwood (2012a) and (2012b).

3.3 Agent's program

The agent's program is:

$$\max_{c_1, a, s, c_2} p\phi(u_1(c_1) + u_2(c_2) + v(\tau)) + (1-p)\phi(u_1(c_1) + v((1+R)s)), \quad (2)$$

subject to the following constraints:

$$c_1 + a + s = W_0, \quad (3)$$

$$c_2 + \tau = (1+R)s + a \frac{1+R}{p}, \quad (4)$$

$$c_1 > 0, c_2 > 0, \tau \geq 0, a \geq 0, s \geq 0. \quad (5)$$

Equation (2) is the agent's expected utility. With probability p , he lives for two periods and consumes successively c_1 and c_2 and hands down τ to his heirs. Otherwise, he only lives for one period and his savings in the riskless bonds are left to his heir as a bequest. Equations (3) and (4) are the budget constraints of the first and second periods. Finally, conditions in (5) state that consumption has to be strictly positive and transfers, savings and annuity holdings cannot be negative. The agent is therefore not permitted to hand down a debt to his heirs or take resources from them. Moreover, the agent is prevented from issuing annuities.

When deriving the first order conditions from the agent's program, we need to account for the possibility of binding constraints for τ , s and a . Let us denote by U_D and U_A the lifetime felicity obtained when the agent lives for one or two periods:

$$U_D = u_1(c_1) + v((1+R)s),$$

$$U_A = u_1(c_1) + u_2(c_2) + v(\tau).$$

Denoting μ_1 and μ_2 the Lagrange multipliers associated respectively with the first and second period budget constraints (3) and (4), we obtain that the first order conditions of the agent's program (2)–(5) are:

$$(p\phi'(U_A) + (1-p)\phi'(U_D))u'_1(c_1) = \mu_1, \quad (6)$$

$$p\phi'(U_A)u'_2(c_2) = \mu_2, \quad (7)$$

$$\frac{\mu_2}{p} - \frac{\mu_1}{1+R} \leq 0 \quad (= 0 \text{ if } a > 0), \quad (8)$$

$$v'((1+R)s)\phi'(U_D) - \frac{1}{1-p} \left(\frac{\mu_1}{1+R} - \frac{\mu_2}{p} \right) - \frac{\mu_2}{p} \leq 0 \quad (= 0 \text{ if } s > 0), \quad (9)$$

$$p v'(\tau)\phi'(U_A) - \mu_2 \leq 0 \quad (= 0 \text{ if } \tau > 0). \quad (10)$$

Equations (8) to (10) are inequalities, as the optimal values for a , s and τ may correspond to corner solutions. These inequalities become equalities whenever interior solutions are obtained.

3.4 Saving choices

We first consider the case where the function ϕ is linear, as it is usually assumed to be. The results obtained in that case are well known and are discussed for example in Davidoff, Brown and Diamond (2005), and Lockwood (2012a). We formalize these findings in our setup to contrast them later on with results derived when the function ϕ is no longer assumed to be linear.

Proposition 1 (Annuities and savings with linear ϕ) *If ϕ is linear, then the amount in-*

vested in annuities equals the present value of the second period consumption. Everything invested in bonds is left to the heirs (through bequest or inter-vivos transfers). More formally:

$$a = \frac{pc_2}{1+R} \text{ and } (1+R)s = \tau.$$

Proof.

The proof can be found in the appendix and relies on the analysis of equations (8)–(10), in the particular case when ϕ is linear and $\phi'(U_A) = \phi'(U_D) = \phi'(0) = 1$. ■

The above proposition shows that, when ϕ is linear, people should purchase an amount of annuities that will exactly finance their future consumption. Intergenerational transfers, which materialize either through bequest or inter-vivos transfers, are independent of life duration. Riskless savings only help to finance the bequest, but do not contribute at all to financing consumption, regardless of the strength of the bequest motive.

We now consider the case when the agent's preferences exhibit positive lifetime risk aversion, i.e. the case of a strictly concave function ϕ .

Proposition 2 (Optimal annuitization with a concave ϕ) *If ϕ is strictly concave and $c_2 > c_2^*$ at the optimum (i.e., the agent prefers to survive), then:*

- either savings and bequest are null: $s = \tau = 0$,
- or capitalized savings are larger than inter-vivos transfers and the annuities do not fully finance second period consumption:

$$(1+R)s > \tau \text{ and } a < \frac{pc_2}{1+R}.$$

Proof. Let us first remark that $c_2 > c_2^*$ implies $u_2(c_2) > 0$ and $U_A - U_D > v(\tau) - v((1+R)s)$. We distinguish two cases: $s = 0$ and $s > 0$.

– $s = 0$. The budget constraint (4) implies that $a > 0$. From (9) using (8) as an equality, we deduce $v'(0)\phi'(U_D) \leq \frac{\mu_2}{p}$. Suppose that $\tau > 0$. We obtain from the previous inequality and (10) as an equality that $v'(0)\phi'(U_D) \leq v'(\tau)\phi'(U_A)$. Since $U_A - U_D > 0$ and ϕ is increasing and strictly concave, $0 < \phi'(U_A) < \phi'(U_D)$ and thus $v'(0) < v'(\tau)$, contradicting the fact that v is concave. We deduce therefore that $s = \tau = 0$.

– $s > 0$. Suppose that $(1+R)s \leq \tau$. It implies $v(\tau) - v((1+R)s) \geq 0$ and $U_A - U_D > 0$. Moreover, the budget constraint (4) implies $a \frac{1+R}{p} = c_2 + \tau - (1+R)s > 0$. Equations (8)–(10) are equalities and yield:

$$\phi'(U_D)v'((1+R)s) = v'(\tau)\phi'(U_A),$$

which implies that $\frac{\phi'(U_D)}{\phi'(U_A)} = \frac{v'(\tau)}{v'((1+R)s)} \leq 1$ in contradiction with $U_D < U_A$ and ϕ strictly concave. We therefore deduce that $\tau < (1+R)s$, and from the budget constraint that $a < \frac{pc_2}{1+R}$, which ends the proof.

■

As soon as the agent is risk averse with respect to lifetime felicity, and willing to leave some transfer or bequest, he should not completely annuitize his consumption. Riskless savings contribute to financing not only transfers to the heir but also the agent's consumption. Transfers received by the heirs will depend on life duration, shorter lives being associated with greater transfers. The agent, who cannot eliminate the possibility of an early death, achieves some partial self insurance by creating a negative correlation between two aspects he thinks desirable: living long and transferring resources to his heir.

To establish further results about risk aversion and annuity demand, we need to make slightly stronger assumptions regarding the willingness to live and to make transfers. More precisely, we make the following assumption:

Assumption A Denote by $c_2^{**} = \inf\{c_2 | u_2(c_2) > v(c_2)\}$. We assume that c_2^{**} exists and that:

1. $u(c) > v(c)$ for all $c > c_2^{**}$,
2. $\frac{u_1(W_0 - \frac{c_2^{**}}{1+R})}{1+R} < v'(c_2^{**})$,
3. $v'(0) < u'_2(c_2^{**})$.

The consumption level c_2^{**} is the smallest second period consumption level that makes the agent's life worthwhile, once accounting for the possibility of bequeathing to the heir. Below that level of consumption, the agent would rather die and hand down all his wealth. The consumption level c_2^{**} is larger than c_2^* defined in Equation (1), which does not account for the possibility of making intergenerational transfers. The three points of the above assumption can be interpreted as follows. Point 1 simply states that any agent enjoying a second period consumption greater than c_2^{**} would prefer to live than to die and bequeath all this consumption to his heirs. Point 2 means that the bequest motive is sufficiently strong in the sense that if the agent was sure to die after period 1, he would leave at least c_2^{**} to his heirs. The last point states that the bequest motive is not too strong, in the sense that the agent living at the second period and endowed with the survival consumption level c_2^{**} is not willing to make any inter-vivos transfers.

We make a further assumption regarding the functional form of the concave transformation ϕ .

Assumption B The function ϕ is of CARA type: $\phi(x) = -\frac{e^{-\lambda x}}{\lambda}$, where $\lambda > 0$.

We specify the aggregator ϕ to have an exponential functional form, such that the resulting preferences are multiplicative. The parameter λ drives the concavity of the aggregator, and therefore the degree of lifetime risk aversion. The larger the coefficient λ , the more risk averse the agent is. As underlined in Bommier (2013), multiplicative preferences enable elasticity of substitution to be disentangled from risk aversion while remaining in the expected utility framework and retaining the assumption of stationarity. In consequence, choices resulting from these preferences are time-consistent and history independent.¹³

We can now state the following result:

Proposition 3 (Decreasing and null annuity demand) *Under Assumptions A and B, the optimal annuity demand is a decreasing function of the lifetime risk aversion λ . Moreover, there exists $\lambda_0 > 0$, such that for all λ greater than λ_0 , the optimal annuity purchase is null.*

Proof. The proof is in two steps. In the first step, we prove by contradiction that there exists a level of risk aversion λ_0 such that the annuity demand is null. Since satisfying the first-order conditions (6)–(10) allows for several possible corner solutions for s and τ , we show that in each case, we arrive at a contradiction. In the second step, we prove that an increase in lifetime risk aversion λ implies a smaller annuity demand a . Again, due to possible corner solutions, we need to consider several cases.¹⁴ ■

Under Assumptions A and B, we are able to derive two forceful conclusions concerning annuity demand. First, the annuity demand is decreasing with lifetime risk aversion. More risk averse agents prefer to purchase fewer annuities. They are more reluctant to take the risk of dying young without leaving a significant amount of bequest. Moreover, the demand for annuities not only diminishes with lifetime risk aversion but also vanishes for sufficiently large levels of lifetime risk aversion. Accounting for lifetime risk aversion may then provide an explanation for the annuity puzzle that holds even when assuming a perfect annuity market.

4 A calibrated model

In this section, we extend our model to a large number of retirement periods so as to calibrate it using realistic mortality patterns and make predictions relating to agents' saving behavior. The section is split into four parts. The first one details the structure of the extended model, and the method to solve it. We also explain how the model compares to the standard additive model, which is considered as a benchmark. The second part describes how both the additive and the

¹³The issues of time inconsistency and history independence do not arise in the two-period framework that is considered in the current Section. However they would do so in the N -period extension considered in Section 4.

¹⁴The formal proof is relegated to the Appendix B.

multiplicative models are calibrated. The third part provides the results derived from the calibrated models, while the last one proceeds with a sensitivity analysis.

4.1 The N -period model extension

4.1.1 The setting

We extend our setup to N periods. As with the two period model, we normalize the retirement date to the date 0 of the model.¹⁵ Mortality remains the sole risk faced by the agent and $p_{t+1|t}$ denotes the probability of remaining alive at date $t+1$ while being alive at date t . Thus, $1 - p_{t+1|t}$ denotes the probability of dying at the end of period t . The agent is alive at date 0, so that: $p_{0|-1} = 1$. We denote by $m_{t|0}$ (resp. $p_{t|0}$) the probability of living exactly (resp. at least) until date t . These probabilities relate to each other as follows:

$$m_{t|0} = (1 - p_{t+1|t}) \prod_{k=1}^t p_{k|k-1} \quad \text{and} \quad m_{0|0} = 1 - p_{1|0},$$

$$p_{t|0} = \prod_{k=1}^t p_{k|k-1} \quad \text{and} \quad p_{0|-1} = 1.$$

The agent is endowed with wealth W_0 when he retires at date 0. In addition to his wealth, he receives a constant periodic income y , while he is alive. This income can be interpreted as an exogenous pension benefit. In order to smooth resources over time and states of nature, we assume that the agent can trade two kinds of financial products: bonds and annuities. A bond is a security of price 1 which pays $1 + R$ in the subsequent period, either to the bond holder or, if he dies, to his heirs. The riskless rate of interest R is constant and exogenous. An annuity is a financial product, which pays off one consumption unit every period following the purchase date, as long as the annuity holder is alive. We assume that the annuity market is perfect, and that the pricing is actuarially fair. This implies that the price π_t of one annuity purchased at date t can be expressed as the present value of the single amount paid every period, conditional on the agent being alive:

$$\pi_t = \sum_{k=1}^{\infty} \frac{p_{t+k|t}}{(1+R)^k} = (1 + \pi_{t+1}) \frac{p_{t+1|t}}{1+R}. \quad (11)$$

The number of annuities purchased (or sold back) at age t is denoted a_t , while the number of bonds held is s_t . We assume that agents can sell back the annuities they hold at any time, which means that the flow of annuity purchase a_t can be positive or negative. However, they cannot issue annuities and cannot therefore hold a negative stock of annuities: $\sum_{k=0}^t a_k \geq 0$ should hold for all

¹⁵We do not consider endogenous retirement decisions. This aspect is formalized in Chai, Horneff, Maurer, and Mitchell (2011).

$t \geq 0$. As agents cannot leave negative transfers to their heirs, we impose that $s_t \geq 0$ for all t . From now on, we refer to the income y as public annuities, contrasting it with private annuities (a_t). We refer to the quantity of bonds (s_t) as being the riskless savings of the agent.

We do not explicitly introduce inter-vivos transfers in this N -period setting as they would be redundant with transfers made through bequest. Indeed, given that what will matter is the present value of transfers received by the heir, making an inter-vivos transfer of δ at time t is equivalent to changing s_τ to $s_\tau + \delta(1+R)^{\tau-t}$ at all periods $\tau \geq t$.

4.1.2 The multiplicative specification

As for the two period model, we assume that preferences are weakly separable, but we allow for lifetime risk aversion. The agent cares for the present value of the bequest he hands down to his heirs. Precisely, we assume that leaving an amount of bequest w_t in period t provides a felicity $v(\frac{w_t}{(1+R)^t})$. Thus, an agent who dies at time t and holds s_t bonds, leaves a bequest $w_{t+1} = (1+R)s_t$. The heir receives that amount in period $t+1$, which provides the agent a felicity $v(\frac{s_t}{(1+R)^t})$. Therefore, living for t periods, with a stream of consumption $(c_k)_{0 \leq k \leq t}$, and a bond holding s_t at death, provides the following felicity:

$$U(c, s) = -\frac{1}{\lambda} \exp \left(-\lambda \left(\sum_{k=0}^t u(c_k) + v \left(\frac{s_t}{(1+R)^t} \right) \right) \right).$$

As in the previous section (Assumption B), we assume that the aggregator is exponential, where $\lambda > 0$ drives the lifetime risk aversion. We call such a model the multiplicative model,¹⁶ so as to contrast it with the standard additive model that will be precisely specified in Section 4.1.4.

The agent maximizes his expected intertemporal utility by choosing his consumption stream $(c_t)_{t \geq 0}$, his bond saving $(s_t)_{t \geq 0}$ and annuity purchase $(a_t)_{t \geq 0}$, subject to per period budget con-

¹⁶The utility function $U(c, s)$ may also be written as

$$U(c, s) = -\frac{1}{\lambda} e^{-\lambda v \left(\frac{s_t}{(1+R)^t} \right)} \prod_{k=0}^t e^{-\lambda u(c_k)},$$

where the multiplicative structure is explicit.

straints. The agent's program can therefore be expressed as follows:

$$\max_{c,s,a} -\frac{1}{\lambda} \sum_{t=0}^{\infty} m_{t|0} \exp \left(-\lambda \left(\sum_{k=0}^t u(c_k) + v \left(\frac{s_t}{(1+R)^t} \right) \right) \right), \quad (12)$$

$$\text{s.t. } W_0 + y = c_0 + s_0 + \pi_0 a_0, \quad (13)$$

$$y + (1+R)s_{t-1} + \sum_{k=0}^{t-1} a_k = c_t + s_t + \pi_t a_t \text{ for } t \geq 1, \quad (14)$$

$$c_t \geq 0, s_t \geq 0, \sum_{k=0}^t a_k \geq 0. \quad (15)$$

It is noteworthy that there is no exogenous time discounting in this model. Time discounting is endogenous and stems from the combination of mortality risk and lifetime risk aversion (see Bommier (2006) or equation (29) later on).

The first order conditions of the previous program can be expressed as follows:

$$u'(c_t) \sum_{k=t}^{\infty} m_{k|0} \exp \left(-\lambda \sum_{j=0}^k u(c_j) - \lambda v \left(\frac{s_k}{(1+R)^k} \right) \right) = \mu_t, \quad (16)$$

$$\frac{m_{t|0}}{(1+R)^t} v' \left(\frac{s_t}{(1+R)^t} \right) \exp \left(-\lambda \sum_{k=0}^t u(c_k) - \lambda v \left(\frac{s_t}{(1+R)^t} \right) \right) = \mu_t - (1+R)\mu_{t+1} \quad (17)$$

(the previous equality holds for $s_t > 0$ and shifts to \leq if $s_t = 0$),

$$\pi_t \mu_t = \sum_{k=t+1}^{\infty} \mu_k \quad (\text{the equality becomes } \geq \text{ if } \sum_{k=0}^t a_k = 0). \quad (18)$$

In the previous equations, the parameter μ_t is the Lagrange multiplier of the budget constraint of date t , or the shadow cost of unit of extra consumption at date t . Since equation (18) also means that $\mu_t \pi_t = \mu_{t+1}(1 + \pi_{t+1})$, we obtain the following intertemporal relationship for μ_t :

$$p_{t+1|t} \mu_t = \mu_{t+1}(1+R) \text{ if } \sum_{k=0}^t a_k > 0. \quad (19)$$

Equation (19) states that the shadow cost of the budget constraint at date $t+1$ is equal to the discounted shadow cost of date t , where the discount takes the probability of dying into account.

From now on, we assume that there exists $T_M < \infty$, such that the probability of remaining

alive after T_M is null: $p_{T_M+1|T_M} = 0$. Plugging equation (19) into (16) and (17) leads to:

$$\begin{aligned} u'(c_t) \sum_{k=t}^{T_M} m_{k|0} e^{-\lambda \sum_{j=t+1}^k u(c_j) - \lambda v\left(\frac{s_k}{(1+R)^k}\right)} \\ = \frac{1+R}{p_{t+1|t}} u'(c_{t+1}) \sum_{k=t+1}^{T_M} m_{k|0} e^{-\lambda \sum_{j=t+1}^k u(c_j) - \lambda v\left(\frac{s_k}{(1+R)^k}\right)} \text{ if } \sum_{k=0}^t a_k > 0, \end{aligned} \quad (20)$$

$$\begin{aligned} = \frac{m_{t|0}}{(1+R)^t} v'\left(\frac{s_t}{(1+R)^t}\right) e^{-\lambda v\left(\frac{s_t}{(1+R)^t}\right)} \\ + (1+R)u'(c_{t+1}) \sum_{k=t+1}^{T_M} m_{k|0} e^{-\lambda \sum_{j=t+1}^k u(c_j) - \lambda v\left(\frac{s_k}{(1+R)^k}\right)} \text{ if } s_t > 0. \end{aligned} \quad (21)$$

The first intertemporal Euler equation (20) is valid for every date t between 0 and $T_M - 1$. It sets as being equal the marginal cost of saving one unit of good today to the marginal cost of consuming one unit more tomorrow. The second Euler equation (21) is true for all dates t between 0 and T_M and equalizes the marginal cost of saving one unit more today to the marginal benefit of one additional unit requested tomorrow.

4.1.3 Implementation

In order to solve the model, we take advantage of the choice of an exponential function ϕ which provides a recursive structure to the agent's utility function. As a consequence, the first order conditions (20) and (21) of date t are independent of any past variables and a backward algorithm can be readily implemented. We start from a guess for the final value of consumption c_{T_M} at date T_M . The backward resolution of the model then yields a unique wealth endowment, compatible with that terminal of level c_{T_M} . We then search for the value of c_{T_M} such that the associated wealth endowment corresponds to the desired initial wealth W_0 .¹⁷

4.1.4 Additive specification

In order to highlight the role of lifetime risk aversion, we consider a benchmark model, in which the intertemporal utility of the agent is a sum of discounted instantaneous felicities. The discount parameter $\beta > 0$ represents the agent's exogenous time preference. This model is very similar to those of De Nardi (2004), De Nardi, French and Jones (2010), Pashchenko (2013), Lockwood (2012b) and (2012a), and Ameriks, Caplin, Laufer and Van Nieuwerburgh (2011). More precisely,

¹⁷All computational codes (in Matlab) are available upon request to authors.

using the same notations as before, the agent's program can be expressed as follows:

$$\max_{c,b,a} \sum_{t=0}^{\infty} \beta^t p_{t|0} u(c_t) + m_{t|0} v\left(\frac{s_t}{(1+R)^t}\right), \quad (22)$$

$$\text{s.t. } W_0 + y = c_0 + s_0 + \pi_0 a_0, \quad (23)$$

$$y + (1+R)s_{t-1} + \sum_{k=0}^{t-1} a_k = c_t + s_t + \pi_t a_t \text{ for } t \geq 1, \quad (24)$$

$$c_t \geq 0, \quad s_t \geq 0, \quad \sum_{k=0}^t a_k \geq 0. \quad (25)$$

In contradistinction to the multiplicative model, we refer to this model as the additive model.

The agent's program yields the following first order conditions:

$$u'(c_t) = \beta(1+R)u'(c_{t+1}) \text{ if } \sum_{k=0}^t a_k \geq 0, \quad (26)$$

$$u'(c_t) - p_{t+1|t}\beta(1+R)u'(c_{t+1}) = \frac{1-p_{t+1|t}}{(1+R)^t} v'\left(\frac{s_t}{(1+R)^t}\right) \text{ if } s_t \geq 0. \quad (27)$$

Equations (26) and (27) imply that $\beta^t(1+R)^t u'(c_t) = v'\left(\frac{s_t}{(1+R)^t}\right) = \beta^{t+1}(1+R)^{t+1} u'(c_{t+1}) = v'\left(\frac{s_{t+1}}{(1+R)^{t+1}}\right)$. From these equalities, it is straightforward to deduce that $\frac{s_t}{(1+R)^t} = \frac{s_{t+1}}{(1+R)^{t+1}}$ as long as we have an interior solution. The discounted value of saving is constant over age. This means that the heir enjoys a bequest whose present value is independent of his parents' life duration. As a result, riskless saving only aims at leaving bequest, while private annuities fully finance consumption. The agent's budget constraint at any date t can be simplified to $y + \sum_{k=0}^{t-1} a_k = c_t + \pi_t a_t$, in which the bond saving quantity does not intervene. As a result, in the additive model, saving in riskless bonds and purchasing private annuities are two independent decisions, which fulfill two independent purposes. This is not the case in the multiplicative model, where private annuities and riskless savings are nested decisions, with both contributing to finance consumption.

4.2 Calibration

We need to calibrate both the multiplicative and the additive models. First of all, we specify our felicity functions u and v . We assume that the agent has a constant intertemporal elasticity of substitution, which means that $-\frac{u'(c)}{c u''(c)}$ is constant, or equivalently that:

$$u(c) = u_0 + \frac{c^{1-\sigma}}{1-\sigma},$$

where the parameter $\sigma > 0$ is the inverse of the intertemporal elasticity of substitution and u_0 a constant. Since u is normalized by a zero felicity for death ($u(d) = 0$), we cannot impose u_0 to be equal to zero. This constant u_0 determines how wide the felicity gap is between being alive and dead, and will have an impact on the optimal consumption and saving plans in the multiplicative model.

Regarding the felicity derived from bequest, we assume that it has the following form:

$$v(w) = \frac{\theta}{1 - \sigma} \left(y_0 + \frac{w}{\psi} \right)^{1 - \sigma}. \quad (28)$$

This functional expression represents a kind of altruism, and accounts for the fact that bequest only comes in addition to other resources the heirs may dispose of. The parameter θ drives the intensity of altruism. With $y_0 > 0$, bequests are a luxury good, as reported in the data (e.g., in Hurd and Smith (2002)). Moreover, the value $v'(0)$ is finite, so that agents bequeath only when their wealth is large enough. This functional form has been chosen for example in De Nardi (2004), De Nardi, French, and Jones (2010), Lockwood (2012a) and (2012b) and Ameriks, Caplin, Laufer and Van Nieuwerburgh (2011).

Regarding our calibration, we proceed in two ways: (i) we fix exogenously some parameters to values that seem reasonable and (ii) we choose some parameter values to match given quantities, like the endogenous rate of time discounting, the value of a statistical life and the average bequest.

4.2.1 Exogenous calibration

First of all, we normalize date 0 of the model as corresponding to the age of 65, assuming that people retire at that age. Mortality data are US 2000 mortality data from the Human Mortality Database. In the data, the maximal age is 110 years. People alive at the age of 65 will live at most for 45 years. This implies that $T_M = 45$ and $p_{46|45} = 0$.

We posit the exogenous rate of return of savings to be equal to 3.00%, which is close to the historical value of the riskless short term interest rate approximated by the three-month T-bond.

We also exogenously calibrate some preference parameters. First, for both functions u and v , we adopt $\sigma = 2$ corresponding to a standard value of 1/2 for the intertemporal elasticity of substitution. Second, for the parameters y_0 and ψ entering the function v , we follow Lockwood's (2012a) approach. The idea is that $y_0 + \frac{w}{\psi}$ represents the per-period consumption of the heir, such that $\left(y_0 + \frac{w}{\psi} \right)^{1 - \sigma}$ is proportional to his lifetime utility. For this, y_0 is set equal to the periodic income y and ψ is interpreted as an actualization parameter which would reflect how bequest may impact consumption. In order to take a plausible value for ψ , we consider that the agent's heir fully annuitizes the bequest. In the model, the agent retires at the age of 65 where life

expectancy is about 18 years. The coefficient ψ must therefore take into account the fact that the real bequest at the age of 65 needs to be capitalized for 18 years on average. Assuming that the age difference between parents and children is approximately 27, the discount factor ψ reflects the value of annuities at the age of 56.¹⁸ We deduce that $\psi = \frac{\pi_{56}}{(1+R)^{18}} = 9.39$ where π_{56} is the value of an annuity at the age of 56.

Finally, we choose the agent's wealth W_0 to be normalized to 1. The present value of the agent's income $N = \sum_{k=0}^{T_M} \frac{p_{k+1|0}}{(1+R)^k} y$ is set equal to W_0 . The quantity N can also be interpreted as the agent's wealth, which has already been annuitized. The non-annuitized wealth W_0 is thus equal to half of total wealth. Lockwood (2012a) considers the cases, where the non-annuitized wealth equals one third and two thirds of the total wealth without any major impact on his conclusions.

4.2.2 Evaluated parameters

We still have to calibrate the following parameters: u_0 driving the gap in felicity between being alive and dead, the strength of bequest motive θ and the lifetime risk aversion λ (in the multiplicative model) or the exogenous time discount β (in the additive model). The calibration aims to replicate three "observable" quantities: the average bequest, the value of a statistical life (VSL) and the rate of time discounting at the retirement age of 65 that we note ρ_0 . Before providing targets for these quantities, we explain how they are defined.

Average bequest. We define the average bequest \bar{w} as the expected discounted value of bequest:

$$\bar{w} = \sum_{t=0}^{\infty} m_{t|0} \frac{w_{t+1}}{(1+R)^{t+1}} = \sum_{t=0}^{\infty} m_{t|0} \frac{s_t}{(1+R)^t},$$

where w_{t+1} is the bequest amount left by an agent dying at the end of period t .

Rate of time discounting. Conventionally, the rate of time discounting ρ_0 at the retirement age (date 0) is defined by:

$$\rho_0 = \left. \frac{\frac{\partial EU}{\partial c_0}}{\frac{\partial EU}{\partial c_1}} \right|_{c_0=c_1} - 1.$$

This quantity is interpreted as being the rate of change of marginal utility, in which we offset the consumption effect. The relationship between the rate of discounting and the parameters depends on the structure of the model. To avoid possible confusion, we use different notations, respectively ρ_0^{mul} (for the multiplicative case) and ρ_0^{add} (for the additive model) when referring to the rate of time discounting but using expressions relating to the structure of the model. Simple calculation

¹⁸See for example the report of Livingston and Cohn (2010) on American motherhood.

leads to the following expressions:

$$\rho_0^{mul} = \frac{m_{0|0} \exp(-\lambda v(s_0))}{\sum_{t=1}^{\infty} m_{t|0} \exp\left(-\lambda \sum_{k=1}^t u(c_k) - \lambda v\left(\frac{s_t}{(1+R)^t}\right)\right)}, \quad (29)$$

$$\rho_0^{add} = \frac{1 - p_{1|0}}{\beta p_{1|0}} - 1. \quad (30)$$

Value of life. The value of a statistical life VSL_0 at the retirement age can be expressed as the opposite of the marginal rate of substitution between the mortality rate and consumption at that age. Noting $q_{1|0} = p_{1|0}^{-1} - 1$ the mortality rate at the retirement age, we define VSL as follows:

$$VSL_0 = - \frac{\frac{\partial EU}{\partial q_{1|0}}}{\frac{\partial EU}{\partial c_0}}.$$

The quantity VSL_0 corresponds to the quantity of consumption an agent would be willing to relinquish to save one statistical life. Our definition of VSL is similar to Johansson's (2002).

Again, although the notion of VSL is independent of the choice of one particular model, we will introduce specific notations when working with specific models. Formulas providing VSL_0 in the multiplicative and additive cases are given by:

$$VSL_0^{mul} = p_{1|0} \frac{\exp(-\lambda u(c_0) - \lambda v(s_0)) - \sum_{t=0}^{\infty} m_{t|0} \exp\left(-\lambda \sum_{k=0}^t u(c_k) - \lambda v\left(\frac{s_t}{(1+R)^t}\right)\right)}{\lambda u'(c_0) \sum_{t=0}^{\infty} m_{t|0} \exp\left(-\lambda \sum_{k=0}^t u(c_k) - \lambda v\left(\frac{s_t}{(1+R)^t}\right)\right)}, \quad (31)$$

$$VSL_0^{add} = p_{1|0} \frac{-(u(c_0) + v(s_0)) + \sum_{t=0}^{\infty} \left(p_{t|0} \beta^t u(c_t) + m_{t|0} v\left(\frac{s_t}{(1+R)^t}\right)\right)}{u'(c_0)}. \quad (32)$$

Benchmark calibration. In the benchmark calibration, we consider the three following targets. First, the average bequest is equal to 20% of the initial wealth W_0 . This value approximately matches the calibration of Lockwood (2012a), who reports an average bequest of 22.6% of the non-annuitized wealth. Second, the rate of time discounting at age 65 equals 5%. This rate of discount generates a consumption rate of growth of -0.1% per year at the age of 65. A decrease in consumption is indeed reported in most studies using micro-level data to assess the consumption profile per age (Japelli (1999) and Fernández-Villaverde and Krueger (2007) among others). Third, the value of a statistical life at age 65 equals 500 times the annual consumption. This fits in with the range of estimates provided in Viscusi and Aldy (2003).

Our benchmark calibration is finally summed up in Table 1. We will investigate the sensitivity of our findings to various values of calibration in the robustness section.

In table 1, the coefficient λ may seem small, but this coefficient should not be interpreted without considering the value of instantaneous utility, with which it multiplies. If we take a con-

Calibration			
Multiplicative model		Additive model	
Exogenous Parameters			
σ		2.0	
W_0		1.0	
$N = \sum_{k=0}^{T_M} \frac{p_{k+1 0}}{(1+R)^k} y$		1.0	
R		3.00%	
y_0		y	
ψ		9.39	
Estimated Parameters			
u_0	157.72	u_0	315.84
λ	4.81×10^{-4}	β	0.969
θ	4.523	θ	4.715

Table 1: Benchmark calibration

sumption c equal to the average agent's consumption, then $\lambda u(c)$ equals 0.072 per year, indicating a coefficient of risk aversion with respect to life duration of 0.072 per year.¹⁹ Moreover, this lifetime risk aversion generates a reasonable relative risk aversion with respect to wealth. It is found to be equal to 0.84 at age 65 in the multiplicative model, not far from the value of 0.74 obtained in the additive one.²⁰

4.3 Results

Our results aim at discussing both the strength of the demand for annuities, and the role annuities would play for consumption smoothing if markets were perfect. Before exposing our results in detail, we want to highlight that even in this extended set-up our main theoretical findings of Proposition 3 still hold. In particular, the annuity demand still decreases with the risk aversion parameter λ and is null (i.e. $a_k = 0$ for all k) for a sufficient large λ .²¹

We expose our results in three steps. We first investigate how much an individual would be willing to pay to have access to a perfect annuity market. This is a standard way of measuring the welfare impact of annuities. Second, we explain to what extent individuals would rely on annuities to finance their consumption if annuities were available at fair prices. Last, we look at the consequences in terms of consumption smoothing.

¹⁹With such a risk aversion with respect to life duration, an agent would be indifferent between living 80 years for sure, or living 78 years or 82.34 years with equal probability.

²⁰The relative risk aversion with respect to wealth at the age of 65 is $-W_0 \frac{\partial^2 EU_{65}}{\partial W_0^2} / \frac{\partial EU_{65}}{\partial W_0}$, where EU_{65} is the expected lifetime utility at the age of 65.

²¹Keeping unchanged the other parameters of our benchmark calibration (Table 1), we find that people never purchase annuities when λ is larger than 0.0133.

4.3.1 Willingness-to-pay for annuities

In order to measure the strength of demand for annuities, we compute the “willingness-to-pay for annuities” (WTP, hereafter), which is defined as the fraction of the non-annuitized wealth an agent would be likely to relinquish to gain access to the private annuity market, rather than being in a world where these annuities do not exist. In other words, an agent endowed with the non-annuitized wealth W_0^{na} and without access to an annuity market would be as well off as an agent endowed with the wealth $W_0 = (1 - WTP) \times W_0^{na}$ but having access to a perfectly fair annuity market. The larger the WTP, the more valuable the annuity market for the agent. This measure is conventional in the literature, and was used for example by Mitchell, Poterba, Warshawsky, and Brown (1999) or more recently by Lockwood (2012a).

	Multiplicative model	Additive model
Target values used for calibration		
Rate of time discounting ρ_0	5.00%	
Value of statistical life VSL	$500 \times c_0$	
Average bequest \bar{w}	$20\% \times W_0$	
Results		
Willingness-to-pay WTP	1.22%	6.86%
Share of consumption financed by:		
Public annuities $\%c/y$	56.42%	55.68%
Private annuities $\%c/a$	9.46%	44.32%
Riskless savings $\%c/b$	34.12%	0.00%

Table 2: Results for the benchmark calibration

The first line in Table 2 shows the difference in WTP between the two models. In both cases, the WTP is positive. The fact that we observe a positive WTP in the multiplicative model means that with a reasonable calibration, lifetime risk aversion is not significant enough to deliver the zero annuity result of Proposition 3. The difference in prediction between the additive and multiplicative models is however quite substantial. While the additive model predicts a WTP of 6.86%, it is more than five times smaller with the multiplicative model, where the WTP is only 1.22%. Although both models were calibrated to provide the same average amount of bequest, the lack of annuities is less penalizing for agents with multiplicative preferences than for those with additive preferences.

4.3.2 Annuities and consumption financing

The second set of results we present relies on the role of annuities in consumption smoothing in the case when individuals have access to a perfect credit market. More precisely, at any age t we break down the agent’s consumption into three components reflecting the respective role of public annuities, riskless savings and private annuities. Indeed, the budget constraint imposes that:

$$c_t = y + [(1 + R)s_{t-1} - s_t] + \left[\sum_{k=0}^{t-1} a_k - \pi_t a_t \right], \quad (33)$$

which means that consumption at age t is financed through public annuities y , the decumulation of riskless savings $(1 + R)s_{t-1} - s_t$ and finally the decumulation of private annuities $\sum_{k=0}^{t-1} a_k - \pi_t a_t$. In Table 2, we report the average shares of consumption financed by public annuities, private annuities and riskless saving. The average is computed over the agent's lifetime and the survival probabilities are taken into account. The main difference lies in the fact that while the additive model predicts that consumption should be fully financed out of (private or public) annuities, the multiplicative model predicts that more than a third of consumption should be financed by riskless savings. This finding is consistent with empirical studies, such as Johnson, Burman, and Kobes (2004), who report that in 1999, one third of the consumption of US people over 65 was financed by decumulation of their savings.

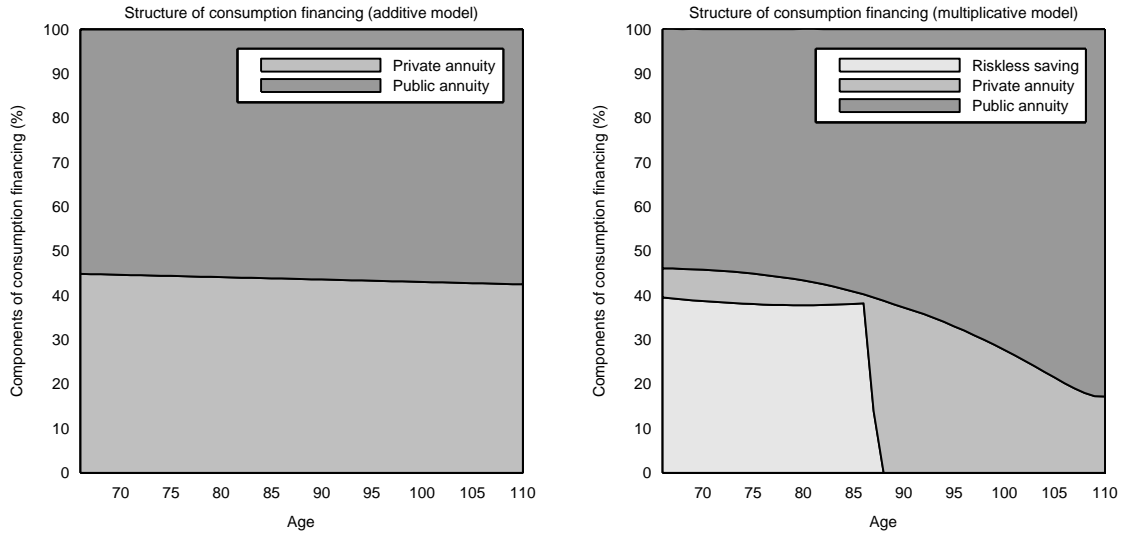


Figure 1: Consumption financing structure

In Figure 1, it is shown how consumption financing varies with age. While in the additive model consumption is fully financed by public and private annuities, we find that in the multiplicative model, private annuities play a significant role only after age 87. Individuals would then ideally postpone their purchase of annuities, which would generate very serious problems of adverse selection in a realistic environment.

Another striking feature of Figure 1 is the sharp decline in the fraction of consumption financed by riskless saving. Equation (33) shows that the plotted expression is equal to $\frac{1}{c_t} [(1 + R)s_{t-1} - s_t]$, which means that it is basically minus the derivative of the stock of riskless savings. According

to Figure 2 hereafter, before age 87, this stock decreases almost linearly, and then it reaches zero where it remains because of the assumed positivity constraint until the end (people rely then on annuities and pension). In continuous time, the (opposite of the) derivative would simply be discontinuous, and jump from a positive value to zero. Thus the “rapid decline” would be vertical. In discrete time, this is not exactly vertical, as the derivative takes two periods to stabilize to zero, which explains the pattern in the fraction of consumption financed by riskless saving, as shown in the right hand side of Figure 1.

4.3.3 Consumption smoothing and bequest profiles

The graphs of Figure 2 reproduce consumption and bequest profiles as a function of age. The consumption profiles decline with age in both models, which is consistent with our calibration choice of 5.00% for the rate of time discounting. This rate implies a decline in consumption at the retirement age of approximately -0.10% . In the additive model, the decline remains constant over time, while it is increasing in the multiplicative model. Such an increasing decline is reported in many empirical studies investigating consumption profiles using micro-economic data (Fernández-Villaverde and Krueger (2007) among others). The multiplicative model generates therefore more realistic consumption data, as discussed in Bommier (2013).

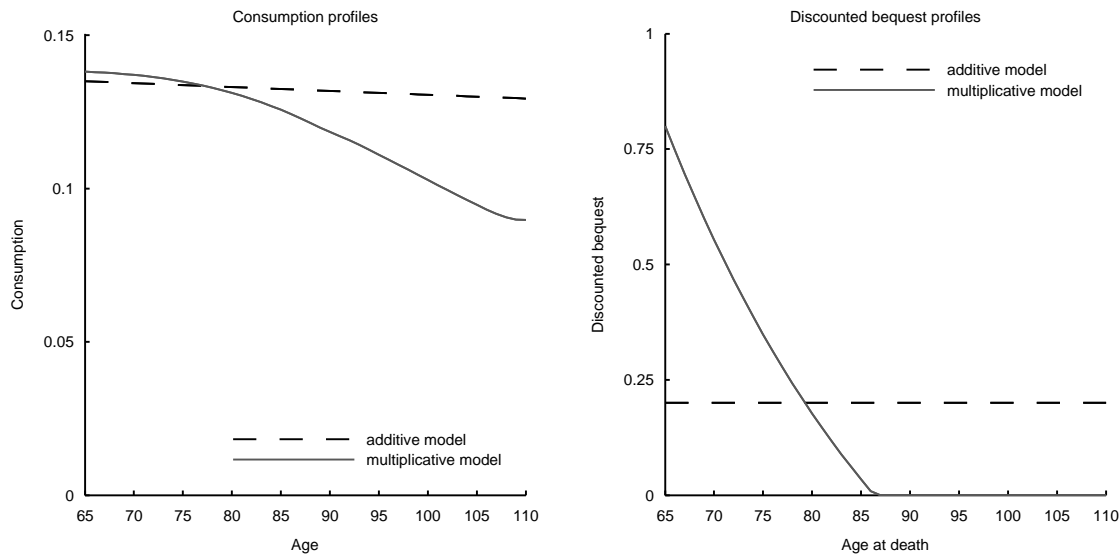


Figure 2: Consumption and bequest profiles

The discounted bequest profiles obtained with each model are also different. As underlined in the theoretical section, in the additive model, the discounted bequest is constant, which means that the present value of what the heir receives is independent of the agent’s life duration. The

multiplicative model provides a different picture, with an amount of bequest that declines with the age at death. The longer an agent lives, the smaller the bequest he leaves to his heirs. Although the agent could fully insure the amount of wealth he leaves to his heirs, being lifetime risk averse makes him choose a strategy that generates a negative correlation between life duration and bequest and avoids him leaving low bequest when dying young. This finding is consistent with the many empirical studies, such as Japelli (1999), that show that agents decumulate wealth as they grow older.²²

4.4 Result robustness

In order to assess the validity of our statement, we check several aspects. First, we study the sensitivity of our results to our calibration choices regarding the VSL, the rate of time discounting and the intensity of bequest motives. Second, we consider other model specifications that can be found in the literature on bequest.

Role of calibration over VSL and rate of time discounting. Our benchmark calibration assumes that the VSL is worth $500 \times c_0$, where c_0 is the agent's consumption at retirement age, while the rate of time discounting is chosen to be equal to 5.00%. To check the sensitivity of our results to these values, we simply rerun simulations for both the additive and the multiplicative models using a wide range of calibrations.²³

First, the VSL is assumed to vary between $200 \times c_0$ and $1000 \times c_0$, while the rate of time discounting remains unchanged. The impact of VSL is very small. An increase of VSL from $200 \times c_0$ and $1000 \times c_0$ has no effect on WTP in the additive model and barely diminishes it in the multiplicative model (it goes from 1.26% to 1.20%). The consequence on consumption smoothing is also practically negligible and the structure of consumption financing is barely affected. We do not reproduce graphs here.

Second, we keep the VSL unchanged to $500 \times c_0$ and the rate of time discounting varies from 4.00% to 6.00% (i.e., the yearly consumption rate of growth at 65 decreases from +0.38% to -0.53%). Results are plotted in Figure 3. As shown by the left-hand side graph, the WTP decreases from 7.9% to 6.1% in the additive model and from 2.2% to 0.8% in the multiplicative one. On the right-hand side graph, we observe that the rate of time discounting barely affects the structure of consumption financing in the additive model. However, in the multiplicative model, the share of private annuities substantially declines with the rate of time discounting (from approx.

²²The interpretation of empirical evidence on age specific wealth profiles should, however, be subject to caution. Indeed, saving decumulation can also be obtained under the assumption of risk neutrality with respect to lifetime felicity if annuities are not fairly priced.

²³We re-estimate the values of u_0 and λ (or β in the additive model) so as to match the different values of VSL and of rate of time discounting.

18% to 3% of the consumption on average). Indeed, in the multiplicative model, time discounting is generated by mortality and lifetime risk aversion. A greater rate of time discounting indicates that agents are in fact more risk averse. They are then more reluctant to purchase annuities, as annuities increase the risk on lifetime utility.

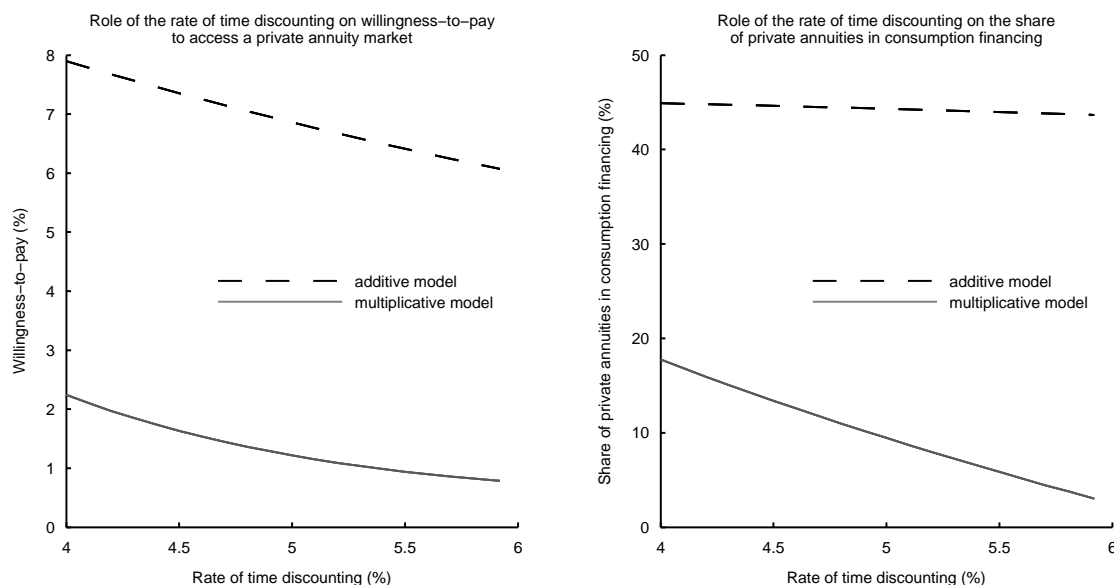


Figure 3: Impact of the rate of time discounting on willingness-to-pay for private annuities and the structure of consumption financing

Role of public annuities. In the benchmark calibration, we assumed that the present value of public annuities y was equal to the non-annuitized wealth W_0 . We now study the impact on our results of the share of public annuities in total wealth, which we vary from 0 to 60%. Results are plotted in Figure 4.

The pattern for WTP displayed on the left-hand side graph of Figure 4 is not surprising since it simply reveals a substitution between public and private annuities. The larger the share of public annuities in total wealth, the less the need for private annuities and thus the smaller the willingness-to-pay. This effect is present and its magnitude is similar for both models.

The impact of public annuities on financing of consumption is plotted in the right-hand side graph of Figure 4. Whenever there is a positive demand for private annuities, public annuities mechanically substitute for private annuities. As the additive model always predicts a positive demand for private annuities, this simple substitution effect is always at play. Public annuities simply crowd out private annuities. With the multiplicative model, the story could be a bit more complex, as the demand for private annuities may become null at some ages. But overall, this has

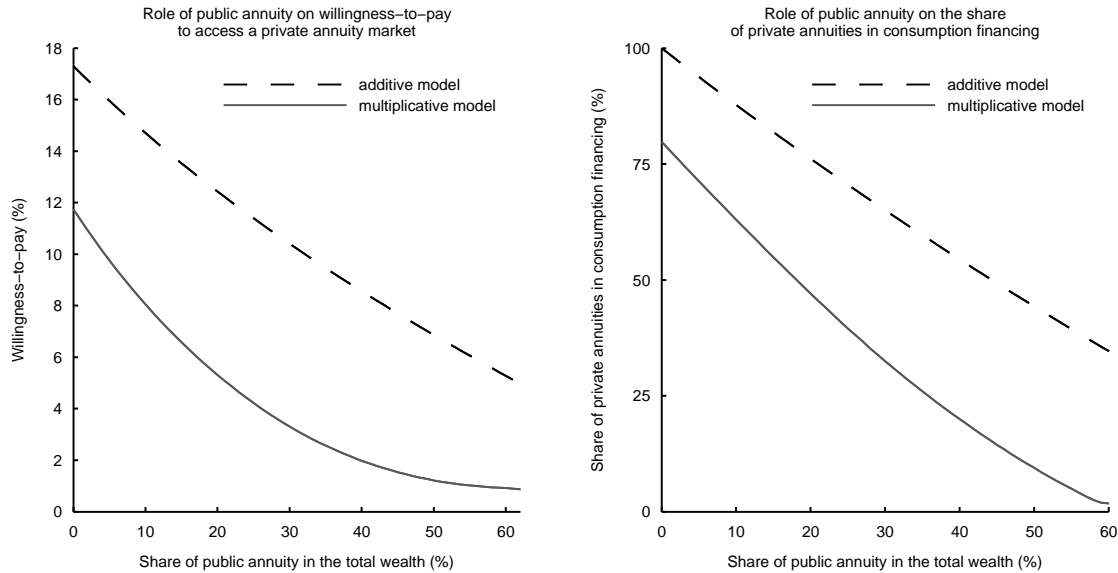


Figure 4: Impact of public annuities on willingness-to-pay for private annuities and on consumption financing

an impact only for high levels of public annuities, and we mainly observe the same crowding out effect as in the additive case.

Role of bequest motives intensity. We now explore the robustness of our findings in relation to the specification of bequest motives. We will discuss how changing the intensity of bequest motives (parameter θ) may impact our results. The benchmark calibration corresponds to an average bequest of 20% of the initial non-annuitized wealth. Graphs in Figure 5 plot our results as a function of the average bequest.

First, the left-hand side graph illustrates the impact of bequest motives on the WTP. As soon as the intensity of the bequest motive becomes significant (average bequest greater than 1% of the non-annuitized wealth), the multiplicative model generates smaller WTP than the additive one. Second, the WTP decrease with the average bequest: the stronger the bequest motive, the more the agent needs to save, the less he cares about annuities. When the average bequest increases, the WTP decreases more rapidly with the multiplicative model than with the additive one. With the multiplicative model, WTP will be below 5% as soon as the average bequest becomes greater than 8% of the non-annuitized wealth. WTP in the additive model becomes below 5% only when the average bequest motive is greater than 35% of initial wealth.

Second, the right-hand side graph of Figure 5 plots the share of private annuities in the financing of consumption. In both models, the share of consumption financed by private annuities decreases with the bequest motive. A stronger bequest motive modifies the agent's trade-off be-

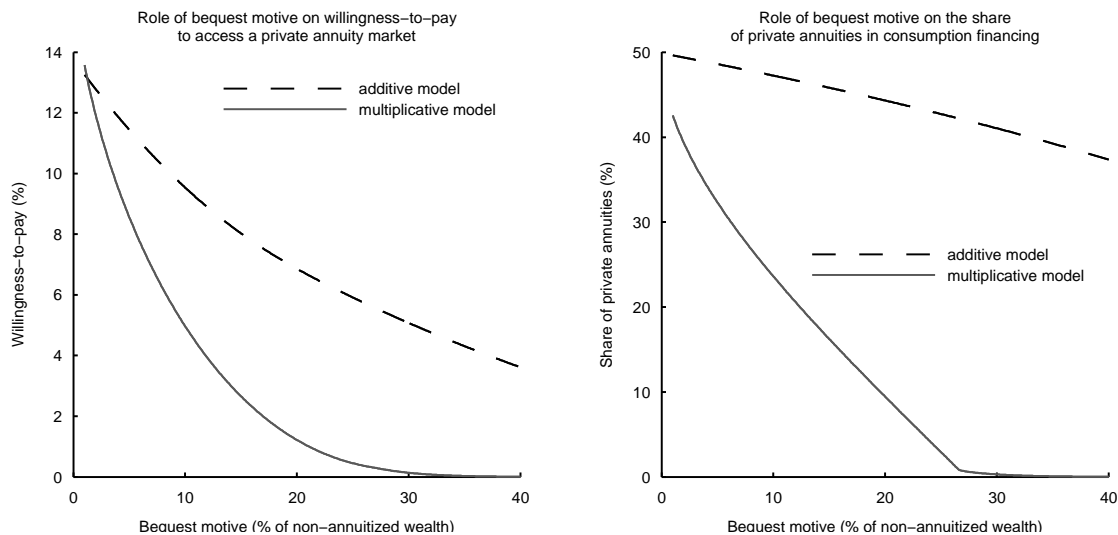


Figure 5: Impact of the bequest motive on willingness-to-pay for private annuities and on consumption financing

tween consumption and bequest. As the agent consumes less, public annuities y thus finance a greater proportion of the agent's consumption and crowd out private annuities. In the multiplicative case, the decline in the share of consumption financed by private annuities is greater than in the additive model. For very strong bequest motives (providing an average bequest of 30% of the non-annuitized wealth and more), the share of private annuities is close to zero, while in the additive model this share never goes below one third.

Alternative bequest specifications. As highlighted in Lockwood (2012b), most functional forms used in the literature to model bequest motives are nested in our parametrization of v in equation (28). However, there is no consensus about how bequests respond to wealth, which depends on the combination of parameters ψ and θ that enter into equation (28). Different combinations of θ and ψ provide different average bequests and different responsiveness levels to changes in wealth. So far, we have not discussed how bequests respond to wealth, since the question we address does not require consideration for heterogeneity in wealth. However, to check that our results do not rely on an implausible wealth elasticity of bequests, we will consider specifications that generate, for a given average bequest, the same wealth elasticities of bequest as in De Nardi (2004), De Nardi, French and Jones (2010), Ameriks, Caplin, Laufer and Van Nieuwerburgh (2011) and Lockwood (2012b).

In practice, in order to obtain various additive reference specifications, we consider pairs of parameters (θ, ψ) that are directly taken from each of the four above mentioned studies.²⁴ The

²⁴More precisely, Lockwood (2012b) converted the three other model estimations (in addition to his own one) into

other parameters such as elasticities of substitution, the rate of time discounting, mortality rates and rate of interest are set as before (Table 1). The ratio of non-annuitized wealth over total wealth is also held constant at one half. In each case, we adjust the initial wealth level W_0 , so as to obtain the same average bequest equal to 20% of the initial wealth, as in our benchmark calibration. The four additive specifications that we consider differ because they assume different wealth elasticities of bequest. In order to illustrate the role of lifetime risk aversion, each of these additive specifications are compared with a corresponding multiplicative model that generates the same amount of average bequests and the same rate of time discounting at age 65. This is done by adjusting the parameters λ (lifetime risk aversion) and θ (intensity of the bequest motives), the other parameters being kept as in the additive specifications. In each case, the constant u_0 is set to generate a VSL equaling 500 yearly consumptions at age 65.

We report in Table 3 the WTP for private annuities, and how consumption is financed for each specification. Table 3 is similar to Table 2 except that instead of using our own calibration for the bequest motive, we now rely on specifications taken from other studies. Exact calibrations can be found in Section C of the Appendix.

Calibration Model	DeNardi		DeNardi et al.		Ameriks et al.		Lockwood	
	Mult.	Add.	Mult.	Add.	Mult.	Add.	Mult.	Add.
Willingness-to-pay (%)	1.81	3.76	1.35	6.69	1.75	4.22	1.56	5.50
Share of consumption financed by:								
Public annuities	56.07	55.56	56.24	55.55	56.13	55.55	56.24	55.56
Private annuities	2.61	44.44	8.96	44.45	2.59	44.45	2.71	44.44
Riskless saving	41.32	0.00	34.79	0.00	41.29	0.00	41.05	0.00

Table 3: Impact of various calibrations for the bequest preferences when the bequest motive is 20% of the non-annuitized wealth

Looking at the additive specifications, we find that the WTP for gaining access to annuities ranges from 3.76% to 6.69% of the initial wealth W_0 . This reflects the heterogeneity in assumptions that can be found in the literature as to the precise form of the bequest motive. In all cases annuities finance about 44% of consumption, the remainder being financed by public annuities.

As previously, the results obtained when introducing lifetime risk aversion strongly contrast with those of the additive specifications. First, the WTP for annuities, though still positive, is much smaller, ranging from 1.35% to 1.81% of initial wealth W_0 . Second, when available, private annuities are used much less. Looking at the different multiplicative specifications, we find that 35% to 41% of consumption is financed by riskless savings, while private annuities are only used to finance from 2.6% to 9% of private consumption. Independent of any market failure, the size of the annuity market is thus found to be much smaller when agents have multiplicative preferences

a common functional form and we in turn adapt his parameters to our functional form (28).

than when they have additive preferences. This result holds for a wide range of bequest motives.

5 Conclusion

The relationship between risk aversion and annuity demand has remained unexplored in the economic literature. However, as soon as we explore the role of risk aversion in a proper way (that is by considering lifetime risk aversion), the demand for annuities is found to decrease with risk aversion. Moreover, annuity demand eventually becomes negative (or vanishes if we add a positivity constraint) if risk aversion is sufficiently large and individuals have a positive bequest motive. A possible reason for the observed low level of wealth annuitization may therefore simply be that individuals are too risk averse to purchase annuities. Intuitively they do not purchase annuities because they do not want to take the risk of dying young without leaving a bequest, which is indeed the worst scenario one can imagine.

Calibration of our model with realistic mortality patterns and preference parameters that seem reasonable indicate that lifetime risk aversion helps in part to explain the annuity puzzle as it generates significantly lower levels of willingness-to-pay for annuities. Nonetheless, reasonable parameter values did not generate a negative demand for annuities. Our calibration suggests that risk aversion alone cannot solve the annuity puzzle. Other elements such as the existence of public pensions, market imperfections, the need for liquidities and rationality biases should also be taken into account in order to end up with a negative demand for annuities. By introducing lifetime risk aversion in the discussion, we add one more piece to the annuity puzzle complementing the other possible explanations suggested so far.

Interestingly enough, corroboration for our theoretical explanation comes from the literature exploring behavioral biases. Indeed, as we discussed previously, several papers including Brown, Kling, Mullainathan and Wrobel (2008), Hu and Scott (2007), Agnew, Anderson, Gerlach and Szykman (2008) and Benartzi, Previtro, and Thaler (2011), provide convincing evidence that the riskiness of annuities was considered to be a major source of concern for agents. Taking this further Brown (2007) points out that people apparently want to buy insurance contracts when utility is low. Our contribution involves showing that this behavioral trait can be reconciled with a standard model of choice under uncertainty (expected utility) when lifetime risk aversion is taken into account. When lifetime risk aversion is significant enough, the state associated with lower utility is also the one with higher marginal utility. The willingness to buy contracts that pay when utility is low is then consistent with rational insurance behavior, which involves purchasing contracts that pay off when the marginal utility is high.

References

- AGNEW, J. R., L. R. ANDERSON, J. R. GERLACH, AND L. R. SZYKMAN (2008): “Who Chooses Annuities? An Experimental Investigation of the Role of Gender, Framing, and Defaults,” *American Economic Review*, 98(2), 418–422.
- AMERIKS, J., A. CAPLIN, S. LAUFER, AND S. VAN NIEUWERBURGH (2011): “The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Bequest and Precautionary Motives,” *Journal of Finance*, 66(2), 519–561.
- ANDERSEN, S., G. HARRISON, M. LAU, AND E. RUTSTROEM (2011): “Multiattribute Utility Theory, Intertemporal Utility and Correlation Aversion,” CEAR Working Paper 2011-04, Georgia State University.
- BENARTZI, S., A. PREVITERO, AND R. H. THALER (2011): “Annuitization Puzzles,” *Journal of Economic Perspectives*, 25(4), 143–164.
- BLEICHRODT, H., AND L. EECKHOUDT (2005): “Saving under Rank-Dependent Utility,” *Economic Theory*, 25(2), 505–511.
- BOMMIER, A. (2006): “Uncertain Lifetime and Intertemporal Choice: Risk Aversion as a Rationale for Time Discounting,” *International Economic Review*, 47(4), 1223–1246.
- (2007): “Risk Aversion, Intertemporal Elasticity of Substitution and Correlation Aversion,” *Economics Bulletin*, 4(29), 1–8.
- (2013): “Life Cycle Preferences Revisited,” *Journal of European Economic Association*, 11(6).
- BOMMIER, A., A. CHASSAGNON, AND F. LEGRAND (2012): “Comparative Risk Aversion: A formal Approach with Applications to Saving Behaviors,” *Journal of Economic Theory*, 147(4), 1614–1641.
- BROWN, J. R. (2007): “Rational and Behavioral Perspectives on the Role of Annuities in Retirement Planning,” NBER Working Paper 13537, National Bureau of Economic Research.
- BROWN, J. R., J. R. KLING, S. MULLAINATHAN, AND M. V. WROBEL (2008): “Why Don’t People Insure Late-Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle,” *American Economic Review*, 98(2), 304–309.
- CHAI, J., W. HORNEFF, R. MAURER, AND O. S. MITCHELL (2011): “Optimal Portfolio Choice over the Life Cycle with Flexible Work, Endogenous Retirement, and Lifetime Payouts,” *Review of Finance*, 15(4), 875–907.
- COOK, P. J., AND D. A. GRAHAM (1977): “The Demand for Insurance and Protection: The Case of Irreplaceable Commodities,” *Quarterly Journal of Economics*, 91(1), 143–156.
- DAVIDOFF, T., J. R. BROWN, AND P. A. DIAMOND (2005): “Annuities and Individual Welfare,” *American Economic Review*, 95(5), 1573–1590.
- DE NARDI, M. (2004): “Wealth Inequality and Intergenerational Links,” *Review of Economic Studies*, 71(3), 743–768.
- DE NARDI, M., E. FRENCH, AND J. B. JONES (2010): “Why Do the Elderly Save? The Role of Medical Expenses,” *Journal of Political Economy*, 118(1), 39–75.
- DORFLEITNER, G., AND M. KRAPP (2007): “On Multiattributive Risk Aversion: Some Clarifying Results,” *Review of Managerial Science*, 1(1), 47–63.

- DRÈZE, J., AND A. RUSTICHINI (2004): “State-Dependent Utility and Decision Theory, 2: Extensions. Dordrecht,” in *Handbook of Utility Theory*, ed. by P. H. S. Barbera, and C. Seild, pp. 839–892. Kluwer Academic Publishers.
- DRÈZE, J. H., AND F. MODIGLIANI (1972): “Consumption Decisions under Uncertainty,” *Journal of Economic Theory*, 5(3), 308–335.
- EDEN, B. (2008): “Substitution, Risk Aversion and Asset Prices: An Expected Utility Approach,” Working Paper 0803, Department of Economics, Vanderbilt University.
- EPSTEIN, L. G., AND S. M. TANNY (1980): “Increasing Generalized Correlation: A Definition and Some Economic Consequences,” *Canadian Journal of Economics*, 13(1), 16–34.
- EPSTEIN, L. G., AND S. E. ZIN (1989): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 57(4), 937–969.
- FERNÁNDEZ-VILLAVARDE, J., AND D. KRUEGER (2007): “Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data,” *The Review of Economics and Statistics*, 89(3), 552–565.
- FINKELSTEIN, A., AND J. POTERBA (2002): “Selection Effects in the United Kingdom Individual Annuities Market,” *Economic Journal*, 112(476), 28–50.
- (2004): “Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market,” *Journal of Political Economy*, 112(1), 183–208.
- GAZZALE, R. S., AND L. WALKER (2011): “I’ll Cross that Bridge if I Get to It: Focusing on the Near Future,” Working Paper, University of Toronto.
- HORNEFF, W. J., R. H. MAURER, O. S. MITCHELL, AND M. Z. STAMOS (2010): “Variable Payout Annuities and Dynamic Portfolio Choice in Retirement,” *Journal of Pension Economics and Finance*, 9(2), 163–183.
- HU, W.-Y., AND J. S. SCOTT (2007): “Behavioral Obstacles in the Annuity Market,” *Financial Analysts Journal*, 63(6), 71–82.
- HURD, M. D., AND J. P. SMITH (2002): “Expected Bequests and their Distribution,” NBER Working Paper 9142, National Bureau of Economic Research.
- INKMANN, J., P. LOPES, AND A. MICHAELIDES (2011): “How Deep is the Annuity Market Participation Puzzle?,” *Review of Financial Studies*, 24(1), 279–319.
- JAMES, E., AND X. SONG (2001): “Annuities Markets Around the World: Money’s Worth and Risk Intermediation,” CeRP Working Paper 16/01, Center for Research on Pensions and Welfare Policies.
- JAPPELLI, T. (1999): “The Age-Wealth Profile and the Life-Cycle Hypothesis: A Cohort Analysis with Time Series of Cross-Sections of Italian Households,” *Review of Income and Wealth*, 45(1), 57–75.
- JOHANSSON, P.-O. (2002): “On the Definition and Age-Dependency of the Value of a Statistical Life,” *Journal of Risk and Uncertainty*, 25(3), 251–263.
- JOHNSON, R. W., L. E. BURMAN, AND D. I. KOBES (2004): “Annuitized Wealth at Older Ages: Evidence from the Health and Retirement Study,” Final report to the Employee Benefits Security Administration of the U.S. Department of Labor, The Urban Institute.

- KIHLSTROM, R. E., AND L. J. MIRMAN (1974): "Risk Aversion with many Commodities," *Journal of Economic Theory*, 8(3), 361–388.
- KIMBALL, M. S., AND P. WEIL (2009): "Precautionary Saving and Consumption Smoothing Across Time and Possibilities," *Journal of Money, Credit and Banking*, 41(2-3), 245–284.
- KOPCZUK, W., AND J. P. LUPTON (2007): "To Leave or Not to Leave: The Distribution of Bequest Motives," *Review of Economic Studies*, 74(1), 207–235.
- LIVINGSTON, G., AND D. COHN (2010): "The New Demography of American Motherhood," A Social and Demographic Trends Report, Pew Research Center.
- LOCKWOOD, L. M. (2012a): "Bequest Motives and the Annuity Puzzle," *Review of Economic Dynamics*, 15(2), 226–243.
- (2012b): "Bequest Motives and the Choice to Self-Insure Late-Life Risks," Working Paper, Northwestern University.
- MILEVSKY, M. A., AND V. YOUNG (2007): "Annuitization and Asset Allocation," *Journal of Economic Dynamics and Control*, 31(9), 3138–3177.
- MITCHELL, O. S., J. M. POTERBA, M. J. WARSHAWSKY, AND J. R. BROWN (1999): "New Evidence on the Money's Worth of Individual Annuities," *American Economic Review*, 89(5), 1299–1318.
- PANG, G., AND M. J. WARSHAWSKY (2010): "Optimizing the Equity-Bond-Annuity Portfolio in Retirement: The Impact of Uncertain Health Expenses," *Insurance: Mathematics and Economics*, 46(1), 198–209.
- PASHCHENKO, S. (2013): "Accounting for Non-Annuitization," *Journal of Public Economics*, 98, 53–67.
- PONZETTO, G. (2003): "Risk Aversion and the Utility of Annuities," CeRP Working Paper 31/03, Center for Research on Pensions and Welfare Policies.
- RICHARD, S. F. (1975): "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions," *Management Science*, 22(1), 12–21.
- SINCLAIR, S. M., AND K. A. SMETTERS (2004): "Health Shocks and the Demand for Annuities," Technical Paper 2004-9, Congressional Budget Office.
- VAN DEN HEUVEL, S. J. (2008): "Temporal Risk Aversion and Asset Prices," Working Paper 2008-37, Federal Reserve Board.
- VAN DER PLOEG, F. (1993): "A Closed-Form Solution for a Model of Precautionary Saving," *Review of Economic Studies*, 60(2), 385–395.
- VISCUSI, W. K., AND J. E. ALDY (2003): "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and Uncertainty*, 27(1), 5–76.
- YAARI, M. E. (1965): "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 32(2), 137–150.
- (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–105.
- YOGO, M. (2009): "Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing and Risky Assets," NBER Working Paper 15307, National Bureau of Economic Research.

Appendix

A Proof of Proposition 1

When ϕ is linear, $\phi'(U_A) = \phi'(U_D) = \phi'(0) = 1$. First order conditions (8)–(10) and the budget constraint can now be expressed as follows:

$$u'_2(c_2) \leq \frac{u'_1(c_1)}{1+R} \quad (= \text{ if } a > 0), \quad (34)$$

$$v'((1+R)s) \leq \frac{1}{1-p} \left(\frac{u'_1(c_1)}{1+R} - u'_2(c_2) \right) + u'_2(c_2) \quad (= \text{ if } s > 0), \quad (35)$$

$$v'(\tau) \leq u'_2(c_2) \quad (= \text{ if } \tau > 0), \quad (36)$$

$$c_2 + \tau = (1+R)s + a \frac{1+R}{p}. \quad (37)$$

Let us show that in any case $(1+R)s = \tau$.

1. $s = 0$. The budget constraint (37) implies that $a > 0$: (34) is therefore an equality. (35) implies then that $v'(0) \leq u'_2(c_2)$. Suppose that $\tau > 0$: we deduce from (36) that $v'(0) \leq v'(\tau)$, which contradicts that v is concave and non-linear. Thus, $(1+R)s = \tau = 0$.
2. $s > 0$. From (35), which is an equality, together with (34) and (36), we deduce that $v'((1+R)s) \geq u'_2(c_2) \geq v'(\tau)$ and $\tau \geq (1+R)s > 0$. The budget constraint (37) implies that $a > 0$. (34), as (35) and (36) are therefore equalities: we deduce that $v'((1+R)s) = v'(\tau)$ and $(1+R)s = \tau$.

We always obtain $(1+R)s = \tau$, and thus also $a = \frac{pc_2}{1+R}$.

B Proof of Proposition 3

First, we show that if λ is large enough then $a = 0$. Let us assume that for all λ , $a > 0$ and see that we obtain a contradiction. When a is interior, the FOC (6)–(10) together with the budget

constraint can be rewritten:

$$\left[pe^{-\lambda U_A} + (1-p)e^{-\lambda U_D}\right] u'_1(c_1) = \mu_1 \quad (38)$$

$$e^{-\lambda U_A} u'_2(c_2) = \frac{\mu_1}{1+R} \quad (39)$$

$$v'((1+R)s)e^{-\lambda U_D} \leq \frac{\mu_1}{1+R} \quad (40)$$

$$v'(\tau)e^{-\lambda U_A} \leq \frac{\mu_1}{1+R} \quad (41)$$

$$W_0 = c_1 + \frac{p(c_2 + \tau)}{1+R} + s(1-p) \quad (42)$$

Because of the possible existence of corner solutions for s and τ we have to consider several cases. We will see that in all possible cases we end up having a contradiction.

1. Assume that $s = \tau = 0$. We have $U_A - U_D = u_2(c_2)$ and (42) becomes $c_1 = W_0 - \frac{pc_2}{1+R}$. (38)–(40) give:

$$\left[p + (1-p)e^{\lambda u_2(c_2)}\right] u'_1\left(W_0 - \frac{pc_2}{1+R}\right) = (1+R)u'_2(c_2) \geq (1+R)v'(0)e^{\lambda u_2(c_2)} \quad (43)$$

Suppose $c_2 < c_2^*$. The inequality (43) implies, since $u(c_2) < 0$ and $p + (1-p)e^{\lambda u_2(c_2)} < 1$:

$$u'_1(c_1) > (1+R)u'_2(c_2) \quad (44)$$

For c_2^* , we derive from Assumption A and $c_2^* \geq c_2^{**}$ the reverse relationship $u'_1\left(W_0 - \frac{pc_2^*}{1+R}\right) < (1+R)u'_2(c_2^*)$. We deduce:

$$u'_2(c_2) > u'_2(c_2^*) \geq \frac{u'_1\left(W_0 - \frac{pc_2^*}{1+R}\right)}{1+R} > \frac{u'_1\left(W_0 - \frac{pc_2}{1+R}\right)}{1+R} = \frac{u'_1(c_1)}{1+R} \quad (45)$$

Inequalities (44) and (45) are incompatible, and we deduce $c_2 \geq c_2^*$ and $u_2(c_2) \geq 0$.

We deduce from (43) that $u'_2(c_2) \geq v'(0)$. We also have:

$$v'(0) \leq \left[1 - p + pe^{-\lambda u_2(c_2)}\right] \frac{1}{1+R} u'_1\left(W_0 - \frac{pc_2}{1+R}\right) \leq \frac{1}{1+R} u'_1\left(W_0 - \frac{pc_2}{1+R}\right)$$

Suppose that $c_2 < c_2^{**}$. Under Assumption A:

$$v'(0) \geq v'(c_2^{**}) > u'_1\left(W_0 - \frac{c_2^{**}}{1+R}\right) \geq u'_1\left(W_0 - \frac{pc_2^{**}}{1+R}\right) \geq u'_1\left(W_0 - \frac{pc_2}{1+R}\right),$$

which contradicts $v'(0) \leq \frac{1}{1+R} u'_1\left(W_0 - \frac{pc_2}{1+R}\right)$. We deduce that $c_2 \geq c_2^{**}$.

The optimal second period consumption exists if there exists a solution $\geq c_2^{**}$ to the following equation:

$$p + (1-p)e^{\lambda u_2(c_2)} = (1+R) \frac{u_2'(c_2)}{u_1' \left(W_0 - \frac{p c_2}{1+R} \right)}$$

The LHS is an increasing function of c_2 , while the RHS is a decreasing one. A solution exists if and only if $p + (1-p)e^{\lambda u_2(c_2^{**})} \leq (1+R) \frac{u_2'(c_2^{**})}{u_1' \left(W_0 - \frac{p c_2^{**}}{1+R} \right)}$. If $u_2(c_2^{**}) > 0$, we can always find a λ such that the optimum does not exist. For the optimum to exist, we thus necessarily have: $c_2^* = c_2^{**} = 0$ and $u_2(0) = 0$.

Let $A > 0$ be arbitrarily large. Since $\lim_{c_2 \rightarrow 0} u_2'(c_2) = +\infty$, it is always possible to choose $\varepsilon > 0$ small enough such that:

$$(1+R) \frac{u_2'(\varepsilon)}{u_1' \left(W_0 - \frac{p \varepsilon}{1+R} \right)} > p + (1-p)e^{\frac{u_2(\varepsilon)}{2} A} > p + (1-p)e^A$$

$\frac{u_2(\varepsilon)}{u_2(\varepsilon/2)}$ is greater than 1, but also smaller than 2 in the vicinity of 0. Indeed, by l'Hospital's rule, its limit is the limit of $2 \frac{u_2'(\varepsilon)}{u_2'(\varepsilon/2)}$, which is smaller than 2 because u_2 is increasing and concave.

We can now always choose $\lambda \geq 0$ such that:

$$p + (1-p)e^{\lambda u_2(\varepsilon)} > (1+R) \frac{u_2'(\varepsilon)}{u_1' \left(W_0 - \frac{p \varepsilon}{1+R} \right)} > p + (1-p)e^{\lambda u_2(\varepsilon/2)} > p + (1-p)e^A \quad (46)$$

Indeed, the first two inequalities simply mean that:

$$\frac{1}{u_2(\varepsilon/2)} > \ln \left(\frac{(1+R)u_2'(\varepsilon)}{(1-p)u_1' \left(W_0 - \frac{p \varepsilon}{1+R} \right)} - \frac{p}{1-p} \right)^{-1} \lambda > \frac{1}{u_2(\varepsilon)}$$

The construction of A insures the last inequality since:

$$\lambda u_2\left(\frac{\varepsilon}{2}\right) > \frac{u_2(\varepsilon/2)}{u_2(\varepsilon)} \ln \left(\frac{(1+R)u_2'(\varepsilon)}{(1-p)u_1' \left(W_0 - \frac{p \varepsilon}{1+R} \right)} - \frac{p}{1-p} \right) > \frac{u_2(\varepsilon/2)}{u_2(\varepsilon)} \frac{u_2(\varepsilon)}{u_2(\varepsilon/2)} A = A$$

Eq. (46) means that c_2 exists and $\frac{\varepsilon}{2} < c_2 < \varepsilon$ together with $\lambda u_2\left(\frac{\varepsilon}{2}\right) > A$. From (43), we

deduce that:

$$\begin{aligned} v'(0) &\leq \left(1 - p + pe^{-\lambda u_2(c_2)}\right) u_1' \left(W_0 - \frac{pc_2}{1+R}\right) \\ &\leq (1 - p + pe^{-A}) u_1' \left(W_0 - \frac{p\varepsilon}{1+R}\right) \end{aligned}$$

The RHS can be made arbitrarily close to $(1 - p) u_1'(W_0)$, by choosing A large and ε small enough. We therefore have $v'(0)$, which can be strictly smaller than $u_1'(W_0)$, which contradicts Assumption A.

As a result, we cannot have $s = \tau = 0$.

2. Assume that $\tau > 0$ and $s = 0$. Eq. (41) is an equality and we deduce from (39)–(41):

$$u_2'(c_2) = v'(\tau) \geq v'(0)e^{\lambda(U_A - U_D)},$$

which implies $U_A - U_D = u_2(c_2) + v(\tau) < 0$, and $c_2 \leq c_2^* \leq c_2^{**}$. Under Assumption A:

$$u_2'(c_2) \geq u_2'(c_2^{**}) > v'(0) \geq v'(\tau)$$

This contradicts $u_2'(c_2) = v'(\tau)$. We cannot have $\tau > 0 = s$.

3. Assume that $s > 0$ and $\tau = 0$. Eq. (40) is an equality and we deduce from (38)–(41):

$$u_2'(c_2) = v'((1+R)s)e^{\lambda(U_A - U_D)} = \left(p + (1-p)e^{\lambda(U_A - U_D)}\right) \frac{1}{1+R} u_1'(c_1) \geq v'(0) \quad (47)$$

This implies that $U_A - U_D = u_2(c_2) - v((1+R)s) > 0$. We deduce from the preceding inequality and from the budget constraint stating that $c_1 = W_0 - \frac{pc_2}{1+R} - s(1-p) \leq W_0 - \frac{pc_2}{1+R}$:

$$u_2'(c_2) \geq \frac{1}{1+R} u_1'(c_1) \geq \frac{1}{1+R} u_1' \left(W_0 - \frac{pc_2}{1+R}\right)$$

Suppose that $c_2^{**} \geq c_2$. From the budget constraint $c_2 = (1+R)s + a \frac{1+R}{p} \geq (1+R)s$. From (47) and $c_2^{**} \geq (1+R)s$, we deduce that $v'(c_2^{**}) \leq v'((1+R)s) \leq \frac{1}{1+R} u_1'(c_1)$. Moreover, $c_1 = W_0 - \frac{pc_2}{1+R} - s(1-p) \geq W_0 - \frac{pc_2^{**}}{1+R} - \frac{c_2^{**}}{1+R}(1-p) = W_0 - \frac{c_2^{**}}{1+R}$. We therefore deduce that $v'(c_2^{**}) \leq \frac{1}{1+R} u_1'(W_0 - \frac{c_2^{**}}{1+R})$, which contradicts Assumption A. We thus have $c_2 \geq c_2^{**}$.

Using (47), we deduce that $(1+R)s$ is a function of c_2 that we denote h and which is defined as:

$$v'(h(c_2)) e^{-\lambda v(h(c_2))} = u_2'(c_2) e^{-\lambda u_2(c_2)}$$

First case. The function $c_2 \mapsto u_2(c_2) - v(h(c_2))$ is increasing. c_2 is defined as:

$$p + (1 - p)e^{\lambda(u_2(c_2) - v(h(c_2)))} = (1 + R) \frac{u_2'(c_2)}{u_1' \left(W_0 - \frac{pc_2 + (1-p)h(c_2)}{1+R} \right)}$$

Using the same proof strategy as in the case $s = \tau = 0$, we first obtain $c_2^{**} = c_2^* = 0$ and then obtain a contradiction with Assumption A.

Second case. The function $c_2 \mapsto u_2(c_2) - v(h(c_2))$ is decreasing (in the vicinity of c_2^{**} – otherwise, the previous argument applies). First remark that it imposes $c_2^{**} > 0$. Otherwise, we have for all $c_2 \geq 0$ sufficiently small: $h'(c_2)v'(h(c_2)) \geq u'(c_2)$. After integration ($u_2(c_2^{**}) = u_2(c_2^*) = 0$), we get a contradiction with $u_2(c_2) - v(h(c_2)) \geq 0$.

c_2 is still defined as:

$$p + (1 - p)e^{\lambda(u_2(c_2) - v(h(c_2)))} = (1 + R) \frac{u_2'(c_2)}{u_1' \left(W_0 - \frac{pc_2 + (1-p)h(c_2)}{1+R} \right)}$$

$u_2(c_2) - v(h(c_2)) = (u_2(c_2) - v(c_2)) + (v(c_2) - v(h(c_2)))$ is the sum of two positive terms (since $c_2 \geq (1 + R)s$) and is strictly positive id $c_2 > c_2^*$. If $v(c_2^{**}) - v(h(c_2^{**}))$ is strictly positive, it is sufficient to choose λ large enough to get a contradiction (c_2 never exists – choose $\lambda > \min_{c_2 \geq c_2^{**}} (u_2(c_2) - v(h(c_2)))^{-1} (1 + R) \frac{u_2'(c_2^{**})}{u_1' \left(W_0 - \frac{pc_2^{**} + (1-p)h(c_2^{**})}{1+R} \right)}$). Therefore, $u_2(c_2^{**}) = v(c_2^{**}) = v(h(c_2^{**}))$.

$c_2 \mapsto u_2(c_2) - v(h(c_2))$ is decreasing in the vicinity of c_2^{**} , but positive. This implies that $u_2(c_2) = v(h(c_2))$ in the vicinity of c_2^{**} , which contradicts the definition of c_2^{**} .

4. Assume that $s > 0$ and $\tau > 0$.

In that case:

$$v'((1 + R)s)e^{\lambda(U_A - U_D)} = v'(\tau) = u_2'(c_2) = \frac{1}{1 + R} \left(p + (1 - p)e^{\lambda(U_A - U_D)} \right) u_1'(c_1) \quad (48)$$

Suppose that $c_2 < c_2^{**}$. We then have: $u_2'(c_2^{**}) < u_2'(c_2) = v'(\tau) \leq v'(0)$, which contradicts Assumption A. We therefore have $c_2 \geq c_2^{**} \geq c_2^*$.

The budget constraint gives $c_2 + \tau = (1 + R)s + (1 + R)a/p > (1 + R)s$. We deduce $v(c_2) + v(\tau) \geq v((1 + R)s)$ (v being concave with $v(0) = 0$ is sub-additive. For all $\lambda \in [0, 1]$, $v(\lambda c) = v(\lambda c + (1 - \lambda)0) \geq \lambda v(c) + (1 - \lambda)v(0) = \lambda v(c)$. Then $v(c_2) + v(\tau) = v((c_2 + \tau) \frac{c_2}{c_2 + \tau}) + v((c_2 + \tau) \frac{\tau}{c_2 + \tau}) \geq v(c_2 + \tau) \geq v((1 + R)s)$).

We have $U_A - U_D = u(c_2) + v(\tau) - v((1+R)s) = (u_2(c_2) - v(c_2)) + (v(c_2) + v(\tau) - v((1+R)s)) \geq 0$ since both terms are positive due to our preceding remark and $c_2 \geq c_2^{**}$.

We deduce that $v'(\tau) \geq v'((1+R)s)$ and $(1+R)s \geq \tau$.

As previously, it is straightforward that τ and $(1+R)s$ are increasing functions of c_2 that we denote respectively $\psi(c_2)$ and $h(c_2)$. c_2 solves:

$$p + (1-p)e^{\lambda(u_2(c_2) + \psi(c_2) - v(h(c_2)))} = (1+R) \frac{u_2'(c_2)}{u_1' \left(W_0 - \frac{p(c_2 + \psi(c_2)) + (1-p)h(c_2)}{1+R} \right)}$$

The proof is now similar to that in 3) and we again obtain a contradiction.

We have shown that if λ is large enough, then $a = 0$. We now prove that an increase in lifetime risk aversion λ implies a smaller annuity demand a . Due to possible corner solutions, we distinguish several cases.

1. We assume that $s, \tau, a > 0$.

$$\begin{aligned} u_2'(c_2) &= \frac{1}{1+R} \left(p + (1-p)e^{\lambda(U_A - U_D)} \right) u_1'(c_1) = v'((1+R)s) e^{\lambda(U_A - U_D)} = v'(\tau) \\ W_0 &= c_1 + \frac{p(c_2 + \tau)}{1+R} + (1-p)s \\ W_0 &= c_1 + s + a \\ c_2 + \tau &= (1+R)s + (1+R) \frac{a}{p} \end{aligned}$$

We have $c_2 \geq c_2^{**}$.

We assume an increase of λ starting from an optimum where $c_1, c_2, \tau, a, s > 0$. For the sake of simplicity, we note $\Delta U = U_A - U_D \geq 0$ and $S = (1+R)s$.

We have:

$$\begin{aligned} \frac{\partial \tau}{\partial \lambda} &= \frac{u_2''(c_2)}{v''(\tau)} \frac{\partial c_2}{\partial \lambda} \\ (1+R)u''(c_2) \frac{\partial c_2}{\partial \lambda} &= (p + (1-p)e^{\lambda \Delta U}) u_1''(c_1) \frac{\partial c_1}{\partial \lambda} + (1-p)e^{\lambda \Delta U} u_1'(c_1) \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \\ u_2''(c_2) \frac{\partial c_2}{\partial \lambda} &= e^{\lambda \Delta U} \left(v''(S) \frac{\partial S}{\partial \lambda} + v'(S) \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \right) \\ -(1+R) \frac{\partial c_1}{\partial \lambda} &= p \left(1 + \frac{u_2''(c_2)}{v''(\tau)} \right) \frac{\partial c_2}{\partial \lambda} + (1-p) \frac{\partial S}{\partial \lambda} \\ \frac{\partial \Delta U}{\partial \lambda} &= u'(c_2) \left(1 + \frac{u_2''(c_2)}{v''(\tau)} \right) \frac{\partial c_2}{\partial \lambda} - v'(S) \frac{\partial S}{\partial \lambda} \end{aligned}$$

We drop arguments and to limit ambiguity, we note $v'_\tau = v''(\tau)$ and $\Gamma = 1 + \frac{u''_2(c_2)}{v''(\tau)} \geq 1$.

$$\begin{aligned}\frac{u''_2}{u'_2} \frac{\partial c_2}{\partial \lambda} &= \frac{u''_1}{u'_1} \frac{\partial c_1}{\partial \lambda} + \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \\ \frac{u''_2}{u'_2} \frac{\partial c_2}{\partial \lambda} &= \frac{v''}{v'} \frac{\partial S}{\partial \lambda} + \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right)\end{aligned}$$

$$\begin{aligned}\left(\frac{u''_2}{u'_2} + \frac{p}{1+R} \Gamma \frac{u''_1}{u'_1} \right) \frac{\partial c_2}{\partial \lambda} &= -\frac{1-p}{1+R} \frac{u''_1}{u'_1} \frac{\partial S}{\partial \lambda} + \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \left(\lambda \left(u'_2 \Gamma \frac{\partial c_2}{\partial \lambda} - v' \frac{\partial S}{\partial \lambda} \right) + \Delta U \right) \\ \frac{u''_2}{u'_2} \frac{\partial c_2}{\partial \lambda} &= \frac{v''}{v'} \frac{\partial S}{\partial \lambda} + \left(\lambda \left(u'_2 \Gamma \frac{\partial c_2}{\partial \lambda} - v' \frac{\partial S}{\partial \lambda} \right) + \Delta U \right)\end{aligned}$$

$$\begin{aligned}\left(\frac{u''_2}{u'_2} + \frac{p}{1+R} \Gamma \frac{u''_1}{u'_1} - \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \lambda u'_2 \Gamma \right) \frac{\partial c_2}{\partial \lambda} + \frac{1-p}{1+R} \left(\frac{u''_1}{u'_1} + \lambda u'_1 \right) \frac{\partial S}{\partial \lambda} &= \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \Delta U \\ \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \frac{\partial c_2}{\partial \lambda} + \left(-\frac{v''}{v'} + \lambda v' \right) \frac{\partial S}{\partial \lambda} &= \Delta U\end{aligned}$$

We deduce:

$$\begin{aligned}a \frac{\partial c_2}{\partial \lambda} + b \frac{\partial S}{\partial \lambda} &= \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \Delta U \\ c \frac{\partial c_2}{\partial \lambda} + d \frac{\partial S}{\partial \lambda} &= \Delta U\end{aligned}$$

and:

$$\begin{aligned}\frac{ad - bc}{\Delta U} \frac{\partial c_2}{\partial \lambda} &= d \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} - b \\ \frac{ad - bc}{\Delta U} \frac{\partial S}{\partial \lambda} &= -c \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} + a\end{aligned}$$

$$\begin{aligned}\frac{ad - bc}{\Delta U} \frac{\partial c_2}{\partial \lambda} &= \frac{(1-p)u'_2}{p + (1-p)e^{\lambda \Delta U}} \left(\left(-\frac{v''}{v'} + \lambda v' \right) \frac{1}{v'} - \frac{1}{u'_1} \left(\frac{u''_1}{u'_1} + \lambda u'_1 \right) \right) \\ \frac{ad - bc}{\Delta U} \frac{\partial S}{\partial \lambda} &= -\left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} + \left(\frac{u''_2}{u'_2} + \frac{p}{1+R} \Gamma \frac{u''_1}{u'_1} - \frac{(1-p)e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} \lambda u'_2 \Gamma \right)\end{aligned}$$

$$\begin{aligned}\frac{ad - bc}{\Delta U} \frac{\partial c_2}{\partial \lambda} &= \frac{1 - p}{p + (1 - p)e^{\lambda \Delta U}} \left(-\frac{v''}{v'} \frac{u'_2}{v'} - \frac{u''_1}{u'_1} \frac{u'_2}{u'_1} \right) \\ \frac{ad - bc}{\Delta U} \frac{\partial S}{\partial \lambda} &= \frac{p}{p + (1 - p)e^{\lambda \Delta U}} \left(\frac{u''_2}{u'_2} + \right) \frac{u'_2}{u'_1} \Gamma \frac{u''_1}{u'_1}\end{aligned}$$

We compute the determinant $\Lambda = ad - bc$:

$$\begin{aligned}\Lambda &= \left(\frac{u''_2}{u'_2} + \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} - \frac{1 - p}{1 + R} \frac{u'_1}{v'} \lambda u'_2 \Gamma \right) \left(-\frac{v''}{v'} + \lambda v' \right) \\ &\quad - \frac{1 - p}{1 + R} \left(\frac{u''_1}{u'_1} + \lambda u'_1 \right) \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \\ &= -\frac{v''}{v'} \left(\frac{u''_2}{u'_2} + \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} - \frac{1 - p}{1 + R} \frac{u'_1}{v'} \lambda u'_2 \Gamma \right) - \frac{1 - p}{1 + R} \frac{u''_1}{u'_1} \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \\ &\quad + \lambda v' \left(\frac{u''_2}{u'_2} + \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} - \frac{1 - p}{1 + R} \frac{u'_1}{v'} \lambda u'_2 \Gamma \right) - \frac{1 - p}{1 + R} \lambda u'_1 \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \\ &= -\frac{v''}{v'} \left(\frac{u''_2}{u'_2} + \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} - \frac{1 - p}{1 + R} \frac{u'_1}{v'} \lambda u'_2 \Gamma \right) - \frac{1 - p}{1 + R} \frac{u''_1}{u'_1} \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \\ &\quad + \lambda v' \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} + \frac{\lambda}{1 + R} (p e^{-\lambda \Delta U} + 1 - p) u'_1 \frac{u''_2}{u'_2} - \frac{1 - p}{1 + R} \lambda u'_1 \frac{u''_2}{u'_2} \\ &= -\frac{v''}{v'} \left(\frac{u''_2}{u'_2} + \frac{p}{1 + R} \Gamma \frac{u''_1}{u'_1} - \frac{1 - p}{1 + R} \frac{u'_1}{v'} \lambda u'_2 \Gamma \right) - \frac{1 - p}{1 + R} \frac{u''_1}{u'_1} \left(\frac{u''_2}{u'_2} - \lambda u'_2 \Gamma \right) \\ &\quad + \frac{\lambda}{1 + R} p e^{-\lambda \Delta U} \left(\Gamma u'_2 \frac{u''_1}{u'_1} + u'_1 \frac{u''_2}{u'_2} \right)\end{aligned}$$

From the last expression, it can readily be deduced that $\Lambda < 0$ and:

$$\begin{aligned}-\frac{\Lambda}{\Delta U} \frac{\partial c_2}{\partial \lambda} &= -\frac{1 - p}{p + (1 - p)e^{\lambda \Delta U}} \left(-\frac{v''}{v'} \frac{u'_2}{v'} - \frac{u''_1}{u'_1} \frac{u'_2}{u'_1} \right) < 0 \\ -\frac{\Lambda}{\Delta U} \frac{\partial S}{\partial \lambda} &= \frac{p}{p + (1 - p)e^{\lambda \Delta U}} \left(-\frac{u''_2}{u'_2} - \frac{u'_2}{u'_1} \Gamma \frac{u''_1}{u'_1} \right) > 0\end{aligned}$$

which means that $\frac{\partial c_2}{\partial \lambda} < 0$ and $\frac{\partial S}{\partial \lambda} > 0$.

For a , we have:

$$\begin{aligned}c_2 + \tau &= S + \frac{1 + R}{p} a \\ \frac{1 + R}{p} \frac{\partial a}{\partial \lambda} &= \Gamma \frac{\partial c_2}{\partial \lambda} - \frac{\partial S}{\partial \lambda} < 0 \\ - (p + (1 - p)e^{\lambda \Delta U}) \frac{\Lambda}{\Delta U} \frac{1 + R}{p} \frac{\partial a}{\partial \lambda} &= p \frac{u''_2}{u'_2} + (1 - p) \Gamma \frac{v''}{v'} \frac{u'_2}{v'} + \Gamma \frac{u''_1}{u'_1} \frac{u'_2}{u'_1}\end{aligned}$$

We finally have $\frac{\partial a}{\partial \lambda} < 0$.

2. We assume that $s, a > 0$, and $\tau = 0$. Previous equations are still valid, but with $\Gamma = 1$. We also obtain that $\frac{\partial a}{\partial \lambda} < 0$.

Moreover the condition for non-interior τ is $v'(0) \leq u'(c_2)$ remains true for a small increase in λ since $\frac{\partial c_2}{\partial \lambda} < 0$.

3. We assume that $\tau = s = 0$ and $a > 0$. We have:

$$\left[p + (1-p)e^{\lambda u_2(c_2)} \right] u_1' \left(W_0 - \frac{p c_2}{1+R} \right) = (1+R)u_2'(c_2) \geq (1+R)v'(0)e^{\lambda u_2(c_2)}$$

Deriving relative to λ yields:

$$\begin{aligned} (1+R)u_2'' \frac{\partial c_2}{\partial \lambda} &= -\frac{p}{1+R} \left(p + (1-p)e^{\lambda u_2(c_2)} \right) u_1'' \frac{\partial c_2}{\partial \lambda} + (1-p)e^{\lambda u_2(c_2)} u_1' \left(u_2 + \frac{\partial c_2}{\partial \lambda} \right) \\ &\quad \left((1+R)u_2'' + p u_2' \frac{u_1''}{u_1'} - (1-p)e^{\lambda u_2(c_2)} u_1' \right) \frac{\partial c_2}{\partial \lambda} = (1-p)e^{\lambda u_2(c_2)} u_1' u_2 \end{aligned}$$

It is straightforward that $\frac{\partial c_2}{\partial \lambda} < 0$ and from $c_2 = (1+R)a/p$, we deduce that $\frac{\partial a}{\partial \lambda} < 0$

4. We assume that $s > 0$ and $\tau > 0$ but $a = 0$. We show that an increase in λ implies a slacker constraint on a .

$$\begin{aligned} u_2'(c_2) &\leq \frac{1}{1+R} (p + (1-p)e^{\lambda \Delta U}) u_1'(c_1) \\ u_2'(c_2) &= v'(\tau) \\ p u_2'(c_2) + (1-p)v'(S)e^{\lambda \Delta U} &= \frac{1}{1+R} (p + (1-p)e^{\lambda \Delta U}) u_1'(c_1) \\ W_0 &= c_1 + \frac{c_2 + \tau}{1+R} = c_1 + \frac{S}{1+R} \\ c_2 + \tau &= S \end{aligned}$$

The derivation relative to λ yields:

$$\begin{aligned} \frac{\partial(c_2 + \tau)}{\partial \lambda} &= \Gamma \frac{\partial c_2}{\partial \lambda} \\ p u_2'' \frac{\partial c_2}{\partial \lambda} + (1-p) \left(v_S'' \frac{\partial S}{\partial \lambda} + v_S' \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \right) e^{\lambda \Delta U} &= \frac{p + (1-p)e^{\lambda \Delta U}}{1+R} u_1'' \frac{\partial c_1}{\partial \lambda} \\ &\quad + \frac{(1-p)e^{\lambda \Delta U}}{1+R} u_1' \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \\ \frac{\partial S}{\partial \lambda} &= -(1+R) \frac{\partial c_1}{\partial \lambda} = \Gamma \frac{\partial c_2}{\partial \lambda} \\ \frac{\partial \Delta U}{\partial \lambda} &= u_2' \frac{\partial c_2}{\partial \lambda} + v_\tau' \frac{\partial \tau}{\partial \lambda} - v_S' \frac{\partial S}{\partial \lambda} \end{aligned}$$

We define Λ as:

$$\begin{aligned}\Lambda &= p u_2'' + (1-p)v_S''\Gamma e^{\lambda\Delta U} + (1-p)\left(v_S' - \frac{1}{1+R}u_1'\right)e^{\lambda\Delta U}\lambda\Gamma (u_2' - v_S') + \frac{p+(1-p)e^{\lambda\Delta U}}{1+R}u_1''\frac{\Gamma}{1+R} \\ &= p u_2'' + (1-p)v_S''\Gamma e^{\lambda\Delta U} - p(1-p)\frac{(u_2' - v_S')^2 e^{\lambda\Delta U}}{p+(1-p)e^{\lambda\Delta U}}\lambda\Gamma + \frac{p+(1-p)e^{\lambda\Delta U}}{1+R}u_1''\frac{\Gamma}{1+R} < 0\end{aligned}$$

The second expression comes from the FOC, and $v_S' - \frac{1}{1+R}u_1' = -\frac{p(u_2' - v_S')}{p+(1-p)e^{\lambda\Delta U}}$. We deduce:

$$\begin{aligned}\Lambda \frac{\partial c_2}{\partial \lambda} &= -(1-p)\left(v_S' - \frac{1}{1+R}u_1'\right)e^{\lambda\Delta U}\Delta U = \frac{p(1-p)e^{\lambda\Delta U}\Delta U}{p+(1-p)e^{\lambda\Delta U}}(u_2' - v_S') \quad (49) \\ \frac{\partial S}{\partial \lambda} &= -(1+R)\frac{\partial c_1}{\partial \lambda} = \Gamma \frac{\partial c_2}{\partial \lambda} \\ \frac{\partial \Delta U}{\partial \lambda} &= \Gamma (u_2' - v_S') \frac{\partial c_2}{\partial \lambda} = \Gamma (u_2' - v_S')^2 \frac{p(1-p)e^{\lambda\Delta U}\Delta U}{p+(1-p)e^{\lambda\Delta U}} \frac{1}{\Lambda}\end{aligned}$$

We have $u_2' - v_S' \geq 0$. We deduce that c_2 and S decrease, while c_1 increases. Indeed, otherwise, $u_2' = v_S' < v_S'$ and $\tau \geq S$, which contradicts $c_2 = S - \tau \geq 0$.

We now want to check that the condition on a is more binding, i.e. that we have:

$$\begin{aligned}u_2'' \frac{\partial c_2}{\partial \lambda} &\leq e^{\lambda\Delta U} \left(v_S'' \frac{\partial S}{\partial \lambda} + v_S' \left(\lambda \frac{\partial \Delta U}{\partial \lambda} + \Delta U \right) \right) \\ (u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) \frac{\partial c_2}{\partial \lambda} &\leq e^{\lambda\Delta U} v_S' \Delta U\end{aligned}$$

If $u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')$ is positive, the result is true. If it is negative, we divide the previous inequality by the definition (49) of $\frac{\partial c_2}{\partial \lambda}$, and we get:

$$\begin{aligned}\frac{1}{\Lambda} (u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) &\leq -\frac{v_S'}{(1-p)\left(v_S' - \frac{1}{1+R}u_1'\right)} \\ (1-p)\left(v_S' - \frac{1}{1+R}u_1'\right) (u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) &\leq -\Lambda v_S' \\ v_S' (\Lambda + (1-p)u_2'' - (1-p)v_S''\Gamma e^{\lambda\Delta U} - (1-p)\lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) &\leq \\ \frac{1-p}{1+R}u_1' (u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) & \\ v_S' \left(u_2'' - \frac{1-p}{1+R}\lambda u_1' e^{\lambda\Delta U} \Gamma (u_2' - v_S') + \frac{p+(1-p)e^{\lambda\Delta U}}{1+R}u_1''\frac{\Gamma}{1+R} \right) &\leq \\ \frac{1-p}{1+R}u_1' (u_2'' - v_S''\Gamma e^{\lambda\Delta U} - \lambda v_S' e^{\lambda\Delta U} \Gamma (u_2' - v_S')) & \\ v_S' \left(u_2'' + \frac{p+(1-p)e^{\lambda\Delta U}}{1+R}u_1''\frac{\Gamma}{1+R} \right) &\leq \frac{1-p}{1+R}u_1' (u_2'' - v_S''\Gamma e^{\lambda\Delta U})\end{aligned}$$

It is equivalent to (using FOC for u'_1):

$$\begin{aligned}
v'_S \left(v''_S \Gamma e^{\lambda \Delta U} + \frac{p + (1-p)e^{\lambda \Delta U}}{1+R} u''_1 \frac{\Gamma}{1+R} \right) &\leq u''_2 \left(\frac{1-p}{1+R} u'_1 - v'_S \right) \\
v'_S \left(v''_S \Gamma e^{\lambda \Delta U} + \frac{p + (1-p)e^{\lambda \Delta U}}{1+R} u''_1 \frac{\Gamma}{1+R} \right) &\leq u''_2 \left(\frac{p(1-p) u'_2 + (1-p)^2 v'_S e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} u'_1 - v'_S \right) \\
v'_S \left(v''_S \Gamma e^{\lambda \Delta U} + \frac{p + (1-p)e^{\lambda \Delta U}}{1+R} u''_1 \frac{\Gamma}{1+R} \right) &\leq p u''_2 \frac{(1-p) u'_2 - v'_S - (1-p) v'_S e^{\lambda \Delta U}}{p + (1-p)e^{\lambda \Delta U}} u'_1
\end{aligned}$$

This relationship is always true since the LHS is negative, while the RHS is positive. Indeed, we know that $u'_2 \leq v'_S e^{\lambda \Delta U}$, and the RHS is greater than $-u'_1 v'_S u''_2 \geq 0$.

5. The case of $a = \tau = 0 < s$ is the same as the previous one with $\Gamma = 1$.

C Calibrations for alternative bequest specifications

We provide here calibrations for measuring the impact of the alternative bequest specifications. All calibrations generate a value of average bequest of 20% of the non-annuitized wealth W_0 and a rate of time discounting of 5.00%. The value of a statistical life is 500 consumptions at 65.

In all cases, the following parameters are fixed:

Parameters	σ	ψ	R
Values	2.0	1.0	3.00%

Calibrations lie in the following table:

Lockwood (2010)				De Nardi, French and Jones (2010)			
Multiplicative model		Linear model		Multiplicative model		Linear model	
Exogenous Parameters				Exogenous Parameters			
$W_0 = N = \sum_{k=0}^{T_M} \frac{p_{k+1 0}}{(1+R)^k} y$		158.66		$W_0 = N = \sum_{k=0}^{T_M} \frac{p_{k+1 0}}{(1+R)^k} y$		454.05	
y_0		183.01		y_0		354.93	
Estimated Parameters				Estimated Parameters			
u_0	1.018	u_0	1.77	u_0	0.355	u_0	0.69
λ	0.071	β	$\frac{1}{1+3\%}$	λ	0.204	β	$\frac{1}{1+3\%}$
θ	2.71×10^{-3}	θ	3.05×10^{-3}	θ	4.04×10^{-4}	θ	4.27×10^{-4}

De Nardi (2004)				Ameriks et al. (2011)			
Multiplicative model		Linear model		Multiplicative model		Linear model	
Exogenous Parameters				Exogenous Parameters			
$W_0 = N = \sum_{k=0}^{T_M} \frac{p_{k+1 0}}{(1+R)^k} y$		218.92		$W_0 = N = \sum_{k=0}^{T_M} \frac{p_{k+1 0}}{(1+R)^k} y$		43.76	
y_0		513.0		y_0		83.03	
Estimated Parameters				Estimated Parameters			
u_0	0.75	u_0	1.43	u_0	3.74	u_0	7.15
λ	0.096	β	$\frac{1}{1+3\%}$	λ	0.019	β	$\frac{1}{1+3\%}$
θ	1.17×10^{-3}	θ	1.37×10^{-3}	θ	0.031	θ	0.036

Table 4: Calibrations for alternative bequest motives