Health Cost Risk, Incomplete Markets, or Bequest Motives - Revisiting the Annuity Puzzle*

Kim Peijnenburg[†] Theo Nijman[‡] Bas J.M. Werker[§]

October 25, 2011

Abstract

It is well known that most rational life-cycle models predict much larger annuitization levels than those observed empirically. We examine the relative importance of three leading explanations for low annuity demand: health cost risk, incomplete annuity menus, and bequest motives. We find that high health cost risk can potentially explain very low annuity demand, while incomplete annuity menus and realistic bequest motives cannot. We find that the timing of the health cost risk is important. If out-of-pocket medical expenses can already be sizeable early in retirement, empirically observed low annuitization levels are optimal. In case health cost risk early in retirement is low, individuals can better save out of their annuity income to build a buffer for health cost shocks at later ages. Empirical evidence shows that in the US for many individuals health cost risk is indeed substantial early in retirement. Incomplete annuity markets do not reduce predicted annuity levels to the empirical levels, as agents are better of buying nominal annuities, save out of this income, and invest that in equity. Very high bequest motives can explain the low empirically observed annuity levels, but generate savings behavior inconsistent with the data.

Keywords: Asset allocation, retirement, life-cycle portfolio choice, annuity, savings
JEL classification: D14, D91, G11, G23

*We thank Andrew Ang, Peter Broer, Agar Brugiavini, Monika Bütler, Katie Carman, Norma Coe, Martijn Cremers, Nadine Gatzert, Donna Gilleskie, Rob van de Goorbergh, Thijs van der Heijden, David Hollanders, John Bailey Jones, Frank de Jong, Ralph Koijen, David Love, Ronald Mahieu, Raimond Maurer, Roel Mehlkopf, Olivia Mitchell, Eduard Ponds, Ralph Rogalla, Lisanne Sanders, Kent Smetters, Ralph Stevens, Marno Verbeek, Annie Yang, and seminar participants at Tilburg University, University at Albany, University of North Carolina at Chapel Hill, The Wharton School, European Economic Association, German Finance Association, IFID conference and the Netspar Pension Workshop for helpful comments and suggestions. Part of this research was conducted while Kim Peijnenburg was visiting the Wharton School of the University of Pennsylvania, and she gratefully acknowledges their hospitality. Bas Werker kindly acknowledges support from the Duisenberg School of Finance and Kim Peijnenburg from All Pensions Group (APG). This paper is the combined version of two papers previously circulated under the names ”Health Cost Risk: A Potential Solution to the Annuity Puzzle” and ”The Annuity Puzzle Remains a Puzzle”. The most recent version of this paper is available at www.kimpeijnenburg.com.

[†]Corresponding author. Bocconi University, IGIER, and Netspar. Via Roentgen 1, 20136 Milan, Italy. Phone: +39 02 5836 2723. Email: kim.peijnenburg@unibocconi.it.

[‡]Finance and Econometrics Group, CentER, Tilburg University and Netspar, Tilburg, the Netherlands, 5000 LE. Phone: 31-13-4662342. Email: Nyman@TilburgUniversity.nl.

[§]Finance and Econometrics Group, CentER, Tilburg University, Netspar, and Duisenberg School of Finance, Tilburg, the Netherlands, 5000 LE. Phone: 31-13-4662532. Email: Werker@TilburgUniversity.nl.

1 Introduction

As a consequence of an ageing population in developed countries, much attention (both by policymakers and academics) is directed towards providing and optimizing financial security during retirement. In this respect, the most important risks the elderly face are longevity risk and health risk. Annuities provide a life-time income until death, thus insuring people against longevity risk. However, in reality a relatively small amount of individuals voluntarily purchases annuity products when they reach retirement. A vast amount of literature focuses on trying to explain this annuity puzzle. At least a dozen potential explanations have been put forward, with several of them providing a reason for less than full annuitization, but generally not the very low voluntary annuitization rates that we observe empirically. Hence several reasons need to be combined to generate empirically plausible annuitization levels, but usually at the expense of creating additional puzzles. In this paper we explore the relative importance of three leading explanations for low annuity demand: health cost risk, incomplete annuity menus, and bequest motives. Furthermore, we determine whether one of these explanations can account for the empirically observed low annuity demands.

Prior research has shown that in simple stylized settings full annuitization of available wealth upon retirement is optimal for individuals who only face uncertainty about their time of death. Yaari (1965) shows that risk averse agents with intertemporally separable utility who are only exposed to longevity risk, and with no desire to leave a bequest, find it optimal to hold their entire wealth in annuities if these are actuarially fair. Annuities are attractive, as they generate a mortality credit that cannot be captured otherwise. This mortality credit is provided in return for giving up one's wealth after death.[1] After the seminal paper by Yaari (1965), a large literature arised that tried to explain the annuity puzzle and found that less than full annuitization can be generated, but not the very low levels seen in the data. Davidoff *et al.* (2005) found that the annuity puzzle is even deeper than previously thought as all one needs is complete markets and no bequests. Furthermore, even in the case of incomplete markets (modeled by assuming habit formation preferences, while only a real annuity is available) which results in a mismatch between desired and available income paths, the optimal annuity demand is higher than empirically observed. In contrast to Davidoff *et al.* (2005), we model incomplete annuity markets via a non-availability of real annuities, while agents prefer a flat real income. In our paper we explore three reasons that might lower annuity demand to the empirically observed levels: (1) health cost risk, (2) incomplete annuity markets, and (3) bequest motives. The intuition behind the influence that these factors might have on annuity demand is as follows. Out-of-pocket medical expenses can lower the optimal annuity demand as

[1]On top of the risk-free rate an annuity provides a mortality credit, because the survivors receive the assets of the deceased. For example, if p is the survival probability and r is the risk-free rate, than the gross return on a one-period annuity is (1 + r)/p, which is larger than the risk-free rate.

they raise the need for liquidity and hence give incentives for precautionary saving (De Nardi *et al.* (2010), Dynan *et al.* (2004), and Palumbo (1999)). As a consequence, uncertain medical costs can reduce the attractiveness of annuities since they impair the ability to smooth consumption in case of high and unexpected health costs. The second factor in our analysis is that annuity menus are typically incomplete. In many cases only nominal annuities are available rather than annuities which hedge inflation risk or which give exposure to equity markets. Hence if only nominal annuities are sold, agents still incur inflation risk and, on top of that, the nominal income in real terms is decreasing with age while agents usually prefer a flat consumption pattern. Such incomplete annuity menus have been found to result in large welfare costs (Horneff *et al.* (2008a) and Koijen *et al.* (2010b)). Third, bequest motives can reduce the attractiveness of annuities, since the wealth allocated to annuities is not bequeathed upon death.

We use a comprehensive stochastic life cycle model from retirement onwards to study the aforementioned reasons for reduced annuity demand. Retirees optimally choose the fraction of wealth annuitized at retirement and follow optimal consumption and asset allocation strategies afterwards, facing capital markets risk and inflation risk. Recently developed numerical methods are used to solve the model.

This paper contributes to the household economics literature in three ways. Our first contribution is that when comparing the relative importance of health cost risk, incomplete annuity markets, and bequest motives we find health cost risk to have a much larger effect on annuity levels. While in the benchmark case health cost risk lowers the annuity demand from 100% to 50% of *total* wealth (implying no additional annuities out of *liquid* wealth), the annuity demand is about 95% of total wealth when assuming either incomplete annuity markets or bequest motives.[2]

Furthermore, we find that the optimal annuity demand depends crucially on the timing of the health cost risk, namely the health cost risk *early* in retirement. The amount of out-of-pocket medical expenses after about 5 years following the annuitization decision is mostly irrelevant for optimal annuity demand. In case the health cost risk is moderate early in retirement, it is optimal for agents to annuitize all wealth and save out of the annuity income to build a sufficient buffer for high out-of-pocket medical expenses later in retirement. If instead out-of-pocket expenses can already be high early in retirement, agents keep a certain amount of wealth liquid, because they do not have enough time to build a buffer to be able to smooth consumption in case of a health cost shock. Hence if an agent perceives his or hers health cost risk to be high early in retirement, this can deter an agent from annuitizing their wealth. We explore this by examining the optimal annuity demand for two different specifications of health costs estimated in Ameriks *et al.* (2011) and De Nardi *et al.* (2010). The paper by Ameriks *et al.* (2011) examines a similar question as our paper, while De Nardi *et al.* (2010) focus on precautionary savings due to health expenses.

[2]Total wealth is pre-annuitized wealth plus liquid financial wealth.

Among other contributions, Ameriks *et al.* (2011) calculate the willingness to pay for an annuity which increases the fraction of total wealth annuitized from 55% to 70% for a fairly wealthy female. We expand on Ameriks *et al.* (2011) by determining the optimal annuity levels instead of willingness to participate in the annuity market. Furthermore, we explore annuity demand for heterogenous investors, which is particularly important since the empirically observed annuity levels vary sizeably with wealth levels.

Thus, as our third contribution, we compare the predicted annuity levels with the empirically observed annuitization levels as a function of wealth and find a close match. In reality, less wealthy agents have a higher fraction of total wealth annuitized (due to high pre-annuitized wealth levels). So the empirically observed fraction of total wealth annuitized is decreasing in total wealth, which we find to be close to the optimal pattern of annuitization. Hence besides proposing a possible solution for the annuity puzzle, we find that the empirical annuity pattern for varying total wealth levels, the annuity-wealth profile, is not far from the optimal pattern when taking into account that agents face (or perceive) high health cost risk early in retirement. We present empirical evidence of high health cost risk early in retirement for the average/median 65-year old. Naturally, health cost risk differs per individual, hence for some agents it will still be optimal to annuitize, while for agents facing ”average” health cost risk, the risk is so high that it deters them from annuitizing their liquid financial wealth.

The reason for exploring health cost risk is twofold. Previous papers (Pang & Warshawsky (2010), Sinclair & Smetters (2004), and Turra & Mitchell (2008)) that explore the impact of health cost risk on annuity demand have either used a small dataset (2 waves) to capture the expenses, not taken into account the correlation between medical expenses and survival probabilities, or only took into account long term care costs. Recent advances have been made by Ameriks *et al.* (2011) and De Nardi *et al.* (2010) to model health cost risk comprehensively, which are the specifications we employ in this paper. Second, previous papers did not calculate the predicted annuity demand for a wide range of wealth levels and explicitly compare them to observed annuity demand.

Similar to our paper, Pang & Warshawsky (2010) examine the effect of health risk on the annuitization decision and find that early in retirement it is optimal to annuitize nothing of your wealth and that from age seventy onwards the optimal annuitization fraction increases with age. This pattern is contrasting with what is observed in reality; agents do not keep buying additional annuities as they get older. Moreover, they find that health cost risk actually increases the demand for annuities later in life. The difference in results is due to their model setup, namely that additional annuities can be bought every year. Pang & Warshawsky (2010) state that annuities represent a specific asset class with its own unique risk and return profile. They model the annuitization decision essentially as a portfolio allocation decision between bonds, equity, and annuities. Since the mortality credit increases with age, an annuity bought at a later age earns a higher return than an annuity bought at age 65. In that case individuals find it optimal to first invest in equity to receive the equity risk premium, but eventually annuities crowd out equity. Health costs are an additional risk factor which drives households to shift demand from risky to riskless assets, namely from equity to bonds and annuities. Then as a consequence of the superiority of annuities over bonds, annuity demand increases due to health costs. Horneff *et al.* (2008a) and Horneff *et al.* (2008b) also find that the optimal annuitization level increases with age. The difference between our study and those mentioned above is that we assume that the annuitization decision takes place at retirement. Several arguments can be given to motivate this choice. First of all in several countries the decision whether to annuitize your pension account or take a lump sum is, due to the tax legislation, to take place at retirement. Furthermore, mandatory annuitization of (a fraction of) wealth at younger ages reduces adverse selection costs that are generated when the annuity date can be chosen. A third reason for our assumption of a single conversion opportunity at retirement is that in reality people make such financial decisions very infrequently rather than annually. Finally Agarwal *et al.* (2009) show that the capability of individuals to make financial decisions declines dramatically at higher ages, hence it seems optimal to make these decisions at younger ages when a person is still able to do so. Adding to these reasons is the complexity of solving such a life cycle model.

As in our paper, Davidoff *et al.* (2005) examine the effect of incomplete annuity markets on annuity demand. They find that low annuity purchases can only be reconciled by a large mismatch between the desired consumption path and available annuity income paths. In their paper they determine the optimal demand for a real annuity, when the optimal real consumption pattern is not flat. They assume a habit formation utility function, which creates the mismatch between the desired real consumption path and available income path (flat). While incomplete annuity markets can explain the lack of full annuitization, they cannot explain the low levels of annuitization found in reality. Our paper examines the impact of incomplete annuity markets, but approaches it from a different angle. We assume a desire for a smooth consumption path in real terms and show that, even if only nominal annuities are available, (almost) full annuitization is still optimal. Our paper extends on the work of Davidoff *et al.* (2005) as it is a more practically relevant calibration of incomplete annuity markets, since in reality many insurers do not offer inflation-indexed annuities (Brown (2007)), while individuals usually prefer a flat real consumption stream. If average annual inflation is 2% the real value of the nominal annuity income is halved in 35 years, thus creating a large mismatch between the desired income path and the available income path.

Among others, Lockwood (2011) and Inkmann *et al.* (2011) explore the impact of bequest motives and find that bequest motives alone can lower annuity demand, but not to empirically observed low levels. Lockwood (2011) examines several specifications for bequest motives and shows that combined with a load some specifications can lower the optimal demand sizeably below full annuitization. Furthermore, Inkmann *et al.* (2011) combine the many possible reasons for lower annuity demand and see whether that can explain low annuity levels. Our approach is not to bundle many potential reasons, but to examine the relative importance of the three most prevalent reasons. Since several papers find bequest motives to be important for annuity demand, we want to revisit this particular explanation.

In this paper we ignore a number of other potential drivers of annuity demand. The reason is that these explanations are less likely to lower the optimal annuity demand sufficiently to match the data. One of these explanations is the presence of loads on annuity prices. However, Mitchell *et al.* (1999) show that loads would have to be unrealistically high to be able to explain the observed low annuity demands. Furthermore, Brown *et al.* (2008a) find that about 3 out of 5 survey respondents state to favor the lump-sum to a social security annuity if it is actuarially fair. Family composition (Brown & Poterba (2000) and Kotlikoff & Spivak (1981)) can be a potential other reason. While Brown & Poterba (2000) shows that the utility gains from annuitization decrease when taking into account couples, the gains are such an order of magnitude (between 10% and 30% for agents with 50% of total wealth pre-annuitized) that a substantial fraction of couples should still fully annuitize.

Several papers have combined different factors for low annuity demand, but in such cases many times new puzzles are created. For instance, Dushi & Webb (2004) show that the combination of high loads, pre-existing annuities, and risk sharing within couples can explain observed low levels of annuitization. However, such a model predicts that annuity demand of singles should be much higher than that of couples and that individuals should annuitize after the death of their spouse. Neither is consistent with the data. These extensive rational frameworks did not find a definitive answer to the annuity puzzle, giving rise to behavioral explanations. For example framing of the annuity choice (Agnew *et al.* (2008), Brown *et al.* (2008b), and Gazzale & Walker (2009)), mental accounting (Hu & Scott (2007)), and complexity of the annuity product (Brown (2007)). Surely part of the explanation for the puzzle lies in the behavioral area, however a better understanding of the rational potential reasons for low annuity demand and their relative importance is vital. Recent advances in estimating models for health cost risk provide new insights into the importance of this factor, which is shown in this paper. Furthermore, when formulating policy regarding stimulating annuity demand (for certain types of agents) it is vital to have a better notion of what the optimal level is.

The remainder of the paper is organized as follows. Section 2 describes the individual's preferences, the setup of the financial market, and the benchmark parameters. Details on the two models for health cost risk that we use are also given in Section 2. The main comparison between the effect of health cost risk, incomplete annuity markets, and bequest motives on annuity demand are presented in Section 3. Section 4 elaborates on the influence of health cost risk on annuity demand and Section 5 and 6 deal with the other two potential solutions to the annuity puzzle: incomplete annuity menus and bequest motives, respectively. Section 7 concludes.

2 The retirement phase life cycle model

2.1 Individual's preferences and constraints

We consider a life-cycle investor during retirement with age $t \in 1, ..., T$, where $t = 1$ is the retirement age and T is the maximum age possible. The individual's preferences are presented by a time-separable, constant relative risk aversion utility function over real consumption, C_t. More formally, the objective of the retiree is to maximize

$$V = E_0 \left[\sum_{t=1}^{T} \beta^{t-1} \left(\prod_{s=1}^{t} p_s \right) \frac{C_t^{1-\gamma}}{1-\gamma} \right], \tag{1}$$

where β is the time preference discount factor, γ denotes the level of risk aversion, and C_t is the real amount of wealth consumed at the beginning of period t. The probability of surviving to age t, conditional on having lived to period $t-1$, is indicated by p_t. We denote the nominal consumption as $\overline{C_t} = C_t \Pi_t$, where Π_t is the price index at time t.

The individual invests a fraction w_t in equity, which yields a gross nominal return R_{t+1} in year $t+1$. The remainder of liquid wealth is invested in a riskless bond and the return on this bond is denoted by R_t^f. The intertemporal budget constraint of the individual is, in nominal terms, equal to

$$W_{t+1} = \max(W_t + Y_t - H_t - \overline{C_t}, 0)(1 + R_t^f + (R_{t+1} - R_t^f)w_t), \tag{2}$$

where W_t is the amount of financial wealth at time t, Y_t is the annual nominal annuity income, and H_t are health costs. The timing of decisions is as follows. At retirement the agent decides which fraction of wealth to invest in annuities. Subsequently, the individual receives annuity income and incurs health costs. After this exogenous shock, the agent decides how much to consume and subsequently invests the remaining liquid wealth. In case the annuity income plus wealth at the beginning of the period is insufficient to pay the health expenses and consume, the individual receives a low minimum consumption level, C_{min}, since almost all western countries have a minimum consumption floor. We perform analyses to show the influence of the minimum consumption floor on the optimal annuity choice in Section 4.6. The decision frequency for the optimal consumption and asset allocation is annually.

The individual faces a number of constraints on the consumption and investment decisions.

First, we assume that the retiree faces borrowing and short-sales constraints

$$w_t \geq 0 \text{ and } w_t \leq 1. \tag{3}$$

Second, we impose that the investor is liquidity constrained

$$\overline{C_t} \leq W_t, \tag{4}$$

which implies that the individual cannot borrow against future annuity income to increase consumption today. The reason we impose this restriction is that in reality it is difficult to get a loan, especially for an elderly person.

2.2 Financial market

The asset menu of an investor consists of a riskless one-year nominal bond and a risky stock. The return on the stock is lognormally distributed with an annual mean nominal return μ_R and a standard deviation σ_R. We assume the nominal interest rate is generated by a Vasicek model, to account for long term mean reversion. The real yield is equal to the nominal yield minus expected inflation and an inflation risk premium.

We model inflation, because in our analysis we want to examine optimal annuitization levels in a world with inflation where only nominal annuities are available. For the instantaneous *expected* inflation rate we assume

$$\pi_{t+1} = \pi_t + a_\pi(\pi_t - \mu_\pi) + \epsilon^\pi_{t+1}, \tag{5}$$

where a_π is the mean reversion parameter, μ_π is long run expected inflation, and the error term $\epsilon^\pi_t \sim N(0, \sigma^2_\pi)$. Subsequently the price index Π follows from

$$\Pi_{t+1} = \Pi_t \exp(\pi_{t+1} + \epsilon^\Pi_{t+1}), \tag{6}$$

where $\epsilon^\Pi_t \sim N(0, \sigma^2_\Pi)$ are the innovations to the price index. We assume there is a positive relation between the expected inflation and the instantaneous short interest rate, that is the correlation coefficient between ϵ^r_t and ϵ^Π_t is positive. The parameters we use are described in Section 2.4.

We consider single-premium immediate life-contingent annuities with real or nominal payouts. Consequently, the annuity income is given by

$$Y = PR_0 A^{-1}, \tag{7}$$

where PR_0 is the premium and A is the annuity factor. The single premium is equal to the present value of expected benefits paid to the annuitant and we assume an actuarially fair annuity. The annuity factor, A, is thus equal to

$$A = \sum_{t=1}^{T} \exp(-tR_0^{(t)}) \prod_{s=1}^{t} p_s, \tag{8}$$

where $R_0^{(t)}$ is the time zero yield on a zero coupon bond maturing at time t. The interest rate term structure that is applied is either nominal or real depending on the type of annuity. The annuity factor for a variable annuity payout is similar to equation (10), but R(t) 0 is equal to the assumed interest rate (AIR), which is fixed. The annual annuity income depends on the return of the portfolio backing the annuity and the AIR determines whether, in expectation, the annuity payout stream increases or decreases over time. The method used to solve our life-cycle problem is described in the Online Appendix.

2.3 Health cost models for out-of-pocket expenditures

Several papers in the literature estimate out-of-pocket medical expenses, though the estimated dynamics for health cost risk differs substantially. For this reason we take the estimates for the process of health expenses from two prominent papers in the literature and determine the optimal annuity demand. In this manner we can, as a first step, disentangle what characteristics of health costs are the main determinant of annuity demand. We examine two different models for health costs: (1) De Nardi *et al.* (2010) and (2) Ameriks *et al.* (2011).[3] Both models vary according to how they specify the stochastic process for health costs, survival probabilities, and the dataset and/or period employed. The details of both health cost models are presented in the Online Appendix.

A key feature of both models is that health costs and survival probabilities are negatively correlated, which is in contrast to the specification used in Pang & Warshawsky (2010) to explore the impact of health cost risk on annuity demand.[4] Both the medical expenditures and survival probabilities depend on the health status of the agent. So in case the agent is in a bad health status, his expected medical expenses are higher *and* his life expectancy is lower. This is particularly important when examining the effect of health costs on annuity demand. Namely the negative correlation between survival probabilities and life expectancy can make annuities relatively more attractive, because after having incurred large health expenses, the agent is more likely to die, which could make the depletion of wealth due to the medical expenses less costly in utility terms.

[3]We also performed the analysis for two additional health cost models estimated by Scholz *et al.* (2006) and French & Jones (2004). These results can be obtained upon request from the authors.

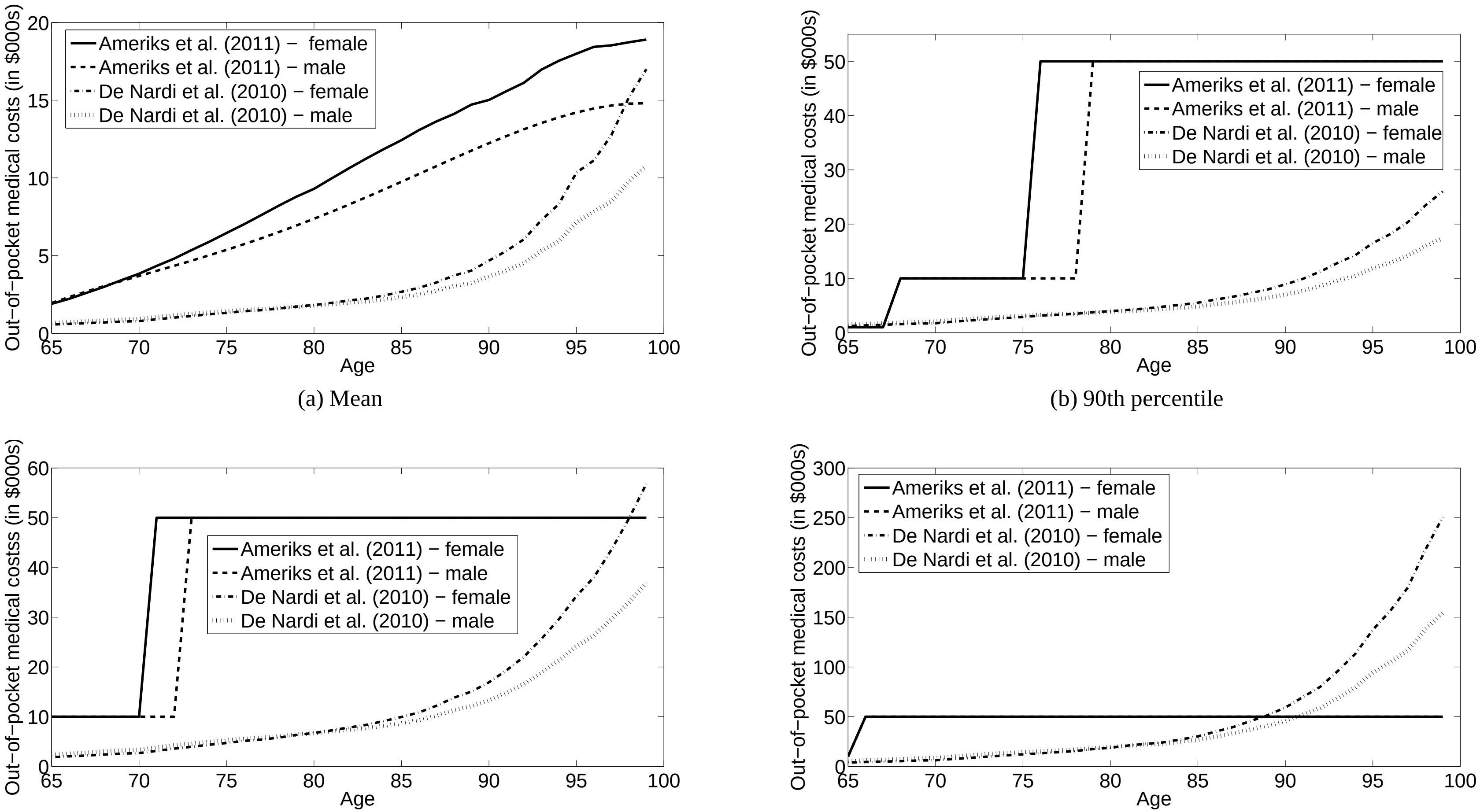
[4]They also use the model by De Nardi *et al.* (2010) but to avoid tracking the health status, they take the average mortality rates and health expenses across people in each income decile.

Figure 1 displays the mean and quantiles of medical expenses for the two health cost specifications. Most importantly, we see that the pattern of health costs over the life cycle differs substantially between the two models, as well as the amount of health costs over the entire life. Panel (a) shows the mean, and we see that the average health costs from the De Nardi *et al.* (2010) model increases substantially with age. This pattern also holds for the three quantiles that are displayed in Panel (b) to (d). The health costs according to Ameriks *et al.* (2011) show a different shape compared to De Nardi *et al.* (2010). When focusing on the 99th percentile, we see that the Ameriks *et al.* (2011) specification implies large health cost risk early in retirement, but less risk later in retirement compared to De Nardi *et al.* (2010). Furthermore, the shape of the curves differ because the health costs according to the Ameriks *et al.* (2011) specification take on only discrete levels ($1000, $10,000, and $50,000). Health costs for males are a bit lower than for females, in both models. In the next section we show that the tail of the health costs, in particular in the first years after retirement, is important for determining the optimal annuity demand. Therefore, we will mostly refer to the two health cost models according to the level of the health cost risk early in retirement:

- Low health cost risk early in retirement = De Nardi *et al.* (2010)
- High health cost risk early in retirement = Ameriks *et al.* (2011)

2.4 Benchmark parameters

We do not estimate the parameters ourselves, but employ common parameters used within the life-cycle literature. As in Pang & Warshawsky (2010) and Yogo (2009), we set β, the time preference discount factor, equal to 0.96. The risk aversion coefficient γ is 5. We determine the optimal annuity demand for a range of initial total wealth levels, but to illustrate the consumption and savings decisions, we use a benchmark wealth level of $350,000. This is approximately equal to the average total wealth level for a single person U.S. household (Dushi & Webb (2004)), where total wealth consists of pre-annuitized wealth and liquid financial wealth. The minimum consumption level guaranteed by the government is set equal to $7000 annually. Ameriks *et al.* (2011) note that the payments under the governments Supplemental Security Income are about $7000 per year and they estimate the consumption floor to be $5700.



(a) Mean

(b) 90th percentile

(c) 95th percentile

(d) 99th percentile

Figure 1: Simulated annual out-of-pocket health costs from age 65 to 100

This graph displays the mean, the 90th, 95th, and 99th percentile of health costs for two models; (1) Ameriks *et al.* (2011) and (2) De Nardi *et al.* (2010).

The equity return is assumed to be lognormally distributed and in accordance with historical stock returns we assume a mean annual nominal return of 8% and an annual standard deviation of 20%. The mean instantaneous short rate is set equal to 4%, the standard deviation to 1%, and the mean reversion parameter to -0.15. The inflation risk premium to determine the real yield is 0.5%. The correlation between the instantaneous short rate and the expected inflation is 0.4. Mean inflation is equal to 2% and the standard deviation of the instantaneous inflation rate is equal to 1.3%. The standard deviation of the price index equals 1.3% and the mean reversion coefficient equals -0.15. Time ranges from $t = 1$ to time T, which corresponds to age 65 and 100 respectively.

3 Main comparison of the impact of health risk, incomplete markets, and bequest motives

In this section we show the impact of health cost risk, incomplete annuity markets, and bequest motives on optimal annuity demand. In the subsequent sections we will elaborate on these findings and explore in more depth under which circumstances our results hold. In the literature a lot of attention is devoted to explaining the low empirically observed annuity levels: the annuity puzzle. However, whether there is really a ”puzzle” depends on the wealth level of the individual. The empirically established annuity levels as fraction of total wealth for high wealth levels can be as low as 50%, compared to much higher levels for less wealthy individuals. Dushi & Webb (2004) report the pre-annuitized fraction of wealth at age 65 of a single female for various wealth levels, which is displayed in Figure 2. The solid line presents the empirical annuitization levels, which are decreasing in the wealth level. When we compare these empirical annuitization levels with the optimal levels from our model described in Section 2, we see that two explanations can potentially account for low voluntary annuity demand; (1) high health cost risk early in retirement and (2) high bequest motives. Both of these explanations can lower annuity demand to the empirically observed levels, and, on top of that, they predict a pattern of annuitization as a function of wealth which is similar to the empirical levels. However, when agents face low health cost risk early in retirement, 100% annuitization is still optimal. This remains the case if annuity markets are incomplete, hence agents can only purchase nominal annuities, and when agents face moderate bequest motives. Details for the intuition behind these findings are presented in the subsequent three sections.

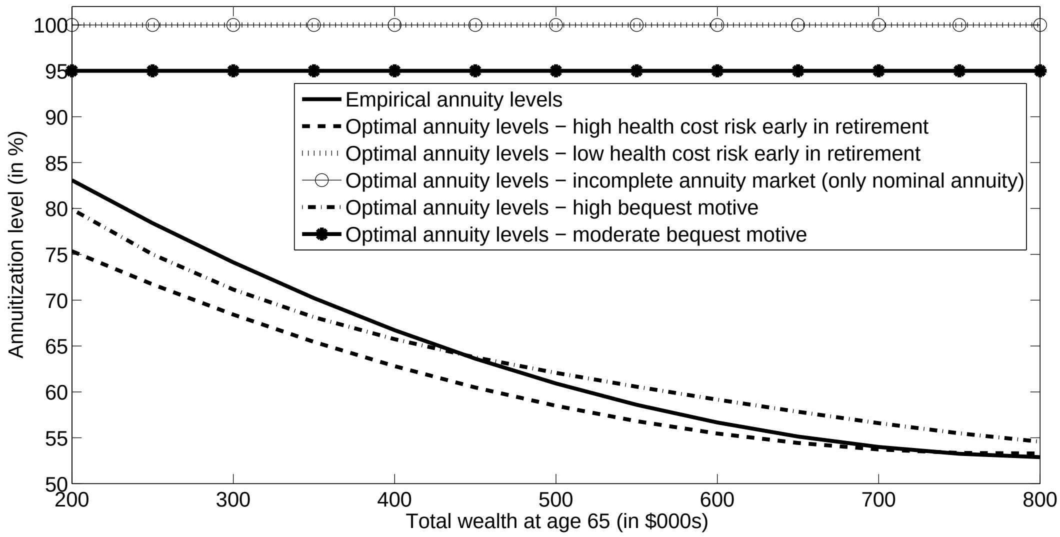


Figure 2: Comparing empirical annuitization levels with optimal annuitization levels.
We display the annuitization levels for a single female at age 65 estimated in Dushi & Webb (2004). They use data from the HRS to estimate the fraction of wealth annuitized. We present the fraction annuitized as percentage of the sum of liquid financial wealth and pre-annuitized wealth. Liquid financial wealth includes financial assets, IRA's, and DC pensions. Pre-annuitized wealth includes social security and DB pensions.

4 Annuity levels and health cost risk

For several decades the annuity puzzle has been explored. Many advances have been made and three of the prevailing reasons found are health cost risk, incomplete annuity menus, and bequest motives. While most papers conclude that the optimal annuity demand is indeed less than 100% of total wealth, in many cases the predicted annuitization level is not as low as in reality and an explicit comparison to the empirically observed annuity levels is not performed. This is were this paper comes in. We explore in great depth the three different explanations separately (Chapter 4, 5, and 6) to compare their relative importance and examine whether either of the explanations has the potential to explain the annuity puzzle.

Full annuitization is optimal in a world where individuals only face longevity risk (Yaari (1965)). However, this result might no longer hold if individuals face substantial health cost risk which raises liquidity needs. In case an agent has insufficient liquid financial wealth to pay the health costs, he or she can only consume a very low level; the minimum consumption level. Furthermore, since health costs are positively autocorrelated, it is more likely that the low consumption levels will persist. In order to study the impact of health cost risk, we focus in Section 5.1 on optimal annuity demand and savings decision when agents face high health cost risk early in retirement (Ameriks *et al.* (2011)) and the impact of low health cost risk early in retirement (De Nardi *et al.* (2010)) will be explored in Section 4.2.

4.1 Optimal annuity demand and savings with *high health cost risk early in retirement*

In Figure 3 we present (for our benchmark specification) the certainty equivalent consumption for various annuitization levels, adopting optimal post-retirement consumption and asset allocation strategies. The dotted line presents the certainty equivalent consumption for a female who does not face out-of-pocket medical expenses. In accordance with previous literature, we find that in that case (almost) full annuitization is optimal. The welfare gains from optimal annuitization compared to no annuitization are substantial: the certainty equivalent consumption increases from $15,000 to $22,000. Davidoff *et al.* (2005) and Mitchell *et al.* (1999) also find high welfare gains. Our goal is to determine whether full annuitization remains optimal if individuals face substantial medical expense risk resulting from the health cost specification of Ameriks *et al.* (2011) with high health cost risk early in retirement. The solid line shows the optimal annuitization level in case the agent faces health costs; the optimal annuity demand is decreased to 65% of total wealth. The benefits of insurance against longevity risk and receiving the mortality credit are outweighed by the (initial) reduction in liquidity. We present the optimal annuity fraction as a function of total wealth, where total wealth consists of liquid financial wealth and pre-annuitized wealth. In most instances, an agent has a certain amount pre-annuitized in the form of social security and/or defined benefit pension wealth. So if an individual already has 70% of total wealth pre-annuitized, he or she should not buy additional annuities.[5]

The previous results also hold for males. The optimal annuity demand is reduced substantially due to out-of-pocket medical expenses, but to a slightly lesser extent than for females. This is not surprising since males face lower out-of-pocket medical expense risk. Furthermore, we see in the figure that the certainty equivalent consumption for males is substantially higher than for females, for all annuitization levels. The reason is that for males both health costs are lower and the annuity income is larger. The income differs since both are actuarially fair for each group and calculated separately. As male life expectancy is lower, the annuity is cheaper.

In Figure 4 we present the median optimal consumption and wealth paths for three cases: (dotted line) no annuitization, (solid line) optimal annuitization (=65%) with health costs, and (dashed line) optimal annuitization (=95%) without health costs. Figure 4a shows that in case of no annuitization, the optimal consumption path is decreasing over time. This reflects the fact that if the longevity risk in the real consumption level is not hedged, agents do not plan much consumption at ages where the probability is high that one will have passed away. If agents face no health cost risk and buy real annuities (dashed line), then inflation risk is hedged and the planned consumption

[5] Our benchmark total wealth equals $350,000 and we will present the optimal annuity demand for different wealth levels in Section 4.4.

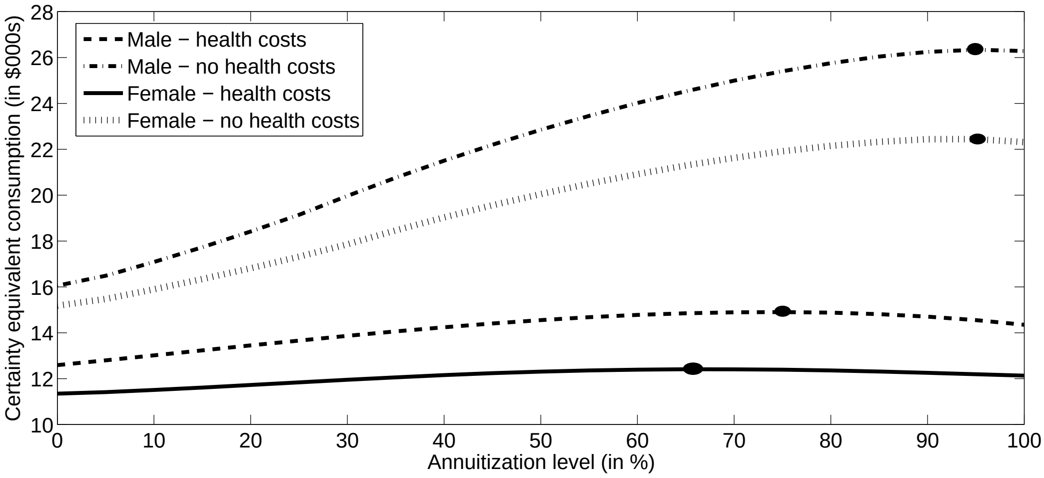
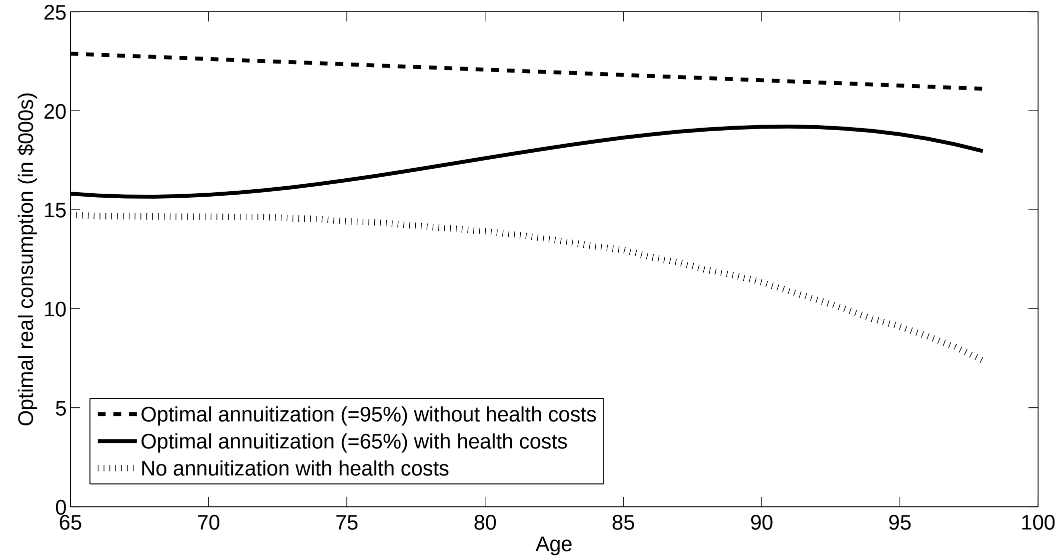


Figure 3: Optimal annuitization levels with high health cost risk early in retirement
The figure displays the certainty equivalent consumption for the life-cycle model with and without medical expenses for males and females. The case without medical expenses does include health status uncertainty and longevity uncertainty. So the difference is whether or not the agent needs to pay medical costs out-of-pocket. The optimal annuitization strategy is the level that generates the highest certainty equivalent consumption. The health cost specification employed is from Ameriks *et al.* (2011).

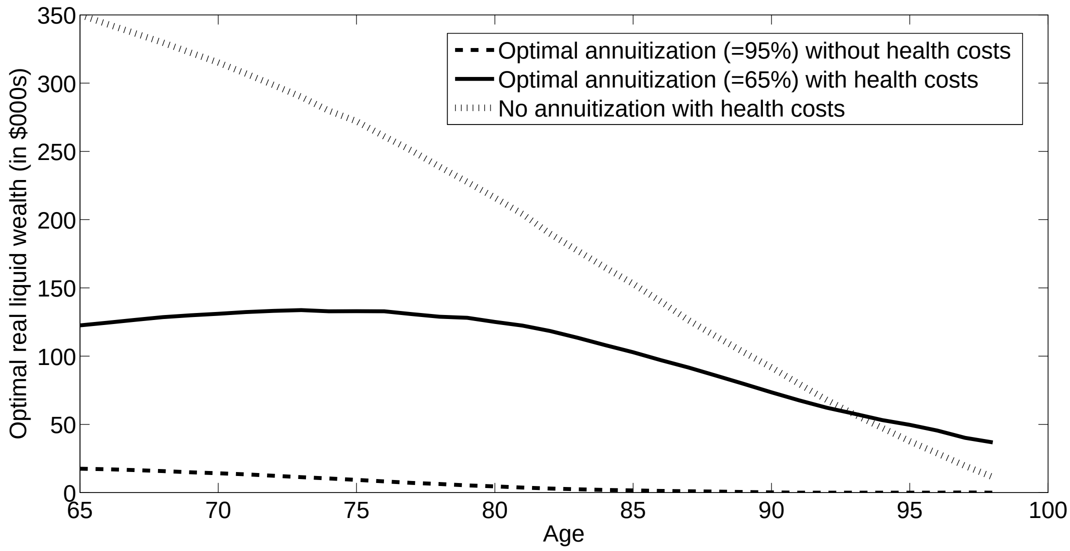
path is approximately flat in real terms (in our specification the time preference parameter and interest rates coincide approximately). However, we see that if an individual faces out-of-pocket medical expenses (solid line), the median consumption path is lower. This is because the individual has to pay the medical expenses and wants to keep a certain amount of wealth liquid to be able smooth consumption in case of high health costs. The consumption at younger ages is slightly lower, because the agent saves to increase the buffer further, while at later ages the individual starts dissaving, thus increasing consumption.

The optimal liquid wealth trajectories are displayed in Figure 4b. If no annuities are bought (dotted line), the median optimal wealth trajectory is decreasing over time. Individuals slowly dissave out of their liquid wealth. If the agent faces health cost risk and invests optimally in a real annuity he keeps a substantial amount of wealth liquid; about $120,000 until age 80. After that age he slowly dissaves out of the wealth buffer. These high levels of precautionary savings are in accordance with Palumbo (1999), De Nardi *et al.* (2010), and Love *et al.* (2009), who show that out-of-pocket medical expenses induce individuals to hold large amounts of precautionary savings. If the agent does not face out-of-pocket medical expenses and annuitizes (almost) his entire wealth, the wealth levels over the life cycle are low (dashed line).

Two stylized facts about the age-wealth trajectory in the data are matched: substantial precautionary savings and slow dissaving at later ages (De Nardi *et al.* (2010) and Palumbo (1999)). Furthermore, in reality many retirees die with large positive amounts of wealth, which is confirmed by our findings: agents keep a buffer for health costs until late in life. Hence, high out-of-pocket



(a) Real consumption



(b) Real liquid wealth

Figure 4: Optimal consumption and wealth paths over the life cycle
Panel (a) displays the median optimal real consumption when agents do not face health costs and annuitize optimally (dashed line), when agents do face health costs and annuitize optimally (solid line), and when agents face health costs but do not annuitize (dotted line). Panel (b) displays the optimal *liquid* real wealth for the same situations. The optimal levels are for a female. The health cost specification employed is from Ameriks *et al.* (2011).

medical expense risk early in retirement can help to simultaneously explain the low observed annuity levels as well as precautionary savings.

4.2 Optimal annuity demand with *low health cost risk early in retirement*

In Section 3 in Figure 2 we displayed the optimal annuity demand in case the agent faces low health cost risk early in retirement (dotted line), and found that full annuitization is optimal.[6] These results are in sharp contrast with the sizeable decrease in optimal annuity demand found in the previous section, when agents face health costs according to the specification by Ameriks *et al.* (2011). The main driver of the variation in results is the difference in timing of the health cost risk, more specifically, the health cost risk *early* in retirement. For the Ameriks *et al.* (2011) health cost specification, health cost risk early in retirement is high, while for the De Nardi *et al.* (2010) specification the risk is low. This can be seen from Figure 1d: for the Ameriks *et al.* (2011) specification there is a 1% probability to incur health costs of $50,000 already at age 66. In contrast, the health cost risk implied by the specification of De Nardi *et al.* (2010) is low early in retirement, but high at later ages.[7] In case of low health cost risk early in retirement, the retiree has enough time to build a large buffer out of the annuity income to insure against health costs later in life.

Pang & Warshawsky (2010) also use the health cost model of De Nardi *et al.* (2010), but find that annuity demand increases due to these health costs. The reason for this seemingly contrasting result is that they do not model annuitization as a one-time decision that is made at retirement, but instead, they optimize annually over the equity-bond-annuity portfolio. Pang & Warshawsky (2010), in effect, modeled the annuitization decision as a portfolio allocation decision. Health costs are an additional risk factor which drives households to shift demand from risky to riskless assets, namely from equity to bonds and annuities. As a consequence of the superiority of annuities over bonds, annuity demand increases due to health costs. Similar modeling assumptions and findings are presented in Pashchenko (2010).

To provide some additional proof for the importance of the timing of health cost risk, we calculate the optimal annuity demand for agents facing health cost risk according to the Ameriks *et al.* (2011) specification from age 70 onwards, but not from age 65 to 69. Hence, they do not

[6]We conducted a myriad of robustness tests and found full annuitization to be optimal in every case. This result holds for both males and females, both when annuity income is calculated actuarially fair for both groups separately and via the average survival probability for males and females. Furthermore, we find that this result holds if individuals can only invest in a nominal annuity. In addition, we tested whether our results hold when taking into account high end-of-life health costs. Among others, Werblow *et al.* (2007) find that proximity to death is a more important determinant of health costs than age. In the health cost specifications we employed this proximity to death effect is not incorporated explicitly. To test whether our results still hold, we include a time-to-death effect according to the findings in Werblow *et al.* (2007). Namely, we alter the medical costs of the De Nardi *et al.* (2010) specification, by increasing the expenses in the year before death with a factor between 2 and 3, depending linearly on the age at death. The optimal annuity demand when we added the end-of-life costs did not change, which is intuitive since the health costs have mostly only risen at advanced ages.

[7]The subsequent analysis is also performed for two additional health cost models estimated by Scholz *et al.* (2006) and French & Jones (2004). The results confirm that indeed if health cost risk is high early in retirement, optimal annuity demand is reduced.

face health cost risk during the first five years of retirement.[8] We find that in this case the optimal annuity demand is increased from 65% to 90%, which is accordance with the finding that the driver of annuity demand is health cost risk *early* in retirement.

4.3 Summary statistics on high health cost risk early in retirement

In this section we present summary statistics to show that for the average person in the US medical expense risk can be high early in retirement. To that end we need to make a distinction between long term care costs and other health costs. The largest health cost risk that agents face are long term care costs, since nursing home costs are very high and "in need of long term care" is a highly persistent health status. Therefore, when assessing the level of risk early in retirement, utilization rates of nursing homes during those ages are particularly relevant. Brown & Finkelstein (2007) estimate that the probability that a 65-year old is in a nursing home or assisted living at age 70 is 0.7% and 0.5% for respectively females and males, which is a very substantial risk. Furthermore, these numbers are conditional on being eligible for purchasing long term care insurance.[9] Hence the risk of going to a nursing home early in retirement is much higher than this 0.7%, because this number is based on the least risky 65-year olds. Murtaugh *et al.* (1995) estimate that about 12% to 23% of the 65-year olds would be rejected for private long term care insurance. Furthermore, the costs associated with living in a nursing home amount to about $75,000 a year for a semi-private room. Only 4% of long term care costs are covered by private insurance and 25% by medicare. Medicaid is available for elderly individuals with no assets, but this is a poor substitute for private care (Ameriks *et al.* (2011) and Brown & Finkelstein (2007)).

Health cost risk excluding long term care is also sizeable. The Medical Expenditure Panel Survey (MEPS) contains data on out-of-pocket medical expenses for the U.S. non-institutionalized population. Using MEPS data from 2006-2008, we find that there is a 1% probability of having health costs (excluding long term care costs) higher than $9,000.

When comparing the health cost processes that we use, note that the model by Ameriks *et al.* (2011) matches the empirical utilization levels of long term care at different ages and find that health cost risk is already high early in retirement. On the other hand, De Nardi *et al.* (2010) find low health cost risk early in retirement. However, this can be (partially) attributed to the dataset that is used to estimate the health cost process in De Nardi *et al.* (2010). The individuals in the AHEAD dataset, which is a part of the HRS, are non-institutionalized and over 70 at the start of the survey in 1994. Therefore the estimation of the health cost process is based on a relatively healthy

[8]We assume agents face zero health costs from age 65 to 69, but the results do not change if we assume that they face a constant health cost equal to the mean health cost at those ages.

[9]The distinction on being eligible for purchasing long term care insurance is not relevant for our paper. This is however the only estimate of transition probabilities into a nursing home from age 65 to 70 that we could find in the literature.

subsample of the population, namely non-institutionalized, and understates the health cost risk early in retirement, but not so much at later ages. In addition, De Nardi *et al.* (2010) note that their estimates potentially underestimate the out-of-pocket expenses for another reason. The expenses reported in the HRS do not include payments made by medicaid when agents could not pay for the out-of-pocket costs themselves. So when simulating the out-of-pocket expenses, a substantial part of these out-of-pocket costs is in the simulations not actually payed "out-of-pocket", in case agents did not have enough wealth and received the minimum consumption level.

4.4 Optimal annuitization for various wealth levels

In this section we determine the optimal annuity demand for different wealth levels using the model with high health cost risk early in retirement (Ameriks *et al.* (2011)). The results are displayed in Figure 5. We see that annuitization levels exhibit a slight U-shaped pattern. If total wealth at retirement is low, optimal annuity levels are higher. For a total wealth level of $200,000 at age 65, it is optimal to have more than 75% of total wealth in annuities.[10] This is because the difference between the normal consumption level and the minimum consumption level is not that high. In numbers; a wealth level of $200,000 can generate an annuity income of about $12,000, which differs only $5000 from the minimum consumption level. If an individual is hit by a large health cost shock and receives the government guaranteed consumption level, this is not so costly in utility terms, because the fall in consumption is not that high.

For intermediate wealth levels, the fall in utility is larger if hit by a health shock, because the difference between the normal (annuity) income and the minimum consumption level is higher. For this reason it is optimal to reduce annuity demand to be able to smooth consumption and prevent consuming only the minimum level. For higher wealth levels, the optimal demand rises slightly. If the wealth level is higher it is easier for agents to build up a buffer fast to insure against health shocks. Our finding of a non-linearity in optimal annuitization levels is analogue to Ameriks *et al.* (2011) who find a non-linearity in savings motives. As above, they explain that agents with high wealth and income levels have less incentives to save, because they have enough income to pay their medical expenditures. While, on the other hand, poor individuals can never save enough to be able to afford the high medical costs. However the incentives to save are especially high for intermediate income and wealth levels. Summarizing, optimal annuity demand depends critically on two factors: (1) the cumulative health costs in the first retirement years, and (2) the savings ability (wealth) in the first years to cover these costs.

[10] We present the annuitization levels as a fraction of *total* wealth, where total wealth is pre-annuitized wealth plus liquid financial wealth. Assume an agent has a social security income and defined benefit pension income, which amounts to a net present value of $150,000. If he or she has in addition liquid wealth of $50,000 (so 75% pre-annuitized), then the retiree should not annuitize the liquid wealth, since he or she has already 75% pre-annuitized of total wealth.

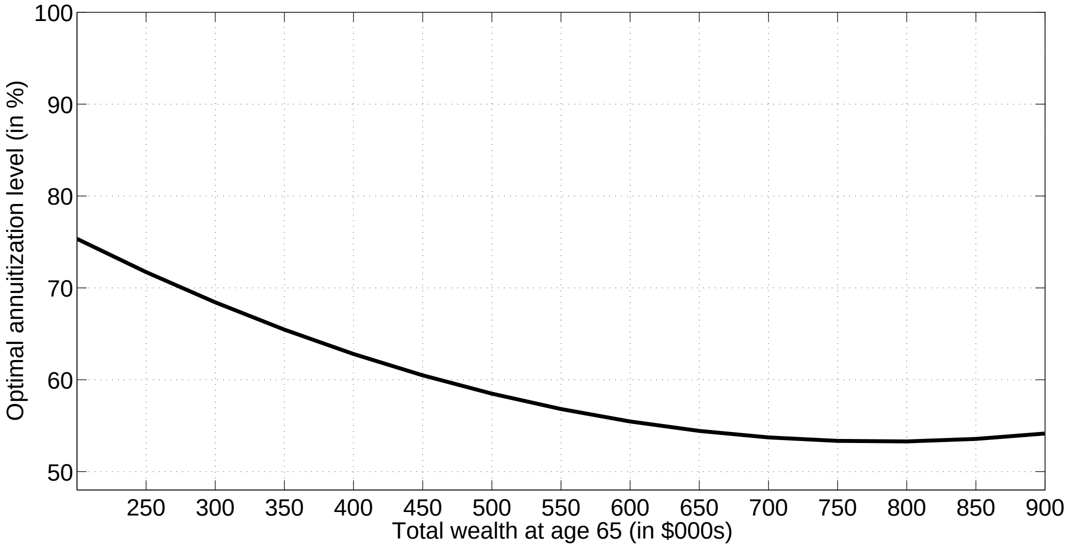


Figure 5: Optimal annuity levels using the high health cost risk early in retirement model for varying initial wealth levels at age 65
The figure displays the optimal annuity demand for a 65-year old for varying wealth levels. The numbers are in thousands of dollars.

Note that the health cost model which we labeled ”high health cost risk early in retirement”, is the specification of Ameriks *et al.* (2011). They estimate the willingness to pay (WTP) for an annuity with a price of $85,000 which generates an income of $5000 per year, for a healthy 62-year old female who has about 55% of wealth pre-annuitized.[11] They find a WTP for this annuity of 0.94.[12] We extend on Ameriks *et al.* (2011) by (1) estimating optimal annuity levels, (2) performing this analysis for varying wealth levels, and (3) comparing these results to the empirical annuity levels (Section 4.5).

The optimal annuity demand for varying levels of liquid financial wealth and pre-annuitized wealth is displayed in Table 1. The optimal annuity levels are presented in two different formats: the numbers without brackets display the (1) optimal percentage in annuities as a fraction of total wealth (total wealth is pre-annuitized wealth plus liquid financial wealth, which is how we displayed the results in previous sections) and the numbers between brackets show the (2) optimal percentage in annuities as a fraction of liquid financial wealth. First of all, we see that the optimal annuitization level as a fraction of total wealth is in all cases below 100%. If an agent has $250,000 of wealth pre-annuitized in the form of social security and/or DB pension wealth, and $250,000 of

[11]The income from pre-annuitized wealth corresponds to a wealth level of about $375,000 and her liquid wealth is $300,000. Hence total wealth is $675,000. In effect the agent is choosing between annuitizing 55% of wealth or annuitizing almost 70% of wealth.

[12]The willingness to pay reflects the load on top of the actuarially fair price that the individual is willing to pay for this product. Hence a WTP of 0.94 means that the individual would even need a 6% bonus to hold the annuity.

Table 1: Optimal annuitization levels (%) for varying pre-annuitized and liquid financial wealth levels when agents face high health cost risk early in retirement

This table reports the optimal annuity levels (in %) in a real annuity. The pre-annuitized wealth level represents the net present value of social security income and/or DB pension income. The number without brackets is the optimal annuity demand as a fraction of total wealth and the number between brackets is the optimal annuity demand as a fraction of liquid wealth. For instance, if the pre-annuitized wealth is \$250k and liquid wealth is \$250k, then 50% of total wealth is pre-annuitized. If then the optimal annuity level is 58%, this means that the optimal annuity demand as a percentage of liquid wealth is 16%. The rest of the parameters are as in the benchmark case.

	Liquid financial wealth				
Pre-annuitized wealth	\$50k	\$150k	\$250k	\$350k	\$450k
\$200k	72	66	60	56	54
	(0)	(21)	(28)	(31)	(34)
\$250k	69	63	58	55	53
	(0)	(1)	(16)	(23)	(27)
\$300k	66	60	56	54	53
	(0)	(0)	(3)	(15)	(22)
\$350k	63	58	55	53	53
	(0)	(0)	(0)	(6)	(16)
\$400k	60	56	54	53	54
	(0)	(0)	(0)	(0)	(13)
\$450k	58	55	53	53	54
	(0)	(0)	(0)	(0)	(8)

liquid financial wealth, he optimally annuitizes 58% of his total wealth. Naturally, when displayed as a percentage of liquid financial wealth it is lower, 16%. Hence the agent should annuitize 16% of his \$250,000 of liquid financial wealth. Note that if the level of pre-annuitized wealth compared to liquid financial wealth is high (lower/left corner of table), the agent should optimally stay out of the voluntary annuity market. Furthermore, for agents with high liquid financial wealth levels, it is optimal to annuitize on a voluntary basis (annuitization as a fraction of liquid financial wealth) on top of the pre-annuitized wealth levels. This is in accordance with the data presented in Inkmann *et al.* (2011), who find that wealthier individuals tend to participate more in the voluntary annuity market. For the low health cost risk early in retirement we find that full annuitization remains optimal. The corresponding graphs and tables are available upon request.

4.5 Comparing the predicted annuity demand with empirical annuity levels

In previous sections we showed that high health cost risk early in retirement decreases optimal annuity demand. Adding to Ameriks *et al.* (2011) we determine the optimal annuity demand for varying wealth levels and we show that the predicted annuity demand matches the empirical levels closely. These results are presented in Figure 6. We see that the optimal annuity demand is a bit below or equal to the observed annuity levels, for all wealth levels. Hence medical expense

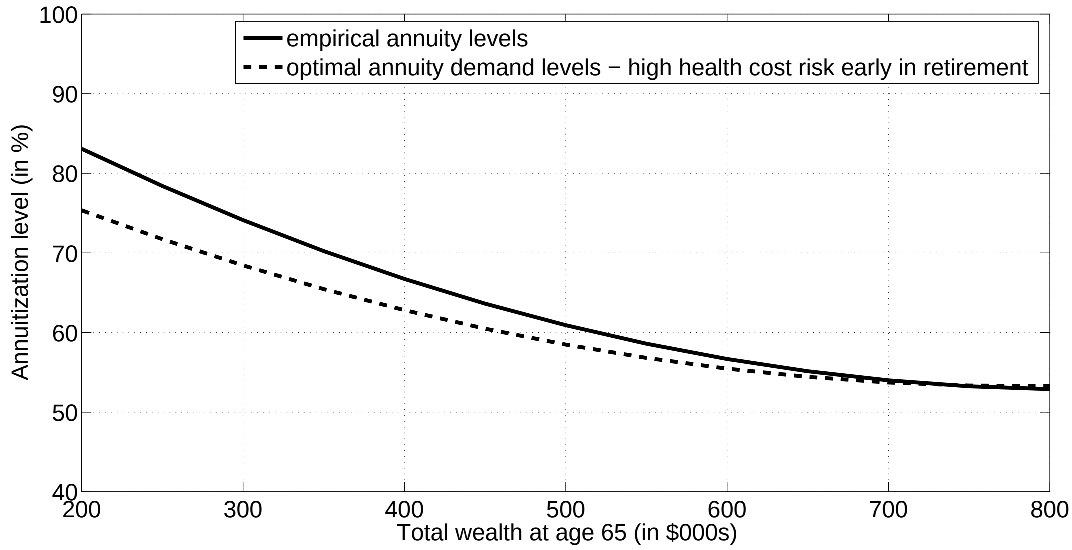


Figure 6: Comparing empirical annuitization levels with optimal annuitization levels.
We display the annuitization levels for a single female at age 65 estimated in Dushi & Webb (2004). They use data from the HRS to estimate the fraction of wealth annuitized. We present the fraction annuitized as percentage of the sum of liquid financial wealth and pre-annuitized wealth. Liquid financial wealth includes financial assets, IRA's, and DC pensions. Pre-annuitized wealth includes social security and DB pensions.

risk, that is their subjective expectation, might help to explain the annuity puzzle. Furthermore, the health costs cannot only help explain the annuity puzzle, but also the empirical relationship between annuity levels and wealth. If agents face health cost risk it is optimal for low wealth households to hold a large fraction of total wealth in annuities, compared to high wealth households, who should optimally annuitize less. Thus the average annuity-wealth profile empirically observed can potentially be explained by high health cost risk early in retirement. Note however that the empirically observed annuity level that we show is the average annuity level and individuals with less than average health cost risk will still want to annuitize (part) of their liquid financial wealth, while individuals in a bad health condition at age 65, might want annuity levels even below their pre-annuitized amount.

4.6 Minimum consumption level

In Figure 7, we show the impact of varying minimum consumption levels on our results. We see that, for a given wealth level, the optimal annuity demand is lower for lower minimum consumption levels. The reason for this finding is that if an agent incurs large health costs, the drop in utility is larger for lower minimum consumption levels. Under those circumstances, an individual is induced to hold a larger amount liquid to be able to smooth consumption and avoid receiving the minimum consumption level.

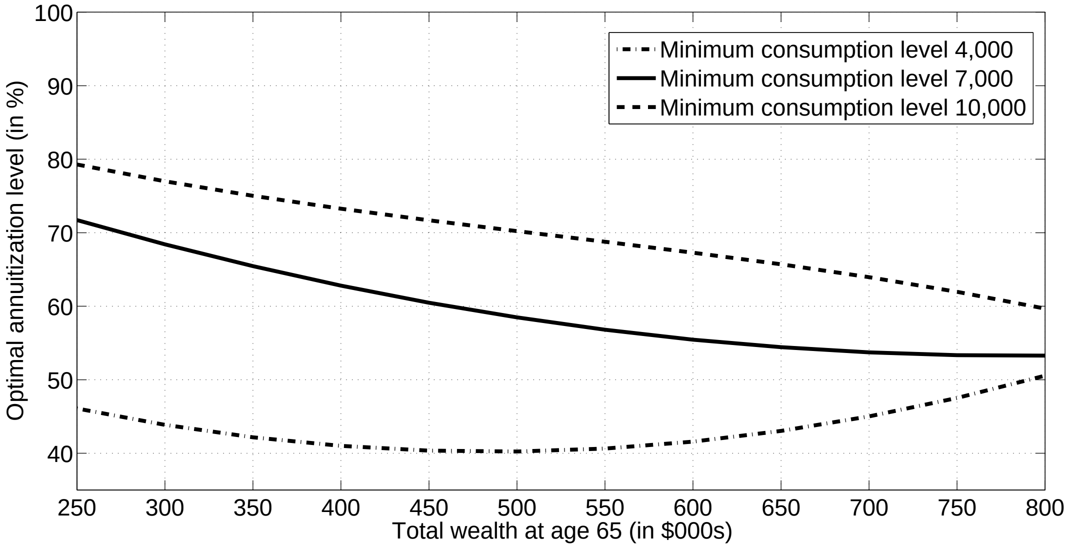


Figure 7: Optimal annuity demand for varying minimum consumption level assuming the high health cost risk early in retirement specification
The figure displays the optimal annuity demand for a 65-year old for varying wealth levels and minimum consumption levels. The numbers are in thousands of dollars.

5 Annuity Levels with Incomplete Annuity Markets

As shown by Davidoff *et al.* (2005) full annuitization is optimal if the annuity market is complete. Furthermore, they conclude that even in the case of market incompleteness, annuities are still valuable and agents should annuitize above their pre-annuitized levels (social security insurance and DB pension wealth). Davidoff *et al.* (2005) calibrate annuity market incompleteness by assuming agents maximize a habit formation utility function and thus usually favor a non-flat consumption profile, while agents can only purchase real annuities which generate a flat income pattern in real terms. In 1 out of 5 calibrations, they found that the optimal annuitization level falls sizeably, from 100% to 75%. As incomplete annuity markets is one of the leading explanations of low annuity demand and it can generate a substantial drop in annuity levels, we revisit this particular explanation. However, we model incomplete annuity markets as the non-availability of real annuities, while agents prefer a flat real consumption pattern. Average inflation of 2% halves the real value of a nominal annuity in 35 years, which can potentially have a large influence on the utility of such an annuity to an agent. Furthermore, a nominal annuity does not provide inflation risk protection nor does it give exposure to the equity market. This will decrease the value of the annuity sizeable and can potentially reduce the annuity levels to the empirically observed low levels.

Previously we showed in Figure 2 the optimal levels of annuitization if annuity markets are incomplete, conditional on optimal consumption and asset allocation strategies. In all cases (almost) full annuitization is optimal, hence the fact that the annuity market is incomplete does not have a material impact on the optimal annuitization level, given that we allow saving from annuity income. The benefits of insurance against longevity risk and the mortality credit outweigh the reduction in liquidity and less ability to get equity exposure at short horizons.

Furthermore, we explore the welfare gains from adding a variable annuity to the menu. We assume that agents can allocate the optimal amount to a variable and a real annuity, compared to only a real annuity. We find that the welfare gains are between 1.3% and 1.5%. Hence adding a variable annuity to the menu does not lead to a large increase of welfare if agents save out of their annuity income to invest in equity. The reason is that individuals can just use these savings to get equity exposure and real annuities provide a much better hedge against inflation risk than equity-linked annuities. In contrast, Koijen *et al.* (2010b) find an optimal allocation of 40% to variable annuities. However, they do not include equity in the post-retirement asset menu. Hence, the only way in which agents can get equity exposure is via a variable annuity.

6 Annuity levels and bequest motives

The desire to leave a bequest might induce agents to annuitize less, because in case of early death, the retiree may not have had sufficient time to build up enough wealth to bequeath. If an agent dies at more advanced ages, the individual saves out of the annuity income to leave a bequest. Several papers have explored the impact of bequest motives on optimal annuity levels with different findings. Two empirical papers on this topic are Brown (2001) and Inkmann *et al.* (2011). Brown (2001) finds that individuals with children leave no significantly different bequest compared to individuals without children.[13] In contrast, Inkmann *et al.* (2011) combine different potential motives for low annuity demand and find that they need a strong bequest motive to match the data. Furthermore, Lockwood (2011) shows that bequest motives alone cannot fully explain the low participation in annuity markets observed in the data. The reason why we revisit bequest motives as a potential explanation for low annuity demand is that many papers explore bequest motives, but not in a relatively parsimonious way by identifying the pure bequest effect without confounding with other explanations. Moreover, comparing these results to the findings for health cost risk and incomplete annuity markets allows us to quantify the relative importance of the three prevailing reasons.

[13] The result even goes in the opposite direction. Agents with no children die with larger amounts of wealth.

6.1 Optimal annuitization strategies at retirement

Following Ameriks *et al.* (2011) and De Nardi *et al.* (2010), we model the bequest motive as follows. An agent derives utility from leaving a bequest B_t:

$$v(B_t) = \frac{\bar{w}}{1-\gamma}\left(\phi + \frac{B_t}{\bar{w}}\right)^{1-\gamma} \tag{9}$$

where $\bar{w}$ is the strength of the bequest motive and ϕ is the prevalence in the population of an bequest motive. ϕ determines the curvature of the bequest motive and hence the extent to which bequests are a luxury good. The optimal bequest in a simplified version of the model provides a better understanding of the meaning of the bequest parameters. In a riskless world the optimal solution can be obtained analytically: Assume an agent starts with an amount of wealth W, does not face longevity risk, and the time preference discount rate is zero. Each year the individual consumes C for T years and derives utility equal to $C^{1-\gamma}/(1-\gamma)$. At death, the retiree leaves a bequest B equal to $(W - CT)$ and derives utility from bequest equal to $(\bar{w}/(1-\gamma))\,(\phi + B_t\,\bar{w})^{1-\gamma}$. The agent chooses C optimally. The resulting optimal consumption is $C = (W + \bar{w}\phi)/(\bar{w} + T)$ and the optimal bequest is $B = \bar{w}(C - \phi)$. Hence the agent leaves a bequest to cover $\bar{w}$ years of spending for the heir at an annual expenditure level $(C - \phi)$, the amount by which his own optimal annual consumption exceeds the threshold ϕ. If W is too low to ensure an income stream for the heir higher than ϕ for $\bar{w}$ years, no bequest is left.

Table 2: Optimal annuitization levels (%) for varying levels of bequest motive
This table reports the optimal annuity levels (in %) in a *real* annuity. The rest of the parameters are as in the benchmark case. The effect of both parameters on the optimal annuity demand is in opposite directions. A higher strength of the bequest motive $\bar{w}$ gives an incentive to annuitize less, while a higher luxury good parameter ϕ increases the incentives to annuitize more.

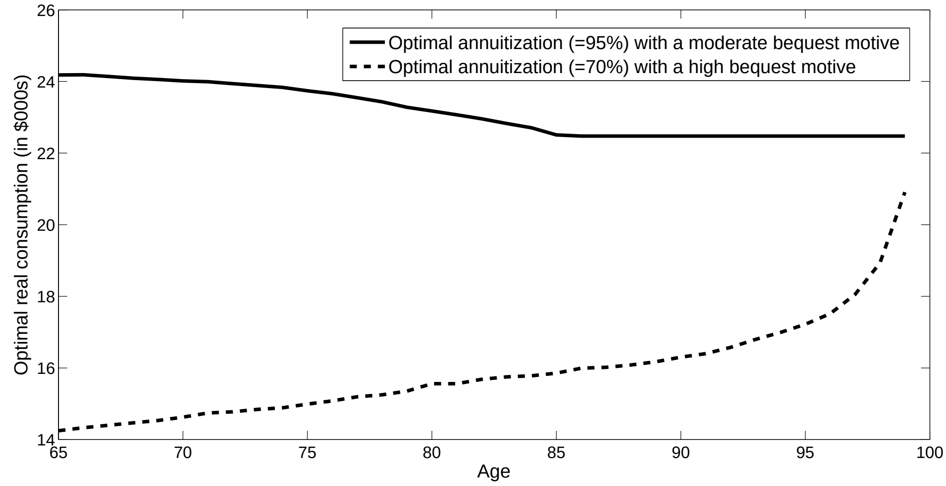
	$\bar{w}$ strength of bequest motive						
ϕ	2	7	12	17	22	27	32
8,000	90	85	80	75	70	70	70
16,000	90	90	85	85	85	80	80
24,000	95	95	95	90	90	90	90
32,000	95	95	95	95	95	95	95
40,000	95	95	95	95	95	95	95

Table 2 shows the optimal annuity levels for different values of the two bequest parameters ($\bar{w}$ and ϕ). Note that a higher strength of the bequest motive $\bar{w}$ gives an incentive to annuitize less, while a higher luxury good parameter ϕ gives an incentive to annuitize more. We see that for many levels of a bequest motive, high annuitization levels remain optimal. Only when the luxury parameter ϕ is equal or lower than \$16,000, annuity levels below 90% can be optimal (depending also on the level of $\bar{w}$). De Nardi *et al.* (2010) find a strength of the bequest motive ($\bar{w}$) equal to 2.4 and a luxury parameter ϕ of 36,000 which implies that full annuitization is optimal. However, Ameriks *et al.* (2011) estimate a strength parameter $\bar{w}$ of 32 and a luxury parameter ϕ of 7,500, resulting in an optimal annuity demand of 70%. Interpreting these bequest parameters estimated in Ameriks *et al.* (2011) using a simplified riskless world, like mentioned above, would imply that agents want to leave a bequest of about $162,000 and consume in total during their retirement $188,000 (to put this into perspective: in such a situation they bequeath almost 13 times their annual income.).

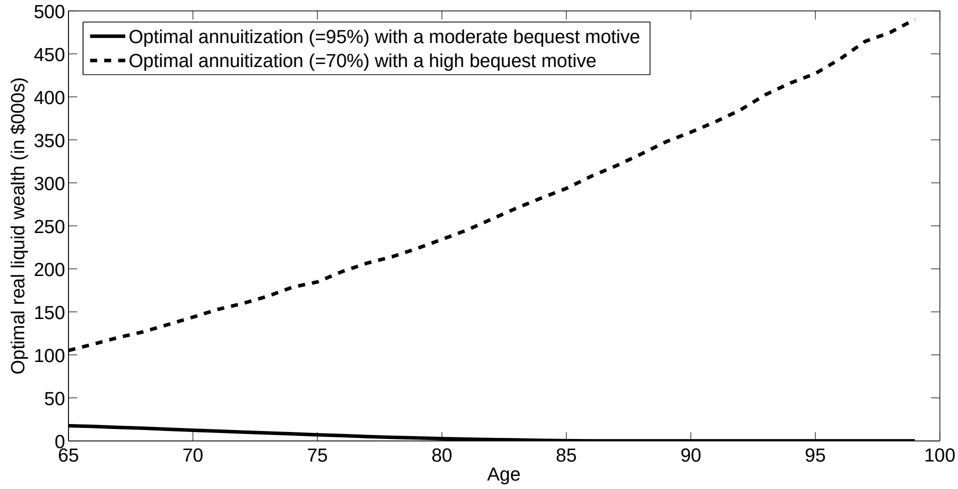
The optimal consumption and wealth levels for the bequest parameters estimated in De Nardi *et al.* (2010) (=modest bequest motive) and for the parameters in Ameriks *et al.* (2011) (=high bequest motives) are displayed in Figure 8. When an agent faces moderate bequest motives (solid line) he consumes approximately his entire annuity income and slowly consumes out of the small amount of wealth that he has liquid. If the bequest motives are much stronger (dashed line) an individual optimally annuitizes 70% of his total wealth which generates an annual real annuity income of about $16,500. The agent optimally saves out of his annuity income to build up a bequest. The median real wealth level at age 80 is about $250,000 and this savings level increases quickly with age. This predicted savings pattern is extremely high and inconsistent with the data.

6.2 Optimal annuitization for various wealth levels

In the previous section we showed the impact of bequest motives on optimal annuity levels for an agent with the benchmark total wealth level; $350,000. Table 3 displays the annuity levels for a total wealth level at retirement of $200,000 and $500,000. To the best of our knowledge, this paper is the first to explicitly explore how wealth levels influence the impact of bequest motives on optimal annuity demand. We see that the predicted annuity levels are lower for lower levels of total wealth. If the luxury parameter ϕ is equal to 24,000, the agents bequest motive only kicks in if his own consumption level is above this threshold ϕ (if $\bar{w} \longrightarrow \infty$ all additional consumption above ϕ is saved to leave as a bequest). Hence if the total wealth level is $200,000, the agent will never consume above this threshold and full annuitization is optimal. Focussing on a luxury parameter ϕ equal to $8,000, we see that still the optimal annuity level is decreasing in the wealth level, for every strength of the bequest motive, $\bar{w}$. The intuition is clear when focussing on the simplified case, a riskless world. If an agent has a wealth level of $200,000 his optimal bequest amount is $20,000, while if an agent has a wealth level of $400,000, his optimal bequest amount is $100,000. The main risk that an agent with a high bequest motive faces is that he dies before having enough time to save out of his annuity income to have a sufficient amount to bequeath. However, in this simple example we see that if the wealth of an agent doubles (from $200,000 to $400,000), the



(a) Real consumption



(b) Real liquid wealth

Figure 8: Optimal consumption and wealth paths over the life cycle
Panel (a) displays the median optimal real consumption when agents have moderate bequest motives (solid line) and when agents have high bequest motives (dashed line). Panel (b) displays the optimal *liquid* real wealth when agents have moderate bequest motives (solid line) and when agents have high bequest motives (dashed line). The optimal levels are for a female.

optimal bequest is increased from $20,000 to $100,000. Hence he has to save relatively much more out of his annuity income, and thus runs sizeably more risk of dying too soon and not having saved enough out of the annuity income.

The main question we want to answer in this section is; can bequest motives lower the optimal annuity levels to the very low levels empirically observed. What we find is that when bequest motives are very high, this can explain the low empirical annuitization levels, for all wealth levels.

However, these very high bequest motives predict savings behavior over the life cycle which is inconsistent with the data. Compared to the performance of health cost risk in explaining the data, bequest motives appear to do less well because (1) very high, less plausible bequest motives are required to match the data on annuitization and (2) these very high bequest motives generate savings behavior which is not in accordance with the data.

Table 3: Optimal annuitization levels (%) for varying levels of bequest motive: wealth 200,000 and wealth 500,000

This table reports the optimal annuity levels (in %) in a *real* annuity. The rest of the parameters are as in the benchmark case. The effect of both parameters on the optimal annuity demand is in opposite directions. A higher strength of the bequest motive $\bar{w}$ gives an incentive to annuitize less, while a higher luxury good parameter ϕ increases the incentives to annuitize more.

	$\bar{w}$ strength of bequest motive													
	wealth 200,000							wealth 500,000						
ϕ	2	7	12	17	22	27	32	2	7	12	17	22	27	32
8,000	95	90	85	85	80	80	80	90	80	75	70	65	65	60
16,000	95	95	95	95	95	95	95	95	85	85	80	80	75	75
24,000	95	95	95	95	95	95	95	95	90	85	85	85	85	80
32,000	95	95	95	95	95	95	95	95	95	90	90	90	85	85
40,000	95	95	95	95	95	95	95	95	95	95	95	95	95	95

7 Conclusion

In this paper we analyze three explanations for the annuity puzzle and establish their relative importance. We find that *health cost risk* can potentially explain the very low annuity levels observed empirically. In contrast, realistic *bequest motives* and *incomplete annuity markets* cannot. Assuming an agent has a very high bequest motive may help explain the very low empirically observed annuitization levels, but it generates extremely high savings levels which is in contrast to the data.

Medical expenses increase the need for liquidity, inducing households to annuitize less and keep wealth liquid. In the literature, several health cost specifications are estimated all implying a different process for out-of-pocket medical expenses. We employ two health cost models to disentangle an important driver of annuity demand: *health cost risk early in retirement*. If health costs can already be high in early retirement it is not optimal to annuitize all wealth, since the retiree cannot save enough during the first years of retirement to cover these potential expenses. The only manner with which to create a buffer against these expenses in early years, is to reduce the annuitization level. In contrast, if the medical expense risk is only moderately high, it is optimal to fully annuitize and subsequently save sizeable amounts out of the annuity income to build up a buffer. We find empirical evidence that many individuals face high health cost risk early in retirement. Extending Ameriks *et al.* (2011), we show that optimal annuity demand varies with initial wealth levels. When faced with high health cost risk early in retirement, agents with a low wealth level find it optimal to annuitize a large fraction of total wealth while retirees with higher wealth levels optimally annuitize less. We compare these optimal annuity demands with the empirically observed annuitization demands for varying wealth levels and find a similar pattern. Both the empirically observed and the optimal annuity levels are decreasing in total wealth (when health costs can be high early in retirement). In addition, the optimal wealth profile over the life cycle for an agent with median wealth is similar to empirically observed paths: agents have sizeable precautionary savings due to out-of-pocket medical expenses and they start dissaving late in retirement.

A possible direction for further research could be to explore the impact of heterogeneity of (perceived) health cost risk on annuitization decisions. For many individuals health cost risk early in retirement can be high. However, for several groups health cost risk early in retirement will deviate from the ”average” risk, for instance wealthy individuals tend to live longer and are in better health than less wealthy agents. Furthermore, we do not know whether the actual health cost risk coincides with the perceived risk of individuals. Hence for some groups high/full annuitization might still be optimal. In addition, cross-country differences in out-of-pocket expenditure risk and annuitization decisions can be used to further explore the impact of health cost risk on annuity levels.

References

Agarwal, S., Driscoll, J.C., Gabaix, X., & Laibson, D. 2009. The Age of Reason: Financial Decisions over the Life-Cycle with Implications for Regulation. *Brookings Papers on Economic Activity*.

Agnew, J.L., Anderson, L.R., Gerlach, J.R., & Szykman, L.R. 2008. Who Chooses Annuities? An Experimental Investigation of the Role of Gender, Framing, and Defaults. *American Economic Review*, **98**(2), 418–422.

Ameriks, J., Caplin, A., Laufer, S., & Van Nieuwerburgh, S. 2011. The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives. *Journal of Finance*, **66**(2), 519–561.

Brook, Robert., Ware, John.E., Rogers, William.H., Keeler, Emmett.B., Davies, Allyson.R., Donald, Cathy.A. Goldberg, George.A., Lohr, Kathleen.N., Masthay, Patricia.C., & Newhouse, Joseph.P. 1983. Does Free Care Improve Adults' Health? Results From a Randomized Controlled Trial. *New England Journal of Medicine*, **309**(23), 1426–1434.

Brown, J.F., Casey, M.D., & Mitchell, O.S. 2008a. *Who Values the Social Security Annuity? New Evidence on the Annuity Puzzle*. NBER Working Paper No. 13800.

Brown, J.R. 2001. Private Pensions, Mortality Risk, and the Decision to Annuitize. *Journal of Public Economics*, **82**(1), 29–62.

Brown, J.R. 2007. *Rational and Behavioral Perspectives on the Role of Annuities in Retirement Planning*. NBER Working paper No. 13537.

Brown, J.R., & Finkelstein, A. 2007. Why Is the Market for Long-Term Care Insurance So Small? *Journal of Public Economics*, **91**(10), 1967–1991.

Brown, J.R., & Poterba, J.M. 2000. Joint Life Annuities and Annuity Demand by Married Couples. *Journal of Risk and Insurance,* **67**(4), 527–553.

Brown, J.R., Kling, J.R., Mullainathan, S., & Wrobel, M.V. 2008b. Why Don't People Insure Late-Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle. *American Economic Review*, **98**(2), 304:309.

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

De Nardi, M., French, E., & Jones, J.B. 2010. Why Do The Elderly Save? The Role of Medical Expenses. *Journal of Political Economy*, **118**(1), 39–75.

Dushi, I., & Webb, A. 2004. Household Annuitization Decisions: Simulations and Empirical Analyses. *Journal of Pension Economics and Finance*, **3**(2), 109–143.

Dynan, K.E., Skinner, J., & Zeldes, S.P. 2004. Do the Rich Save More? *Journal of Political Economy*, **112**(2), 397–444.

French, E., & Jones, J.B. 2004. On the Distribution and Dynamics of Health Care Costs. *Journal of Applied Econometrics*, **19**(6), 705–721.

Gazzale, R.S., & Walker, L. 2009. *Behavioral Biases in Annuity Choice: An Experiment*. Williams College Economics Department Working Paper Series. Available at SSRN: http://ssrn.com/abstract=1370535.

Horneff, W.J., Maurer, R.H., Mitchell, O.S., & Dus, I. 2008a. Following the Rules: Integrating Asset Allocation and Annuitization in Retirement Portfolios. *Insurance: Mathematics and Economics*, **42**(1), 396–408.

Horneff, W.J., Maurer, R.H., & Stamos, M.Z. 2008b. Life-cycle Asset Allocation with Annuity Markets. *Journal of Economic Dynamics & Control*, **32**(11), 3590–3612.

Hu, W.Y., & Scott, J.S. 2007. Behavioral Obstacles in the Annuity Market. *Financial Analysts Journal*, **63**(6), 71–82.

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

Koijen, R.S.J., Nijman, T.E., & Werker, B.J.M. 2010b. *Optimal Annuity Risk Management*. forthcoming Review of Finance.

Kotlikoff, L.J., & Spivak, A. 1981. The Family as an Incomplete Annuities Market. *Journal of Political Economy*, **89**(21), 372–391.

Lockwood, Lee. 2011. Bequest Motives and the Annuity Puzzle. *forthcoming Review of Economic Dynamics*.

Love, David.A., Palumbo, Michael.G., & Smith, Paul.A. 2009. The Trajectory of Wealth in Retirement. *Journal of Public Economics*, **191**(1-2), 191–208.

Mitchell, O.S., Poterba, J.M., Warshawsky, M.J., & Brown, J.R. 1999. New Evidence on the Money's Worth of Individual Annuities. *American Economic Review*, **89**(5), 1299–1318.

Murtaugh, C., Kemper, P., & Spillman, B. 1995. Risky Business: Long Term Care Insurance Underwriting. *Inquiry*, **32**(3), 271–284.

Palumbo, M.G. 1999. Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle. *Review of Economic Studies*, **66**(2), 395–421.

Pang, G., & Warshawsky, M. 2010. Optimizing the Equity-Bond-Annuity Portfolio in Retirement: The Impact of Uncertain Health Expenses. *Insurance: Mathematics and Economics*, **46**(1), 198–209.

Pashchenko, Svetlana. 2010. *Accounting for Non-Annuitization*. Working paper, University of Virginia.

Scholz, J.K., Seshadri, A., & Khitatrakun, S. 2006. Are Americans Saving "Optimally" for Retirement. *Journal of Political Economy*, **114**(4), 607–643.

Sinclair, S.H., & Smetters, K.A. 2004. *Health Shocks and the Demand for Annuities*. Technical Paper Series 2004-9, Congressional Budget Office.

Turra, C.M., & Mitchell, O.S. 2008. *The Impact of Health Status and Out-of-Pocket Medical Expenditures on Annuity Valuation* . Oxford University Press. Chap. in John Ameriks and Olivia S. Mitchell, editors, Recalibrating Retirement Spending and Saving, pages 227–250.

Werblow, A., Felder, s., & Zweifel, P. 2007. Population Ageing and Health Care Expenditures: A School of ”Red Herrings”? *Health Economics*, **16**(10), 1109–1126.

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

Yogo, M. 2009. *Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing and Risky Assets*. NBER Working Paper No. 15307.