6-2020

Essays in Retirement Economics

Gunnar Poppe Yanez
The Graduate Center, City University of New York

Essays in Retirement Economics

by

Gunnar Poppe Yanez

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2020

This manuscript has been read and accepted by the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor Sangeeta Pratap

Date

Chair of Examining Committee

Professor Christos Giannikos

Date

Executive Officer

Professor Frank Heiland

Professor Sangeeta Pratap

Professor Wim Vijverberg

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Essays in Retirement Economics

by

Gunnar Poppe Yanez

Adviser: Professor Sangeeta Pratap

This dissertation consists of three chapters.

Chapter 1 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 percent 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.

Chapter 2 Early claiming of Social Security benefits imply a reduction in the annuity value these payments offer to individuals who retire before the normal retirement age. To understand the prevalence of this behavior, in this paper I investigate the effect of mortality learning on early benefits take-up and early retirement by building a life-cycle model where individuals reduce their longevity uncertainty as they age. As individuals are more certain of their lifespans, the annuity value provided by Social Security benefits is less appealing, and consequently, early claiming can be optimal. Using data from the Health and Retirement Study to calibrate the model, I find that mortality learning is an important element in explaining this phenomena: early benefits take-up is 37.4 percent lower in a counterfactual scenario where individuals do not learn about their mortality. The impact of this result on a basic policy aimed at discouraging early retirement is discussed.

Chapter 3 The Annuity Puzzle has been studied in the economic literature for over 50 years. This chapter provides a summary of the main findings.

Acknowledgments

I have been very fortunate to receive the mentorship and guidance of Dr. Sangeeta Pratap throughout the process of writing this dissertation, from the idea conception to the fleshing out of the models. Her help and patience have been invaluable. I would also like to acknowledge the help and dedication I received from the members of my committee, Dr. Frank Heiland and Dr. Wim Vijverberg, without their comments and expertise I would not have been able to develop the chapters of my dissertation.

I would also like to acknowledge my friends support throughout all these years: Pablo Crespo, Evan Warshaw, Zhendong Zhao, Ernesto Garcia, and Antonio Cusato. Thank you for encouraging me to thrive on.

I also want to thank and acknowledge my family. To my wife Margaret who supported me throughout all these years. To my little baby girl Gabriela who filled my heart with joy ever since she was born. To my mother for all her support during these years.

Finally I want to dedicate this work to my father, my role model. I hope someday to have a fraction of the knowledge and wisdom he had.

Contents

List of Tables

List of Figures

Chapter 1

Mortality Risk Learning and the Demand for Annuities

1.1 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 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.

[1]Friedman 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.

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

[2] Finding 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.

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 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 the 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

[3]Reichling 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.

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 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 considerably 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.

1.2 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.

[4]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.

1.2.1 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 $\widehat{\delta}^i_k$. I refer to $\widehat{\delta}^i_k$ as idiosyncratic subjective frailty since it will allow me to define later an idiosyncratic *subjective* survival probability $P_t(\widehat{\delta}^i_k)$. Additionally, I assume individuals are aware of the heterogeneity across the population with the existence of a set of individual frailty types $\Delta = [\underline{\delta}, \overline{\delta}]$, which allows them to know there is a well defined cumulative distribution $F \in \Gamma(\Delta)$.

Suppose $\rho_t(\widehat{\delta}^i_k)$ 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 $\widehat{\delta}^i_k$. We have then $P_t(\widehat{\delta}^i_k) = \prod_{\imath=0}^{t} \rho_\imath(\widehat{\delta}^i_k)$. 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 $P_t(\widehat{\delta}^i_k)$, 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

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

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

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

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

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.

1.2.2 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 $e_t^i \sim N(0, \sigma_e^2)$:

$$s_t^i = d_t(\delta^i) + e_t^i \tag{1.1}$$

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.

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 $\sigma^2_{\delta^i}$,

$$\widehat{\delta}^i_t = \widehat{\delta}^i_{t-1} + G_t[s^i_t - \gamma t^\gamma \widehat{\delta}^i_{t-1}] \tag{1.2}$$

$$\sigma^2_{\delta^i,t} = \sigma^2_{\delta^i,t-1} - G_t \gamma t^\gamma \sigma^2_{\delta^i,t-1} \tag{1.3}$$

where G_t is the Kalman gain at time t given by

$$G_t = \frac{\gamma t^\gamma \sigma^2_{\delta^i,t-1}}{(\gamma t^\gamma)^2 \sigma^2_{\delta^i,t-1} + \sigma^2_e} \tag{1.4}$$

As in Guvenen (2007), to initiate the filtering process using equations (2) and (3) we must specify the initial values $\widehat{\delta}^i_0$ and $\sigma^2_{\delta^i,0}$, as these represent the information with which individuals enter the economy.

Smoothness, i.e., $\widehat{\delta}^i_{t+1} \approx \widehat{\delta}^i_t$, 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 $[s^i_t - \gamma t^\gamma \widehat{\delta}^i_{t-1}]$ 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^2_e. Therefore, if 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 $[s^i_t - \gamma t^\gamma \widehat{\delta}^i_{t-1}]$.

To illustrate this point, for the same frailty δ^i I simulate 1000 learning path realizations which are then amplified by noise variance $\sigma^2_e = 1$ (low noise), a 1000 by noise variance $\sigma^2_e = 10$ (medium noise), a 1000 by noise variance $\sigma^2_e = 100$ (high noise), and a 1000 by noise variance $\sigma^2_e = 1000$, all while keeping fixed a suitable set of the rest of parameters. In order to elicit the general smoothness due to each σ^2_e, I calculate the standard deviation of the realizations of $\widehat{\delta}^i_t$ 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 1.1.

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 $\sigma^2_e = 1$ to $\sigma^2_e = 10$, 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 1.1 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 1.1(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 1.1(b). Finally, in Figure 1.1(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.

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-cyle each δ^i is not observed by its correspondent individual. The process of learning δ^i then implies that the distribution of learned frailties $\widehat{\delta}^i_t$ 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.

1.3 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.

1.3.1 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.

1.3.2 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

$$q_t(\widehat{\delta}^i_{t-1}) = \frac{\rho_{t+1}(\widehat{\delta}^i_{t-1})}{(1+r)} + \frac{\prod_{\iota=t+1}^{t+2} \rho_\iota(\widehat{\delta}^i_{t-1})}{(1+r)^2} + \cdots + \frac{\prod_{\iota=t+1}^{T} \rho_\iota(\widehat{\delta}^i_{t-1})}{(1+r)^{T-t}} \tag{1.5}$$

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 $\widehat{\delta}^i_{t-1}$ 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 $\phi_t(\widehat{\delta}^i_{t-1}, \widehat{\delta}^i_t)$ has to take into account current and previous frailty beliefs:

$$\phi_t(\widehat{\delta}^i_{t-1}, \widehat{\delta}^i_t) = \frac{1 + q_{t+1}(\widehat{\delta}^i_t)}{q_t(\widehat{\delta}^i_{t-1})} - 1. \tag{1.6}$$

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

1.3.3 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 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

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

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

s.t.

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

where

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

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.

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, $(t_{RET-1} - t)w = (T - t_{RET})b$.

1.4 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.

1.4.1 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.

1.4.2 Simulated Method of Moments

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

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

and

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

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

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

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

where $g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$ is a 30 $\times$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 $\times$30 identity matrix. The calibrated parameters can be found in Table 1.2.

1.4.3 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 1.3 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.

1.5 Results

Figure 1.2 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 1.3 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.

Figure 1.4 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 uncapable 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 1.3, 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.

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.

1.5.1 Robustness

Figure 1.5 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 1.5(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 1.5(b), the reason being due to a delay in frailty learning. Lastly, higher annuitization also takes place when the initial frailty variance belief $\sigma^2_{\delta^i,0}$ in Figure 1.5(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 $\sigma^2_{\delta^i,0}$ would reflect) will just delay the learning of a wide idiosyncratic frailty distribution.

1.5.2 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 1.6 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.

1.6 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.

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

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

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

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

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

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

which by using the result above implies that

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

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.

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

where

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

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 5.

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 during different ages throught 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 and 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: (σ_e^2 , $\sigma_{\delta^i,0}^2$, $\widehat{\delta}_0^i$, γ,σ_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.

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 (σ_e^2 , $\sigma_{\delta^i,0}^2$, $\widehat{\delta}_0^i$, γ,σ_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.

1.7 Tables and Figures

Table 1.1: 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

Table 1.2: 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

Table 1.3: 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

Table 1.4: Bequest Parameters

Parameter	Value
Φ_1	-9.5
Φ_2	11.6
τ_b	0.10
x_b	40*$\bar{w}$

Table 1.5: 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

Table 1.6: 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 1.7: 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

Figure 1.1: Learning Simulation: Convergence to True Value by Noise

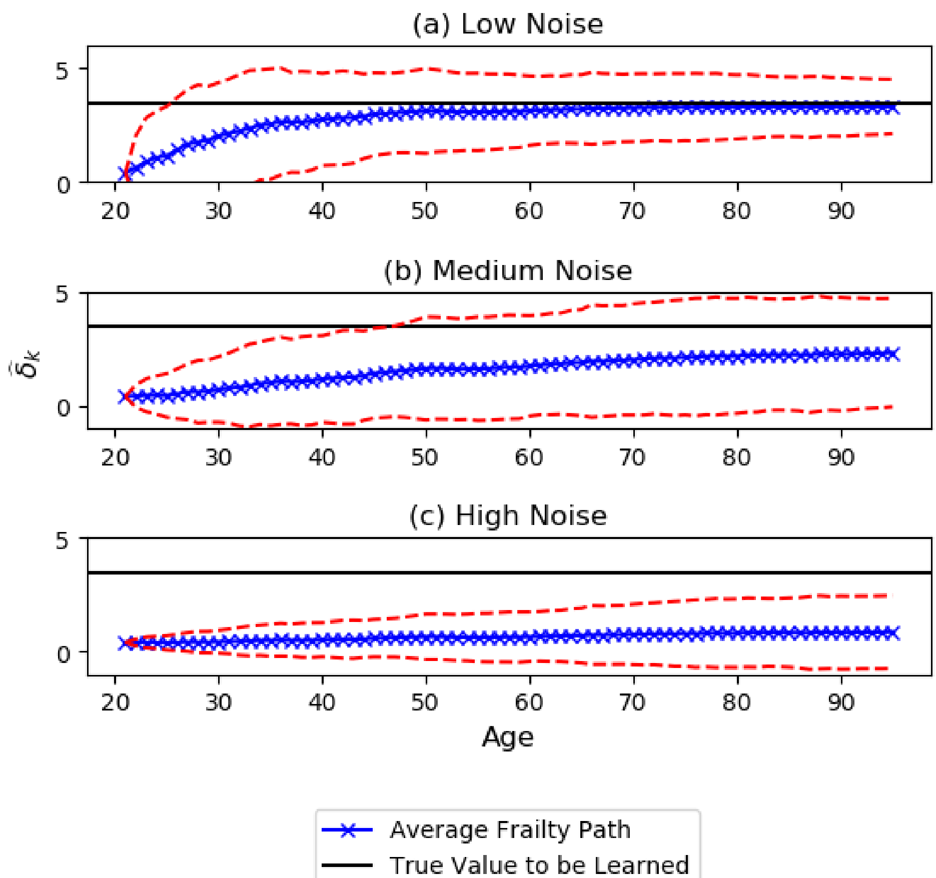
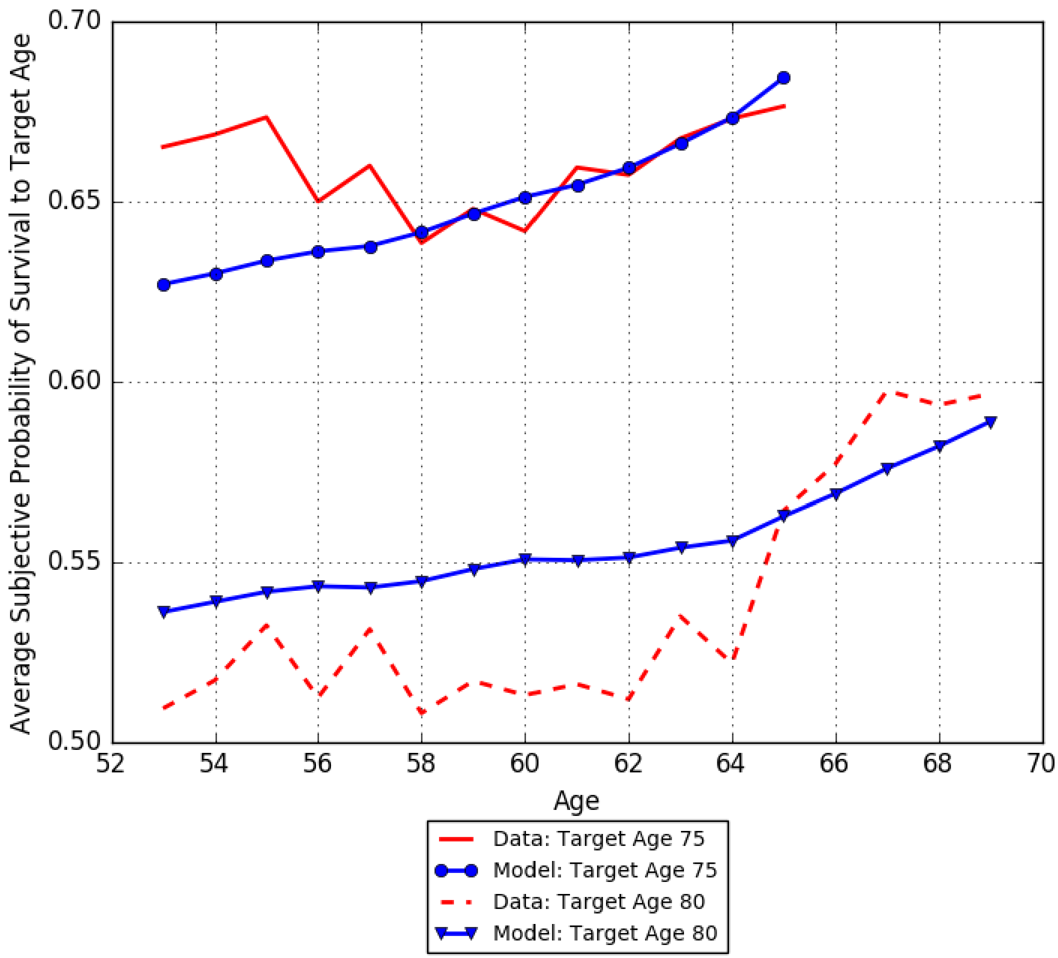


Figure 1.2: Average Subjective Beliefs, Data and Model Predicted



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

Figure 1.3: Percentage of Annuities offering a Lower Return than the Risk-Free Interest Rate

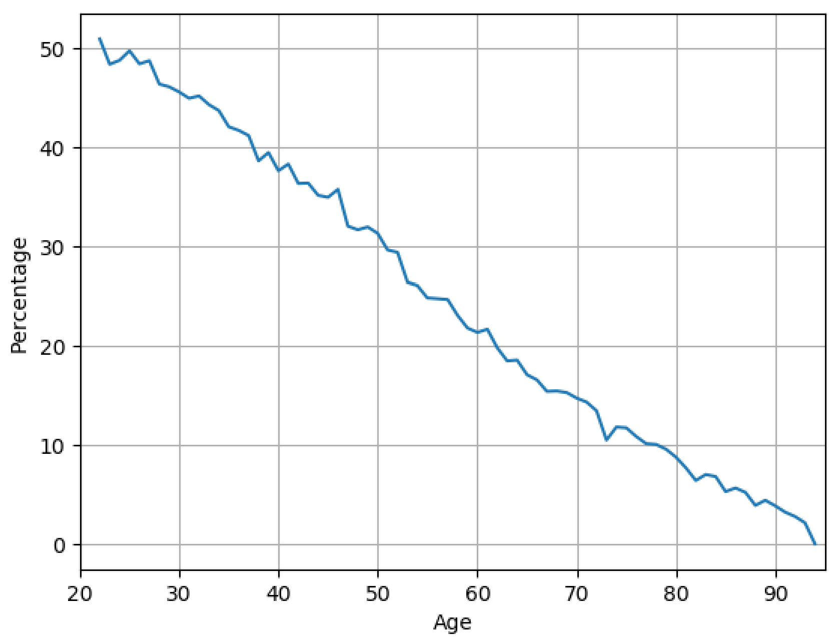


Figure 1.4: Annuitization of Wealth, Base Model

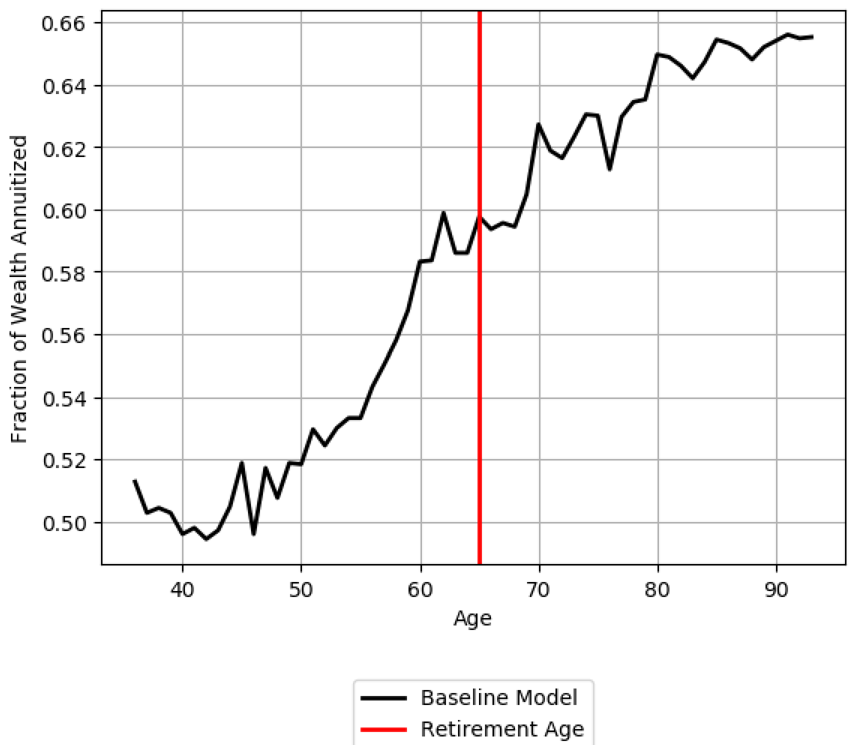


Figure 1.5: Robustness, Parameters Comparison

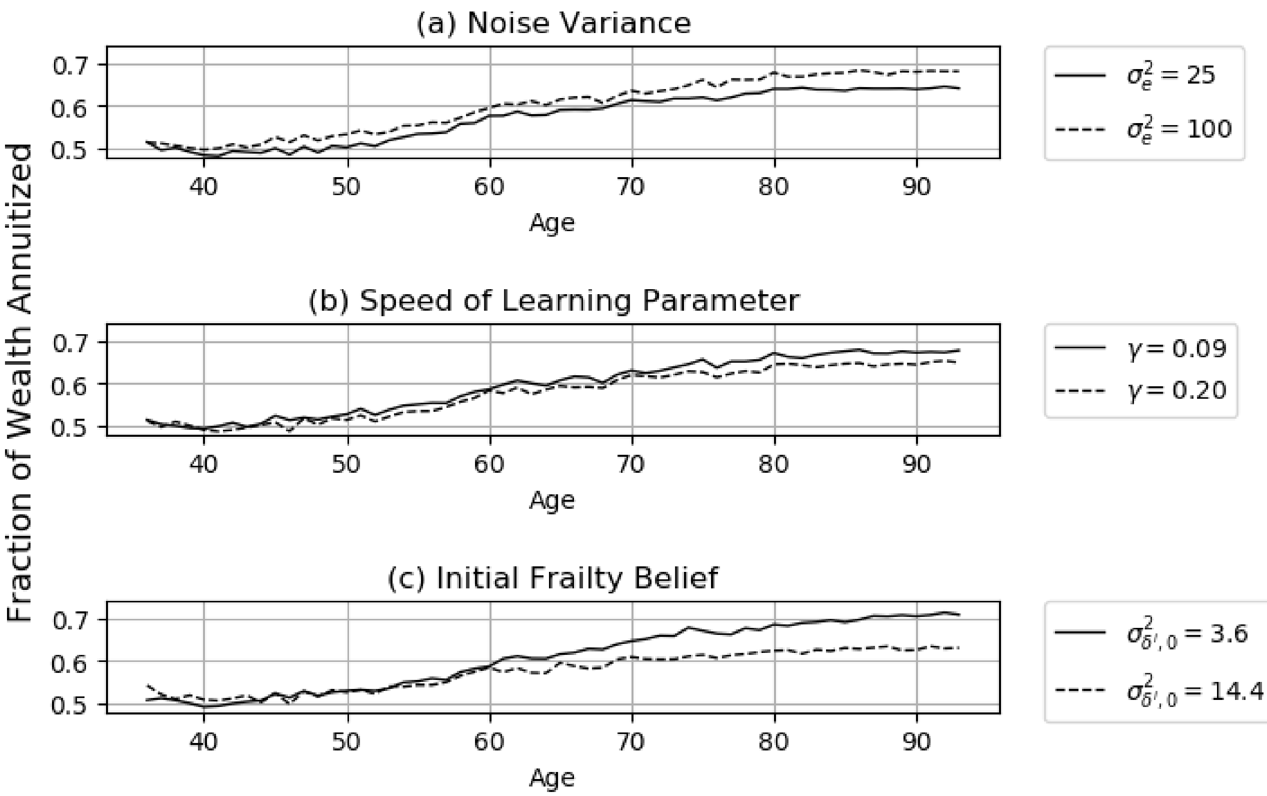
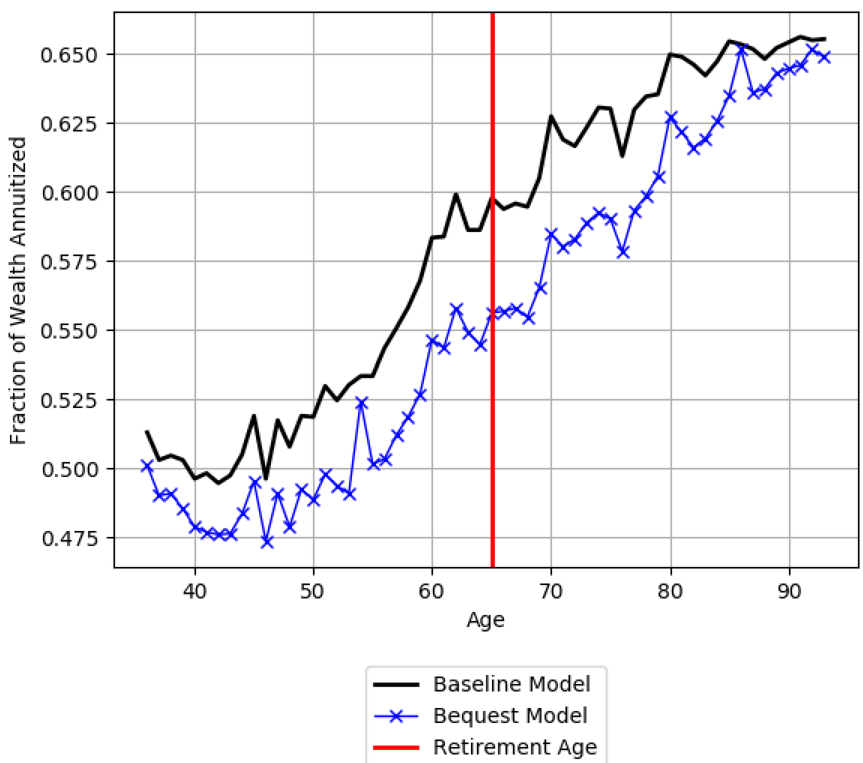


Figure 1.6: Annuitization of Wealth, Bequest Model Comparison



Chapter 2

Early Retirement and Longevity Uncertainty

2.1 Introduction

The decision to retire early and claim reduced Social Security benefits, that is, before full benefit retirement age, is an important subject of study in retirement finance literature. Even though the Social Security Administration incentivizes the delaying in claiming Social Security benefits (Knoll and Olsen, 2014), and that delaying benefits claims has largely proven to be an optimal strategy (Coile et. al., 2002), claiming benefits at the earliest possible age (62 years old) is a prevalent decision for early retirees: from 1985 to 2005 above 50% of all first time beneficiaries were 62 years old for both men and women.[1] Furthemore, in 2013 the average retirement age for men in the U.S. was 64 while for women it was 62 (as reported without decimal points by Munnell, 2015).[2] Design problems in the SSA program

[1]See Muldoon and Kopcke (2008) for more details. Additionally, the Social Security Bulletin, OASDI Monthly Statistics (1970 - 2006) shows that the percentage of men claiming before the normal retirement age has increased from 36% in 1970 to 70.5% in 2006.

[2]This study also provides detailed data on retirement profiles and labor force throughout the last decades. In particular, men's labor force participation for ages 55-64 has been steadily declining since the end of the second world war until the 1980s, and never recovered to be above 75%.

that can incentivize early claiming (Bentez-Silva and Heiland, 2007, 2008; Coile and Gruber, 2007; Munnell and Soto, 2015), and generous early retirement provisions (Dorn and Sousa-Poza, 2010; Mitchell and Fields, 1984), are among the most important factors explaining early benefits take-up in the U.S. Nevertheless, at the core of this decision, heterogeneity in life expectancies play a central role, since the reduction in benefits for early claims are actuarially fair only for a given life expectancy measure, i.e., an average measure (assuming a constant and common discount rate among individuals).[3] Heterogenous effects of single-rule policies then are important to study, as pointed out by Liebman (2002), Duggan and Soares (2002), and more recently by Li et. al. (2014).

This paper studies the effect of individuals' idiosyncratic mortality learning on early claiming of Social Security benefits after an early exit of the labor force, two combined decisions that I refer to as early retirement. Given that Social Security benefits provide an annuity service meant to hedge longevity risk, a higher level of resolved longevity uncertainty (i.e., mortality learning) will reduce the value that full-age Social Security benefits bring to individuals. Early retirement then can be optimal for individuals whose marginal utility of leisure becomes higher due to this reduction in the utility of the present value of full retirement-age Social Security benefits. In other words, as an individual is more certain about how long she will live, she is more able to plan the future use of her existing savings, diminishing the need to accumulate more wealth—like Social Security benefits—towards the end of her life; this situation in turn, allows her to enjoy more leisure during the early retirement years.

For this purpose, I model mortality learning as the structural formation of subjective survival probabilities, specifically, as a learning mechanism of objective survival probabilities. Therefore, this model assumes that individuals act upon their beliefs of what their

[3]See Heiland and Yin (2014) for a description of the actuarially fair determination of different benefits of the Social Security system.

survival chances are, rather than an econometrician's calculation of them. In this manner, this process 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. Concretely, I model the learning process about a survival distribution parameter, i.e., a frailty parameter, implying that individuals improve the knowledge of their idiosyncratic survival probability distributions as they age. To estimate the parameters of this learning process, I minimize then the distance between the model's predicted average subjective survival probability per age and its empirical counterpart in the Health and Retirement Study (HRS) data.

Following this, I build a standard heterogeneous agents life-cycle model which, embedded with these mortality learning dynamics, endogenously allows individuals to retire early, that is, to claim Social Security benefits early and exit the labor force before the normal retirement age.[4] The virtue of this model then is to account for both cross-sectional heterogeneity and across time, given that an individual's position in the frailty population distribution is unique and that all individuals refine their beliefs of their idiosyncratic frailties as they age. This scenario allows me to study how learning about idiosyncratic mortality affects aggregate retirement behavior, providing a new explanation for the empirical high proportion of beneficiaries who claim social security benefits early.

Using the HRS data (RAND file, 10th wave, 2010) to isolate the effect of mortality learning on the early retirement decision, in a baseline model I first target the fraction of new early retired individuals from the total eligible population (newly retired plus labor force), for the correspondent early retirement ages. Subsequently, I compute the same fraction in a model that does not allow individuals to learn about their idiosyncratic mortality beyond age

[4]It is of course possible to work while retired and receive benefits at the same time. The Social Security Administration will deduct \$1 in benefits for every \$3 earned above a different limit, and by 2020, this limit was set to \$48,600 per year. For tractability and simplicity purposes this scenario is not considered in the present study.

61, i.e., a no-learning model, but which is identical to the baseline model otherwise. The net effect on early retirement decisions then comes from the difference in the uncertainty about idiosyncratic survival probabilities during these crucial years in each model. Specifically, given that in the no-learning model mortality uncertainty is greater, the annuity value that full retirement-age Social Security benefits offer is greater in comparison to that of the baseline model. Therefore, we should expect this fraction of new early retirees, for the correspondent ages before the normal retirement age, to be lower than that of the baseline model; in other words, the choice to retire early—with reduced benefits—should be lower in an environment where individuals do not learn about their mortality.

The baseline results indicate that mortality learning is important for explaining early retirement: the fraction of retirees, for ages 62 to 65, falls by 37.4% in the no-learning model. Furthermore, I examine how mortality learning discourages retirement differently by wealth, and find that the wealthier an individual is the greater the effect of mortality learning on the early retirement decision will be. Lastly, I conduct an experiment where I use a simple alternative benefit reduction schedule that increases the benefit per age progressively (an increasing benefit model), while keeping constant the actuarial present value of future benefits for the average life expectancy in the baseline model. The goal of such experiment is to examine the reduction in the new retirees fraction of the population due to the reduction in the benefits for ages 62 to 65, in both the baseline and the no-learning model. I find that the net effect of this policy is greater in the baseline model, i.e., where individuals learn about their mortality, suggesting that policies aimed at discouraging early retirement should take into account mortality learning.

The fundamental difference of this paper with other past studies analyzing early retirement which also use subjective mortality data, is the interpretation given to these. I assume that the reported subjective survival probabilities in the HRS are the result of a learning process of idiosyncratic mortalities, and as a consequence, a dynamic interpretation is given to these data, by expecting older individuals to have a more accurate assessment of their survival chances compared to younger individuals. In previous works, like Sun and Webb (2011), by recovering annual survival probabilities from subjective mortality beliefs in the HRS data, it is concluded that optimal retirement strategies using subjective survival probabilities do not differ significantly from those that use survival probabilities from standard life tables. In contrast, in the present paper, the culprit for early retirement to analyze is the underlying mortality learning process of the individual, and subjective survival probabilities data are instead used to calibrate a mortality learning process that is mostly unobservable, given that we do not count with data on subjective survival beliefs for all ages, at all ages. This use of subjective beliefs data then, and most importantly, the mortality learning interpretation embedded in them, is what also marks the difference of the present paper from other works like Hurd et. al. (2004) and Gan et. al (2005).

In what follows of the paper, section 2 describes the mortality model and the heterogenous life cycle model used, section 3 describes the empirical strategy, section 4 explains the results of simulations and experiments, and section 5 concludes.

2.2 Model

In this section I develop a model in which individuals slowly learn about their true survival probabilities, which in turn are determined by their own—to be learned—true frailty. As in Poppe-Yanez (2018), the model follows closely the one developed by Guvenen (2007) concerning learning dynamics about individuals' characteristics.

2.2.1 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 $\hat{\delta}^i_k$. I refer to $\hat{\delta}^i_k$ as idiosyncratic subjective frailty since it will allow me to define later an idiosyncratic *subjective* survival probability $P_t(\hat{\delta}^i_k)$. Additionally, I assume individuals are aware of the heterogeneity across the population with the existence of a set of individual frailty types $\Delta = [\underline{\delta}, \overline{\delta}]$, which allows them to know there is a well defined cumulative distribution $F \in \Gamma(\Delta)$.

Suppose $\rho_t(\hat{\delta}^i_k)$ 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 $\hat{\delta}^i_k$. We have then $P_t(\hat{\delta}^i_k) = \prod_{\imath=0}^{t} \rho_\imath(\hat{\delta}^i_k)$. 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 $P_t(\hat{\delta}^i_k)$, 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

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

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

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

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

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

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

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 $P_t(\hat{\delta}^i)$ 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 survival probabilities will come from the aggregate probabilities of a Social Security Administration life table.

2.2.2 Mortality Learning

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 m_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 $e_t^i \sim N(0, \sigma_e^2)$:

$$m_t^i = d_t(\delta^i) + e_t^i \tag{2.1}$$

This framework—plus the specification of $d_t(\delta^i)$ 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 observation m_t, she still has imperfect information about her frailty δ^i.

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 $\sigma^2_{\delta^i}$,

$$\widehat{\delta}^i_t = \widehat{\delta}^i_{t-1} + G_t[m^i_t - \gamma t^{\gamma} \widehat{\delta}^i_{t-1}] \tag{2.2}$$

$$\sigma^2_{\delta^i,t} = \sigma^2_{\delta^i,t-1} - G_t \gamma t^{\gamma} \sigma^2_{\delta^i,t-1} \tag{2.3}$$

where G_t is the Kalman gain at time t given by

$$G_t = \frac{\gamma t^{\gamma} \sigma^2_{\delta^i,t-1}}{(\gamma t^{\gamma})^2 \sigma^2_{\delta^i,t-1} + \sigma^2_e}$$

As in Guvenen (2007), to initiate the filtering process using equations (2) and (3) we must specify the initial values $\widehat{\delta}^i_0$ and $\sigma^2_{\delta^i,0}$, as these represent the information with which individuals enter the economy.

At this point it is 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-cyle 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.

2.2.3 Recursive Formulation

In this section, early retirement means to collect Social Security benefits (according to a pre-determined schedule), and exit the labor force completely (no part-time working available), as explained above. Also, by continuing working, an individual cannot claim any benefits. For the early retirement decision the individual has to take into account the different benefit amounts and the current resolved longevity uncertainty she has.

I formulate the life-cycle model with endogenous early retirement as a dynamic programming problem an individual faces. In this context, the model below is for each individual i, and although all the variables are defined at the individual level, from here on only frailty belief $\widehat{\delta}^i$ will be indexed by i. Additionally, rather than formulating a model rich in retirement incentives, as in for example Benitez et. al. (2009), for ages 62 to 65 I let early retirement be affected by the relative importance of leisure (with respect to consumption), which is exogenously determined. That is, different incentives are embedded when the individual decides to stop working and increase leisure. In this manner, the objective of the model is to isolate the effect of mortality learning, through frailty $\widehat{\delta}_t^i$, on early retirement. I also assume that it is not possible to delay retirement past age 66, nor is possible to retire before age 62.

With forward-looking individuals then, the analysis is focused on these four years starting at age 62. At any given period $62 \leq t \leq T$, an individual observes the realization of the noisy signal m_t and forms a belief of her frailty $\widehat{\delta}_t^i$. With this information, the individual decides between consumption c_t, savings k_{t+1}, and for ages 62 to 65, in a dichotomous form, whether to participate in the labor force, having leisure $l = l^w$, or to retire early, having leisure $l = l^r$, such that $l^r > l^w$. Retirement in this environment is an absorbing decision, that is, individuals cannot return to the labor market once they left it.[6] For every age $t < T$ an individual receives a flow utility from consumption, and furthermore, the forward-looking utility at period t the individual subjectively believes she has is

$$U_t = \sum_{s=t}^{T} \beta^s P_s(\hat{\delta}_t^i) u_s(c_s, l_s) \tag{2.4}$$

where β is a discount factor, and

$$u_t(c, l) = \frac{\left(c^{\eta_t} l^{1-\eta_t}\right)^{1-\varsigma}}{1-\varsigma}$$

where ς is the coefficient of relative risk aversion, and η_t is the relative importance of consumption with respect to leisure. Furthermore, when an individual dies, her wealth is taxed at the 100 percent, and there is no accidental bequest redistribution to survivors.

The value function for this decision is

$$\begin{aligned}
V_t^o(k_t, \widehat{\delta}_t^i, m_t, t_0) \quad &= \quad \max_{\{r,w\}} \left\{ V_t^r(k_t, \widehat{\delta}_t^i, m_t, t_0), V_t^w(k_t, \widehat{\delta}_t^i, m_t) \right\} \\
for \quad & \quad V_{t-1}^o = V_{t-1}^w \\
and \quad & \quad 62 \le t \le 65.
\end{aligned}$$

where V_t^r is the value function corresponding to the decision of retiring, and V_t^w is the value function of the decision to continuing participating in the labor force.

Depending on the age chosen to collect Social Security benefits $y_c(t_0)$ and retire, the

[6]In principle this model could be extended to study delayed retirement as well, yet for the present computational constraints I focus only on early retirement.

recursive problem when the individual has decided to retire is

$$V_t^r(k_t, \widehat{\delta}_t^i, m_t, t_0) = \max_{c_t, k_{t+1}} \left\{ u_t(c_t, l_t = l^r) + \beta\rho_{t+1}(\widehat{\delta}_t^i) E\left[V_{t+1}^r(k_{t+1}, \widehat{\delta}_{t+1}^i, m_{t+1}, t_0) \mid \widehat{\delta}_t^i \right] \right\}$$

$$k_{t+1} + c_t \quad = \quad (1+r)k_t + y_c(t_0)$$

$$\begin{aligned} \widehat{\delta}_t^i &= \widehat{\delta}_{t-1}^i + G_t[m_t - (\gamma t^\gamma)\widehat{\delta}_{t-1}^i] \\ \sigma_{\delta^i,t}^2 &= \sigma_{\delta^i,t-1}^2 - G_t(\gamma t^\gamma)\sigma_{\delta^i,t-1}^2 \\ k_{t+1} &\geq 0 \end{aligned}$$

where

$$\begin{aligned} G_t &= \frac{(\gamma t^\gamma)\sigma_{\delta^i,t-1}^2}{(\gamma t^\gamma)^2\sigma_{\delta^i,t-1}^2 + \sigma_e^2} \\ m_t &= (\gamma t^\gamma)\delta^i + e_t^i \\ y_c(t_0) &= \begin{cases} s \times [1 - 0.067 * (66 - t_0)] & for\ 63 \leq t_0 \leq 66 \\ s \times 0.75 & for\ t_0 = 62 \end{cases} \end{aligned}$$

where Social Security benefits are a function $y_c(t_0)$ of the age t_0 at which the individual starts collecting these benefits, and where s is the normal retirement age Social Security (full) benefit that can only be collected without penalties at age 66. This benefit function expresses the penalization the Social Security program imposes on early retirement for individuals born in the period between 1943-1954: a 6.7% decrease if the individual decides to retire at age 64, 13.3% decrease if the individual decides to retire at age 63, and so on.

Note how the expectation in the value function V_t^r—and in the value function V_t^w described below—is taken over the possible next frailty beliefs. Therefore the randomness of this model is correctly interpreted as randomness about (knowledge of) the survival distribution itself, given that, from a subjective point of view, it can be different next period thanks to the learning process. This feature makes the model structural in the formation of subjective survival probabilities, since a reduced form model would not account for how subjective beliefs are formed.

On the other hand, when the individual chooses to continue working, she continues receiving wage w. Recursively,

$$V_t^w(k_t, \widehat{\delta}_t^i, m_t) = \max_{c_t, k_{t+1}} \eta_t \left\{ u_t(c_t, l_t = l^w) + \beta \rho_{t+1}(\widehat{\delta}_t^i) E\left[V_{t+1}^o(k_{t+1}, \widehat{\delta}_{t+1}^i, m_{t+1}) \mid \widehat{\delta}_t^i \right] \right\}$$

$$\begin{aligned} k_{t+1} + c_t &= (1+r)k_t + w \\ \widehat{\delta}_t^i &= \widehat{\delta}_{t-1}^i + G_t[m_t - (\gamma t^\gamma)\widehat{\delta}_{t-1}^i] \\ \sigma^2_{\delta^i,t} &= \sigma^2_{\delta^i,t-1} - G_t(\gamma t^\gamma)\sigma^2_{\delta^i,t-1} \\ k_{t+1} &\geq 0 \end{aligned}$$

where

$$\begin{aligned} G_t &= \frac{(\gamma t^\gamma)\sigma^2_{\delta^i,t-1}}{(\gamma t^\gamma)^2\sigma^2_{\delta^i,t-1} + \sigma_e^2} \\ m_t &= (\gamma t^\gamma)\delta^i + e_t^i \end{aligned}.$$

This recursive model then can map uncertainty about mortality, through the subjective survival probability of the next period $\rho_{t+1}(\widehat{\delta}_t^i)$, into the decision to claim Social Security benefits or not through the value function $V_t^o(k_t, \widehat{\delta}_t^i, m_t)$. To conclude, the individual must be retired at age 66 if she has not done so already.

For each period $62 \leq t \leq 65$, I compute the fraction of new retired individuals; these are the parameters I target for the empirical strategy explained in the next section.

2.3 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 life-cycle model with endogenous retirement.

2.3.1 Data

Data for Mortality Learning

I use the subjective survival chances responses of the Health and Retirement Study survey years 2000 to 2010. 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?"

I use this wide sample in an effort to maximize the number of respondents, with different ages, for the same target ages. These questions about survivorship have a larger sample for the target ages of 75 and 80. 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. For more details of the sample size see Table 2.1.

Data for Retirement

For the sample of retirees and individuals in the labor force for ages before normal retirement age, I only use the 10th wave of the Health and Retirement Study data (2010), and I filter for cohorts of individuals born in the period between 1943-1954, that is, individuals who have a normal retirement age of 66. This HRS wave is ideal because these cohorts that share the same normal retirement age, and which are able to retire in 2010 from ages 62 to 66, were interviewed in the same survey.

In this HRS survey, individuals are considered part of the labor force if they are currently working, unemployed and looking for work, temporarily laid off, on sick period, or on other leave. The HRS questionnaire conveniently distinguishes individuals who dropped out of the labor force due to disability, so I exclude them for any calculation.

For ages 62 to 65, weighted respondent-level answers were used for the calculation of the fraction of (non-disability) retirees. As explained in the previous section, in the model I calculate new retirements per age, and given that respondents who decided to retire previous to the survey would also be included in as currently being retired, an implied measure must be calculated. The results can be seen in Table 2.2.

Consistent with previous literature that finds peak retirement frequencies at extreme ages, we can see that the peaks of retirement here are at the lowest age and one year before the normal retirement age.

2.3.2 Simulated Method of Moments for the Mortality Learning Process

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

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

and

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

we have 13 first moments for $\rho_{75}(\hat{\delta}_k^i)$ and 17 first moments for $\rho_{80}(\hat{\delta}_k^i)$, 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

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

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

where $g(\sigma_e^2, \sigma_{\delta^i,0}^2, \widehat{\delta}_0^i, \gamma, \sigma_A^2)$ is a 30 $\times$1 matrix measuring the distance between the sample moments and the model moments concerning $\rho_{75}(\hat{\delta}_k^i)$ and $\rho_{80}(\hat{\delta}_k^i)$. Since the objective of the simulated method of moments is that of calibration, W is simply set to be a 30 $\times$30 identity matrix; the choice of W would matter for estimating the variance of the parameters, which is not what is pursued in this study. The calibrated parameters can be found in Table 2.3.

2.3.3 Life-Cycle Model Calibration

As mentioned above, targeting the fraction of new early retirees reported in the HRS data, as shown in Table 2.2, I calibrate the relative importance of leisure η_t for ages 62 to 65, along with the leisure when participating in the labor force and when retired, l^w and l^r respectively. That is, since we are interested in the effect of the mortality learning process on the early retirement decision, the rest of the parameters are chosen so the model—embedded with this exogenous process of mortality learning—can reproduce the observed probabilities of early retirement. Consequently, the rest of the parameterization, along with the relationship between l^w and l^r, just serve a scaling purpose. Following 2016's Merryl Lynch retirement study, the latter are scaled to be $l^w = 0.5l^r$. Using data from the Bureau of Labor Statistics, this study estimated that in 2015 retirees 65 years old and older enjoyed, collectively, 126 billions hours of leisure time, while this same measure for the working population between ages 25 and 64 was 71.2 billions hours—approximately half the total leisure hours of the retired population.[7]

The individual's wage w is set to be the probability of working given that an individual is participating in the labor force during retirement ages p^w, times average wage $\bar{w}$,

$$w = p^w \times \bar{w}.$$

Using the 10th wave of HRS data I calculate $p^w = 0.91$. Following Hosseini (2015), the average wage $\bar{w}$ and the Social Security benefit s are in turn calibrated to target a replacement ratio of 0.45.

[7] Merrill Lynch, (2016), "Leisure in Retirement: Beyond the Bucket List A Merrill Lynch Retirement Study conducted in partnership with Age Wave". Available at https://agewave.com/wp-content/uploads/2016/05/2016-Leisure-in-Retirement_Beyond-the-Bucket-List.pdf

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,

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

where P_t^{1950} denotes the surviving probability to age t for a cohort born in 1950 (see Figure 2.1). This profile of survival probabilities is chosen because of its relevancy for the HRS surveys made between the years 2000 through 2010, a timeframe for which a 1950 cohort would be between 50 to 60 years old. Also, it is important to note that as 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 plausible to assume that all individuals are aware of this aggregate measure.

Table 2.4 summarizes the rest of the parameters being calibrated for the dynamic programming computation. All these calibrated parameters, jointly with the SMM estimated values of section 3.2, determine the baseline model to be used.

Solution Algorithm

To compute the solution we first solve the learning model to use the resulting subjective frailty beliefs later on in the life-cycle model. The steps to take to this end are the following:

1. Using a tentative set of parameters of the learning model, simulate the learning path of $\mathbb{Q}$ possible frailties, using equations (2) and (3), for $\mathbb{N}$ individuals.

2. Repeat step 1 adjusting the parameters until these minimize the Simulated Method of Moments criteria, as explained in section 3.2. The implied distribution of subjective frailty beliefs determines the region of the state space in the life-cycle model that will be visited recursively.

3. With each frailty belief elicited per age, solve the dynamic programming model detailed in section 2.3 with an appropriate asset grid. For the early retirement ages, we can now keep track of the fraction of the $\mathbb{N}$ individuals who choose to retire.

2.4 Results

Assessing the estimation of the mortality learning process, Figure 2.2 shows the 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.

To analyze the impact of learning during ages 62 to 65, I calculate the fraction of retired individuals in the baseline model, calibrated to match the net retirement fraction found in the HRS data, and in a no-learning model, where learning stops at age 61. In the latter model then, all subjective probabilities used at ages $t > 61$ are fomed at age $t = 61$. That is,

$$P_t(\hat{\delta^i_{61}})\ for\ t > 61.$$

Without mortality learning during these years, but keeping the rest of the model parameters the same as the baseline's, we can observe the difference in retirement decisions. The results are shown in Table 2.5. On average, during these four years, the new retirees fraction drops from 0.195 in the baseline model to 0.12 in the learning model, implying a 37.4% reduction in the fraction of retirees. The differences per year also show that the peaks are preserved at extreme ages, and that overall the shape of a retirement curve is preserved.

In Figure 2.3 we can observe the effect of mortality learning per top wealth individuals. Wealthier individuals tend to retire more than the rest of the individuals, but are also more discouraged to retire if they stop learning about mortality during these ages, that is, the differential in the fraction of retirees per wealth classification is greater for wealthier individuals. The impact of mortality learning is greater for wealthier individuals given that they can afford to disregard their income more than the rest of the individuals. In both models, poor individuals cannot afford to retire.

2.4.1 Robustness

Table 2.6 shows variations above and below of two key parameters of the baseline model. A double noise variance implies the individual learns about her actual frailty rather slowly, and therefore, the individual finds the need to hedge more against longevity risk compared to the baseline model. As a result, she would like to increase the annuity value Social Security benefits offer, for which she will be less likely to retire. With the same logic, if the learning noise variance is half the original value, the individual will be more likely to retire. Similarly, the individual values more the annuity value of the benefits if the initial frailty variance is half the actual frailty variance, making her less likely to retire than in the baseline model. 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 will just delay the learning of a wide idiosyncratic frailty distribution.

2.4.2 Policy Experiment

I conduct a policy experiment that, even though it is objectively extreme and unrealistic, it serves as an illustration of how altering the benefit payment schedule could discourage early retirement. Taking as reference the benefit reduction for the standard individual $P(\delta^{std} = 1)$, I design a new schedule of Social Security payments that are in comparison reduced for the early years, between ages 62 and 65, but increased for the latter, in this case the last six used in the life cycle model, between ages 90 and 95. The goal of this experiment is to see the effect of a policy discouraging early retirement taking into account the process of mortality learning—all while keeping the same actuarially present value of the benefits for the standard individual, that is, the standard individual will not be discouraged to behave differently. This extreme policy experiment then serves to highlight the one-rule policy effects of the Social Security administration when individuals have heterogeneous discount factors.

There are two developments that take place after this policy is implemented in a mortality learning environment. First, as a direct effect of this policy, the reduction of benefits for the first years will discourage early retirement for some individuals who were not discounting the future enough to be discouraged of early retirement before the policy implementation. After the policy is implemented, the present value of the stream of benefits for these individuals falls, and as a consequence, they prefer to continue working. On the other hand, a second situation takes place due to the mortality learning dynamics. Some individuals resolve some of their longevity uncertainty enough that they would have been encouraged to retire early before the policy implementation, but now, with the new benefit schedule, they too prefer to continue working. Therefore, the difference in the new retirees fraction will always be greater in a learning environment than in a no-learning environment.

Figure 2.4 shows the net effect of this policy on the fraction of new retirees in both the baseline model and in the no-learning model. Although the fraction differentials in both models range from 0.01 to 0.04, we can see that such a simple policy experiment would have a more prevalent effect in a learning environment. We can conclude then that policies aimed at discouraging early retirement should take into account mortality learning when formulating these objectives.

2.5 Conclusions

This paper shows how the process of mortality learning influences the decision to retire early or not. In this sense, mortality beliefs, as a dynamic product of mortality learning, are of central importance given that individuals act on what they believe is the nature of their mortality, and therefore, any decision concerning annuity benefits must take into account these mortality beliefs. Early retirement, understood as early claiming of Social Security benefits after an early exit of the labor force, is an annuity decision of this nature given that by claiming these benefits early the annuity value they offer is lower.

To this end, first I calibrate the parameters of a mortality learning process using subjective survival probabilities data from the Health and Retirement Study. In this environment, 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 between ages 62 and 65, must decide recursively to participate in the labor market, and therefore delay the claiming of Social Security benefits, or retire early with reduced benefits, and implicitly reduce the annuity value these offer. The results show that mortality learning is important for the decision of early retirement: if individuals were not learning about their mortality, early retirement would be 37.4% lower. Furthermore, the effect of mortality learning is robust to important learning parameters variations. I also find that this result is important for policies aiming at discouraging early retirement, given that more individuals will be discouraged to retire in a learning environment than in a no-learning environment, that is, the discouragement of policies will also affect individuals who initially were not planning to retire early, but who eventually would have decided to if there was not going to be a policy implementation.

Appendix (Gender Comparison)

The exercise of this paper was also performed for men and women separately. The results did not throw different results per gender due to the similarities in the shapes of of the learning frailty paths for both genders, which on their own conform with the aggregate measure (all individuals) used in the paper. Figure 2.5 and Figure 2.6 show the similar average subjective survival probabilities to target ages 75 and 80, respectively, by gender. Because the shape of this subjective survival probabilities in the life-cycle are similar on the three measures (men, women, all), the generation of frailty learning path is the same for each measure. That is, the simulated method of moments calibration throws the same learning frailty parameters for men, women, and both combined. Although women seem to have consistently a higher level of subjective survival probability for both target ages, this is not enough to alter the frailty learning parameters because the learning mechanism is more concerned with the evolution of beliefs. In terms of the learning mechanism, if there was a difference in the evolution of beliefs for men and women, then the distance between the average subjective survival beliefs for both target ages (75 and 80) should be significantly different for each gender at any age.

2.6 Tables and Figures

Table 2.1: 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

Table 2.2: Fraction of Retired Individuals, New and Total (HRS)

Age	Total Retired	Implied New Retired
62	0.27	0.27
63	0.40	0.17
64	0.44	0.07
65	0.60	0.27

Table 2.3: 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

Table 2.4: Life-Cycle Parameters

Parameter	Description	Value
T	Age of Certain Death	95
t_R	Retirement Age	65
τ	Social Security Tax	0.3
ς	Coefficient of Relative Risk Aversion	3
r	Annual Risk-free Interest Rate	0.04
β	Discount Factor	0.99
l^w	Leisure when Working	50
l^r	Leisure when Retired	100
η_{62}	Relative Importance of Leisure, Age 62	0.54
η_{63}	Relative Importance of Leisure, Age 63	0.49
η_{64}	Relative Importance of Leisure, Age 64	0.43
η_{65}	Relative Importance of Leisure, Age 65	0.44

Table 2.5: Fraction of Retired Individuals per Model, Mortality Learning Effect

Age	Baseline Model (Target)	Non-learning Model	Percentage Change
62	0.271	0.202	-25.1%
63	0.169	0.089	-47%
64	0.069	0.007	-89.2%
65	0.275	0.189	-30.1%
Average	0.195	0.12	-37.4%

Table 2.6: Fraction of Retired Individuals, Robustness

Age	Baseline	Noise Variance		Initial Variance Belief	
		Double	Half	Double	Half
62	0.27	0.14	0.39	0.42	0.11
63	0.17	0.02	0.31	0.34	0.001
64	0.07	0.001	0.24	0.28	0
65	0.27	0.13	0.41	0.44	0.09

Figure 2.1: Life Table Survival Probabilities Profiles, 1950 Cohort

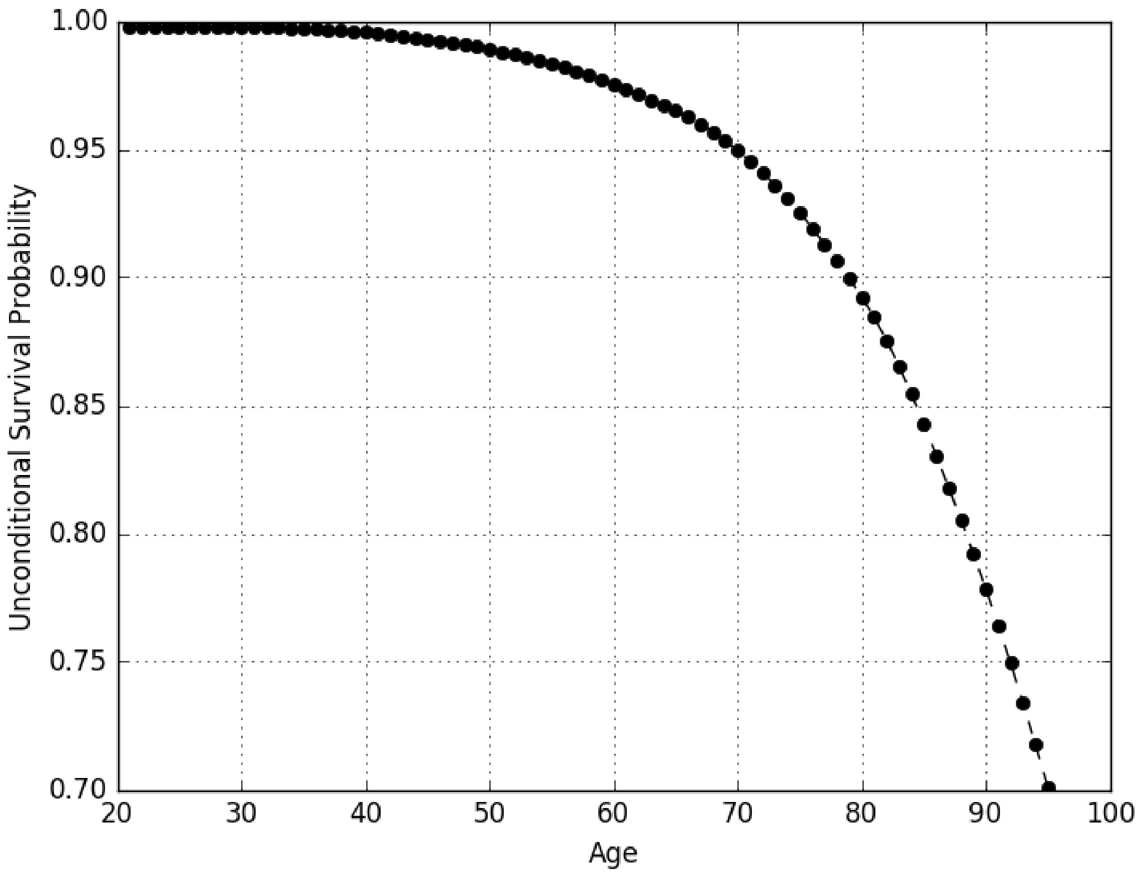
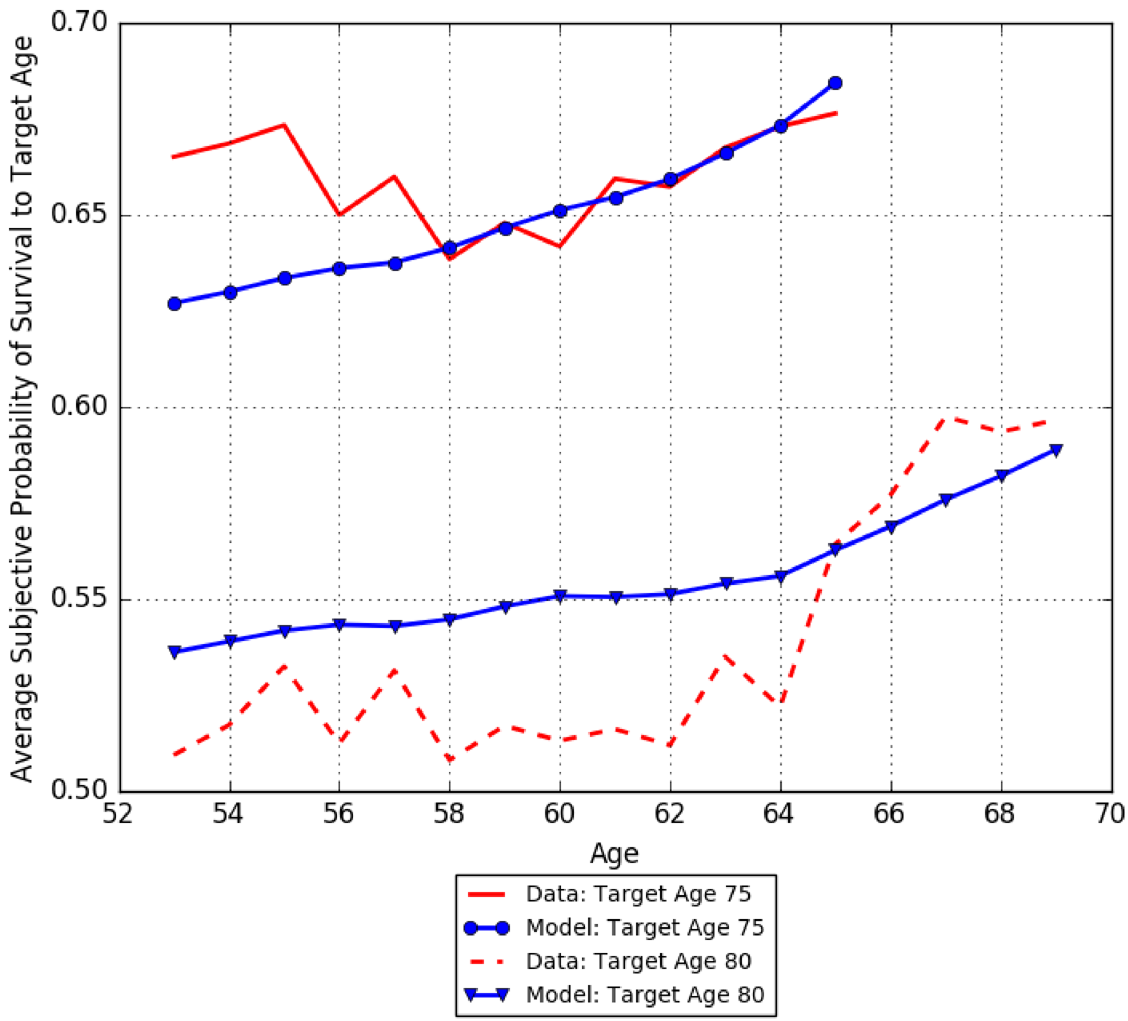


Figure 2.2: Average Subjective Beliefs, Data and Model Predicted



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

Figure 2.3: The Effect of Learning Ages 62-65: Retirees Fraction Ratio by Wealth Distribution

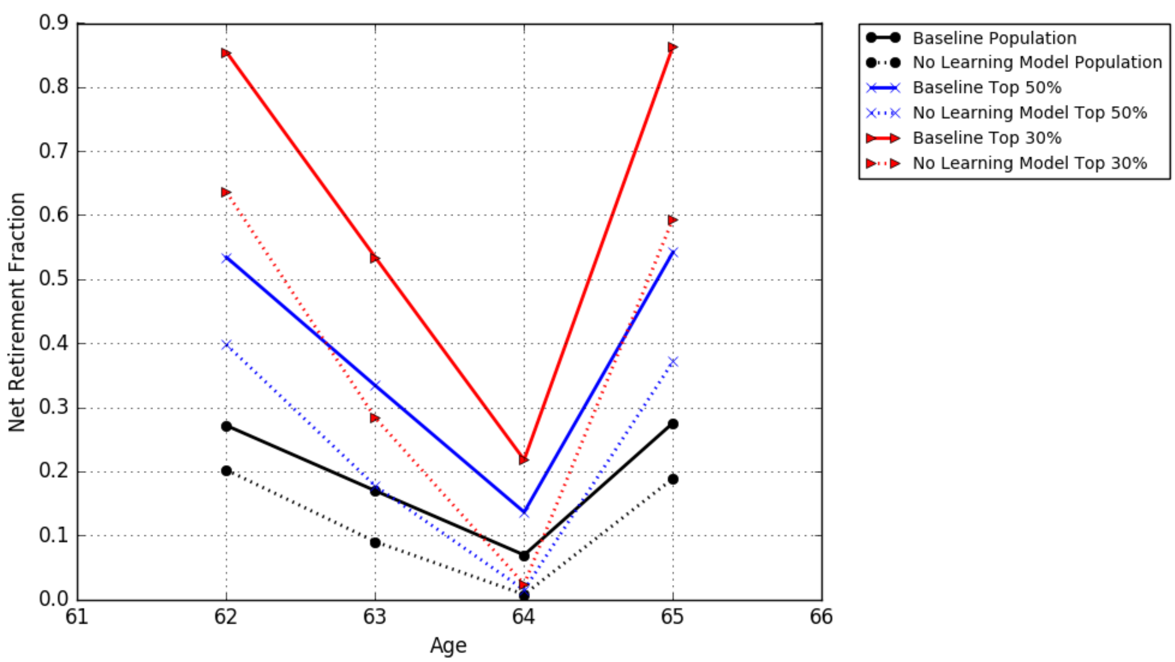


Figure 2.4: The Effect of an Alternative Benefit Schedule per Environment

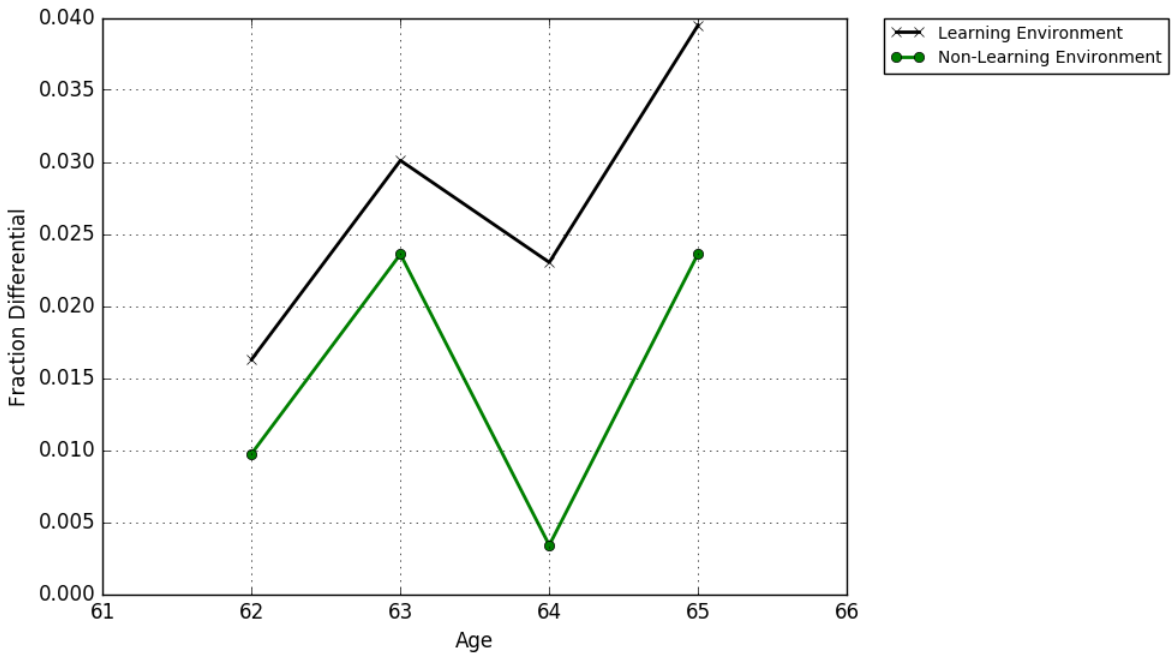


Figure 2.5: Average Subjective Survival Belief to Age 75, per Gender

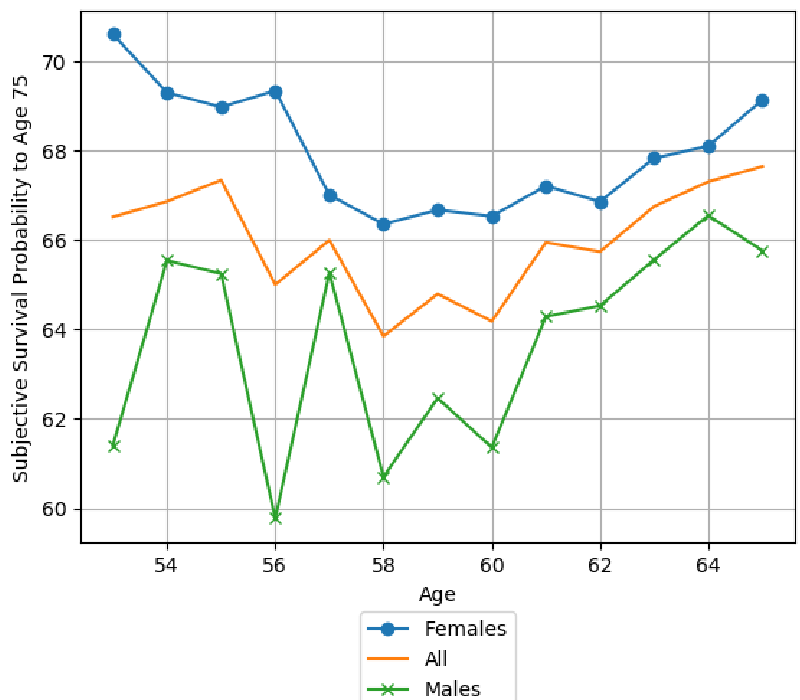
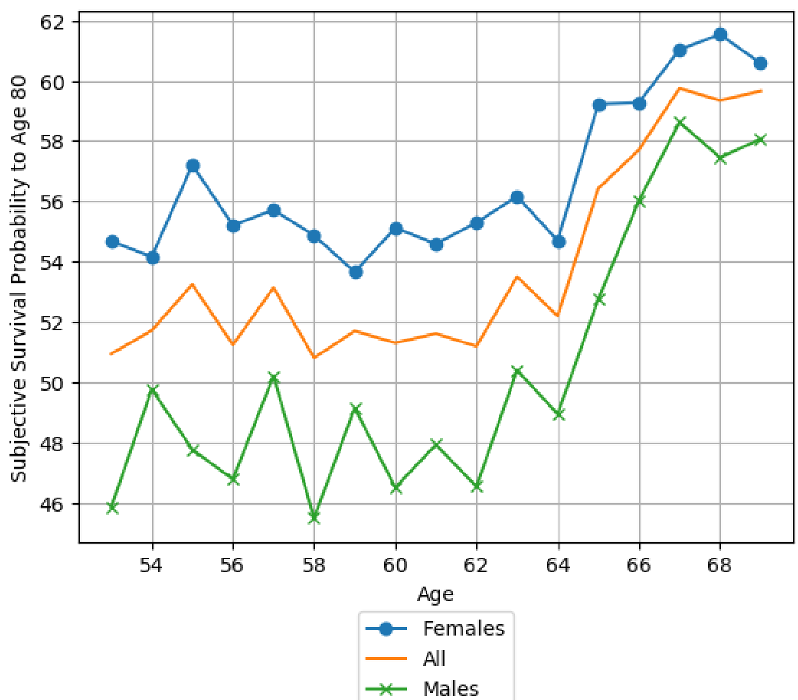


Figure 2.6: Average Subjective Survival Belief to Age 80, per Gender



Chapter 3

Literature Review: Why is the Demand for Annuities so low?

3.1 Introduction

The annuity puzzle in the life-cycle and retirement literature refers to the discrepancy between the theoretical prediction of a high demand for annuities, especially by retirees and individuals near the end of life, and the empirical low holdings of these assets. For over fifty years this puzzle was examined in its variants, and most of the economic literature produced has documented its robustness across many theoretical scenarios. It was only until recent new developments that an agreement of a theoretical plausible reason for low annuitization was found.

In this paper I review the essential theoretical literature studying the annuity puzzle, from its early understanding to the latest developments in the life-cycle literature addressing it.

3.2 Annuities Dominance

Broadly defined, an annuity is designed to make payments to the owner of this asset for the remaining years of her life. The party in charge of making the payments, i.e., an insurance company, sells to the potential owner an annuity contract priced at the present value of these future payments. In this manner, by their nature, annuities play an important role in life-cycle savings and consumption behavior, given the possibility to use annuities as a savings vehicle by an individual—a process usually referred to as annuitization. We can consider this income certainty the main benefit of an annuity: the individual's risk of not having income during the last years of life, i.e., the longevity risk, is completely hedged. On the other hand, the cost of this agreement for an individual would be the loss of all wealth after death; there would be no possibility of bequeathing.

In 1965, Yaari showed the theoretical result that rational individuals should annuitize their entire wealth, and even though this conclusion was reached using a standard life-cycle model where bequests were not considered, this analysis clearly showed how uncertainty about lifespans should be translated into full annuitization. To understand this result intuitively we must go back Modigliani's (1966) life-cycle hypothesis, which demonstrated that lifetime utility is maximized when consumption is smoothed over the life-cycle. In other words, considering a very basic scenario where a lifespan is known with certainty, an optimal saving behavior is one that allows for the same consumption during retirement and working years—mainly because it is assumed there is no income during retirement years. Annuities would be the optimal tool in this setting when we assume instead that lifespans are uncertain (which for modeling purposes, is translated into assigning a survival probability to each year of life).

If individuals do not know in advance the age of death then they do, in fact, face a longevity risk, which in this case is tantamount to the risk of outliving their own savings.

As mentioned above, it is straightforward to see why annuities are traditionally thought to be able to hedge this longevity risk, as they provide payments until death regardless of the age at which it occurs. The main finding by Yaari (1965) is that, in a basic scenario with no bequests, fairly priced annuities state-wise dominate risk-free bonds if an individual has to choose between these two assets during the life-cycle. To see why, consider an annuity that pays 1$ in perpetuity, starting the period after it was bought. The actuarially fair price q_t of this annuity at age t is

$$q_t = \frac{\rho_{t+1}}{(1+r)} + \frac{\prod_{\imath=t+1}^{t+2} \rho_\imath}{(1+r)^2} + \cdots + \frac{\prod_{\imath=t+1}^{T} \rho_\imath}{(1+r)^T},$$

where ρ_t is the conditional survival probability to age t, and r is the risk-free interest rate bonds provide. Suppose that this annuity can be reversed in a recursive fashion, i.e., the individual can sell it back every period based on the same pricing it has been bought (actuarially fair). The one-period gross return would then be,

$$1 + \phi = \frac{1 + q_{t+1}}{q_t}, \tag{3.1}$$

where ϕ is the net annuity return, we can express the current price as a function of the future price using the actuarial formulation,

$$q_t = \frac{\rho_{t+1}}{(1+r)} + \frac{\rho_{t+1}}{(1+r)} \left[q_{t+1} \right],$$

which implies

$$1 + \phi = \frac{(1+r)}{\rho_{t+1}},$$

showing the gross annuity return dominance since $\rho_{t+1} < 1$.

Effectively, this simple result had profound implications for retirement finance research, even though it does not inform us much about the interaction between risk aversion, consumption smoothing, and the demand for annuities. It only informs us that, due to the higher return of annuities, it is preferable to invest an individual's wealth in annuities than in risk-free bonds.

Regarding other simplifying assumptions, in the standard life-cycle model used by Yaari the market for annuities is complete (meaning that to sell them, reverse them, or the repricing of annuities by individuals is possible), utility of consumption is inter-temporally separable (Von Neumann-Morgenstern utility form), and there is absence of a planner's distributive tool (a Social Security program). Yet the most important assumption of this basic model is the lack of any kind of bequeathing behavior from individuals. In this regard, it is easy to see why a bequest motive would deter annuitization, at least in a theoretical sense, given that this is the only situation where an individual has a purpose for the accumulated wealth after death. Moreover, it is also clear to observe how the assumption of a fairly priced annuity implies that this is a competitive market, such that the insurer pools all individual longevity risks, and therefore those who die earlier subsidize those who die last. This subsidy is also known as "mortality credit".

It is important to note that for much of the subsequent literature a Social Security program was included in the models studying annuitization. The interpretation is that, even in the presence of this annuity-like benefit, the entire (remaining) wealth should be privately annuitized, because the dominance of the annuities return is for the entire wealth, and not just by a fraction that would be spent on what social benefits cover (Social Security benefits, Medicare).

3.2.1 Essential Empirical Evidence

The annuity puzzle is a puzzle not because the data doesn't replicate the full annuitization result of Yaari, rather because the empirical evidence shows a polar result of almost no annuitization of wealth by retirees. Even though there was a wide acceptance that Yaari's model was not realistic enough, Modigliani (1986) famously talked about the importance of the lack of annuitization when in his Nobel acceptance speech he stated:

> "It is a well known fact that annuity contracts... are extremely rare. Why this should be so is a subject of considerable current interest and debate."

During the 1980s, after a series of studies, Friedman and Warshawsky (1990) reported that only 2% of the elderly population in the Retirement History Survey held annuities. They characterized this situation as a "puzzle" given that the annuities market in the United States was well developed by then. They also pointed that a smoothing-consumption behavior was present for short run fluctuations, so the fact that in the long run older individuals tended to not dissave did not conform with the life-cycle hypothesis, unless individuals had a strong bequest reason—which on its own could be an explanation for the lack of annuitization. More recently, using data 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.

More evidence reflecting the low demand of annuities, mainly found in survey data about retirement products, can be found in Benartzi et. al. (2011). Also, from a more aggregate point of view, the Investment Company Institute reported that by June 2015 annuity holdings totaled only 8.6% of total U.S. retirement assets.[1]Furthermore, the low level of annuitization is true not only for the United States, but also for many other countries. See for example Butler and Teppa (2007) for its documentation in Switzerland, Vidal-Melia (2003) in Spain, and Sakamoto (2010) for its analysis in Japan.

3.3 Robustness of the Basic Model Prediction

Davidoff, Brown, and Diamond (2005) modeled in a life-cycle setting many hypotheses—purported after Yaari's work—aimed to explain the annuity puzzle. Their work proved

[1] Investment Company Institute (2015).

that the demand for annuities in a scenario that incorporates many frictions, in contrast to the standard life-cycle used by Yaari, still predicts a high level of annuitization of wealth. Specifically, in an Arrow-Debreu setting, they find that high—if not complete—levels of annuitization of wealth are the result of optimal investment (and therefore still theoretically predicted) if just two conditions are met: 1) no bequest motives, and 2) those who survive are paid a rate of return greater than other risk-matching assets. This clarification was important because it delimited and identified these two possible (theoretical) explanations for which annuitization could be low as the only mechanisms that would allow low annuitization as optimal investment. In this line of reasoning, they emphasized that complete annuitization is achieved regardless of any type of exponential discounting, expected utility axioms, intertemporal separability, or even actuarially fair annuities. They also introduced the notion of annuity return comparability, in order to compare this asset with a risk-free asset, as used for a simplified explanation in Section 2. Regarding this type of robustness found, Davidoff, Brown, and Diamond claim that "the general theory itself is insufficient to answer questions about the optimal fraction of annuitized wealth",[2] and thus, only this type of simulations can shed light regarding the demand for annuities.

Partial annuitization, yet still high, was the result in some settings of these simulation exercises. For example, this occurs in a scenario when an individual may desire a consumption path for which only a variable annuity return would be suitable, but only a constant annuity return is available in the market, i.e., an incomplete annuity market. Additionally, this study also examines the role of liquidity, emphasizing the fact that annuities are not typically a liquid asset, which may make them undesirable in the presence of expenditures shocks that cannot be insured. In both situations, since an optimal path for consumption would never be achieved anyway (due to the restrictions imposed) their model does not predict complete annuitization of wealth. Yet, the most important result of this exercise is that predicted

[2] Davidoff, Brown, and Diamond (2005, p. 1574)

annuitization of wealth remains considerably high in both scenarios: a rational individual should annuitize between 75 percent and 90 percent of her wealth—even when we assume a heavy discounting factor and a habit formation utility form.

Furthermore, Davidoff, Brown, and Diamond (2005) conclude:

> "These results suggest that lack of annuity demand may arise from behavioral considerations, and that some mandatory annuitization may be welfare increasing."

Effectively, there exists a behavioral approach research to the annuity puzzle due to the inability of predicting low annuitization in traditional models. This literature will not be covered in the present document.

3.4 Dual Role of Health Shocks

Many subsequent studies just confirmed or found the robustness of the results found in Davidoff, Brown, and Diamond (2005). See for example Turra and Mitchell (2008), and Peijnenburg, Nijman and Werker (2013). Brown et al. (2008, p. 304) state: "As a whole, however, the literature has failed to find a sufficiently general explanation of consumer aversion to annuities."

The advent of life-cycle heterogeneous agents models provided a new framework to study consumption behavior during retirement years. Precautionary savings were able to explain why decumulation of savings did not empirically take place (De Nardi, French, and Jones, 2006). Moreover, the identification of medical expenses shocks by Palumbo (1999) clarified that, during the life-cycle, health shocks were an important factor for precautionary savings.

Reichling and Smetters (2015) advanced the discussion for the annuity puzzle by giving a more realistic role to the stochastic nature of health. To understand their result, it is important to note that if health shocks were to act as any other expenditures shocks, then a high level of annuitization result would not be discarded, as Davidoff, Brown, and Diamond (2005) demonstrated. That is, even though it might be easier to assume that annuities' lack of liquidity makes them undesirable in the presence of health shocks, which later convert into medical expenditures shocks, one must also realize that annuities can function as long-term-care insurance if bought earlier in life. Reichling and Smetters realize this latter insurance-like use of annuities may not be as strong if the expectations of survival to advanced years are low—a situation that may precisely arrive after a health shock. Hence, an observed health shock may have a dual role to deter the demand for annuities: on the one hand, it creates a medical expenditure shock that needs to be covered using liquid assets (unlike annuities), and on the other hand, the realization of the health shock itself may limit the longevity of the individual, diminishing the need to purchase a longevity insurance such as an annuity.

With this intuition in mind these authors introduce the concept of stochastic mortality. To understand these concepts and results, consider the conditional (and subjective) survival probability to age t, as a function of health h determined at time t, $\rho_t(h_t)$, and assume that the individual's health change is governed by a Markov process. Using the same analysis framework from Section 2, suppose now we have to take into account a health shock from time t to time $t+1$. Assume that, ex post, this shock is so prominent that it reduces all remaining (subjective) survival probabilities such that $q_{t+1}(h_{t+1}) < q_{t+1}(E(h_{t+1}))$, i.e., the realization of the health shock is lower than what its expectation predicted. Consequently, if an individual buys an annuity using the expectation of health, that is, before the shock, then we can use this inequality and show that

$$\begin{aligned}
\frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[q_{t+1}(h_{t+1})\right] &< \frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[q_{t+1}(E(h_{t+1}))\right] \\
\frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)} + \frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[q_{t+1}(h_{t+1})\right] &< \frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)} + \frac{\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[q_{t+1}(E(h_{t+1}))\right] \\
\frac{q_t(E(h_{t+1}))\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[1+\phi_t(h_{t+1})\right] &< \frac{q_t(E(h_{t+1}))\rho_{t+1}(E(h_{t+1}))}{(1+r)}\left[1+\phi_t(E(h_{t+1}))\right]
\end{aligned}$$

where $1+\phi_t(h_{t+1}) = \frac{(1+r)}{\rho_{t+1}(h_{t+1})}$ is the annuity return for time t calculated using health at time $t+1$, and $1+\phi_t(E(h_{t+1})) = \frac{(1+r)}{\rho_{t+1}(E(h_{t+1}))}$ using expectation $E(h_{t+1})$. Using Equation 1, this inequality implies that the gross annuity return will *not* always be higher than the risk-free return, that is,

$$\begin{aligned}
1+\phi_t(h_{t+1}) &< 1+\phi_t(E(h_{t+1})) \\
1+\phi_t(h_{t+1}) &< \frac{1+r}{\rho_{t+1}(E(h_{t+1}))}
\end{aligned}$$

The interpretation of this result is that, due to a negative health shock, there is a possibility the annuity price is recalculated downwards, such that the return calculated is not strictly higher than the risk-free return anymore.

This situation then is one where the survival distribution itself is stochastic. In other words, the distribution governing the survival probabilities as a function of health changes, such that all survival probabilities $\rho_{t+1}(h)$ change as well. Therefore, stochastic health, as Reichling and Smetters claim, is translated into stochastic mortality—if a negative health shock is sufficiently strong to decrease the remaining survival probabilities, then the utility an annuity provides is automatically reduced, as the individual does not need to hedge longevity risk as much (she is more likely to die sooner). So, as mentioned above, to this situation we add the other known effect of a health shock: medical expenses that would inevitably require liquidity to be covered. As a total effect, Reichling and Smetters report optimal levels of annuitization for around 30 percent wealth, without assuming particularly heavy discount factors nor a habit formation utility form. This work then constitutes the first plausibly theoretical prediction of low levels of annuitization.

Even though the main mechanism for the optimal low level of annuitization (stochastic mortality) is identified and examined, Reichling and Smetters use an all-encompassing model—which even includes bequest motives. In contrast with the detailed multi-model simulation study of Davidoff, Brown, and Diamond (2005), there is no clear isolation of the effect that a reduction in longevity expectations may have on annuitization, nor of the liquidity effects of the health shock, which affects income not only through medical expenditures, but also through reduced earnings. Furthermore, this situation does not allow one to hypothesize other plausible reasons outside their model that might lower the optimal demand for annuities, as their all-encompassing model includes the most important potential culprits for low annuitization previously identified in the literature. As a result, 30 percent annuitization of wealth as a theoretical prediction would still seem too high if we consider that all plausible reasons lowering the demand for annuitites have been considered.

3.5 General Equilibrium and Annuitization

The latest strand of life-cycle literature addressing the annuity puzzle takes a completely different approach. Rather than searching for reasons for which an individual would not want to annuitize in partial equilibrium, this literature questions whether annuitization is welfare-improving in a general equilibrium environment. In a broad sense, what this literature has concluded is that it is not socially optimal when all individuals annuitize their wealth; and this is due to adverse general equilibrium effects. The implication of this new approach is that, while almost the entire literature has been focused on partial equilibrium model predictions, a general equilibrium model is able to explain why annuitization of wealth is not a social desirable outcome, and therefore, if we interpret this social optimality rule as a model prediction, then the low demand for annuities observed in the data conforms economic theory.

To understand this approach, one must take as reference an economy where annuitization does not take place, and then compare the changes in welfare when annuitization is introduced. In an overlapping generations model, that is, a model that constitutes a suitable framework to study the implications of retirement behavior, and where prices of factors are computed taking into account supply and demand, the general equilibrium effects of individuals annuitizing their wealth can be various. In this regard, starting from a scenario without annuities, this literature acknowledges that when individuals start engaging in annuity contracts, the most important feature to study is the wealth of those who choose to annuitize .

Heijdra, Mierau, and Reijnders (2014) find two potential mechanisms for which changes in wealth due to annuitization, in general equilibrium, are undesirable. First, as a consequence of wealth not being saved and instead being transferred to other individuals through annuity contracts, then a decline in capital accumulation will follow, i.e., people annuitizing are not, by definition, saving. As a result, those who are not retirees yet, and therefore still depend on their wages, will be worse off when compared to a situation where nobody annuitizes their wealth. Because of general effects this would lower wages due to the (now) lower capital usage in the economy (complementary inputs). This particular point has been previously considered by Kotlikoff, Shoven, and Spivak (1986), but more recent work showed the demand for annuities can still be modeled assuming there is no consequence for the total amount of capital in the economy, e.g. Reichling and Smetters (2015). The second mechanism, on the other hand, is considered to affect capital accumulation more prominently, and has been the focus of other recent studies as well. Specifically, in an annuity contract, the wealth transfer from those who die first to those who die last —a.k.a. the mortality credit—constitutes an accidental bequest type of wealth transfer that is not taking place. The attention then is centered on the potential younger recipients of accidental bequests who, because older individuals engage in annuity contracts, will not receive this wealth transfer anymore. Heijdra, Mierau, and Reijnders note that this intergenerational transfer is especially important since it induces beneficial savings effects for the entire economy. In conclusion, it is this decline in savings, causing a decline in capital, the reason why welfare declines in the economy once annuities are introduced.

It is important to emphasize that this general equilibrium analysis takes into account the fact that individuals do in fact prefer to annuitize. That is, even though it is detrimental to annuitize as a society, it is still optimal to annuitize from an individual point of view, i.e., in partial equilibrium. Feigenbaum, Gahramanov, Tang (2013) do a careful study of this situation of lack of accidental bequests in general equilibrium due to annuitization. They note that at the core of the problem lies the concept of pecuniary externalities (McKean, 1958), that is, externalities that operate through prices (rather than resources) due to aggregate behavior, which is why the general equilibrium scenario is very important, allowing to account for the lack of aggregated transfers that fail to occur when *everybody* in the economy annuitizes their wealth. Furthermore, they note that in partial equilibrium the accidental bequests received are fixed. Therefore, if individuals were to coordinate on a strategy of not annuitizing any wealth, then society welfare would be increased. In other words, this setting constitutes a Prisoners Dilemma game—the socially optimal outcome is not a Nash equilibrium. Moreover, Feigenbaum, Gahramanov, Tang assert that if individuals were to coordinate, this coordination would be irrational— they call it "optimal irrational behavior". By acknowledging the low levels of annuitization in the data, they state:

> "the failure of households to annuitize is, remarkably, a real-world example of almost everyone behaving irrationally in a manner that improves their welfare

via macroeconomic dynamics".

3.6 Conclusions

The annuity puzzle not only remained a puzzle for many years in economic literature, but it also opened up the discussion of what should be analyzed in partial equilibrium or general equilibrium. While there is little doubt that all mechanisms that could potentially explain low levels of annutization have been considered, it is still open to debate the assertion that coordination to avoid general equilibrium pecuniary externalities is the reason for why we do not observe higher levels of annuitization in the data. Nevertheless, this relatively new approach constitutes a building block on which more models can be developed on—specially to study the mechanisms through which accidental bequests are being redistributed in reality (estate taxation, public goods, etc.), as these would constitute a coordination device for which individuals choose to no annuitize their wealth.

Bibliography

[1] Bell, F. C. and M. L. Miller (2005). "Life tables for the United States social security area 1900-2100." Office of the Chief Actuary, Social Security Administration

[2] Benartzi, Shlomo, Alessandro Previtero, and Richard H. Thaler (2011) "Annuitization puzzles." Journal of Economic Perspectives 25.4: 143-64.

[3] Benitez-Silva, H., & Heiland, F. (2007). "The social security earnings test and work incentives." Journal of Policy Analysis and Management, 26(3), 527-555.

[4] Benitez-Silva, H., & Heiland, F. (2008). "Early claiming of social security benefits and labour supply behaviour of older Americans." Applied Economics, 40(23), 2969-2985.

[5] Brown, Jeffrey R., Arie Kapteyn, Erzo FP Luttmer, and Olivia S. Mitchell (2013). "Cognitive constraints on valuing annuities." No. w19168. National Bureau of Economic Research.

[6] Brown, Jeffrey R., Jeffrey R. Kling, Sendhil Mullainathan, and Marian V. Wrobel (2008). "Why Don't People Insure Late-Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle." American Economic Review, 98(2): 304–09

[7] Brugiavini, Agar (1993). "Uncertainty resolution and the timing of annuity purchases." Journal of Public Economics 50.1 : 31-62.

[8] Butler, Monika, and Federica Teppa (2007). "The choice between an annuity and a lump sum: Results from Swiss pension funds." Journal of Public Economics 91.10 : 1944-1966.

[9] Caliendo, Frank N., Nick L. Guo, and Roozbeh Hosseini (2014). "Social security is NOT a substitute for annuity markets." Review of Economic Dynamics 17.4 : 739-755.

[10] Coile, C., & Gruber, J. (2007). "Future social security entitlements and the retirement decision." The Review of Economics and Statistics, 89(2), 234-246.

[11] Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond (2005). "Annuities and individual welfare." The American economic review 95.5 : 1573-1590.

[12] De Nardi, Mariacristina, Eric French, and John Bailey Jones (2006). "Differential mortality, uncertain medical expenses, and the saving of elderly singles." No. w12554. National Bureau of Economic Research.

[13] De Nardi, Mariacristina (2004). "Wealth inequality and intergenerational links." The Review of Economic Studies 71.3 : 743-768.

[14] Dorn, D., & Sousa-Poza, A. (2010). " 'Voluntary'and 'involuntary'early retirement: an international analysis." Applied Economics, 42(4), 427-438.

[15] Duggan, J. E., & Soares, C. J. (2002). "Actuarial nonequivalence in early and delayed social security benefit claims." Public Finance Review, 30(3), 188-207.

[16] Feigenbaum, James, Emin Gahramanov, and Xueli Tang (2013). "Is it really good to annuitize?" Journal of Economic Behavior & Organization 93 : 116-140.

[17] Feigenbaum, James, Frank N. Caliendo, and Emin Gahramanov (2011). "Optimal irrational behavior." Journal of Economic Behavior & Organization 77.3 : 285-303.

[18] Feldstein, Martin (2005). "Rethinking social insurance." American Economic Review 95.1: 1-24.

[19] Fields, G. S., Mitchell, O. S., & Mitchell, O. S. (1984). "Retirement, Pensions, and Social Security." MIT Press.

[20] Franks, Peter, Carolyn M. Clancy, Marthe R. Gold, and Paul A. Nutting (1993). "Health insurance and subjective health status: data from the 1987 National Medical Expenditure survey." American Journal of Public Health 83, no. 9 : 1295-1299.

[21] Friedman, Benjamin M., and Mark J. Warshawsky (1990). "The cost of annuities: Implications for saving behavior and bequests." The Quarterly Journal of Economics 105.1 : 135-154.

[22] Gan, Li, Hurd, M. D., & McFadden, D. L. (2005). "Individual subjective survival curves." In "Analyses in the Economics of Aging", (pp. 377-412). University of Chicago Press.

[23] Gan, Li, Guan Gong, Michael Hurd, and Daniel McFadden (2015). "Subjective mortality risk and bequests." Journal of Econometrics 188, no. 2: 514-525.

[24] Guvenen, Fatih (2007). "Learning your earning: Are labor income shocks really very persistent?" The American Economic Review 97.3 : 687-712.

[25] Haveman, R. H., & Wolfe, B. L. (1984). "Disability transfers and early retirement: A casual relationship?" Journal of Public Economics, 24(1), 47-66.

[26] Heijdra, Ben J., Jochen O. Mierau, and S. M. Reijnders (2014). "A tragedy of annuitization? Longevity insurance in general equilibrium." Macroeconomic Dynamics 18.7 : 1607-1634.

[27] Heiland, Frank, and Na Yin (2014) "Have we finally achieved actuarial fairness of social security retirement benefits and will it last?" Michigan Retirement Research Center Research Paper 2014-307.

[28] Heimer, Rawley, Kristian Ove R. Myrseth, and Raphael Schoenle (2015). "YOLO: Can Subjective Life-Expectancies Explain Household Investment Puzzles?" Available at SSRN 2669484 .

[29] Hosseini, Roozbeh (2015). "Adverse selection in the annuity market and the role for social security." Journal of Political Economy 123.4 : 941-984.

[30] Huggett, Mark (1993). "The risk-free rate in heterogeneous-agent incomplete-insurance economies." Journal of Economic Dynamics and Control, 17(5-6), 953-969.

[31] Hurd, M. D., Smith, J. P., & Zissimopoulos, J. M. (2004). "The effects of subjective survival on retirement and social security claiming." Journal of Applied Econometrics, 19(6), 761-775.

[32] Hurd, Michael D., and Kathleen McGarry (1995). "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 : S268-S292.

[33] Hurd, Michael D., and Kathleen McGarry (2002) "The predictive validity of subjective probabilities of survival." The Economic Journal 112.482: 966-985. Investment Company Institute (2015). The U.S. Retirement Market, Second Quarter (2015).

[34] James, Estelle, Guillermo Martinez, and Augusto Iglesias (2006) "The payout stage in Chile: who annuitizes and why?" Journal of Pension Economics and Finance 5.02 : 121-154.

[35] Jeske, Karsten, and Sagiri Kitao (2009) "US tax policy and health insurance demand: Can a regressive policy improve welfare?" Journal of Monetary Economics 56.2 : 210-221.

[36] Knoll, Melissa AZ, and Anya Olsen (2014). "Incentivizing delayed claiming of Social Security retirement benefits before reaching the full retirement age." Soc. Sec. Bull. 74: 21.

[37] Kotlikoff, Laurence J., John Shoven, and Avia Spivak (1986). "The effect of annuity insurance on savings and inequality." Journal of Labor Economics 4.3, Part 2 : S183-S207.

[38] Li, Q., Pantano, J., He, D., & Casanova, M. (2014). "Social Security Benefits, Life Expectancy and Early Retirement." 2014 Meeting Papers (No. 824). Society for Economic Dynamics.

[39] Liebman, J. B. (2002). "Redistribution in the current US social security system". In 'The distributional aspects of social security and social security reform', (pp. 11-48). University of Chicago Press.

[40] Manton, Kenneth G., Eric Stallard, and J. Vaupel (1981). "Methods for comparing the mortality experience of heterogeneous populations." Demography 18.3 : 389-410.

[41] McKean, Roland N. (1958). "Efficiency in government through systems analysis with emphasis on water resource development." Efficiency in government through systems analysis with emphasis on water resource development. .

[42] Merrill Lynch, (2016). "Leisure in Retirement: Beyond the Bucket List A Merrill Lynch Retirement Study conducted in partnership with Age Wave." Available at https://agewave.com/wp-content/uploads/2016/05/2016-Leisure-in-Retirement_Beyond-the-Bucket-List.pdf

[43] Modigliani, Franco. (1986). "Life cycle, individual thrift, and the wealth of nations." Science 234.4777: 704-712.

[44] Muldoon, D., & Kopcke, R. W. (2008). "Are people claiming social security benefits later?" Challenge, 1501, 1990s.

[45] Munnell, A. H. (2015). "The average retirement age–an update." Notes, 1920, 1960-1980.

[46] Munnell, A. H., & Soto, M. (2005). "Why Do Women Claim Social Security Benefits So Early?" Center for Retirement Research at Boston College.

[47] Palumbo, Michael G.(1999). "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle." Review of Economic Studies 66(2): 395421.

[48] Panis, Constantijn WA, and Michael J. Brien (2016)."Innovations and trends in annuities: Qualifying Longevity Annuity Contracts (QLACs)."

[49] Pashchenko, Svetlana, and Ponpoje Porapakkarm (2013). "Quantitative analysis of health insurance reform: Separating regulation from redistribution." Review of Economic Dynamics 16.3 : 383-404.

[50] Peijnenburg, Kim, Theo Nijman, and Bas JM Werker (2016). "The annuity puzzle remains a puzzle." Journal of Economic Dynamics and Control 70 : 18-35.

[51] Reichling, Felix, and Kent Smetters. (2015). "Optimal annuitization with stochastic mortality and correlated medical costs." The American Economic Review 105.11 : 3273-3320.

[52] Sakamoto, Junichi. (2010). "Annuity Markets in Japan." Pension Research Council Working Paper No. 2010-21 . Available at SSRN: https://ssrn.com/abstract=1707475 or http://dx.doi.org/10.2139/ssrn.1707475

[53] Siegrist, J., Wahrendorf, M., Von dem Knesebeck, O., Jurges, H., & Borsch-Supan, A. (2007). "Quality of work, well-being, and intended early retirement of older employees—baseline results from the SHARE Study." The European Journal of Public Health, 17(1), 62-68.

[54] Sun, Wei, and Anthony Webb. (2011). "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 .

[55] Vaupel, James W., Kenneth G. Manton, and Eric Stallard. (1979). "The impact of heterogeneity in individual frailty on the dynamics of mortality." Demography 16.3 : 439-454.

[56] Vidal-Melia, Carlos, Ana Lejarraga-Garcia, and Jose E. Devesa-Carpio. (2003). "Defined contribution pensions, married couples, and the 'annuity puzzle'". 6th Italian-Spanish Conference on Financial Mathematics. University of Trieste .

[57] Yaari, Menahem E. (1965). "Uncertain lifetime, life insurance, and the theory of the consumer." The Review of Economic Studies 32.2 : 137-150.