Abstract

We study how the design of options and guarantees can shape the exposure to mortality/longevity risk in life insurance contracts. We develop a model of selective withdrawals driven by exogenous and endogenous factors, offering insights into traditional approaches to the analysis of surrender guarantees and dynamic adverse selection. We show how the mortality risk profile of policyholders can be represented in terms of a frailty process shaped by the relative attractiveness of different contract benefits in different states of the world. We then outline some applications to pricing and optimal contract design in traditional contracts and variable annuities.
1 Introduction

Voluntary policy termination (or surrender) may give rise to substantial costs for insurers, from the partial recovery of acquisition expenses to the loss of future premium income and associated investment returns. Even if surrender risk is properly accounted for in the pricing schedule, and individual contract terminations may result in a profit for the insurer, there are potential costs of adverse selection, as policyholders may exercise the surrender decision in response to private information about their health status. Despite the importance of dynamic adverse selection, very few papers have modeled this phenomenon in life insurance.

On the theory side, standard references are Jones (1998) and Valdez (2001). Jones (1998) develops a setup for selective lapsation based on frailties. The idea is to model the dependence between surrenders and deaths by using a proportional frailty variable. Valdez (2001) develops a test for selective lapsation based on copula models. In both contributions mortality and lapse rates are deterministic and contract design plays no role.

On the empirical side, the few studies we are aware of do not typically find a significant relationship between policy termination and mortality risk (see the surveys of Albert et al. (1999), and Hayes (2008), for example). The main difficulty is the lack of granular policyholder data and the tendency of to run tests based on data resulting from the aggregation of insurance contracts with different net exposure to mortality risk (e.g., endowments and term assurances).

The main purpose of this work is to outline a framework that can be used to model dynamic adverse selection in the presence of different contractual features, such as minimum return guarantees and options to (partially) surrender the policy for its cash value.
recognized by the insurer. We model the dynamics of adverse selection by using a representation of policyholders’ conditional survival probabilities based on a frailty process which arises endogenously from the design of the contract.

We determine the dynamics of the frailty process for several different types of contracts and guarantees, in the presence of both endogenous and exogenous termination. By the former, we mean the exercise of any American claim embedded in an insurance contract. By the latter, we mean termination triggered by a jump process with intensity driven by state variables capturing shocks that policyholders may be facing along different phases of the business/life cycle, such as financial constraints and unexpected medical expenses due to health deterioration. We use our results to draw insights on the optimal design and pricing of traditional contracts and variable annuities, offering a number of practical examples.

The paper is organized as follows. In the next section we discuss the model setup. In section 3 we study the process of dynamic adverse selection and show how it can be captured by an endogenous frailty process. We also derive implications for the optimal design and fair pricing of insurance contracts. In section 4 we consider the case of traditional insurance contracts and of unit-linked contracts, providing several numerical examples of relevant frailty processes. Section 5 concludes. Further remarks and technical details are contained in the appendix.
2 Setup

2.1 Demographic risks

Consider a family of $m$ individuals entering an insurance contract at time $0$. We let $T > 0$ denote the maturity of the contract, and take as given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfying the usual conditions.\footnote{Namely right continuity and $\mathbb{P}$-completeness. Here, $\mathbb{F} := (\mathcal{F}_t)_{t \in \mathcal{T}}, \mathcal{T} := [0,T]$, is an increasing sequence of sigma-fields representing the information available to agents over time.} We model the insured’s random residual life times as stopping times $\tau^1, \ldots, \tau^m$. Assuming that the individuals are homogeneous with respect to health status and other relevant risk characteristics at inception, we let the $\tau^i$’s have the same law at time zero.

Denote by $N^d(t) = (N^d_1(t), \ldots, N^d_m(t))'$ the vector of death indicator processes, meaning that each $N^d_i(t) := 1_{\tau^i \leq t}$ keeps track of whether the $i$-th individual has died or not by each time $t$. We model the components of $N^d(t)$ as the first jumps in the components of an $m$-variate Cox (or conditionally Poisson) process with intensity vector $\mu^d(t) = (\mu^d_1(t), \ldots, \mu^d_m(t))'$, predictable with respect to $\mathcal{G}$, a strict subfiltration of $\mathbb{F}$.

The idea is that $\mathcal{G} := (\mathcal{G}_t)_{t \in \mathcal{T}}$ provides information on demographic and other relevant risk factors, but not on the occurrence of the death times themselves. As policyholders are homogeneous, the $\mu^d_i$’s are assumed to have the same law as seen from time zero. By the conditional Poisson assumption, each insured’s conditional survival probability can be written as

$$P(\tau^i > s | \mathcal{F}_t) = 1_{\tau^i > t} \underbrace{e^{-\int^s_t \mu^d(u) \, du}}_{\mathcal{G}_t}.$$ 

In the following, it will be convenient to define the average force of mortality,

$$\overline{\mu}^d(t) := \frac{\sum_{i=1}^m \mu^d_i(t) (1 - N^d_i(t-))}{\sum_{i=1}^m (1 - N^d_i(t-))},$$
which we will associate with the single jump process \( N_d(t) := 1_{\tau \leq t} \). Here, \( \tau \) can be seen as a stopping time representing the death time of the average survivor in the group of original \( m \) individuals. The process \( N_d \) is Cox with intensity driven by the filtration \( G \vee H \subseteq F \), with \( H = (\vee_{i=1}^m H_t^i)_{t \in \mathcal{T}} \) and \( H_t^i := \sigma(\tau^i \wedge t) \), with the understanding that \( \bar{\tau} = 0 \) on \( \{\tau^{(m)} \leq t\} \), where \( \tau^{(m)} \) is the last death occurrence. On \( \{\tau^{(m)} > t\} \), we can then write the conditional survival probabilities of the average survivor as

\[
P(\tau > s | G_t \vee H_t) = E \left[ e^{-\int_s^t \bar{\tau}^i d\tau^i} | G_t \vee H_t \right]. \tag{2.1}
\]

### 2.2 Insurance contracts

Insurance contracts are characterized by a quadruple \( (P, B^l, B^d, B^w) \), where \( P \) denotes the single premium paid by the insured at inception (the extension to regular premiums is covered later), and \( B^x \) denotes \( G \)-adapted living \((x = l)\), death \((x = d)\), and withdrawal \((x = w)\) benefits. The cumulative gains to the \( i \)-th insured from holding the insurance contract are therefore given by

\[
G^i(t) = \int_0^t (1 - N^i_d(s-))(1 - N^i_w(s-)) dB^l(s) + \int_0^t (1 - N^i_w(s-))B^d(s-)dN^i_d(s) + \int_0^t (1 - N^i_d(s-))B^w(s-)dN^i_w(s)
\]

\[
= B^l(\tau^i-)1_{\tau^i \wedge \theta^i \leq t} + B^l(t)1_{\tau^i \wedge \theta^i > t} + B^d(\tau^i-)1_{\tau^i \leq t}1_{\theta^i > t} + B^w(\theta^i-)1_{\tau^i > t}1_{\theta^i \leq t},
\]

where \( \theta^i \) is a stopping time modelling the withdrawal of the policyholder from the contract, and \( N^i_w(t) := 1_{\theta^i \leq t} \).

Agents have access to a money market account with price process \( M \), yielding instan-
aneously the $\mathbb{G}$-predictable short rate $r$, so that $M(t) = e^{\int_0^t r(s) ds}$. Under no-arbitrage, there is a probability measure $\mathbb{Q}^i$, equivalent to $\mathbb{P}$, under which $G^i$ is a $\mathbb{Q}^i$-martingale after deflation by the money market account. As markets are incomplete, $\mathbb{Q}^i$ is not unique. Hence, every individual in the portfolio can have a different private valuation of the contract. The value of the contract to insured $i$ is

$$V^i(t; \theta^i) = M(t) \mathbb{E}^{\mathbb{Q}^i} \left[ \int_{t}^{\theta^i \wedge T} M^{-1}(s) dG^i(s) \bigg| \mathcal{F}_t \right].$$

By the conditional Poisson assumption, the above can be rewritten as

$$V^i(t; \theta^i) = 1_{\tau^i > t} \tilde{M}(t) \mathbb{E}^{\mathbb{Q}^i} \left[ \int_{t}^{\theta^i \wedge T} \tilde{M}^{-1}(s) d\tilde{G}^i(s) \bigg| \mathcal{G}_t \right] = 1_{\tau^i > t} \tilde{V}^i(t; \theta^i), \quad (2.2)$$

where $\tilde{V}^i(t; \theta^i)$ is the pre-death price of the contract, and the fictitious money market account $\tilde{M}$ and cumulative gains $\tilde{G}^i$ are given by

$$\tilde{M}(t) = e^{\int_0^t (r(s) + \mu^i(s)) ds}, \quad \tilde{G}^i(t) = B^L(t) + \int_0^t B^D(s) \mu^i_d(s) ds + \int_0^t B^u(s-) dN^i_u(s).$$

### 2.3 Surrender decisions

Policyholder $i$ exits the pool of insureds at the stopping time $\sigma^i := \tau^i \wedge \theta^i$. We assume $\theta^i$ to admit the representation

$$\theta^i := \theta^i_* \wedge \theta^i,$$

where $\theta^i_*$ captures the optimal decision to surrender, whereas $\theta^i$ is an exogenous surrender time triggered by the first jump of a Cox process with $\mathbb{G}$-predictable intensity $\lambda^i$ and

6
associated indicator $N^i_s(t) := 1_{\varphi^i_s \leq t}$. In particular, $\theta^i_\ast$ solves the problem

$$V^i(0) := \sup_{\vartheta \in T} V^i(0; \vartheta) = \sup_{\tilde{\vartheta} \in T_{\tilde{\vartheta}}, \tilde{\vartheta} \leq \tau^i \wedge \theta^i_\ast} V^i(0; \tilde{\vartheta}),$$

(2.3)

where $T_{\tilde{\vartheta}}$ denotes the set of stopping times relative to $\tilde{\vartheta}$. We set $\theta^i_\ast = +\infty$ if no endogenous surrender decision takes place.

Relative to $\theta^i_\ast$, the surrender time $\theta^i$ represents a ‘suboptimal’ decision due to exogenous constraints (e.g., financial constraints, out-of-pocket medical expenses, etc.) accelerating the exit from the contract. In our setup, a delay in the surrender decision (relative to an optimal reference level) can be generated by switching from a measure $Q^i$ (or a reference common measure $Q$) to an equivalent measure $Q^\sharp,i$ yielding an endogenous stopping rule $\theta^\sharp,i_\ast$ satisfying $\theta^\sharp,i_\ast \geq \theta^i_\ast$ almost surely (see Bacinello et al., 2010).

It is useful to conclude this section by emphasizing how our setup differs from standard multiple decrement models based on deterministic exit probabilities. Here, exit from the pool of policyholders is triggered by a death event ($\sigma^i = \tau^i$) or a surrender event ($\sigma^i = \theta^i$), whose conditional laws are in general highly dependent. This means, in particular, that the standard factorization

$$P(\sigma^i > s \mid \mathcal{G}_t) = P(\tau^i > s \mid \mathcal{G}_t)P(\theta^i > s \mid \mathcal{G}_t)$$

will only hold in very special cases. An example is when the endogenous surrender decision is ignored (set $\theta^i_\ast = +\infty$), and the intensity $\lambda^i$ is independent of the force of mortality $\mu^i_d$.

\footnote{The case of no exogenous surrender is similarly covered by setting $\theta^i = +\infty$.}
3 Dynamic adverse selection

3.1 Definition

We are interested in understanding when and how the conditional law of the death times of the individuals in the insurance portfolio may diverge from that of the entire population. As a useful benchmark, let us define the average force of mortality of the individuals still in the portfolio at each time \( t \), \( \bar{\mu}_{d|p}(t) \), by setting

\[
\bar{\mu}_{d|p}(t) := \frac{\sum_{i=1}^{m} \mu_{i|d}(t) (1 - N_{i|d}(t))}{\sum_{i=1}^{m} (1 - N_{i|d}(t))}.
\]

In analogy with (2.1), we can express the survival probability of the average surviving policyholder before the last exit time, \( \sigma^{(m)} \), as

\[
P\left(\tau_{p} > s \mid G_{t} \lor \mathcal{H}_{t}\right) = \mathbb{E}\left[ e^{-\int_{s}^{t} \bar{\mu}_{d|p}(u) \, du} \mid G_{t} \lor \mathcal{H}_{t}\right],
\]

where we have used the fact that on \( \{ \sigma^{(m)} > t \} \) the stopping time \( \tau_{p} \) can be thought as coinciding with the first jump of a Cox process with intensity \( \bar{\mu}_{d|p} \).

The distance (to be suitably defined) between the conditional law of \( \tau \) and of \( \tau_{p} \) can be used as a proxy for the strength of dynamic adverse selection induced by the relative attractiveness of different contractual benefits in different states of the world, in particular contingent on the (privately observable) realization of each individual force of mortality \( \mu_{i|d} \). For example, one may associate the inequality

\[
P(\tau_{p} > s \mid G_{t} \lor \mathcal{H}_{t}) \leq P(\tau > s \mid G_{t} \lor \mathcal{H}_{t})
\]

with a deterioration in the average health quality of the policyholders relative to the
overall population. This may be the result of active selection (\( \theta^i \)) or passive selection (\( \theta^j \)), depending on the path of \( \mu^i_d \), \( \lambda^i \), and other relevant state variables over \([0, t]\). Our framework aims precisely at gauging these trade-offs for different contract specifications.

### 3.2 Contract design and endogenous frailty

We now formalize the role of contract design in driving selective withdrawals. Consider an insurance contract \((P, B^l, B^d, B^w; c)\) parameterized by \( c \in \mathcal{C} \), where \( \mathcal{C} \) is the space of parameter configurations affecting the contract design. As a simple example, consider the case of a unit-linked or variable annuity policy, where the minimum guaranteed amount \( A^x(t) \) is offered on benefit of type \( x \in \{l, d, w\} \) at each time \( t \), so that \( B^x(t) = \max\{F(t), A^x(t)\} \), with \( F \) denoting the policy account value, which evolves according to

\[
dF(t) = \begin{cases} 
\frac{dS(t)}{S(t)} F(t) - \phi F(t) \, dt - dB^l(t), & \text{if } F(t) > 0 \\
0 & \text{otherwise}
\end{cases}
\]

Here, \( S \) tracks the value of the unitis or reference index, and \( \phi > 0 \) is the asset management charge (AMC).

As uncertainty unfolds and the \( \mu^i_d \)'s take different paths, policyholders will develop different views on the attractiveness of staying in the contract, depending on the private valuation of the benefits promised and the moneyness of different guarantees. Using the average population mortality as a baseline level, the policyholders’ average force of mortality can always be written as

\[
\bar{\mu}_{dp}(t; c) = \bar{\mu}_d(t) \bar{\eta}(t; c), \quad (3.1)
\]

\[^3\text{In particular, this could be the result of a ratchet, roll-up, or other type of guarantee.}\]
for a suitable $\mathbb{G}$-predictable process $\tilde{\eta} > 0$, which can be interpreted as a dynamic frailty process introducing heterogeneity in the mortality risk profile of the policyholders. This heterogeneity is endogenous, as it is induced by the contractual features summarized by the element $c$ of $C$. An alternative interpretation of (3.1) in terms of an equivalent change of measure is given in the next section. We can apply the same logic at individual policyholder’s level and use the representation

$$
\mu^i_{dp}(t; c) = \overline{\mu}_d(t) \eta^i(t; c),
$$

(3.2)

for some strictly positive, predictable process $\eta^i$.

### 3.3 Endogenous frailty and pricing

As all individuals are ex-ante homogeneous, the initial value of the contract is given by $V^i(0) = V(0)$. However, it will in general be incorrect to assume that the pricing formula (2.2) applies, as $\overline{\mu}_d$ will not reflect the mortality experience of the underlying pool of policyholders. The relevant price would instead be delivered by

$$
V^i(t; \theta^i(c), c) = 1_{t > 0} \hat{M}(t; c) \mathbb{E}^{Q^i(c)} \left[ \int_t^{\theta^i(c) \wedge T} \hat{M}^{-1}(s; c) d\hat{G}^i(s; c) \bigg| \mathcal{G}_t \right] = 1_{t > 0} \hat{V}^i(t; \theta^i(c), c),
$$

(3.3)

where the fictitious money market account and cumulated dividend process are parameterized by $c$ and depend on the force of mortality $\overline{\mu}_d(t) \eta^i(t; c)$. We note that even if a reference common pricing measure $Q = Q^i$ is assumed for the policyholders, the value of the contract at each time $t > 0$ may be different for each insured depending on the evolution of its own force of mortality $\mu^i_{dp}$. The idea can be illustrated by computing the
average survival probabilities as follows:

\[
P[\tau_p > T \mid G_t \vee H_t] = \mathbb{E}[e^{-\int_T^{\tau_p} \hat{\mu}(s; c) \, ds} \mid G_t \vee H_t]
\]

At the individual level, we have

\[
P[\tau_i > T \mid G_t] = \mathbb{E}[e^{-\int_T^{\tau_i} \hat{\mu}_i(s) \, ds} \mid G_t]
\]

We will now define the likelihood process

\[
Z_c = \frac{d\mathbb{P}^c}{d\mathbb{P}} igg|_{F_t} = \prod_{i=1}^{m} (1_{\tau_i > t} + \eta_i(t; c) 1_{\tau_i \leq t}) e^{\int_0^{\tau_i \wedge t} \tilde{\mu}_i(s) \, (\eta_i(s; c) - 1) \, ds}
\]

which is the Radon-Nykodym derivative of \(\mathbb{P}^c\) with respect to \(\mathbb{P}\), in other words

\[
Z_c = \frac{d\mathbb{P}^c}{d\mathbb{P}}
\]

assuming \(\mathbb{E}[Z_c(t)] = 1\). Equation (3.4) can be seen as a change of measure of the death probability from the real world probability \(\mathbb{P}\) to an equivalent probability measure \(\mathbb{P}^c\), dependent on \(c\), which accounts for the actual evolution of mortality for the underlying pool of policyholders. At the aggregate level it results in

\[
P[\tau_p > T \mid G_t \vee H_t] = \mathbb{P}^c[\tau > T \mid G_t \vee H_t].
\]

This result resembles the standard loadings applied by actuaries to reference (pop-
ulation) mortality tables. However that kind of calibration is mostly done to take into account changes in death probability due to longevity risk and moral hazard. Here $Z_c$ is a process that, given the type of contract we are considering, modifies over time the shape of the survival probability curve, improving the risk of longevity or mortality, and its effect is different at any point in time. If the pricing is consistent, we expect that under the new probability measure $\mathbb{P}^c$ the average survival probability of the portfolio is equal to that of the referee population. In our simulation study we will show for endowments a set of contracts for which the endogenous frailty process is always equal one, meaning that for those contracts there are not any changes in the shape of the survival probability. Those are the contracts perfectly priced by the standard pricing approach, for the others is required a change of measure to compute the survival probability, proxied by the endogenous frailty process.

However, equation (3.4) deserves some further comments. The process $\eta_i$ is nothing more that an intensity, different for each individual, which captures the phenomenon of adverse selection just by changing each individual’s mortality intensity. The crucial point lies in the fact that this process arises endogenously from the setup, and it is not specified a priori, i.e. given our assumptions, in a portfolio of homogeneous contracts and homogeneous investors we would experience the phenomenon of adverse selection. This change of measure is a particular case of Biffis et al. (2010), and, under our setting, all the properties and the proofs hold even in this case.

\footnote{It is common thought among actuaries that most of the policyholders, which purchase a life insurance with an high death benefit, think that their survival probability is low, and viceversa. Hence life tables are calibrated in order to reflect this fact, given the type of contract it is being priced.}
4 Simulation study

4.1 Applications to traditional contracts

In this section we want to apply this setting to traditional contracts, to see the behaviour of the insurers to different types of these products. Consider an insurance contract which pays a fixed amount $L$ upon survival at a fixed date $T$, and a fixed amount $D$ in case of death prior $T$. If the policyholder surrenders at time $t < T$ he receives a surrender value $W$, equal to the amount of premiums paid to date less a penalty $\gamma$. The premium rate $P$, paid by the insured, is computed by the insurer using the mortality table of the referee population after underwriting and a technical interest rate $r$. In this simplified setup, we can think at the contractual terms as the vector $c = (L, D, W, P, \gamma, T, r)' \in C$. Since $L, D, W, P, \gamma, T, r \in \mathbb{R}^+$, in this case $C \subseteq \mathbb{R}^7_+$. 

In all the contracts, of our simulations, the benefit upon survival is fixed and equal to reference of 250. The duration of all the contracts is equal to 20 year. The interest rate also is fixed at 3% per year. The penalty upon surrender is equal 5% per year at inception, and decreases over time, becoming negligible after 5-6 years. The benefit upon death is not constant, and varies between 0 and 500. Which means that it varies between 0 and 2 times the benefit upon survival. In Appendix A are reported more details on the implementation of the simulations.

The figures from 1 to 8 show how the average survival probability of those inside the portfolio, and the average frailty process, change over time and among different contracts with different death benefit. We considered separately the case where the surrender time is triggered only by the endogeneous decision, and when it depends on both the endogenous and exogenous decisions.
Figures 1 and 2 show the average survival curves of the portfolio, when only endogenous surrender is allowed. As we can see the average portfolio survival probability decreases as we increase the death benefit, meaning that policyholders with increases in life expectancy are no longer willing to stay in the contract. Interestingly there is a set of contracts for which the average survival curves are identical, specifically they are the contracts with death benefit between 0.528 and 0.62 times the survival benefit (between 132 and 155 in real values).

Figure 3 shows the average frailty process for the same set of contracts. Recalling (3.1), a value of one of the average frailty means that there is no difference between the population and the portfolio average mortality intensity. When the frailty is less than one, the mortality intensity of the portfolio is lower than that of the population, vice versa when the frailty is greater than one. From this figure it is clear that, the contracts with identical survival curves are contracts where the endogenous frailty is one over the whole period. This means that these contracts do not experience the phenomenon of adverse selection, and hence only these contracts are fairly priced by the standard pricing formula. For the contracts with death benefits lower than 0.528 times the survival benefit, the average mortality of the portfolio is lower than that of the population. The reversed happens for the contracts with death benefits greater than 0.62 times the survival benefit. This results are underlined even by the figure 4, which shows the average survival curves of the portfolio divided by the average survival probability of the initial population (taken as a benchmark).

Figures from 5 to 8 show the same arguments, but when we add even exogenous lapse. We can see from these figures that the results do not change much except for the fact that the graphs are less smooth.
4.2 Applications to variable annuities

In this section we study the endogenous frailty process in the index linked policies with roll-up guarantees upon survival, death and surrender. As in the previous application, the interest risk-free rate \( r \) is constant and equal to 3%. We assume for simplicity that no partial withdrawal, or living benefits, are allowed during the duration of the contract, so that the only life benefit is upon accumulation. The generic benefit is \( B^x(t) = \max\{F(t), Pe^{g_x}\} \), with \( x \in \{l, d, w\} \) and where \( g_x \) is the guaranteed rate upon event \( x \). We assume also that each policyholder fully invests the same amount \( \Pi \), in the same index-fund, which evolves as a Geometric Brownian Motion with constant volatility \( \sigma \) and constant drift \( \alpha \), under the real probability measure \( \mathbb{P} \). The contractual terms are the elements of the vector \( c = (g_l, g_d, g_w, \Pi, \phi, \gamma, T, r, \mu, \sigma)' \in C \). Since \( g_l, g_d, g_w, \Pi, \phi, \gamma, T, r, \mu, \sigma \in \mathbb{R}_+ \), we might write \( C \subseteq \mathbb{R}_{10}^+ \). Anyway, we must be aware of the fact that, \( g_j = 0 \) does not mean that the guarantee \( j \) is not present in the contract, because it is the case when a roll-up guarantee coincides with a return of premium guarantee. Asymptotically, the guarantee \( x \) is not present in the contract when \( g_x = -\infty \), for this reason we have \( g_l, g_d, g_w \in \mathbb{R}_+ \cup -\infty \) and \( C \nsubseteq \mathbb{R}_{10}^+ \).

In all the contracts considered, the guaranteed rate upon accumulation is fixed and equal to 2.5% per year. The duration of all the contracts is equal to 20 year. The interest rate also is fixed at 3% per year. The penalty upon surrender is equal 6% per year at inception, and decreases over time, becoming negligible after 5-6 years. The guaranteed rate upon death varies between 0 and 5% per year. Hence it is between 0 and 2 times guaranteed rate upon accumulation. For the guaranteed rate upon surrender we used the value of 0 and 2.5% per year. Hence the guarantee upon survival can be either equal to a return of premium guarantee, or equal to the guarantee upon accumulation. In
Appendix B are reported more details on the implementation of the simulations.

The graphs from 9 to 20 show the results for the change in the average survival probability, and the average frailty process, for different values of the set of guarantees \( C \) considered. As before, we considered separately the case where the surrender time is triggered only by the endogeneous decision, and when it depends on both the endogenous and exogenous decisions.

The figures 9 and 10 show the average survival curves of the portfolio when \( g_w = 0 \). While in the traditional contracts they were monotonic decreasing in the death benefit, now the effect is not monotonic anymore. Infact the average survival probability reaches its maximum when the guaranteed rate upon death is almost 2.5\% (the guaranteed rate upon accumulation), then it decreases. From the average endogenous frailty process, in figure 11, we see that the mortality intensity of the portfolio increases faster for guaranteed rate upon death greater than 2.5\%. Anyway, for all the contracts considered the frailty is above one, meaning that policyholders with greater life expectancy tend to surrender these contracts over time.

The figures 12 and 13 show the average survival curves of the portfolio when \( g_w = g_l = 2.5\% \). Here again, the average survival probability is not monotonic, but it is more clear the negative relationship between survival curves and guarantee upon death. Infact, the average survival curve is decreasing in the guaranteed rate upon death until it reaches a minimum near 3.5\% per year, after it increases slightly. The average endogenous frailty process, figure 14, is not everywhere above one, as for the contracts with \( g_w = 0 \).

In the figures from 15 to 20 we can see what happen to the same contracts if we add even the exogenous surrender. As for the traditional contracts, the shapes of the figures
do not change much.

5 Conclusion

We have seen how the decisions to surrender of each policyholder generate different path of the average mortality intensity for different type of contracts. The surrender decisions shape the aggregate mortality risk profile of the residual policyholder’s and there is high sensitivity of the endogenous mortality exposure to the contract design. These changes may be completely offset by exogenous surrender. Moreover it could be hard to claim that we are experiencing adverse selection even if actual and predicted mortality rates are different. For these reasons we have developed a framework to properly analyze the trade-off between endogenous and exogenous drivers of surrender and adverse selection and we have shown that the decisions to surrender generate an endogenous frailty process which involves a change in each mortality intensity. This change can be mimicked by a change of measure entering the insurer’s pricing functional. Moreover the behaviour of this process strongly depends on contractual features, and hence tests and calibrations should be made case by case.

For this work we have relied to an affine setting to take the computation simple, a future extension could be to see whether the setup is robust with respect to traditional mortality models (e.g. Lee-Carter and Cairne-Black-Dowd). More important, the framework can be used to develop robust approaches to detection/estimation of dynamic adverse selection.
References


Appendix

A Endogenous frailty in traditional contracts

The premium is computed by the insurer such that, under the full filtration \( \mathbb{F} \), it verifies the following equality

\[
P = \frac{D \mathbb{E}^\mathbb{Q} \left[ \int_0^T e^{-rt} dN_d(t) \mid \mathcal{F}_0 \right] + L e^{-rT} \mathbb{Q}(\tau > T \mid \mathcal{F}_0)}{\int_0^T e^{-rt} \mathbb{Q}(\tau > t \mid \mathcal{F}_0) \, dt},
\]

and the following, under the subfiltration \( \mathbb{G} \),

\[
P = 1_{\{\tau > 0\}} \frac{D \mathbb{E}^\mathbb{Q} \left[ \int_0^T e^{-(rt+\Lambda_d(t))} \mu_d(t) \, dt \mid \mathcal{G}_0 \vee \mathcal{H}_0 \right] + L \mathbb{E}^\mathbb{Q} \left[ e^{-(rT+\Lambda_d(T))} \mid \mathcal{G}_0 \vee \mathcal{H}_0 \right]}{\mathbb{E}^\mathbb{Q} \left[ \int_0^T e^{-(rt+\Lambda_d(t))} \, dt \mid \mathcal{G}_0 \vee \mathcal{H}_0 \right]}.
\]

However, while the insurer uses an aggregate \( \mu_d \), each insured \( i \) uses his own force of mortality. The insured decides to lapse at time \( t \) if the surrender value is greater than the expected present value of future payoffs. If we set the surrender value as the sum of
the premium paid less a surrender penalty $\gamma$ (decreasing in time), the policyholder $i$ surrenders if the following inequality holds

$$W = (1 - \gamma) P t \geq D \mathbb{E}^{Q_i} \left[ \int_t^T e^{-r(s-t)} dN_d^i(s) | \mathcal{F}_t \right]$$

$$+ L e^{-r(T-t)} Q^i \left( \tau^i > T | \mathcal{F}_t \right)$$

$$- P \int_t^T e^{-r(s-t)} Q^i \left( \tau^i > s | \mathcal{F}_t \right) ds$$

$$= D 1_{(\tau^i>t)} \mathbb{E}^{Q_i} \left[ \int_t^T e^{-r(s-t)+\Lambda^i_d(s)-\Lambda^i_d(t)} \mu_d(s) ds | \mathcal{G}_t \right]$$

$$+ L 1_{(\tau^i>t)} \mathbb{E}^{Q_i} \left[ e^{-r(T-t)+\Lambda^i_d(T)-\Lambda^i_d(t)} | \mathcal{G}_t \right]$$

$$- P 1_{(\tau^i>t)} \mathbb{E}^{Q_i} \left[ \int_t^T e^{-r(s-t)+\Lambda^i_d(s)-\Lambda^i_d(t)} ds | \mathcal{G}_t \right].$$

We have run a Monte Carlo study to investigate the phenomenon of selective lapsation in our setup in traditional life insurance contract. In this example the risk-free interest rate $r$ is deterministic, hence the only thing we need to simulate is the stochastic force of mortality $\mu^i_d$, and we used an affine process of the following type

$$d\mu_d(t) = -\beta \mu_d(t) dt + \sigma^d dW^d(t),$$

that have been calibrated on the italian male life tables, taken from http://www.mortality.org/. This process is useful since permits to compute the expected survival probability in closed form using the Ricatti’s ordinary differential equations (see Biffis, 2005). For the exogenous force of lapse we assumed a constant intensity $\mu_w$ equal for all the insureds. The coefficients and the parameters of all the models used are
reported in table (1).

We have simulated a portfolio of \( m = 2500 \) insured with age 50 and equal initial force of mortality. All the contracts have duration \( T = 20 \) years. All the insureds have the same law \( \mu_d \) for the force of mortality that was calibrated on the italian male mortality tables. But for each insured we have simulated a different process \( \mu_d^i \) for the force of mortality, with a different sample path. At each point in time we need to compute the expectation of the probability of death of the insured given the actual information available to him. This can be done by solving the Ricatti equations, as explained before, or by using a Least Square Montecarlo approach. Since we work in a Markovian environment this expectation must be a function depending on the current value of the relevant state variables and on the time length \( u \)

\[
\mathbb{P}[\tau^i \geq t + u | \mathcal{G}_t] = \mathbb{E}[e^{-\int_{t}^{t+u}\mu_d(s)ds} | \mathcal{G}_t] = g(X^i(t), u).
\]

We can approximate the function \( g \) with an orthogonal projection of the simulated survival probabilities onto the vector space generated by a finite set of functions \( h(X^i) = (h_1(X^i), \cdots, h_q(X^i))' \) taken from a suitable basis

\[
g(X^i(t), u) \approx \beta(u)' h(X^i(t))
\]

where \( h \) is the chosen set of functions and \( \beta(u) = (\beta_1(u), \cdots, \beta_q(u))' \) is a \( q \) vector of coefficients (depending on \( u \)) obtained through to the following minimization

\[
\min_{\beta(u)} \left\{ \sum_{i=1}^{m} \left( \mathbb{P}[\tau^{(i)} \geq t + u | t] - \beta(u)' h(X^i(t)) \right) \right\}
\]
where $\hat{P}[r^{(i)} \geq t + u | t] = \exp\{-\sum_{s=t}^{t+u} \mu_{d}^{(i)}(s)\}$ is the simulated path of the survival probability from $t$ to $t + u$ for the insured $i$ (see Bacinello et al., 2010).

B Endogenous frailty in variable annuities

The fair asset management charge is set by the insurer such that

$$\phi^* : \Pi = V(0),$$

we have shown in section (3.3) that $V(0) = V^i(0)$, and $V^i(0)$ has been defined defined in section (2.3). The index fund, under the real probability measure $\mathbb{P}$, evolves as a Geometric Brownian Motion with constant volatility $\sigma$ and constant drift $\alpha$

$$\frac{dS(t)}{S(t)} = \alpha dt + \sigma dW(t)$$

Also the fund value of the contract evolves as a GBM

$$\frac{dF(t)}{F(t)} = (\alpha - \phi) dt + \sigma dW(t),$$

with initial condition $F(0) = \Pi$. For the stochastic force of mortality and surrender we used the same processes of the traditional contracts. The coefficients and the parameters of all the models used are reported in table (I).

The generic benefit of the contract is

$$B^x(t; c) = \max\{F(t), P e^{\gamma s}\}, \quad x \in \{l, d, w\}$$
The cumulated gains are
\[
G^i(t; c) = B^i(T; c)1_{(\sigma^i > T)} + B^d(\tau^i - ; c)1_{(\tau^i \leq t)}1_{(\theta^i > t)} + B^w(\theta^i - ; c)1_{(\tau^i > t)}1_{(\theta^i \leq t)}.
\]

Now under no-arbitrage conditions the value of the contract is
\[
V^i(t; \theta^i, c) = e^{rt}E^{\mathbb{Q}^i}\left[ \int_t^{\theta^i \wedge T} e^{-rs}dG^i(s; c) | \mathcal{F}_t \right].
\]

For these contracts we have another source of uncertainty to simulate, which is driven by the variability of the index fund. We can hence compute the optimal stop \( \theta^i_* \) by using the dynamic programming approach. Let us take an arbitrary partition of the time interval \([0, T]\) defined as \( \pi_n = \{0 = t_1 < t_2 < \cdots, t_n = T\} \) and for simplicity let it be the trivial one \( \{0, 1, 2, \cdots, T\} \), we now have a discretized stopping problem. As for a standard American option pricing, we can use the Snell envelope of \( (V^i(t; \theta^i, c))_{t \geq 0} \) and apply the dynamic programming approach to obtain the optimal stopping time. The basic idea is to use a backward procedure which involves, at any time step, a comparison between the payoff provided by the surrender option and the continuation value. The procedure starts at the final time \( T \), at this time the payoff of the contract (given that the policyholder is alive) is
\[
B^i(T),
\]
and the contract expires. At time \( T - 1 \) the payoff of the contract in case of surrender
is \( B^w(T - 1) \) less a penalty \( \gamma \), and the expected continuation value is

\[
\mathbb{E}^Q[e^{-r} \hat{G}^i(T)|\mathcal{G}_{T-1}] = \mathbb{E}^Q[e^{-r} B^d(T)1_{\tau^i \geq T}|\mathcal{G}_{T-1}]
\]

\[
+ \mathbb{E}^Q[e^{-r} B^d(T)1_{T-1 \leq \tau^i < T}|\mathcal{G}_{T-1}],
\]

the policyholder will lapse the contract if

\[
(1 - \gamma(T - 1)) B^w(T - 1) \geq \mathbb{E}^Q[e^{-r} \hat{G}^i(T)|\mathcal{G}_{T-1}].
\]

At time \( T - 2 \) the payoff of the contract, in case of surrender, is \( B^d(T - 1) \), the expected continuation value is

\[
\sum_{t=T-1}^{T} \mathbb{E}^Q[\hat{G}^i(t)|\mathcal{G}_{T-2}] = \mathbb{E}^Q[e^{-2r} \hat{G}^i(T)|\mathcal{G}_{T-2}] + \mathbb{E}^Q[e^{-r} \hat{G}^i(T - 1)|\mathcal{G}_{T-2}]
\]

\[
= \mathbb{E}^Q[e^{-2r} B^d(T)1_{\tau^i \geq T}|\mathcal{G}_{T-2}]
\]

\[
+ \mathbb{E}^Q[e^{-r} B^d(T)1_{T-1 \leq \tau^i < T}|\mathcal{G}_{T-2}]
\]

\[
+ \mathbb{E}^Q[e^{-r} B^d(T - 1)1_{T-2 \leq \tau^i < T-1}|\mathcal{G}_{T-2}],
\]

the policyholder will lapse the contract if

\[
(1 - \gamma(T - 2)) B^w(T - 2) \geq \sum_{t=T-1}^{T} \mathbb{E}^Q[\hat{G}^i(t)|\mathcal{G}_{T-2}].
\]

The procedure continues till time of inception and deliver the optimal stop \( \theta^*_i \) going backward.

Now we consider the computation of the two generic expected benefits, such as

\footnote{For simplicity we will write \( B^x(t) \) instead of \( B^x(t; c) \)}
\[ \mathbb{E}^Q\left[e^{-r(t-s)}B^d(t)1_{\tau_i=t}\mid \mathcal{G}_s \right] \] and \[ \mathbb{E}^Q\left[e^{-r(T-s)}B'(T)1_{\tau_i>T}\mid \mathcal{G}_s \right]. \] If we assume independence between financial and demographic risk factors, it is possible to decompose the continuation value in the product of the expectation of future discounted payoff, and the expectation of the future survival probability, hence

\[
\begin{align*}
\mathbb{E}^Q\left[e^{-r(t-s)}B^d(t)1_{\tau_i=t}\mid \mathcal{G}_s \right] &= \mathbb{E}^Q\left[-r(t-s)B^d(t)\mid \mathcal{G}_s \right]\mathbb{E}^Q\left[1_{t-1 < t < t}\mid \mathcal{G}_s \right] \\
&= e^{-r(t-s)}\mathbb{E}^Q\left[B^d(t)\mid \mathcal{G}_s \right] Q^d[t-1 < \tau_i < t]\mid \mathcal{G}_s],
\end{align*}
\]

and similarly

\[
\begin{align*}
\mathbb{E}^Q\left[e^{-r(T-s)}B'(T)1_{\tau_i>T}\mid \mathcal{G}_s \right] &= e^{-r(T-s)}\mathbb{E}^Q\left[B'(T)\mid \mathcal{G}_s \right] Q^d[\tau_i > T]\mid \mathcal{G}_s].
\end{align*}
\]

We have discussed in Appendix A how to compute the expected survival probability for each policyholder at any point in time, now we need even to compute the expected value of the guarantee at any point in time. To solve this problem we can decomposed the guarantee as the sum of an asset plus an European put option on that asset, and then we priced it using the standard Black-Scholes formula. For example, the generic death benefit at time \( t \) is \( B^d(t) = \max\{F(t), P e^{g_{\Delta t}}\} = F(t) + (P e^{g_{\Delta t}} - F(t))^+ \). We can compute the expectation of this payoff in a closed form solution with the standard
Black-Scholes formula for a put option

\[
e^{-r(t-s)}E^Q[B^d(t)|G_s] = e^{-r(t-s)}E^Q[F(t) + (P e^{g(t)} - F(t))^+|G_s]
\]

\[
= e^{-\phi(t-s)}F(s) + e^{g(t-r(t-s))}P\Phi(-d_2(t, s))
\]

\[
- e^{-\phi(t-s)}F(s)\Phi(-d_1(t, s))
\]

\[
= e^{-\phi(t-s)}F(s)\Phi(d_1(t, s)) + e^{g(t-r(t-s))}P\Phi(-d_2(t, s)),
\]

where

\[
d_1(t, s) = \frac{1}{\sigma\sqrt{t-s}}\ln \frac{F(s)}{P e^{g(t)}} + \frac{\sqrt{t-s}}{\sigma}r + \frac{\sigma^2}{2}
\]

\[
d_2(t, s) = d_1(t, s) - \sigma\sqrt{t-s},
\]

and \(\Phi\) represents the standard normal cumulative distribution function.

We simulated \(N = 10000\) fund path and for each of them we take the same mortality intensity of each policyholder, in order to see how behave each policyholder in each state of the world. For each configuration of guarantees \(c \in C\) considered we computed the average survival probability over all the fund simulations. Basically, for each fund path \(F(t)^{(j)} (j = 1, \ldots, N)\) the \(\mu^j_i\) are conditionally i.i.d. and hence we can compute the average mortality intensity all the portfolio by

\[
\bar{\mu}_{dp}(t) = \frac{\sum_{i=1}^{m} \mu_{id}^j(t) \mathbf{1}_{\sigma^j > t}}{\sum_{i=1}^{m} \mathbf{1}_{\sigma^j > t}}.
\]

Since each \(\bar{\mu}_{dp}^j\) depends on the simulation \((j)\) since, the optimal stops \((\theta_{s(j)}^i)_{i=1,\ldots,m}\)

\((j = 1, \ldots, N)\) depend on the simulated value of the fund. Thus, even \((\sigma^j)_{i=1,\ldots,m}\)

\(^6\)The result is identical for the other benefits.
depend on \((j)\). Contrarily, the \((\tau^i)_{i=1,\ldots,m}\) and the \((\theta^i)_{i=1,\ldots,m}\) are simulated only once and are the same for all the fund paths. After having computed all the \(\bar{\mu}^{(j)}_{dp}\), we have taken the average over all the fund simulations to get the average portfolio force of mortality over all the fund values

\[
\bar{\mu}_{dp}(t) = \frac{\sum_{j=1}^{N} \bar{\mu}^{(j)}_{dp}(t)}{N}.
\]

From it we have computed the portfolio average survival probability and the portfolio average frailty process shown in the figures.

\section*{C Tables}

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(\sigma)</th>
<th>(\beta)</th>
<th>(\sigma^d)</th>
<th>(\mu_w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.045</td>
<td>-0.106</td>
<td>3 \cdot 10^{-4}</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 1: This table summarize the main parameters used for the simulations of the force of mortality, index value and force of lapse. The stock has mean 0.05 and volatility 0.045, the force of mortality has drift \(-0.106\) and volatility \(3 \cdot 10^{-4}\). The exogenous force of lapse is costant intensity equal to 0.04.
D Figures

Figure 1: Survival Curves Traditional Contracts (only endogenous surrender)
Figure 2: Survival Curves Traditional Contracts (only endogenous surrender)

Figure 3: Frailty Processes Traditional Contracts (only endogenous surrender)
Figure 4: Survival Curves Traditional Contracts, ratio with the survival probability of the whole population (only endogenous surrender)

Figure 5: Survival Curves Traditional Contracts (endogenous and exogenous surrender)
Figure 6: Survival Curves Traditional Contracts (endogenous and exogenous surrender)

Figure 7: Frailty Processes Traditional Contracts (endogenous and exogenous surrender)
Figure 8: Survival Curves Traditional Contracts, ratio with the survival probability of the whole population (endogenous and exogenous surrender)

Figure 9: Survival Curves Index-Linked, $g_w = 0$ (only endogenous surrender)
Figure 10: Survival Curves Index-Linked, $g_w = 0$ (only endogenous surrender)

Figure 11: Frailty Processes Index-Linked, $g_w = 0$ (only endogenous surrender)
Figure 12: Survival Curves Index-Linked, $g_w = 2.5\%$ (only endogenous surrender)

Figure 13: Survival Curves Index-Linked, $g_w = 2.5\%$ (only endogenous surrender)
Figure 14: Frailty Processes Index-Linked, $g_w = 2.5\%$ (only endogenous surrender)

Figure 15: Survival Curves Index-Linked, $g_w = 0$ (both endogenous and exogenous surrender)
Figure 16: Survival Curves Index-Linked, $g_w = 0$ (both endogenous and exogenous surrender)

Figure 17: Frailty Processes Index-Linked, $g_w = 0$ (both endogenous and exogenous surrender)
Figure 18: Survival Curves Index-Linked, $g_w = 2.5\%$ (both endogenous and exogenous surrender)

Figure 19: Survival Curves Index-Linked, $g_w = 2.5\%$ (both endogenous and exogenous surrender)
Figure 20: Frailty Processes Index-Linked, $g_w = 2.5\%$ (both endogenous and exogenous surrender)