

Mortality Risk Learning and the Demand for Annuities*

Gunnar Poppe-Yanez[‡]

May 14, 2020

Abstract

The discrepancy between the high demand for annuities predicted by economic theory and the empirical low holdings of these assets, known as the annuity puzzle, is still not completely understood in economic studies of retirement finance. This paper assesses the effect of individuals' mortality risk learning process on annuitization. I isolate this effect by building a life-cycle model in which individuals have imperfect information of their true survival probability distribution, and therefore have to update their beliefs about it in a Bayesian manner. Using data on subjective mortality by the Health and Retirement Study to evaluate the model, the baseline result shows that the demand for annuities can be about 40% lower than full annuitization solely attributable to individuals learning about their true mortality risk—a situation that does not allow for the known strong take-up for annuities to take effect. I further expand the model to have a bequest motive to show how more features that drive down annuitization can be added and interacted with this learning mechanism.

1 Introduction

Understanding the motivations behind retirement financial decisions is key for the proper design of pension and retirement policies. Yaari (1965) showed that in a world with complete annuity markets, intertemporally separable utilities, and uncertain lifespans, it is optimal for risk averse individuals to hold their entire wealth in fair priced annuities. This conclusion is reached because an annuity

*Job Market Paper.

[†]Previously circulated as "Mortality Learning and Optimal Annuitization". I am deeply indebted to Sangeeta Pratap, Frank Heiland, and Wim Vijverberg for their guidance and support throughout this project. I also thank Moshe Arye Milevsky, Lilia Maliar, Julen Esteban-Pretel, and Nuria Rodriguez-Planas for their useful comments.

[‡]Ph.D. Program in Economics, the Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016. Email: gpoppeyanecz@gradcenter.cuny.edu

potentially provides high payments to individuals who reach advanced ages—a feature known as mortality credit. Nevertheless, this theoretical result is not borne out by the data because annuities holdings of retired individuals, in the United States and other developed countries, are very low.¹ This discrepancy is referred to as the annuity puzzle, and in order to have a better understanding of it, subsequent studies examined variants of Yaari’s model by relaxing assumptions and incorporating new features. Nevertheless, most of these new approaches further proved the robustness of Yaari’s result of complete annuitization, or in its defect, sufficiently high levels of it.²

In this paper I study the demand for annuities under a new paradigm of heterogeneity in mortality beliefs. Most of the previous studies that demonstrated how annuities are a dominant savings asset assumed a representative agent’s logical evolution of survival probabilities. In contrast, acknowledging the heterogeneity in mortality beliefs offers a promising alternative to study the demand for annuities, since an individual’s mortality belief evolution is idiosyncratic and not constrained to have a logical process, nor to conform with an econometrician’s estimation of survival probabilities. In this paper, I model mortality beliefs as the result of a process of mortality risk learning, and therefore, due to the changing nature of mortality beliefs, it is possible for a risk averse individual to choose to not annuitize if she updates her survival probabilities. The implications of this approach are relevant given that annuities are thought to function as longevity risk insurance; that is, they are meant to hedge the risk of running out of savings at the end of life. However, in this paper, I show how less knowledge about one’s idiosyncratic mortality risk itself may actually decrease the demand for annuities. Specifically, if individuals perceive a changing distribution determining their survival chances, then the status of annuities as a dominant savings asset disappears: a particular investment in annuities that once seemed optimal, with a particular mortality belief, will not necessarily seem optimal in the future if this belief is revised such that the subjective survival probabilities are updated downwards.

Notwithstanding, for a lifetime beliefs-focused approach to be tractable in an economic model, it is also necessary to assume how these heterogeneous beliefs evolve through time. While cross-sectional survival probabilities for a particular age can be extrapolated in a straightforward manner, the evolution of mortality risk beliefs during the life-cycle must follow a rule in their updating process. To this end, I assume individuals are learning about their mortality risk in an optimal manner, i.e., in

¹Friedman and Warshawsky (1990) report that only 2% of the elderly population in the Retirement History Survey hold annuities. More recently, using data from from the Health and Retirement Study, Hosseini (2015) finds that only about 3% of total retirement wealth in the United States has been privately annuitized.

²Finding a compelling theoretical reason for the lack of annuitization has proven to be a difficult task: see Davidoff, Brown and Diamond (2005). More recently Peijnenburg et al. (2016) show that even in a more general environment full annuitization is the best strategy.

a Bayesian updating form. I assume the evolution of subjective survival probabilities is the product of the process of learning about a survival distribution parameter, which could be interpreted as a frailty parameter. This learning process implies individuals improve the knowledge of their idiosyncratic survival probability distributions as they age, and consequently, tantamount to modeling the formation of subjective survival probabilities as a learning process of objective survival probabilities.

With this framework then it is possible to map subjective mortality beliefs, as a process of learning about idiosyncratic mortality, to a recursive decision of annuitization of wealth. The contribution of this paper is twofold: first I provide a new framework of mortality risk learning capable of generating a distribution of subjective survival probabilities consistent with subjective beliefs data, and secondly, I empirically evaluate a life-cycle model embedded with this learning mechanism to determine the extent to which, longevity uncertainty, embedded in the updating process of subjective mortality beliefs, can account for the observed low levels of annuitization. To obtain the parameters of the learning process I use a Simulated Method of Moments to minimize the distance between the learning mechanism’s predicted average subjective survival probability per age, and its empirical counterpart found in the Health and Retirement Study data. Moreover, I use an asset rebalancing model similar to Reichling and Smetters (2015) to study the effect of the learning process on optimal annuitization.³ As in these authors’ study, I assume that knowledge about individuals’ mortality risk is not private information, and even though this assumption could be supported by Finkelstein and Poterba (2004)—where no evidence of asymmetric information in annuity markets is found—the reason to adopt this demand-focused modeling approach is because the presence of a supply side would only drive annuitization further down.

The timing consideration when pricing annuities is a key aspect of the asset rebalancing model in this paper, as is also the case for the model in Reichling and Smetters (2015). In these authors’ model, individuals would purchase a fairly priced annuity using the expectation of a health shock, that is, before the realization of a health shock bound to determine their actual survival probabilities, generating in this manner the possibility of a lower annuity return. In the present model, individuals would purchase an annuity priced before their subjective survival probability is updated, and therefore, if

³Reichling and Smetters (2015) note that even though an explicit rebalancing of assets is hard to observe in the data, evidence of individuals *selling* annuities in a theoretical sense exists. The purchase of life insurance, combined with the fact that the secondary market for life insurance is growing at a rapid pace, can be considered as equivalent to the selling of annuities by individuals. Furthermore, an actual (yet relatively understudied) secondary annuity market exists (see for example Panis and Brien, 2016), although the correspondent quantity of buyers and sellers is still undocumented to my knowledge.

the new survival belief is lower than the previous one, an annuity investment is no longer optimal. This new paradigm then allows me to study different aspects in the learning process that have a direct impact on annuitization, such as the initial mortality beliefs individuals have when entering the economy, or the impeding noise they face in the learning process. In particular, what determines whether or not an investment in annuities is desired is the downward-smoothness of mortality beliefs (i.e., frailty beliefs) throughout time, given that perfect smoothness would imply no change in subjective survival probabilities beliefs.

Simulation results show that optimal annuitization effectively increases during the life-cycle since individuals are more certain of their mortality as they age, that is, the preference for annuities as a dominant savings asset takes effect once mortality beliefs are no longer being updated. Nevertheless, the baseline result shows that longevity uncertainty at retirement age 65 implies an optimal annuitization about 40% lower than full annuitization, which is a considerable reduction in annuitization compared to previous studies, and critically, such low level of annuitization is solely the product of the mortality learning process. That is, using a demand-focused model. Additionally, I show the robustness of this annuitization profile in the life-cycle by comparing different specifications of the learning process. In this regard, it is important to note that annuitization in this model is the result of the downward-smoothness of beliefs in time (during the learning process), which a priori, would imply that a greater (lower) difficulty to learn objective mortality would imply a higher (lower) profile of annuitization in the life-cycle. Nevertheless, I show that if the level of noise during the learning process is sufficiently high, it would be rational for an individual to ignore any new information in the learning process, which in turn would allow smoothness of frailty beliefs to take place as well.

Contrasting with previous works, the stochastic nature of survival expectations in this model is not explicitly determined by health shocks—a novel feature that avoids the potential problem of having medical expenditures shocks at advanced ages, which in turn would not let the model identify the effect of health-shock-driven subjective survival expectations on optimal annuitization. Furthermore, as studied in Brugiavini (1993), by avoiding the use of health shocks I also avoid the problem of early annuitization: when a health shock reduces the present value of an annuity, individuals would still annuitize assets and pool this risk by annuitizing wealth early in life. In this model, even though subjective mortality is stochastic, early annuitization is not possible because instead of expecting future health shocks, individuals are refining beliefs. Finally the important difference between the present work and past studies using subjective survival probabilities is that, instead of using the reported probabilities directly in a life-cycle model, I use these data to model a structural learning

process of mortality, giving a direct theoretical interpretation to the subjectivity of beliefs. Subjective mortality has captured the attention of different fields thanks to the availability of data from the Health and Retirement Study: Hurd and McGarry (1995, 2002) use subjective mortality data as predictor of actual mortality; Heimer et. al. (2015) show how subjective mortality data can explain retirement savings puzzles; Gan et. al. (2015) use a life-cycle model to show how subjective mortality data is more adequate in the use of this type of models; Sun and Webb (2011) study early claiming of Social Security using subjective mortality data and the implications for medical underwriting.⁴

In what follows of the paper section 2 describes a model of mortality learning, section 3 describes a life-cycle model embedded with these learning dynamics for optimal annuitization, section 4 describes the empirical strategy, section 5 explains the results of simulations, and section 6 concludes.

2 A Model of Mortality Learning

In this section I describe a mechanism in which individuals slowly learn about their true survival probabilities, which in turn are determined by their own—to be learned—true frailty. Concretely, I model mortality learning as an individual's learning process about the exact deviation of her frailty with respect to an average (or standard) frailty, within a population frailty distribution. Given that this criterion defines survival probabilities, I deem subjective knowledge about idiosyncratic frailty tantamount to subjective survival probabilities. In the model, at the beginning of life, all individuals share a common belief of their frailty, but as individuals age they learn about their unique position in the frailty distribution, and therefore improve the knowledge of their actual idiosyncratic survival expectations. This process of learning then generates a distribution of subjective survival probabilities per period that is not necessarily objective, but which in turn can be calibrated with the subjective survival probabilities reported in empirical data.

The learning dynamics section of the model follows closely the one developed by Guvenen (2007) concerning learning dynamics about individuals' characteristics.

⁴Subjective mortality has captured the attention of different fields thanks to the availability of data from the Health and Retirement Study: Hurd and McGarry (1995, 2002) use subjective mortality data as predictor of actual mortality; Heimer et. al. (2015) show how subjective mortality data can explain retirement savings puzzles; Gan et. al. (2015) use a life-cycle model to show how subjective mortality data is more adequate in the use of this type of models; Sun and Webb (2011) study early claiming of Social Security using subjective mortality data and the implications for medical underwriting.

2.1 Formation of Subjective Survival Probabilities

An individual i has a risk of mortality determined by her own idiosyncratic frailty δ^i . Following Vaupel, Manton, and Stallard (1979) and Manton, Stallard, and Vaupel (1981), frailty δ^i determines a (objective) survival probability up to period t , $P_t(\delta^i)$. In other words, frailty δ^i determines the (unconditional) probability of survival $P_t(\delta^i)$ of individual i at each age—the higher the frailty, the lower this probability is at each age.

If any individual knew the true value of her frailty δ^i she would use $P_t(\delta^i)$ to calculate the risk of mortality throughout her life-cycle. In this model instead I assume frailty is unobserved, and therefore, each period k individuals form a belief of their actual frailty $\widehat{\delta}_k^i$. I refer to $\widehat{\delta}_k^i$ as idiosyncratic subjective frailty since it will allow me to define later an idiosyncratic *subjective* survival probability $P_t(\widehat{\delta}_k^i)$. Additionally, I assume individuals are aware of the heterogeneity across the population with the existence of a set of individual frailty types $\Delta = [\underline{\delta}, \bar{\delta}]$, which allows them to know there is a well defined cumulative distribution $F \in \Gamma(\Delta)$.

Suppose $\rho_t(\widehat{\delta}_k^i)$ is the individual's subjective probability of survival at the beginning of period t , conditional on surviving until the end of period $t - 1$, based on period k subjective frailty $\widehat{\delta}_k^i$. We have then $P_t(\widehat{\delta}_k^i) = \prod_{\tau=0}^t \rho_\tau(\widehat{\delta}_k^i)$. The certain end of life at age T implies that $\rho_{T+1} = 0 \forall i$.

In order to pin down the unconditional survival probability as a function of frailty $P_t(\widehat{\delta}_k^i)$, assume individuals are aware of the effect frailty has on the force of mortality, that is, the instantaneous rate of mortality as a function of frailty. I assume this mechanism is identical across individuals, as proposed by Vaupel, Manton, and Stallard (1979), and Manton, Stallard, and Vaupel (1981).⁵ Specifically, let $h_t(\delta^i)$ be the force of mortality of individual i at age t for any δ^i , such that $h'_t(\delta^i) > 0$. Assume that for any two individuals i and j we have

$$\frac{h_t(\delta^i)}{h_t(\delta^j)} = \frac{\delta^i}{\delta^j}.$$

Furthermore, assuming j is the *standard* individual such that $\delta^j = \delta^{std} = 1$, we have

$$h_t(\delta^i) = \delta^i h_t^{std},$$

where h_t^{std} is the force of mortality of the standard individual whose frailty has been normalized to 1. This in turn determines the cumulative mortality hazard $H_t(\delta^i)$ for an individual with frailty δ^i as

$$H_t(\delta^i) = \int_0^t h_s(\delta^i) ds = \delta^i \int_0^t h_s ds \equiv \delta^i H_t^{std}$$

⁵To consider heterogeneity in frailty, this framework is also used by Hosseini (2015).

where H_t^{std} is the cumulative mortality hazard of the the standard individual. $H_t(\delta^i)$ determines the unconditional survival probability $P_t(\delta^i)$ for individual i , which consequently will be a function of H_t^{std} , since

$$P_t(\delta^i) = \exp(-H_t(\delta^i)) = \exp(-\delta^i H_t^{std})$$

This last expression allows me to define now the concept of subjective unconditional survival probabilities.

Definition 1.- A subjective unconditional survival probability based on period k subjective frailty $\widehat{\delta}_k^i$ is defined as

$$P_t(\widehat{\delta}_k^i) = \exp(-H_t(\widehat{\delta}_k^i)) = \exp(-\widehat{\delta}_k^i H_t^{std})$$

To make this model computable we also need to assume individuals are aware of the standard individual with frailty $\delta^{std} = 1$. This lets us interpret $P_t(\widehat{\delta}_k^i)$ as an individual's deviation-of-the-standard belief, that is, the belief about how far her frailty deviates from the standard frailty. This is the key aspect of the definition above.

It is important to note that, in general, there is flexibility regarding who to consider the standard individual. This allows the model to be tractable in the next section. Concretely, the standard individual will be identified as the one who has an objective survival probability identical to the survival data found in life tables.

2.2 How do individuals learn their own frailty?

I assume there is a noisy signal of frailty in order to model the fact that individuals do not observe it directly. In this model, individuals learn about their own frailty in a Bayesian manner through random realizations of the noisy signal. Specifically, I define s_t to be the sum of the signal and the noise. As signal I use a function of frailty $d_t(\delta^i)$, and the noise is defined as $e_t^i \sim N(0, \sigma_e^2)$:

$$s_t^i = d_t(\delta^i) + e_t^i \tag{1}$$

This framework—plus the specification of $d_t^i(\delta^i)$ to be discussed below—allows the individual to fully learn about δ^i at latter periods in life (high values of t). Before that, even though the individual observes the realization of the noisy signal s_t , she still has imperfect information about her frailty δ^i .

2.2.1 Bayesian Learning

Once the object to be learned is formulated the dynamic Bayesian process of updating beliefs can be specified. For this purpose, this learning process can be expressed as a Kalman filtering problem using a state-space representation. Given that the state variable being learned—the unobserved frailty δ^i —is a scalar, there is no need to specify a state equation. The observation equation on the other hand, or the specific way in which what is unobservable affects what is observable, corresponds to the specification of the noisy signal, i.e., equation (1).

Before going into detail of the learning process, let us define $d_t(\delta^i)$ as

$$d_t(\delta^i) = \gamma t^\gamma \delta^i,$$

which uses the parameter γ as a regulator of the speed of learning, given that this formulation is a function of a trend. As mentioned above, this formulation allows for full learning at latter stages of life. The lower the value of γ the longer complete learning will be delayed. A plausible interpretation of this formulation is that aging reveals true frailty at the latter years of life.

With this framework now set we can formulate the Kalman update equations for the optimal (dynamic) learning of δ^i , and its variance $\sigma_{\delta^i}^2$,

$$\widehat{\delta}_t^i = \widehat{\delta}_{t-1}^i + G_t [s_t^i - \gamma t^\gamma \widehat{\delta}_{t-1}^i] \quad (2)$$

$$\sigma_{\delta^i, t}^2 = \sigma_{\delta^i, t-1}^2 - G_t \gamma t^\gamma \sigma_{\delta^i, t-1}^2 \quad (3)$$

where G_t is the Kalman gain at time t given by

$$G_t = \frac{\gamma t^\gamma \sigma_{\delta^i, t-1}^2}{(\gamma t^\gamma)^2 \sigma_{\delta^i, t-1}^2 + \sigma_e^2} \quad (4)$$

As in Guvenen (2007), to initiate the filtering process using equations (2) and (3) we must specify the initial values $\widehat{\delta}_0^i$ and $\sigma_{\delta^i, 0}^2$, as these represent the information with which individuals enter the economy.

Smoothness, i.e., $\widehat{\delta}_{t+1}^i \approx \widehat{\delta}_t^i$, is a result of the optimal learning mechanism itself, and as such there are two possible reasons for a smooth path. First, as it is easy to conjecture, once the true value of frailty has been fully learned, that is, once $[s_t^i - \gamma t^\gamma \widehat{\delta}_{t-1}^i]$ is close to zero, then the learning path will tend to be smooth. Secondly, this smoothness will also happen when the Kalman gain G_t gets closer to zero—a situation that will happen when there exists a high level of noise σ_e^2 . Therefore, if

Table 1. Simulation: Standard Deviation of Learning Path Realizations, Average

Age	$\sigma_e^2 = 1$	$\sigma_e^2 = 10$	$\sigma_e^2 = 100$	$\sigma_e^2 = 1000$
20-95	0.297	0.238	0.109	0.036
60-95	0.119	0.199	0.114	0.039

the individual acknowledges she faces a high level of noise she will then choose to ignore the signal in equation (1), and as a result, she will have a smoothed frailty learning path regardless if her current frailty belief is close to the true value of her frailty or not. Therefore, noise at the time of learning does not have a clear a priori effect on how smooth the learning path will be. Notice also how the learning process does not allow for a smoothed path to become non-smoothed again; for this to occur there would have to exist a late change in G_t or $[s_t^i - \gamma t^\gamma \widehat{\delta}_{t-1}^i]$.

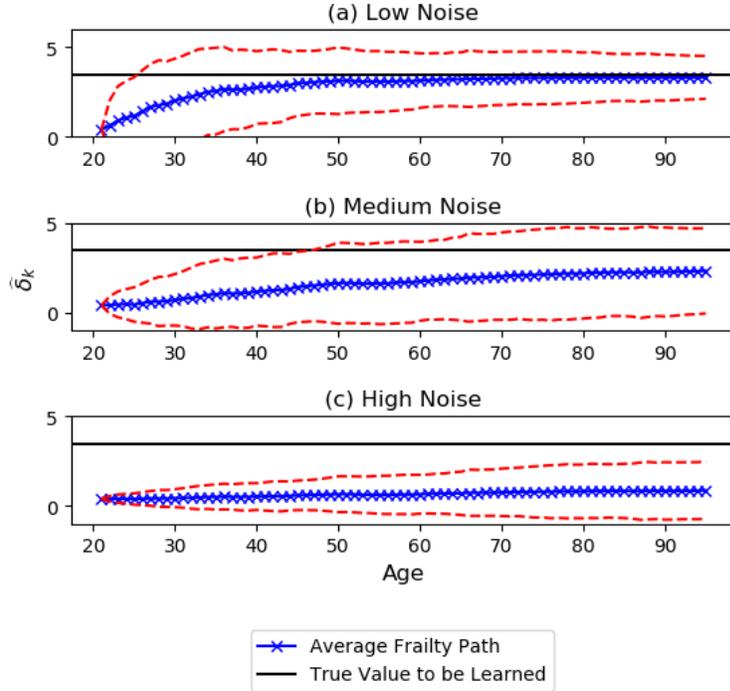
To illustrate this point, for the same frailty δ^i I simulate 1000 learning path realizations which are then amplified by noise variance $\sigma_e^2 = 1$ (low noise), a 1000 by noise variance $\sigma_e^2 = 10$ (medium noise), a 1000 by noise variance $\sigma_e^2 = 100$ (high noise), and a 1000 by noise variance $\sigma_e^2 = 1000$, all while keeping fixed a suitable set of the rest of parameters. In order to elicit the general smoothness due to each σ_e^2 , I calculate the standard deviation of the realizations of $\widehat{\delta}_t^i$ per learning path, that is, a lower value of the standard deviation signals a smoother path, and then I calculate the average of this statistic per simulation. The results can be seen in Table 1.

As explained above, when the standard deviation is calculated for the entire path, from ages 20 to 95, smoothness increases as the noise of the learning mechanism increases, indicating the individual tends to ignore more the signal she receives about her frailty. On the other hand, when we examine the path smoothness from ages 60 to 95, we can observe that the path is less smooth when going from $\sigma_e^2 = 1$ to $\sigma_e^2 = 10$, indicating that less noise can induce too less noise during the last years of age once it has already allowed for the actual frailty value to be learned, or close to be learned.

With the first three simulations (low noise, medium noise, and high noise) we can also observe how noise impedes the learning of the true value. Figure 1 shows the average frailty value learned per age (with the correspondent lower and upper confidence intervals) by type of noise simulation. When the noise is low the average learning path stabilizes near the true value around age 50, as shown in Figure 1(a), yet when the noise increases moderately the average learning path drifts toward the true value but it never reaches it, as shown in Figure 1(b). Finally, in Figure 1(c), we can see that when

the noise is high the average learning path hardly ever moves in time, indicating that the true frailty value is never learned, and that the individual hardly ever updates her beliefs since their initial value.

Figure 1: Learning Simulation: Convergence to True Value by Noise



At this point it is also useful to fix ideas about mortality beliefs and an individual's mortality deviation from the standard individual's mortality. Normalizing the standard individual's frailty to 1, $\delta^{std} = 1$, allows the model to trace the distribution of actual (idiosyncratic) frailties, that is, a distribution assumed to be around $\delta^{std} = 1$. Yet throughout most of the life-cycle each δ^i is not observed by its correspondent individual. The process of learning δ^i then implies that the distribution of learned frailties $\widehat{\delta}_t^i$ at time t is not the same as the distribution of δ^i , and consequently, the distribution of subjective mortality beliefs will not be the same as the actual mortality distribution of individuals, unless δ^i is fully learned by all of them at advanced ages.

3 Optimal Annuitization and Longevity Uncertainty

In this section I describe an asset rebalancing model to study optimal annuitization. Actuarially fair annuity prices are calculated using frailty beliefs, that is, eliciting subjective survival probabilities which will be used in the calculation of the present value of future annuity payments. As mentioned above, in this paper we only study the demand of annuities, and for this purpose, the assumption of no

private information in the annuity market allows these actuarially fair annuity prices (as calculated by individuals) to be used in the model, i.e., insurers would still pool individual longevity risks using annuity prices that conform with the frailty beliefs of the individuals. As individuals learn about their own frailties, the smoothness of their learning paths determines optimal annuitization as an endogenous decision that crucially depends on this learning process. This also implies that for this demand-focused model annuities markets can be interpreted as complete: to sell annuities, and repricing them, is possible; and furthermore, asset rebalancing can be performed without transaction costs.

3.1 Individuals

The economy is populated with individuals who enter the labor market when born at age t_1 . Individuals believe there is a specific survival probability at each age, and therefore, each individual forms a subjective survival probability belief each year, which is based on a frailty $\widehat{\delta}_t^i$ belief, as explained in section 2. All individuals retire at period $t = t_{RET}$, and they are certain they cannot survive longer than age T . Furthermore, there are no accidental bequests, that is, if an individual dies her wealth is not redistributed among survivors.

Each individual receives a flow utility from consumption c_t^i , such that the forward-looking utility at period t the individual believes she has is

$$U_t = \sum_{s=t}^T \beta^s P_s(\widehat{\delta}_t^i) u(c_s^i).$$

where $u(c) = \frac{c^{1-\varsigma}}{1-\varsigma}$, and ς is the coefficient of relative risk aversion. β is the discount factor.

3.2 Annuities and Bonds

Wealth a_t can be invested in either risk-free bonds or annuities. Total return TR then will be decided between these two types of returns, and while the risk-free bond return r is constant, the one-period return of the annuity ϕ_t is not, therefore,

$$TR = \theta_t \phi_t + (1 - \theta_t)r$$

where $\theta_t \in \{0, 1\}$ reflects the decision of whether to invest in annuities or not at time t .

Assuming annuities pay a dividend of \$1, and that the risk-free interest rate r is used to discount future cash flows, what individual i believes is the actuarially fair annuity price at time t , based on the frailty belief at time $t - 1$, is

$$q_t(\widehat{\delta}_{t-1}^i) = \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} + \frac{\prod_{i=t+1}^{t+2} \rho_i(\widehat{\delta}_{t-1}^i)}{(1+r)^2} + \dots + \frac{\prod_{i=t+1}^T \rho_i(\widehat{\delta}_{t-1}^i)}{(1+r)^{T-t}} \quad (5)$$

In the same fashion as Reichling and Smetters (2015), it is assumed that this is also the competitive price offered by insurers as knowledge of frailty $\widehat{\delta}_{t-1}^i$ is not private information. As mentioned above, this pricing is supported by the assumptions made for the present demand-focused model.

The lag between the frailty belief and the pricing embeds the idea that subjective mortality will determine the (subjective) return on annuities: the price at time t is set the previous period, such that at the time of determining the sale price of the annuity, an update in frailty beliefs has already occurred. Consequently, the computation of the annuity return $\phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i)$ has to take into account current and previous frailty beliefs:

$$\phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) = \frac{1 + q_{t+1}(\widehat{\delta}_t^i)}{q_t(\widehat{\delta}_{t-1}^i)} - 1. \quad (6)$$

Proposition 1: *The subjective return of an annuity $\phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i)$, as priced as in equation (6), exceeds the risk-free interest rate r if and only if $\widehat{\delta}_t^i \leq \widehat{\delta}_{t-1}^i$ (Proof in Appendix A).*

Proposition 1 states that for the annuity return to be greater than the risk-free interest rate r , it is necessary to have downward-smoothed updated frailty beliefs, i.e., $\widehat{\delta}_t^i \leq \widehat{\delta}_{t-1}^i$. Otherwise, if the individual updates her frailty belief to be higher, then the price of the annuity at time $t+1$, $q_{t+1}(\widehat{\delta}_t^i)$, lowers the annuity return. Furthermore, as mentioned above, it is assumed that individuals are risk averse in this economy.⁶ This framework then is accounting for the effect of the learning mechanism of subjective frailties on the return on annuities; downward-smoothness updating is necessary for the strict preference for annuities to kick in.

3.3 Recursive Formulation

Each period (age) t , an individual receives her previously determined wealth a_t and her disposable current income (net earnings before retirement and Social Security benefits when retired). The price q_t at which she could buy (competitively) an annuity is set according to the previous period, that is, using her previous frailty belief: $q_t(\widehat{\delta}_{t-1}^i)$. Nevertheless, she observes the realization of the noisy signal s_t and consequently forms a current belief of her frailty $\widehat{\delta}_t^i$ which, while determining the price for the next period $q_{t+1}(\widehat{\delta}_t^i)$, settles the one-period annuity return $\phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i)$ should the individual decide to sell before the current period ends. With this information then, and after deciding between

⁶For a discussion of risk neutral individuals, stochastic survival probabilities, and statewise annuity dominance see Reichling and Smetters (2015).

consumption and savings, the individual decides to invest her wealth (savings) in either annuities or risk-free bonds, i.e., $\theta_t \in \{0, 1\}$. Lastly, before the period ends, the agent sells her annuities or bonds holdings, determining this way her wealth for the next period a_{t+1} .

This problem can now be formulated in a recursive manner. The state vector is formed by assets a_t^i , the noisy signal s_t , and the previous frailty belief $\widehat{\delta}_{t-1}^i$. Define V_t^i as the value function of a t year old individual, the dynamic problem is then

$$V_t^i(a_t^i, \widehat{\delta}_{t-1}^i, s_t) = \max_{c_t, a_{t+1}, \theta_t} \left\{ u(c_t^i) + \beta \rho_{t+1}(\widehat{\delta}_t^i) E \left[V_{t+1}^i(a_{t+1}^i, \widehat{\delta}_t^i, s_{t+1}) \mid \widehat{\delta}_t^i \right] \right\}$$

s.t.

$$\begin{aligned} a_{t+1} &= \left(\theta_t \phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) + (1 - \theta_t)r \right) (a_t + w(1 - \tau) - c_t) \quad \text{for } t < t_{RET} \\ a_{t+1} &= \left(\theta_t \phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) + (1 - \theta_t)r \right) (a_t - c_t + b) \quad \text{for } t \geq t_{RET} \\ \widehat{\delta}_t^i &= \widehat{\delta}_{t-1}^i + G_t [s_t^i - \gamma t^\gamma \widehat{\delta}_{t-1}^i] \\ \sigma_{\widehat{\delta}^i, t}^2 &= \sigma_{\widehat{\delta}^i, t-1}^2 - G_t \gamma t^\gamma \sigma_{\widehat{\delta}^i, t-1}^2 \\ a_{t+1}^i &\geq 0 \end{aligned}$$

where

$$\begin{aligned} G_t &= \frac{\gamma t^\gamma \sigma_{\widehat{\delta}^i, t-1}^2}{(\gamma t^\gamma)^2 \sigma_{\widehat{\delta}^i, t-1}^2 + \sigma_e^2} \\ s_t^i &= \gamma t^\gamma \delta^i + e_t^i \\ \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) &= \frac{1 + q_{t+1}(\widehat{\delta}_t^i)}{q_t(\widehat{\delta}_{t-1}^i)} - 1 \end{aligned}$$

and where w is the individual's wage in the current period, and b and τ are the Social Security benefit and Social Security tax, respectively. There is no analytical solution for this problem; numerical methods are necessary to compute the solution of this model.

3.3.1 Social Security

A Social Security benefit is included to track the trade-off between taxing during the working years and income during retirement (this benefit is also deemed as an annuity substitute; see Feldstein, 2005). Therefore, I include a standard form of Social Security program without redistributive roles, that is, I adopt a model of balanced government spending such that Social Security benefits and taxes conform a balanced budget, that is, $(t_{RET-1} - t)w = (T - t_{RET})b$.

4 Empirical Strategy

The empirical strategy consists in first calibrating, through a Simulated Method of Moments, the governing parameters of the frailty learning process using the Health and Retirement Study data on subjective mortality beliefs. Subsequently, the implied learning process of frailty is used to compute the present value of an annuity at each age. In this manner, we can track the decision of the individual of whether to annuitize or not her wealth as she ages.

4.1 Data

I use weighted subjective survival chances responses of the Health and Retirement Study (RAND files), for biennial surveys from 2000 to 2010 (waves 5 to 10). This subjective belief is elicited by asking respondents to assign a numerical value between 0 and 100. In general, the form of the question is

“Next I have some questions about how likely you think various events might be. When I ask a question I’d like for you to give me a number from 0 to 100, where ’0’ means that you think there is absolutely no chance, and ’100’ means that you think the event is absolutely sure to happen.”

with the specific question being:

“(What is the percent chance) that you will live to be *_target age_* or more?”

These questions about survivorship have a larger sample for the target ages of 75 and 80, so I only use these two target ages for the estimation exercise. I restrict the lowest possible age of the respondents to 53 and cap the maximum age to 65 for target age 75, and 69 for target age 80. This gives enough weighted data to see how the expectations of survival for two target ages differ per response. I do not exploit the panel nature of the data, though this can be left for a future exercise. For more details see Appendix C.

It is important to mention that I do not filter corner answers (focal responses of 0% chance or 100% chance) as I deem these responses informative of the individuals beliefs at the time of responding, as econometrically biased or inaccurate as they may be. Furthermore, in this paper subjective survival probabilities are not directly used to build a survival probability distribution (see Gan et al, 2015), rather they are used to infer about the structural mortality learning model of section 3.

Table 2. SMM Calibrated Parameters

Parameter	Description	Value
σ_e^2	Noise Variance	50
$\sigma_{\delta^i,0}^2$	Initial Belief Frailty Variance	7.2
$\widehat{\delta}_0^i$	Initial Common Frailty Belief	0.4
γ	Speed of Learning Regulator	0.11
σ_A^2	Actual Frailty Variance	60

4.2 Simulated Method of Moments

Interpreting each belief as a conditional subjective belief, that is,

$$\rho_{75}(\widehat{\delta}_k^i) \text{ for } k = 53, \dots, 65$$

and

$$\rho_{80}(\widehat{\delta}_k^i) \text{ for } k = 53, \dots, 69$$

we have 13 first moments for $\rho_{75}(\widehat{\delta}_k^i)$ and 17 first moments for $\rho_{80}(\widehat{\delta}_k^i)$, making a total of 30 moments to match for the four parameters of the frailty learning process that we are trying to calibrate (σ_e^2 , $\sigma_{\delta^i,0}^2$, $\widehat{\delta}_0^i$, γ). Additionally, by assuming that individuals know with more accuracy their objective mortality as they age, we can calibrate the variance of the actual frailty distribution σ_A^2 , which would be approximately equal to the variance of frailty beliefs of the oldest individuals.

The calibrated parameters then are given by

$$\hat{b} = \arg \min_b g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)' W g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$$

$$\text{for } b = \sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2$$

where $g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$ is a 30×1 matrix measuring the distance between the sample moments and the model moments concerning $\rho_{75}(\widehat{\delta}_k^i)$ and $\rho_{80}(\widehat{\delta}_k^i)$, and for simplicity W is a 30×30 identity matrix. The calibrated parameters can be found in Table 2.

Table 3. Life-Cycle Parameters

Parameter	Description	Value
T	Age of Certain Death	95
t_R	Retirement Age	65
τ	Social Security Tax	0.3
r	Annual Risk-free Interest Rate	0.04
β	Time Discount Factor	0.99

4.3 Life-Cycle Parameters

Once the path of the state variables $\rho_t(\widehat{\delta}_k^i)$ and s_t is simulated with the calibrated parameters I proceed to solve for the dynamic programming problem of section 3.3 over them. Table 3 summarizes the rest of the parameters being calibrated for the dynamic programming computation.

In order to simplify an already complicated model, the end of life is set to age 95. Additionally, the individual's wage w is set to be constant and calibrated to target the replacement ratio of the average Social Security benefit. Following Hosseini (2015), the targeted replacement ratio for the United States is set to 45%. Social Security tax τ then is calibrated to match the expression described in section 3.3.1.

The standard individual for which frailty is set to unity is taken from the Social Security life tables of Bell and Miller (2005), cohort 1950, that is,

$$P_t(\widehat{\delta}_k^i) = \exp(-\widehat{\delta}_k^i H_t^{std}) = \exp(\widehat{\delta}_k^i \ln(P_t^{1950}))$$

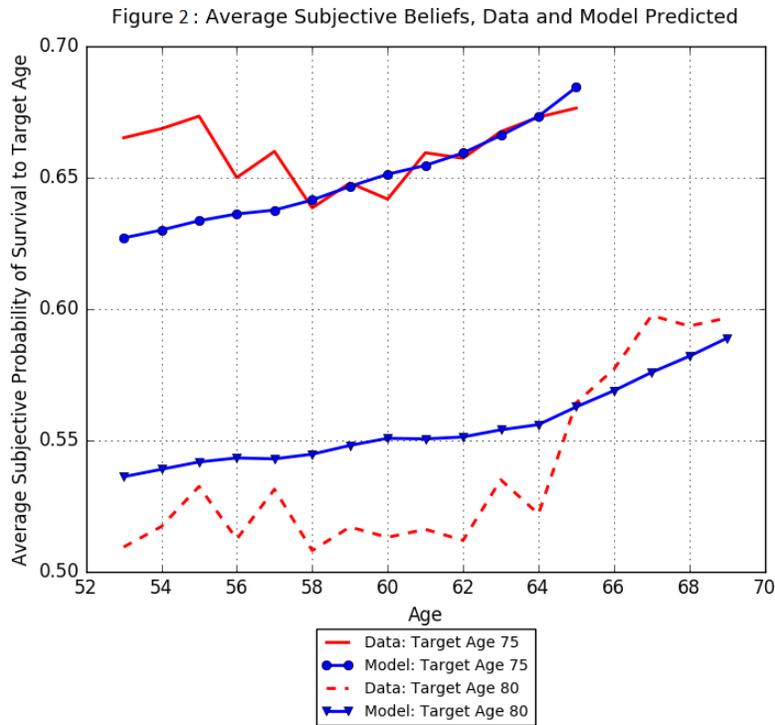
where P_t^{1950} denotes the surviving probability to age t for a cohort born in 1950.

This profile of survival probabilities is chosen because, typically, life tables have been used in the life-cycle literature to infer the survival probability of a representative individual, as in Huggett (1993). It is therefore plausible to assume that all individuals are aware of this aggregate measure.

All these calibrated parameters jointly with SMM estimated values of section 4.2, determine the baseline model to be used.

5 Results

Figure 2 shows the learning model’s predicted average subjective beliefs of survival along with the average subjective beliefs of the data. Considering that the moments to match in the empirical exercise are the average subjective probabilities themselves, and not the distance between the two target ages per age, we can observe that the model replicates well the overall distance per age for both average subjective beliefs. The trajectory of both beliefs paths is also replicated well for the advanced ages in the sample. Figure 3 shows the percentage of annuities offering a lower return than the risk free return per age. The decrease in this percentage as the individual ages is an indication of what the learning effect will be on annuitization: frailty learning paths become downward-smoothed as an individual’s updating process of frailty beliefs confirms her previous beliefs (no more learning), and therefore, the annuity rate of return is greater for more individuals during the last years of life, as explained in section 3.2.



Note: SMM Average Estimated Values of the Learning Process. Data from the HRS, survey years 2000,2002,2004,2006, and 2010

Figure 3: Percentage of Annuities offering a Lower Return than the Risk-Free Interest Rate

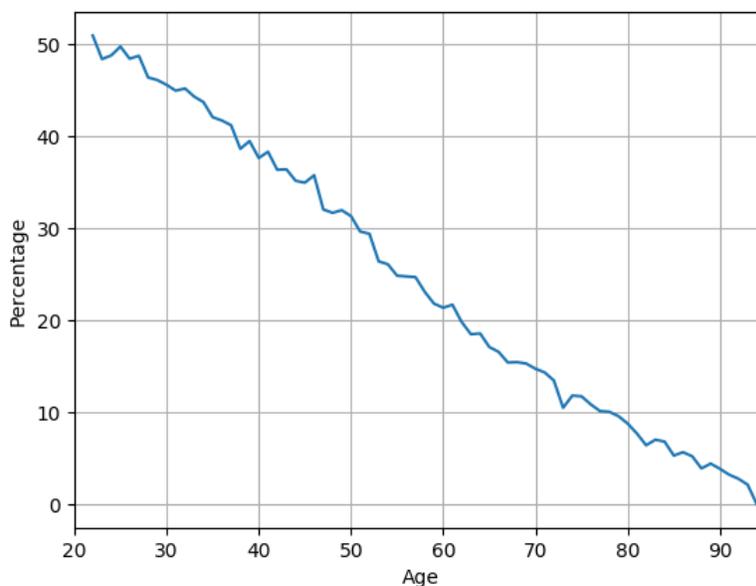
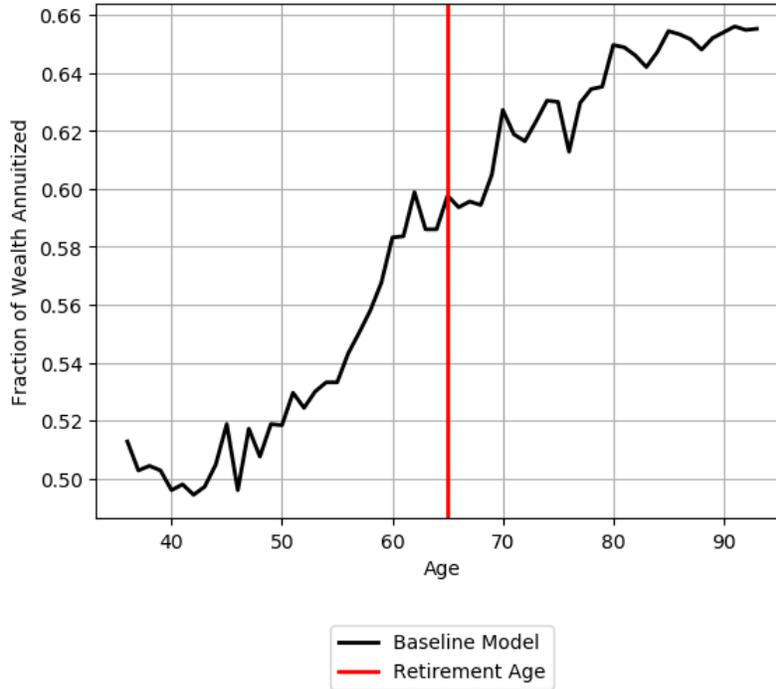


Figure 4 presents the annuitization profile for the baseline model. In conformity with the decreasing percentage of annuities offering a lower return than the risk free return, the effect of the individual learning about her mortality is the increasing annuitization profile. This profile does not imply complete annuitization during retirement years: the percentage of wealth annuitized at retirement age is approximately only 60%, and by age 90 is above 65%. This low level of annuitization is solely due to the individual learning about her true mortality risk, as the learning mechanism determines how downward-smoothed the frailty belief path is (see Proposition 1). The intuition in this scenario is that if an individual does not update her mortality risk beliefs, either because she already knows her true mortality risk or because she is incapable to learn it (for example due to a high level of noise in the learning mechanism), then the strong take-up effect for annuities, as first explored by Yaari (1965), will take effect. But if the individual is actively learning about her mortality risk such that she is constantly updating her beliefs, then at any age she can update her mortality risk beliefs upwards, opening the possibility for an annuity investment to not be optimal anymore.

Moreover, it is also important to notice that, based on the percentage of annuities with a return greater than the the risk free return in Figure 3, it could have been conjectured a 70% level of annuitization at age 65, instead of just 60%. But the actual (lower) level of annuitization is due to risk averse quality of the individual during the life-cycle.

Figure 4: Annuitization of Wealth, Base Model



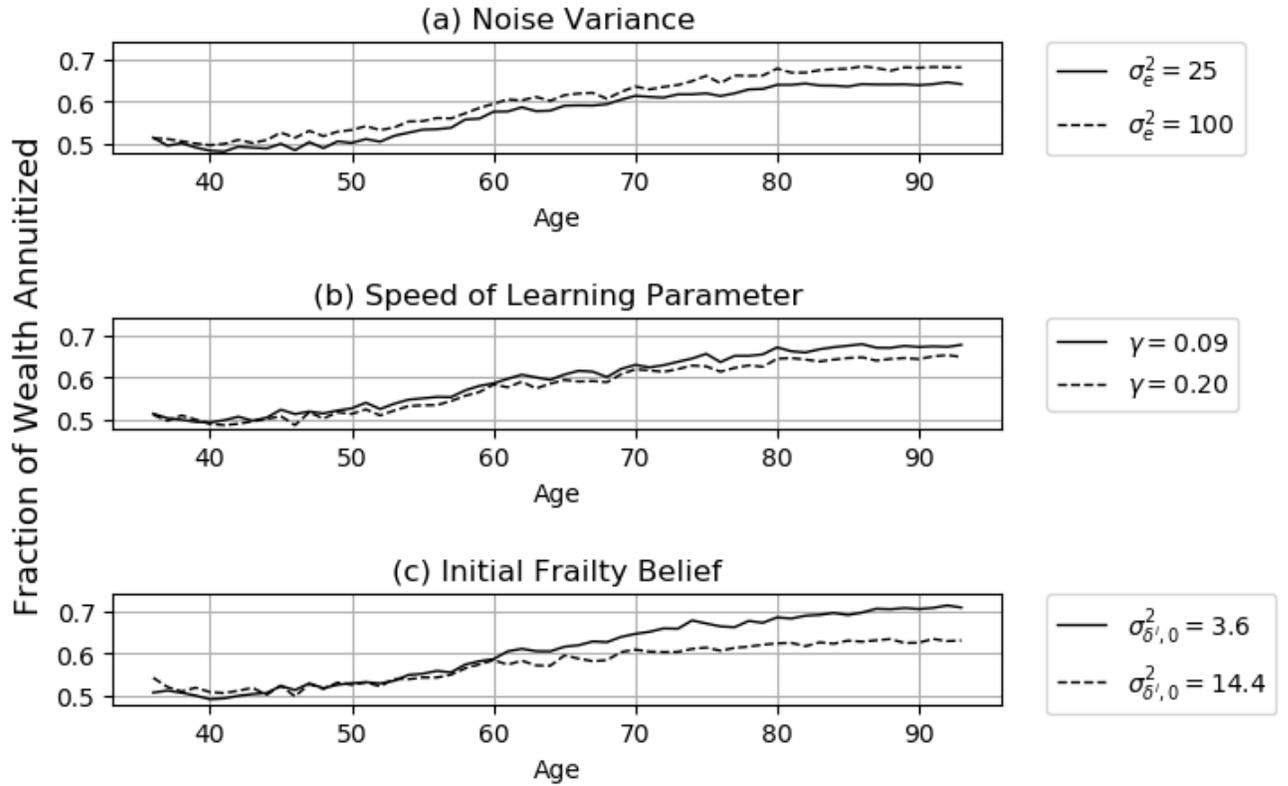
To gauge the importance of this result, one should compare this annuitization profile in the context of past non-complete annuitization results in the literature. While Davidoff, Brown and Diamond (2005) found optimal annuitization of as low as 75% by imposing the assumptions of habit formation and incomplete annuity markets, Peijnenburg et al. (2016) show that market incompleteness in a more general environment can still yield full annuitization. On the other hand, based on a similar mechanism that allows them to model stochastic mortality risk (instead of mortality risk learning), Reichling and Smetters (2015) have found optimal levels of annuitization between 36% and 26%. Yet their model operates in a much richer environment that includes income shocks, multiple transmission channels through which health shocks affect income, and bequest motives. In the present model, the goal is to isolate the subjective mortality beliefs channel, and in this way examining the learning mortality effect on annuitization; the only operating mechanism that determines optimal annuitization is the frailty learning process which, through the resulting subjective survival probabilities, encompass all perceived mortality risks (including negative health shocks), which in turn, directly affect the present value of future annuity payments. Mortality learning in this model does not affect income nor expectations of future expenditures shocks, and furthermore, does not take into account the supply side of the annuities market, which would only drive the level of annuitization down further. There are no other theoretical assumptions that would encourage the individual in the model to disregard annuities, that

is, these results take place in a non-stringent environment of intertemporally separable utilities and complete annuity markets.

5.1 Robustness

Figure 5 shows variations above and below of the parameters for the baseline model. Three parameters shaping the frailty learning mechanism are studied to understand their impact on annuitization. A higher noise variance σ_e^2 , as explained in section 2.2, does not have an a priori effect on annuitization. But as seen in Figure 5(a), in this case, as the individual learns about her actual frailty in a slower manner, her frailty belief path is smoother as she trusts less in the signal she observes. Compared to the baseline model then there will be more downward-smoothed frailty paths, and therefore, the overall level of annuitization in the economy will increase. An increase in annuitization also takes place when the speed of learning parameter γ is lower in Figure 5(b), the reason being due to a delay in frailty learning. Lastly, higher annuitization also takes place when the initial frailty variance belief $\sigma_{\delta^i,0}^2$ in Figure 5(c) is further below the actual frailty variance. This last parameter variation shows the effect of underestimating the frailty variance at the beginning of life, that is, given that learning about one's own mortality is already difficult, a prior uniform belief of the frailty distribution (which is what a low $\sigma_{\delta^i,0}^2$ would reflect) will just delay the learning of a wide idiosyncratic frailty distribution.

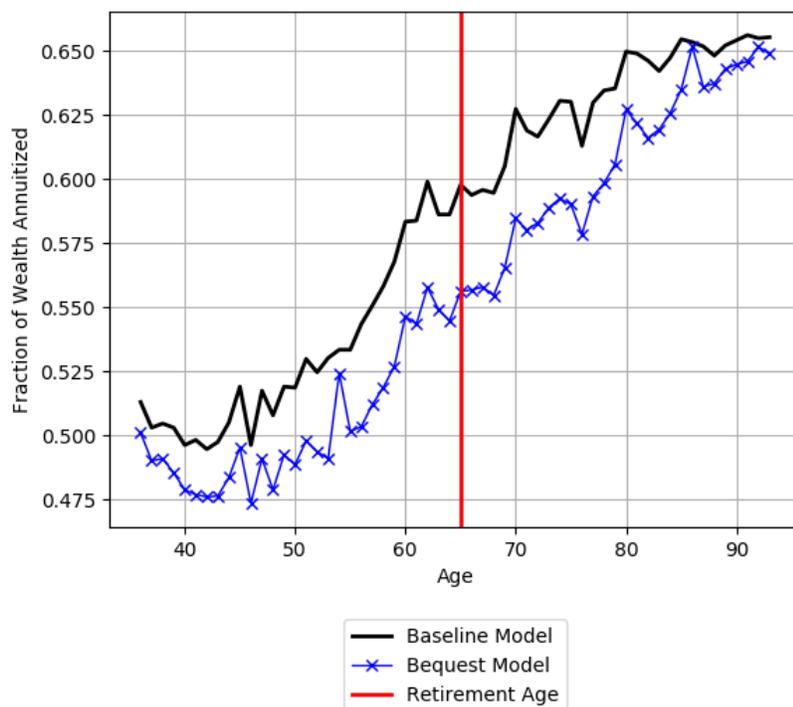
Figure 5 : Robustness, Parameters Comparison



5.2 Bequests Motive Extension

Facing a subjective different probability of death each period, for further examination a bequest motive is added to this framework. Following De Nardi (2004), a calibrated model is described in Appendix B. Examining Figure 6 we can see that bequests have a non-negligible impact on annuitization, that is, on top of the direct effect of mortality risk learning. The reasons why optimal annuitization is lower in the bequest model, as the individual ages, is because the utility of bequeathing is greater as the probability of dying increases. This new feature allows us to see how more details can be added to this mortality learning model, enabling us to drive down further the levels of optimal annuitization.

Figure 6 : Annuitization of Wealth, Bequest Model Comparison



6 Conclusions

This paper shows how subjective mortality beliefs, as the result of the process of learning actual mortality, influence the demand for annuities in the life-cycle. Specifically, as annuities are always preferred when the distribution of survival probabilities does not change, in this paper I showed how uncertainty about the survival distribution, that is, uncertainty about mortality risk, does not allow for such strong take-up of annuities result to take place. This result is significant given that individuals act on what they believe the nature of their mortality is. Subjective survival probabilities are the most important factor for individuals when determining annuities return.

First I calibrated the parameters of a learning process using subjective survival probabilities data from the Health and Retirement Study. Individuals learn in a Bayesian manner about their actual frailty and in the process form subjective beliefs about it. Subsequently, I use the generated subjective survival probabilities to construct a life-cycle model in which individuals must decide recursively to annuitize or not, highlighting in this manner the subjective mortality role for optimal annuitization. The demand-focused model and the assumptions of complete markets and asset rebalancing without costs do not interfere with the mortality risk learning motive for optimal annuitization. As such,

the results show that absent expenditure shocks, or constraints affecting income at advanced ages, optimal annuitization is already incomplete due to the mortality risk learning process. The baseline calibrated model shows that optimal annuitization yields approximately only 60% of annuitized wealth at retirement age. Furthermore I examine how noise in the learning process, learning ability, and initial beliefs affect the levels of annuitization, and I find the results of the base model to be mostly robust to these variations.

This work is meant to be complementary to the already large literature about optimal annuitization. These findings suggest that a large fraction of annuitization, or the absence of it in the data, can be explained if we take into account the formation process of subjective mortality beliefs.

References

- [1] Bell, F. C. and M. L. Miller . "Life tables for the United States social security area 1900-2100." Office of the Chief Actuary, Social Security Administration (2005).
- [2] Brown, Jeffrey R., Jeffrey R. Kling, Sendhil Mullainathan, and Marian V. Wrobel. "Why don't people insure late life consumption: A framing explanation of the under-annuitization puzzle". No. w13748. National Bureau of Economic Research, (2008).
- [3] Brown, Jeffrey R., Arie Kapteyn, Erzo FP Luttmer, and Olivia S. Mitchell. "Cognitive constraints on valuing annuities". No. w19168. National Bureau of Economic Research, (2013).
- [4] Brugiavini, Agar. "Uncertainty resolution and the timing of annuity purchases." *Journal of Public Economics* 50.1 (1993): 31-62.
- [5] Caliendo, Frank N., Nick L. Guo, and Roozbeh Hosseini. "Social security is NOT a substitute for annuity markets." *Review of Economic Dynamics* 17.4 (2014): 739-755.
- [6] Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond. "Annuities and individual welfare." *The American economic review* 95.5 (2005): 1573-1590.
- [7] De Nardi, Mariacristina. "Wealth inequality and intergenerational links." *The Review of Economic Studies* 71.3 (2004): 743-768.
- [8] Feldstein, Martin. "Rethinking Social Insurance." *American Economic Review* (2005).

- [9] Franks, Peter, Carolyn M. Clancy, Marthe R. Gold, and Paul A. Nutting. "Health insurance and subjective health status: data from the 1987 National Medical Expenditure survey." *American Journal of Public Health* 83, no. 9 (1993): 1295-1299.
- [10] Friedman, Benjamin M., and Mark J. Warshawsky. "The cost of annuities: Implications for saving behavior and bequests." *The Quarterly Journal of Economics* 105.1 (1990): 135-154.
- [11] Gan, Li, Guan Gong, Michael Hurd, and Daniel McFadden. "Subjective mortality risk and bequests." *Journal of Econometrics* 188, no. 2 (2015): 514-525.
- [12] Guvenen, Fatih. "Learning your earning: Are labor income shocks really very persistent?." *The American Economic Review* 97.3 (2007): 687-712.
- [13] Heimer, Rawley, Kristian Ove R. Myrseth, and Raphael Schoenle. "YOLO: Can Subjective Life-Expectancies Explain Household Investment Puzzles?." Available at SSRN 2669484 (2015).
- [14] Hosseini, Roozbeh. "Adverse selection in the annuity market and the role for social security." *Journal of Political Economy* 123.4 (2015): 941-984.
- [15] Huggett, Mark. "The risk-free rate in heterogeneous-agent incomplete-insurance economies." *Journal of Economic Dynamics and Control* 17 (1993), 953-969.
- [16] Hurd, Michael D., and Kathleen McGarry. "Evaluation of the subjective probabilities of survival in the health and retirement study." *Journal of Human Resources, Special Issue on the Health and Retirement Study: Data Quality and Early Results, Vol. 30* (1995): S268-S292.
- [17] Hurd, Michael D., and Kathleen McGarry. "The predictive validity of subjective probabilities of survival." *The Economic Journal* 112.482 (2002): 966-985.
- [18] James, Estelle, Guillermo Martinez, and Augusto Iglesias. "The payout stage in Chile: who annuitizes and why?." *Journal of Pension Economics and Finance* 5.02 (2006): 121-154.
- [19] Jeske, Karsten, and Sagiri Kitao. "US tax policy and health insurance demand: Can a regressive policy improve welfare?." *Journal of Monetary Economics* 56.2 (2009): 210-221.
- [20] Manton, Kenneth G., Eric Stallard, and James W. Vaupel. "Methods for comparing the mortality experience of heterogeneous populations." *Demography* 18.3 (1981): 389-410.
- [21] Reichling, Felix, and Kent Smetters. "Optimal annuitization with stochastic mortality and correlated medical costs." *The American Economic Review* 105.11 (2015): 3273-3320.

- [22] Panis, Constantijn WA, and Michael Michael J. Brien. "Innovations and trends in annuities: Qualifying Longevity Annuity Contracts (QLACs)" (2016).
- [23] Pashchenko, Svetlana, and Ponpoje Porapakkarm. "Quantitative analysis of health insurance reform: Separating regulation from redistribution." *Review of Economic Dynamics* 16.3 (2013): 383-404.
- [24] Peijnenburg, Kim, Theo Nijman, and Bas JM Werker. "The annuity puzzle remains a puzzle." *Journal of Economic Dynamics and Control* 70 (2016): 18-35.
- [25] Sun, Wei, and Anthony Webb. "How Do Subjective Mortality Beliefs Affect the Value of Social Security and the Optimal Claiming Age." Center for Retirement Research at Boston College Working Paper 2011-22 (2011).
- [26] Vaupel, James W., Kenneth G. Manton, and Eric Stallard. "The impact of heterogeneity in individual frailty on the dynamics of mortality." *Demography* 16.3 (1979): 439-454.
- [27] Yaari, Menahem E. "Uncertain lifetime, life insurance, and the theory of the consumer." *The Review of Economic Studies* 32.2 (1965): 137-150.

APPENDIX A (Proof of Proposition 1)

As shown in Reichling and Smetters (2015), with *deterministic* survival probabilities fairly priced annuities statewise dominate risk-free bonds. This deterministic quality is translated in the present framework as having smoothed beliefs, in virtue that if $\widehat{\delta}_t^i = \widehat{\delta}_{t-1}^i$, then

$$\begin{aligned} q_t(\widehat{\delta}_{t-1}^i) &= \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} + \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[\rho_{t+2}(\widehat{\delta}_{t-1}^i) + \dots + \frac{\prod_{l=t+2}^T \rho_l(\widehat{\delta}_{t-1}^i)}{(1+r)^{T-t-1}} \right] \\ q_t(\widehat{\delta}_{t-1}^i) &= \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} + \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[q_{t+1}(\widehat{\delta}_{t-1}^i) \right] \\ \frac{(1+r)}{\rho_{t+1}(\widehat{\delta}_{t-1}^i)} q_t(\widehat{\delta}_{t-1}^i) &= 1 + q_{t+1}(\widehat{\delta}_{t-1}^i) \end{aligned}$$

which replaced in equation (6) gives the relation between gross returns as

$$1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_{t-1}^i) = \frac{1+r}{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}$$

implying that the annuity return will always be higher since $\rho_{t+1}(\widehat{\delta}_{t-1}^i) < 1$. On the other hand, if $\widehat{\delta}_t^i > \widehat{\delta}_{t-1}^i$ then the annuity return $\phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i)$ will not generically exceed the risk-free interest rate r since it implies $q_{t+1}(\widehat{\delta}_t^i) < q_{t+1}(\widehat{\delta}_{t-1}^i)$, such that

$$\begin{aligned} \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[q_{t+1}(\widehat{\delta}_t^i) \right] &< \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[q_{t+1}(\widehat{\delta}_{t-1}^i) \right] \\ \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} + \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[q_{t+1}(\widehat{\delta}_t^i) \right] &< \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} + \frac{\rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[q_{t+1}(\widehat{\delta}_{t-1}^i) \right] \\ \frac{q_t(\widehat{\delta}_{t-1}^i) \rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) \right] &< \frac{q_t(\widehat{\delta}_{t-1}^i) \rho_{t+1}(\widehat{\delta}_{t-1}^i)}{(1+r)} \left[1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_{t-1}^i) \right] \end{aligned}$$

which by using the result above implies that

$$\begin{aligned} 1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) &< 1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_{t-1}^i) \\ 1 + \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) &< \frac{1+r}{\rho_{t+1}(\widehat{\delta}_t^i)} \end{aligned}$$

which *does not* imply the annuity return will always be higher than the risk-free return.

APPENDIX B (Bequests Model)

As a baseline value I set the value of the coefficient of relative risk aversion as $\varsigma = 3$. The model including a bequest motive now becomes

$$V_t^i(a_t^i, \widehat{\delta}_{t-1}^i, s_t) =$$

$$\max_{c_t, a_{t+1}, \theta_t} \left\{ u(c_t^i) + \beta \left(\rho_{t+1}(\widehat{\delta}_t^i) E \left[V_{t+1}^i(a_{t+1}^i, \widehat{\delta}_t^i, s_{t+1}) \mid \widehat{\delta}_t^i \right] + [1 - \rho_{t+1}(\widehat{\delta}_t^i)] \Phi(a_{t+1}^i) \right) \right\}$$

s.t.

$$a_{t+1} = \left(\theta_t \phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) + (1 - \theta_t)r \right) (a_t + \bar{w}_t(1 - \tau) - c_t) \quad \text{for } t < t_{RET}$$

$$a_{t+1} = \left(\theta_t \phi_t(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) + (1 - \theta_t)r \right) (a_t - c_t + ben) \quad \text{for } t \geq t_{RET}$$

$$\widehat{\delta}_t^i = \widehat{\delta}_{t-1}^i + G_t [s_t^i - \gamma t^\gamma \widehat{\delta}_{t-1}^i]$$

$$\sigma_{\delta^i, t}^2 = \sigma_{\delta^i, t-1}^2 - G_t \gamma t^\gamma \sigma_{\delta^i, t-1}^2$$

$$a_{t+1}^i \geq 0$$

where

$$\begin{aligned} G_t &= \frac{\gamma t^\gamma \sigma_{\delta^i, t-1}^2}{(\gamma t^\gamma)^2 \sigma_{\delta^i, t-1}^2 + \sigma_e^2} \\ s_t^i &= \gamma t^\gamma \delta^i + e_t^i \\ \phi(\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i) &= \frac{1 + q_{t+1}(\widehat{\delta}_t^i)}{q_t(\widehat{\delta}_{t-1}^i)} - 1 \\ \Phi(a_t^i) &= \Phi_1 \left[1 + \frac{a_t^i - \tau_b * \max(0, a_t^i - x_b)}{\Phi_2} \right]^{1-\varsigma} \end{aligned}$$

The bequest function $\Phi(a_t^i)$ is based on De Nardi (2004): Φ_1 measures the concern about leaving bequests, Φ_2 measures the degree to which bequests are considered a luxury good, and τ_b is the estate tax, which is applicable if the estate exceeds the exemption level x_b . These parameters are calibrated to match the transfer wealth share in the U.S. and are displayed in Table B.1. Optimal annuitization results for the model with bequest motive can be seen in Figure 6.

APPENDIX C (Data and Algorithm)

Data

As mentioned in section 4, this study uses weighted subjective survival chances responses of the Health and Retirement Study (RAND files), for biennial surveys from 2000 to 2010 (waves 5 to 10). This subjective belief is elicited by asking respondents to assign a numerical value between 0 and 100. Table C1 presents the respondents age intervals at the time of the survey, along with the number of observations used from each survey (dataset). The questions about survivorship have a larger sample for the target ages of 75 and 80 which is why these ages are used. Note how these target ages are asked

Table C1. HRS Subjective Mortality Data Used

Dataset Year	Subjective Survival to Age 75		Subjective Survival to Age 80	
	Respondents Age	# obs	Respondents Age	# obs
2000	53-65	7016	53-69	8766
2002	53-65	6230	53-69	8094
2004	53-65	6604	53-69	8512
2006	53-64	5388	65-69	609
2008	53-64	4727	65-69	603
2010	53-64	2219	65-69	603

during different ages through the biennial surveys. Tables C2 and C3 present statistical information for the target ages of 75 and 80, respectively. These tables present the moments I match with the learning model, along with the number of observations and 10th and 80th percentiles. We can observe how the survival probabilities range within the 60s percent chances for target age 75, and within the 50s percent chances for target age 80. There are no discernible one-age jumps even though the subjective chances for target age 80 increase during the ages 65 throughout 69. Between ages 53 and 65 the difference in subjective percentage chances for the target ages is about 12 percentage points.

Algorithm

To compute the solution I first solve the learning model in order to use the implied subjective frailty beliefs in the life-cycle model. It is important to note that the value function form is the same for all individuals. What determines cross-sectional heterogeneity in this model is the distribution of frailties.

The model solution and calibration are similar to the ones developed by Guvenen (2007). The model does not pursue an inferential exercise for the parameters, but instead an efficient way to calibrate these via a simulated method of moments. The steps to follow to this end are the following:

1. First use a guess of the learning parameters that will be calibrated: $(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$.
2. Sample 100 individuals from a standard normal distribution which are then transformed to mimic a sampling of the distribution $N \sim (1, \sigma_A^2)$ by multiplying each occurrence by σ_A and adding 1.

**Table C2. Subjective
Survival Chance to Age 75, by Age**

Age	n	Mean	10th Pct	80th Pct
53	1825	66.51	20	90
54	2322	66.86	20	90
55	2661	67.33	20	90
56	2581	65	20	93
57	2639	65.99	20	90
58	2510	63.84	20	90
59	2741	64.79	20	90
60	2634	64.18	20	90
61	2806	65.94	20	95
62	2696	65.73	20	95
63	2871	66.75	20	95
64	2877	67.30	25	99
65	1814	67.64	25	100

**Table C3. Subjective
Survival Chance to Age 80, by Age**

Age	n	Mean	10th Pct	80th Pct
53	1046	50.94	10	80
54	1224	51.72	10	80
55	1298	53.24	10	80
56	1211	51.24	10	80
57	1339	53.14	10	80
58	1316	50.81	10	80
59	1604	51.70	10	80
60	1763	51.31	10	80
61	2002	51.61	10	80
62	2021	51.19	10	80
63	2114	53.50	10	80
64	2055	52.19	10	80
65	2098	56.42	10	85
66	2097	57.70	10	85
67	2001	59.75	10	90
68	1979	59.35	10	90
69	1961	59.66	10	90

3. For each individual sample from a standard normal distribution $N \sim (0, 1)$ 75 occurrences which are then transformed to mimic a sampling of the distribution $N \sim (0, \sigma_e^2)$ by multiplying each occurrence by σ_e . The 75 occurrences are meant to mimic a shock that every individual faces from age 20 to age 95. Perform this step 100 times for each individual, so for each individual there are 100 different shock paths in a lifetime.
4. For every individual I set a common initial frailty belief that is multiple of actual frailty σ_A^2 , that is $\sigma_{\delta^i,0}^2 = \psi * \sigma_A^2$. With this simplification, calibration of the multiplier ψ is tantamount to a calibration of $\sigma_{\delta^i,0}^2$.
5. For each of the 100 possible learning paths, for each of the 100 individuals, simulate a Kalman filtering learning process using equations 2 and 3, and using common initial beliefs $\sigma_{\delta^i,0}^2$ and $\widehat{\delta}_0^i$. In total there are 10000 possible learning paths of a duration 75 periods (ages).
6. For each frailty belief at time k , elicited in step 5, compute subjective unconditional survival probabilities for all ages t anchored on the standard individual's unconditional survival probabilities, as expressed in

$$P_t(\widehat{\delta}_k^i) = \exp(-\widehat{\delta}_k^i H_t^{std}) = \exp(\widehat{\delta}_k^i \ln(P_t^{1950}))$$

7. For each frailty compute then the theoretical counterpart for the moments to be matched, that is, the average conditional subjective survival probabilities corresponding to the target ages: $\rho_{75}(\widehat{\delta}_k^i)$ and $\rho_{80}(\widehat{\delta}_k^i)$, which make a total of 30 empirical moments.
8. With the theoretical and empirical moments compute the vector of distance $g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$, which together with $W = I$ compute the criteria

$$g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)' W g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$$

9. Using the same samples of a standard normal distribution $N \sim (0, 1)$ from Step 2 and 3, repeat Steps 2 through 8 with a new set of parameters $(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$ until a minimum of the criteria in Step 8 is reached.
10. For each individual, solve the dynamic programming problem described in section 3.3; the value function needs to visit the frailty space generated for each individual. Using $[\widehat{\delta}_{t-1}^i, \widehat{\delta}_t^i]$ solve the annuitization decision problem each period during the life-cycle. Keep track of the amount of wealth each individual decides to annuitize.

11. Aggregate all individuals annuitized wealth each period to compute the fraction of annuitized wealth in the economy.