UNIVERSITÀ DEGLI STUDI DI TORINO

SCUOLA DI MANAGEMENT ED ECONOMIA

DIPARTIMENTO DI SCIENZE ECONOMICO-SOCIALI
E MATEMATICO-STATISTICHE

Corso di Laurea Magistrale in Scienze Statistiche, Attuariali e Finanziarie

Master of Science in Quantitative Finance and Insurance

An Agent-based Model for Lapse Risk in Life Insurance

Supervisor:
Prof. Pietro Terna

Examiner:
Prof. Sergio Margarita

Candidate:
Francesco Checcacci

ACADEMIC YEAR 2014/2015

Ai miei genitori.

Acknowledgments

My thanks go to professor Pietro Terna, my teacher and supervisor, for introducing me to agent-based models and for his guidance in writing my thesis.

Contents

Like most people, actuaries prefer staying in their ‘comfort zones,’ or areas of familiarity, particularly when it comes to research, where most of our efforts are spent digging further into known models and topics. This is important, as a science needs to refine and extend its core knowledge base. But the tendency to stay close to home is fine for the health of a science only as long as the underlying environment itself is relatively stable.

Unfortunately, we are in an unstable, complex, evolving environment. Our employers face serious problems that span insurance, finance, the capital markets, and the economy as a whole. These problems reach across multiple professions—many comfort zones—and thus have no ‘owner profession.’
Being the professionals who, by our own proclamation, ‘make financial sense of the future,’ it is incumbent upon us to step up and play a leadership role in formulating research solutions to these problems.

Shaun Wang and Donald Mango, 2003
(actuaries)

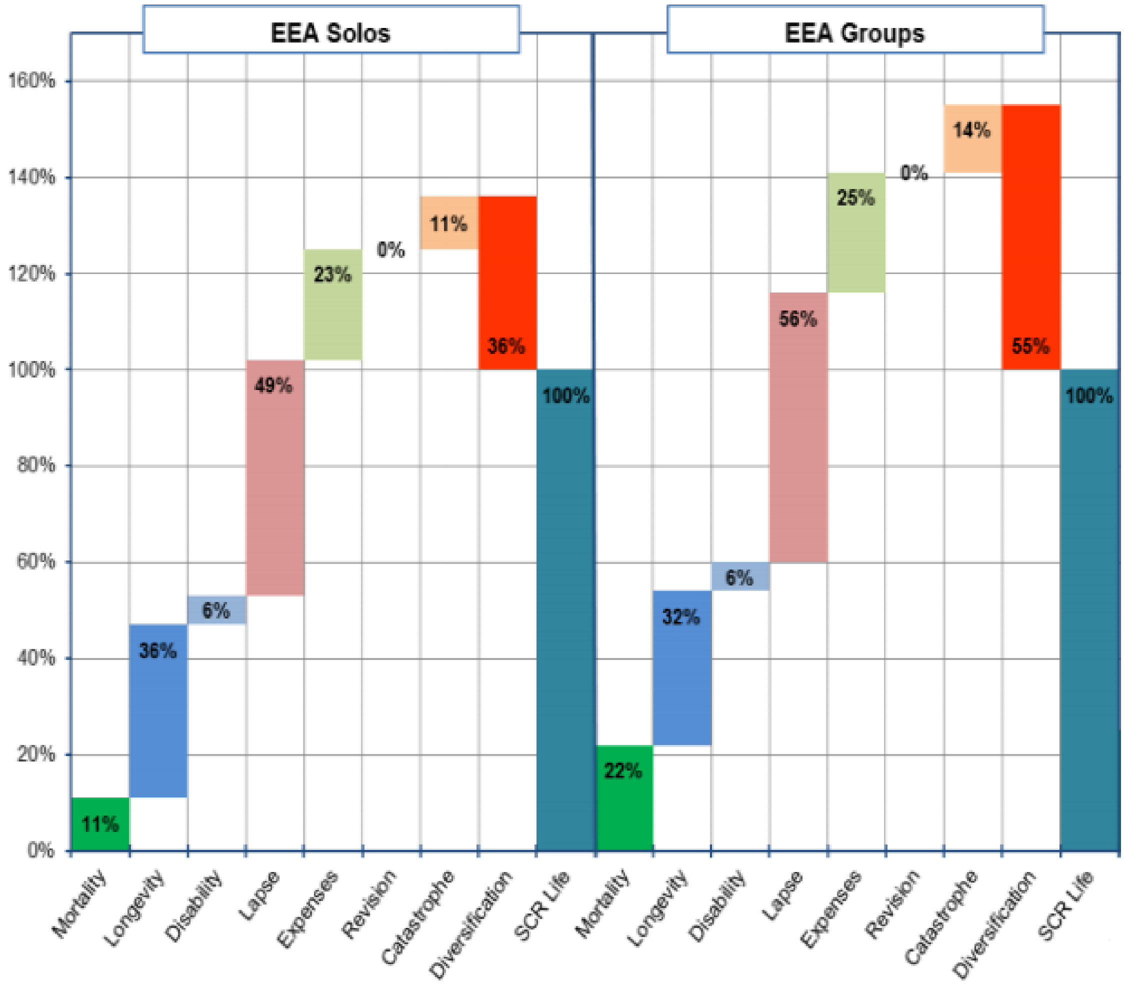
Introduction

Lapse risk is one of the main risks for an insurance company. In this thesis we refer to lapse risk for the life insurance business.

Under the European Union Solvency II Directive, which aims to harmonize the EU insurance regulation, insurance companies will have to provide a Solvency Capital Requirement (SCR) to absorb losses due to any type of risk with a 99.5% confidence level interval (EIOPA, 2013).

Weindorfer (2012) described the composition of the SCR for groups and solos belonging to European Economic Area (EEA) in 2012.

We report the figures for the SCR appointed to cover risks particular to the Life Insurance Business.



Without considering the netting effect of diversification, Lapse Risk alone is responsible for about half of the capital insurance company must provide for in order to cover from risks relative to the life insurance business.

The peculiarity of lapse risk, with respect to other risks, is that it depends on behavior and it is therefore more complex to model and subject to larger confidence intervals than biometric risks.

Agent-based model (ABM) is a tool to model a complex system. Combining data and the rules upon which agents act and interact we can observe how the events unfold.

The main purpose of this thesis is to build an agent-based model that would evaluate the effects of policyholder behavior and management countermeasures on the liquidity of the firm.

The life insurance business is notoriously not as liquidity dependent as the banking system; however The European regulator included a massive lapse scenario, similar to a "bank run", to be accounted for in defining the SCR.

After the concerns rose by the EIOPA, Financial Stability Board and the International Association of Insurance Supervisors on liquidity risk as a possible systemic risk emanating from the financial sector, the Geneva association (2012) published a paper on surrender and liquidity risks as well.

Given this background, in paragraph 1.1 we first describe the various policyholder behaviors and options that can impact the liquidity and solvency of firm and then in paragraph 1.2 we focus on the surrender options and its determinants.

In the second chapter we introduce agent-based models and their applications to life and non-life insurance. In particular, in paragraph 2.1 we highlight the importance of switching from aggregate level data to micro data relative to the single policyholder to model a complex system.

In the third chapter we describe the results of six recent papers which used generalized linear regression to model lapse risk with data relative to the single policyholders.

In particular we bring to the attention of the reader the logit regression by Milhaud et al. (2010) reviewed in paragraph 3.3 as we will use the coefficients provided in their paper to define the lapse probability of the last version of the model in paragraph 4.4.

The model is described as a project diary to account for the modifications I applied in three steps: paragraph 4.2, 4.3 and 4.4.

The first version of the model in paragraph 4.1 is not propaedeutic to understand the following three paragraphs. It is a different model whose purpose is to analyze cash outflows, accounting for lapse risk, allowing for the insurer to influence policyholders in their decision. This first version of the model has not been completed and suggestions on how to develop this model will be provided in the last paragraph.

In paragraph 4.2 we consider lapse and liquidity risks for an insurance company having in its portfolio only pure endowments.

In paragraph 4.3 we introduce other two types of contract and we facilitate the user definition of the relevant features of the contracts: duration, underwriting age and sum insured.

In paragraph 4.4 we introduce a lapse probability of the contract particular for each policyholder dependent on his characteristics and the coefficients provided by Milhaud et al. (2010).

In chapter 5 it follows a comparative analysis of the results of the three models. Further guidance on the results obtained can be found in the conclusions.

In the last paragraph 'further developments' we suggest model extensions necessary to make it an operational tool for insurance companies to evaluate their liquidity and lapse risks.

Chapter 1

Lapse Risk in Life Insurance

1.1 Policyholder Behavior

There are many aspects of a life insurance contract that depends on the behavior of the insured party: lapse and surrender rates for contracts that offer the possibility of an early withdrawal, take up rates for guaranteed annuity option (GAO), the possibility of suspension of the premiums, the option to extend the policy and many other.

The Milliman actuarial and consulting firm conducted a survey on the use of models to deal with the choices made by the policyholder (Clark et al., 2013). They included major insurance companies across Europe, the US and Japan.

Among traditional contracts in the life insurance business in Europe they classify two main types of products: participating business and unit linked contract without guarantees.

The first one contains guaranteed benefit features, like a guaranteed minimum rate, and it may include additional benefit depending on the performance of a pool of assets of the insurance company; the possibility to include additional benefits in the contract is called ‘profit sharing system’.

The second type of contract is a unit linked without guarantees where the value of the investment is linked to an underlying asset, the unit.

For these two traditional types of contract the survey found that policyholder behavior modeling is more common for early surrender of the contract rather than GAO and other features.

It is nonetheless opinion of the authors that companies should model also the take up rates for GAO as the increase in longevity could make the option to annuitize more attractive.

Yaary (1965) showed that a rational consumer, with no bequest motives and access to an actuarially fair annuity, chooses to annuitize all his wealth, yet the majority fails to annuitize even a small portion of their wealth, but this tendency may change in extreme economic condition.

For now take up rates are low and this is why the majority of insurance companies do not model this behavior and their primary concern are lapse and surrender rates.

Before going further it the case to distinguish between the lapse and surrender of a policy. The surrender of a policy implies the will of the insured party to withdraw it, while a policy may lapse due to a failure on the insured party to pay the premiums, although the difference is clear, in much of the literature these terms are used interchangeably.

The European Union Solvency II Directive, which aims to harmonize the EU insurance regulation, tries to quantify “lapse” risk for capital requirement purposes, therefore in this framework lapse risk is defined as (CEIOPS, 2009):

> (...) the risk of loss, or of adverse change in the value of insurance liabilities, resulting from changes in the level or volatility of the rates of policy lapses, terminations, renewals and surrenders.

Capital requirement is not the only reason why insurance companies should model this behavior. Lombardi et al. (2012) sum them up in their paper:

> (a) market insurance products, (b) price products and evaluate product profitability, (c) compensate agents and advisors for acquisition and retention of policyholders, (d) value assets, liabilities, reserve and capital for various economic conditions, and (e) transfer or hedge the risks.

1.2 Emergency Fund, Interest Rate and Policy Replacement Hypothesis

According to the Milliman survey, the first variable used to model surrender rates is the difference between a reference rate and the credited interest rate to the policyholder.

The above is called ‘Interst rate hypothesis’ (IRH) and it is the most widely used method to model surrender rates: if a competitor, or the market in general, provides higher interest rates, surrender rates will increase, vice versa, they will decrease.

This hypothesis matches the analysis of lapse rates in the German insurance industry from 1997 to 2009 made by Kiesenbauer (2012), yet the author found that this hypothesis is better reflected into lapse rates for unit linked contract. These are highly speculative contracts and most of the policyholder assets are allocated in the equity segment, therefore the motive that drives the investment in this type of contract is mainly the pursuit of high returns.

Moudiki and Planchet (2014) for example considers a systematic and unavoidable 2% surrender rate to which they add a conjunctural 5% surrender rate whenever the unit linked to the contract falls below the initial value invested in it.

This theory, however, fails to explain lapse rate for endowment policies where motives are not strictly return driven.

Another credited theory is the ‘Emergency fund hypothesis’ (EFH) (Fier and

Liebenberg, 2013) according to which lapse rates are caused by financial distress. The policyholder that finds himself unable to satisfy his consumption needs starts using the cash value of the policy as an ‘emergency fund’. The policyholder buys the endowment for bequest motives for example, or to provide for his retirement, but he ends up using the fund for a different purpose.

Also this theory has a reflection in the correlation between unemployment and surrender rates experienced in Korea from 1997 to 2002 as pointed in the paper of Haefeli and Ruprecht (2012). In this case policyholders used their life insurance as a substitute to unemployment insurance.

Another theory found in the literature is the ‘Policy Replacement Hypothesis’ (PRH) (Outreville, 1990) according to which a policyholder surrenders a policy in order to buy a new one with better terms.

At the 2011 Society of Actuaries (SoA) meeting in Chicago was presented a survey (LIMRA, 2011) on a sample of households representing the “middle market” (35-100 thousand dollars salary) for life insurance products. Among the top reasons for the purchase of a life insurance product at the fifth place it figures the replacement of another life insurance policy.

Chapter 2

Agent-based Models in the Insurance World

2.1 Aggregate level modeling versus household level data

The analyses described so far are made at an aggregate level. In both cases surrender rates are modeled through the use of macro variable: interest and unemployment rate.

As the paper of Lombardi et al. (2012) describes, aggregate level modeling poses

> (…) little or no differentiation of policyholder behavior based on different socio-demographic, attitudinal or behavioral factors. Such an aggregate level analysis fails to account for the value that different policyholders place on certain features.

Fier and Liebenberg (2013) instead used household-level data to model lapse rates and found that, to a substantial negative income shock corresponds a 25% increased likelihood of surrendering the policy.

Campbell et al. (2014) made another step forward into the field of behavioral economics.
In three paragraphs they make a review of the literature on behavioral economics applied to: purchasing, annuitization and surrender behavior. In these paragraphs they consider the psychological factors which may drive policyholder choices.

The literature on behavioral economics applied to policyholder behavior is scarce and so it is its use in the insurance industry but offers many insights into the decision making process of the policyholder.

In the last paragraph they make some suggestion on how to move from traditional techniques to a better understanding of policyholder behavior and how to modify it.

> We need to move beyond concluding “a random 3 percent of our policyholders will lapse this year” to understand the underlying decisions that led to it and identify which 3 percent of the population it is likely to be, and how their decisions could be influenced, so as to understand what can be done today to change behaviors and how behaviors might evolve in the future under different scenarios.

Once we brought our analysis beyond aggregate level data to policyholder level data to exploit behavioral economics and predictive modeling another problem must be dealt with, quoting from Lombardi et al. (2012):

> (...) predictive modeling relies on historical experience to predictive future experience. Thus, it is not very reliable predicting future experience when there is a fundamental change in the environment.

2.2 Agent-based modeling in the insurance industry

Alan Mills in his paper "Complexity Science: An introduction (and an invitation) for actuaries" (Mills, 2010) suggests the use of Agent Based Simulation (ABM) to model policyholder behavior.
In ABM we have a set of agents, each with their peculiar characteristics, acting and interacting with each other and the environment.

Through ABM we are able to change environmental condition for which we have no data, this allows us to observe, in our model, how agents would act in a setting of circumstances we never experienced.

With regard to the insurance sector, the data the industry has, on policyholder behavior at the micro level, is scarce (Campbell et al. , 2014).

To have a better picture, let me use an example and quote the authors Casti and Jones (2000) and their ABM called "Insurance World".

> Insurance World simulator grew out of the wish on the part of several firms in the catastrophe insurance industry to create a kind of "laboratory" with which to experiment with various risk scenarios. For instance, suppose a force 5 hurricane drops on Miami Beach. Who are the winners and who are the losers in the aftermath of such a major catastrophe? And what kinds of firms spring up to fill the niches left behind by those firms that disappear after such an event? These are the type of questions that are of vital concern to catastrophe insurance firms---but for which no experimentation is possible.

Although this work is related to the non life insurance business and it does not involve policyholder behavior analysis, it studies the effects of an event (which did not happen), not only in catastrophe insurance industry itself but also on the interaction with the capital markets,

> (...) developing market share, repayment of loans, attitudes toward risk, amount of risk assigned to reinsurers and so forth. In this way, various management strategies can be tried in an attempt to find those that lead to prosperity, or at least avoidance of extinction.

The success of this project brought to "Insurance World 2" which considered a broader range of risks and interactions.

Once we are able to recreate the interactions between the agents and environment (thanks to historical data, behavioral economics or other credited theories) we can run the model under different scenarios, stressing some factor or parameter.

At this point we can address another major bias of the insurance industry and modeling practitioners in general, the belief that the purpose of modeling is prediction. Prediction is not the only concern to the insurance industry. Quoting from the paper of Alan Mills:

> Many actuaries believe the primary purpose of modeling is prediction, to predict policyholder events, economic measures like interest rates, and health care expenditure trend rates. But prediction is only one of the many purposes of agent-based models, and – especially considering the futility of trying to predict the long-term behavior of complex systems – not the most important.

Joshua M. Epstein (2008) wrote an interesting paper named “Why Model?” which identifies 16 reasons other than prediction to model a phenomenon. With the help of the literature I will review a few of them.

Explanation. The “Wealth Distribution” (Wilensky, 1998) ABM, which is an adaptation from Epstein and Axtell’s “Sugarscape” model (Epstein and Axtell, 1996), simulates the distribution of wealth in an artificial world where wealth is represented by grains distributed in the world. The model shows the Pareto’s 80-20 law in action: a large number of “poor” people and a few “rich”ones. The model cannot predict the level of wealth of an agent, or if he will run out of grains, it is does not even rely on real data, but it offers some insight on the causes of wealth inequalities and how to reduce it.

To guide data collection. The words of Epstein better consolidate the idea:

> From (Maxwell’s) equations the existence of radio waves was deduced. Only then were they sought … and found! General relativity predicted the deflection of light by gravity, which was only later confirmed by experiment. Without models, in other words, it is not always clear what data to collect!

Mills revisited some of them in an insurance logic:

> Real time crisis management.
> Even in cases where we cannot predict, models can help us understand plausible outcome ranges. Using sensitivity analysis, we can explore a range of parameters to identify salient uncertainties, regions of robustness, important thresholds, and possible outcomes.
> Training.
> Models can help actuaries and executives better understand and develop intuitions about the behavior of the complex systems in which they operate. (...)
> Policy Support.
> Policy proposals are often developed by people who do not understand the complex systems for which they are proposing change. By modeling the behavior of such systems, actuaries can support policy development.

In his final paragraph the paper of Campbell et al. (2014) suggests to use behavioral economics for

> Applying insights from behavioral economics to guide better product design;
> Applying the insights to help “de-bias” policyholders in their decision making processes.

Chapter 3

Generalized Linear Models for Lapse Risk

3.1 Modeling lapse and surrender decision in the literature

Most of the literature on this topic focuses on an analysis of lapse rates made at an aggregate level and explore the EFH and the IRH using macro variables.

Eling and Kochanski (2012) makes a review on more than 50 theoretical and empirical papers. Among the 12 empirical ones, 7 relied on aggregate level data, 5 were based on products and policyholder characteristics. Research based on policyholder level data is limited as this type of data is not easy to obtain from insurance companies. Each of these paper used generalized linear models (GLM) in their attempt.

Of these five papers I am going to review: Cerchiara et al. (2009), Kagraoka (2005), Milhaud et al. (2010) and Eling and Kiesenbauer (2011). Then in paragraph 5.1 and 5.2 I will comment on two papers published after the paper of Eling and Kochanski (2012): Fier and Liebenberg (2013) and Liebenberg et al. (2010). Notice that the second paper was presented at the Risk Theory Society annual seminar in 2010 but it was published on the Journal of Risk and Insurance only in 2012 and therefore was not included in the review by Eling and Kochanski (2012). Finally in paragraph 5.3 I will comment on selected explanatory variables.

Most of the following papers, either tried to model lapse rates, the number of lapses, or the probability of a lapse. The first two are usually modeled with a Poisson regression since such variables are positive real numbers. The third is a probability, between zero and one, and is modeled through a binomial model. The reader should pay attention to which is the variable modeled (sometimes referred to as the response variable or Y).

3.2 Negative Binomial Model.

Kagraoka (2005) distinguish between two types of model to analyze event data, such as the surrender decision of a policy holder: the binomial regression (using for example logit or probit link functions), which models the probability of the event, and count regression which models directly the number of events.

Kagraoka argues that, as the surrender of an insurance contract is a rare event, using the binomial model we would have few observations which correspond to high incidences1, therefore he finds the count regression model to be more appropriate. Kagraoka models the number of surrenders, within a month, of pools of contracts with same attributes of

[1] On this topic see 'Dealing With Zeros in Economic Data' by Humphreys (2013).

the policyholders and same origination month of the contract (contract date). The attributes of the policyholders are: gender and age.

Among count regression models we already named Poisson regression. The problem with Poisson regression is that it makes the following assumption on the dispersion of the response variable:

$$VAR[Y] = E[Y] = \exp(X\beta)$$

Kagraoka uses instead a negative binomial model which allows for a greater dispersion of the variable with respect to its mean, moreover Kagraoka chooses a particular version of the negative binomial regression, called Negative Binomial I (Negbin I)2, which enjoys the following reproductive property as explained in the paper:

> Assume Y1 ~ Negbin I(μ1, σ) and Y2 ~ Negbin I(μ2, σ) are independently distributed. It follows that the random variable Y = Y1+Y2 is negative binomial distributed Y ~ Negbin I(μ1+μ2, σ).

Such property is needed due to the different number of insurance contracts in each pool. As we said before, the variable modeled is the number of surrenders in a pool, which depends on how many active insurance contracts are in the pool. The Negbin I model adjusts the prediction to the number of contract in the pool.

For comparison purposes he also implements a Poisson regression (which enjoys the reproductive property too):

> The intensity of the Negbin I and the Poisson models are parameterized as $\mu_i = n_i \cdot \exp(\beta X_i)$ where $n_i$is the number of active contracts in the i-th pool, X_i is a covariate vector (the explanatory variables) of the i-th pool, and β is a vector of regression coefficients.

The explanatory variables are: the attributes (gender and age), seasonality, time elapsed from the contract date and change in unemployment rate (since the underwriting year).

The data was collected from a single insurance company between 1993 and 2001, the contracts considered are annuities (some other insurance features are added to the contract as in Japan is not possible to sell plain annuity contracts) and the policyholder and the beneficiary are the same person.

Since the author finds the incorporation of the dispersion parameter to be statistically significant, the paper focuses more on the results of the Negbin I model:

> Every explanatory variable is meaningful to model surrender. The unemployment rate difference prominently explains surrender among others, and time elapsed form the contract date follows. The number of surrender increase as

[2]Kagraoka refers to Cameron and Trivedi (2013)

unemployment rate rises from the contract date. (…) Surrender rises as time goes by, takes peak at three year, and gradually falls over three year. This tendency is interpreted as follows. There are two types of policyholders with respect to their financial condition, stable and unstable. The unstable policyholders surrender their insurance contracts as time passes, and the number of surrender increases. There remains fewer unstable policyholders in the pool as time elapses, and the number of unstable policyholders decrease. The number of surrender of female is fewer than that of male. It is considered housewife takes out insurance if the household is affordable. The number of surrender peaks around March[3]. Surrender become fewer as older the age at the contract date.

The coefficients of both regressions are summarized in table 3.

Important results are drawn from a deeper analysis on the relation between unemployment rate and surrenders:

> We investigate whether surrender is retarded or advanced to the change of unemployment rates. We change a lag of unemployment rate from one to six months, and repeatedly estimate the parameters. The log-likelihoods are given in Table 4. The log-likelihoods decrease as the lag is greater, and the best model is without lag. This fact implies that surrender is not retarded to the change of unemployment rates, and policyholders surrender their insurance contracts before their economic conditions get worse.

3.3 Logit Regression

Milhaud et al. (2010) study lapse decisions on a data set of 28506 policyholder provided by the Spanish life insurer AXA Seguros, covering the period from 1990 to 2007. The study focuses on endowment policies.

They used the Classification And Regression Trees (CART) model (Breiman et al.,1984) and a Logit regression. For the CART model, the authors provide a theoretical framework in their paper.

The results.

The first splitting-rule in the classification tree, which is the more discriminant variable for the surrender binary response variable, is the presence of a profit participation benefit in the contract. Policyholders with the profit benefit (PB) option are more likely to surrender their contract. The authors provide three possible explanations for this:

[3]Quoting from the paper: “In Japan, school and business term start in April and an annual income is determined at the end of fiscal year. Many new insurance contracts are taken around April, and surrender and transfer of the insurance contracts occurs in this season. We caution that an unusual quarter period is defined based on the recognition of the seasonal pattern in the surrender; the first quarter starts November and ends in January and so on”

> (…) first people move to a new product which globally offers a higher PB, second a high PB in the first years of the contract enables the policyholder to over perform the initial yield and could lead her to surrender the contract and recover the surrender value, third someone with a PB option simply receives frequent information on it and on the surrender value which can prompt her to surrender.

In agreement with Eling and Kiesenbauer (2011), policyholders who have chosen a periodic premium are more likely to lapse, in particular, the highest surrender activity is experienced for annual and bi-monthly premiums.

With regard to timing of the lapse it is found that lapses are more frequent when the period in which the policyholder would have to pay a penalty in case of surrender stops; thereafter lapse risk increases.

Finally underwriting age and gender coefficients do not seem statistically significant.

3.4 Poisson Models

Cerchiara et al. (2009) considers a data set of life insurance savings policies from 1991 to 2007. They model lapse "experience" using a GLM of Poisson with a logarithmic function linking the linear predictor to the expected value of the variable modeled:

$$\log(E[Y]) = X\beta$$

Where Y is the variable modeled, X is a vector of covariates and$\beta$the coefficients. In their paper, section 3, they provide a theoretical framework for GLMs.

The risk factors they used were: the type of product, the duration of the policy and the calendar year of exposure. It is not clear which is the variable modeled. The author themselves have some doubts about the consistency of the data sets they were provided.

They found that a duration in force of the policy of two and five years positively influences lapses, while it is very unlikely a lapsation of the policy within a year from the underwriting and at the tenth year. They interpret the very low lapse experience after ten years of policy duration

> (…) as a form of automatic policyholder selection: those policyholders who have not surrendered in the first ten years seem particularly unlikely to countenance surrender thereafter

The calendar year of exposure has a high explanatory power: from the -37 % low in 1999 to the +26% high in 2007.

There is a considerable difference in lapse experience among 13 different product groups. It is very interesting to see how some groups of product are not affected from the 2007 crisis or how different durations of the policy do not affect the lapse experience form some groups whereas other groups are more affected. They grouped the products based on some characteristics at their discretion and they explain that their choice may bias the results. They do not share in the paper the characteristics of the product groups.

Eling and Kiesenbauer (2011) were provided the data by a German life insurer including 2.5 million policies for a total of 8.9 million policy years, the largest data among the empirical literature.

They analyze the period between 2000 and 2010, therefore including the stock market plunge in the early 2000 and the 2008 crisis.

The data includes 7 different product categories including traditional insurance products and unit linked ones.

Three types of lapse rates are modeled: early lapse, late lapse and total lapse:

> The early lapse rate is defined as ratio of all lapses without surrender value over new business written. The surrender value and thus this ratio strongly depend on the product design, e.g., term life insurance has a very limited surrender value for the entire contract duration. The late lapse rate is defined as ratio of all lapses with surrender value plus all policies made paid up (i.e., the customer stops or reduces premium payments but does not lapse the contract) over sum insured of the entire portfolio at the beginning of the calendar year.10 Again this ratio depends on the product design. The total lapse rate is given by the aggregated sum insured of early and late lapses divided by the average volume of business in force during the calendar year (i.e., half of the portfolio sum insured at the beginning and at the end of the calendar year).

The lapses are either quantified as: premium volumes, number of contracts or sum insured.

The analyzed variables are: calendar year, contract age, gender and premium payment (single vs. regular). They extend current literature by adding other variables: remaining policy duration, distribution channel and supplementary cover.

They use a Poisson GLM to model lapse rates. First they make a regression with respect to each single variable, then, due to the large amount of data, they are only able to see the interaction effects of two variables at a time. I will review only the analysis without interaction.

Some of the above characteristics change during the life of a contract. The authors deal with this problem in the following way:

> Each contract is split into all possible combinations of considered product and policy(holder) characteristics. We denote such a combination of characteristics as model point. A sample model point is: endowment (product type), 2005 (calendar year), 5 (contract age), 35 (remaining policy duration), 25 (policyholder age), broker (distribution channel), no (supplementary cover), male (policyholder sex), and regular (premium payment). For each model point, the exposure of all contracts in the portfolio needs to be determined, i.e., the time (measured in years on a daily basis) all contracts belong to the corresponding model point. Finally, we determine the number of early and late lapse events for each model point. Each lapsed contract is counted as lapse in the model point which represents the product and policy(holder) characteristics at the lapse date. (…)
> The time each contract belongs to a certain model point u is used to define the exposure (time)e_u. Latter is the sum of the time that each contract belongs to the model point u (in years, possibly zero) taking into account all contracts in the portfolio. Introducing additionally the lapse rate lr_u, we can rewrite the GLM as $E(Y_u) = e_u \cdot lr_u = \exp(\beta X)$. (…)
> The above discussion focuses – strictly speaking – on the consideration of number of contracts. The same line of argument can, however, be applied for lapsed regular premiums and lapsed sum insured assuming that each single Euro can be lapsed or not. Thus, the same modeling approach is used, except that the exposure is measured in Euros instead of years. The exposure of regular premiums and sum insured for the model point of a single contract is determined as product of timely exposure (in years) multiplied by yearly premium and total sum insured, respectively.

In table 4 they report the coefficients of the regression of the total lapse rates, measured in the three different ways we said before, where each covariate is take one at a time (no interaction is considered here). For the Poisson regression the coefficients are easily interpreted by exponentiation of the coefficients:

> (…) a value of 0.15 for endowment (using number of contracts to measure exposure) means that the lapse rate for endowment policies is exp(0.15) = 1.16 times the lapse rate of annuities representing the reference level; in other words, the lapse rate for endowments is 16% higher than the lapse rate of traditional annuities

In the figures that follow they use the same type of graph used by Cerchiara et al. (2009) to effectively represents the coefficients; the reference level is set to 0 in the ordinate axis. The graphs, as table 4, represent the results for the three different lapse measures. Such distinction is particularly important for term life insurance where there is no surrender value and therefore any lapse for these policies is considered an early lapse.

The results.

Little difference of total lapse rates is experienced among the products. Moreover the results contrast the analysis of the less recent paper of Renshaw and Haberman (1986) where the highest lapse rates were experienced for unit linked products which traditionally did not include any type of guarantees. Recently in the German market some kind of guarantees were introduced for unit linked product which, according to Eling and Kiesenbauer (2011), experience the lowest lapse rates.

The results on the coefficients for the calendar year are coherent with those of Cerchiara et al. (2009) with a general increasing trend of total lapse rates nearing the 2008 crisis and a decrease after 2009.

Total lapses are the highest for policies at the beginning of their life and then they decreases. The policyholder quickly figures out if he will be able to afford to pay the premiums and if in general the policy fits his need, therefore either he lapses in the first years of the contract or he will keep the policy until maturity.

Lapse rates are increasing with respect to remaining policy duration. This is the first paper which considers this explanatory variable.

The results and the authors' interpretation on the policyholder age explanatory variable:

> Three age groups can be distinguished: policyholders until age 25, policyholders between 26 and 40, and policyholders older than 40. Policyholders in the middle group have an almost constant lapse rate at the level of the reference age 39. The lapse rate for the youngest policyholders is significantly below, but steadily increasing. Such policies might be initially 'sponsored' by the policyholder's parents. When the family circumstances change (e.g., marriage or birth of children) the needs might change and the insurance premiums are not affordable any longer. Lapse rates for the oldest age group are steadily increasing until age 60, before decreasing again. For products with a savings component, a possible explanation for this effect is that especially people older than 50 might have difficulties to find a new job in case of unemployment. According to the emergency fund hypothesis (see Outreville, 1990), those persons might access their life insurance savings as emergency funds. Other customers in the late 50's might choose to retire early and use their life insurance savings to bridge the gap until the payments from the social pension scheme start.

It is found that to an independent distribution channel (tied agents or brokers) corresponds a lower lapse activity, while to a bank channel corresponds a higher lapse activity. The authors explain that a bank agent may be focused on short term sales having his customers realize later they did not need a policy, while an independent agent should focus on maintaining a long term customer relation.

The inclusion of a supplementary cover in the contract (term life insurance, disability insurance, accident insurance, and surviving dependents insurance) is accompanied by higher lapse rates:

> (…) the premium for policies including additional cover is higher than for stand-

alone policies. In case of financial distress, it is more likely that a policyholder is forced to lapse such a product bundle. Additionally, Pinquet et al. (2011) believe that customers' insufficient knowledge of insurance products can cause lapse.

With respect to gender, women are less likely to lapse an insurance contract in agreement with Kagraoka (2005).

In agreement with Milhaud (2010), single premium policy experience 90% less lapse activity than regular premiums. As explained in Campbell et al. (2014) policies that have become self-supporting may seem “free” and not easily given up. This concept in Behavioral Economics is referred to as “Love of the free”.

For the other paper (Renshaw and Haberman, 1986) I refer you to the review made by Eling and Kochanski (2012).

Other than the papers reviewed by Eling and Kochanski (2012) I found other two papers using GLMs to model lapse decisions. I describe them in the next two sub paragraphs.

3.5 Logistic regression for voluntary lapse decision.

In this sub paragraph I am going to describe the paper of Fier and Liebenberg (2013) called ‘Life Insurance Lapse Behavior’ which specifically models the probability of voluntarily lapse the policy. The authors did not include in their research those households whose lapsation of the policy was due to a third party, for example as a consequence of a work benefit lost after a change of job.

The terminology used in their paper is different from the one given in my introduction. Following Kuo et al. (2003) the term ‘lapse’ identifies both surrender activity (electing to receive the eventual cash value of the policy) and lapse activity (failure to pay the premiums).

The data set used in their work is a set of surveys, each at a two year distance, from 1996 to 2006, collected through the University of Michigan Health and Retirement Study (HRS).

They tested the EFH, through the use of the unemployment and income and wealth shocks household variables. For unemployment they consider a household where a component of which recently lost a job. For the shocks they considered four quartiles of gravity of the shock where the fourth one includes the smallest or positive shocks.

They tested the PRH checking whether or not the household bought a new life insurance product.

They tested whether the lapse of a policy was related with some important life cycle variables: age, marital status and retirement age. Age was divided in quartiles where the first quartile included the youngest observations of the data set (46-62). The age of the households is defined as the age of the household's primary respondent. The marital status variables checked whether the component of the household had recently divorced or lost his/her spouse.

It is important to notice that these variables, which were used for the regression, are dynamic. The dependent variable, the lapse of the policy, is explained through an event which dynamically modifies the household characteristics, the loss of a job, the income and wealth shocks, the purchase of another policy, a divorce, the retirement and loss of the spouse, and not on the current static state of the households (whether he is employed, married, retired and so on)

Another set of static control variables was included: Income, net worth, debt, liquidity, the number of children, employment status and education.

Due to a non linear age lapse relation, the model has been tested with two different specifications. The first captures non linearity by setting the age variable as continuous and making a regression with respect to the age squared. The second specification divides the ages in four quartiles as specified before. The results found were consistent with both model specifications.

The EFH is reflected in the fact that, according to the logistic model, households that suffered a substantial negative income shock were 25% more likely to lapse. By substantial negative shock they refer to the first two level of gravity of the income shocks. The theory is supported also from the positive relation between lapse and first level wealth shocks. No relation was found with the loss of a job but this may be due to the fact that only voluntarily lapse of the contracts were considered.

With regard the PRH they found that around 13.7 percent of households that voluntarily lapsed a policy also purchased a new life insurance policy, while only a 2.4 per cent of those that bought a new life insurance product did not lapse another policy.

Life Cycle variables.

The first three youngest quartiles (46-76) were found to be more likely to lapse with respect to the elderly households which was picked as base line in the logistic regression in the second model specification, while the positive and negative coefficients for the first model specification (age and age squared) suggest an increasing and then decreasing probability of a lapse of the policy with respect to age. Retirement and loss of the spouse life cycle events were found to be positively correlated with the lapse of a policy.

A Third type of analysis was conducted on the sample. It was divided in four sub samples based on the four age categories. It was found that lapse determinants differ for age category. While the PRH holds for all ages, we can reject the null hypothesis that the coefficients for a first level income shock equals zero only for the first two age quartiles, therefore the EFH holds for the youngest quartiles (46-68). Following the same logic it was found that divorce and retirement life events are significant and positively related to the lapse decision only for younger households while to the loss of the spouse corresponds a positive and significant increase in the odd ratio of $e^{0.6229} \simeq 1.86$only for households between 69 and 76:

$$\frac{\frac{P(lapse: widowed)}{1 - P(lapse: widowed)}}{\frac{P(lapse: notwidowed)}{1 - P(lapse: notwidowed)}} = 1.86$$

The above formula can help interpret the coefficients in the multivariate settings (table 5 of their paper).
The paper contains the coefficients of the models and level of significance for each variable used.

3.6 Cragg's two part regression for life insurances drop and purchase

The less recent analysis made by Liebenberg et al. (2010) called 'A Dynamic Analysis of the Demand for Life Insurance' considered two surveys made by the Survey of Consumer Finances (SFC), one made in 1983 and the other in 1989.

The purpose of their study is to capture the effect of life cycle events related to life insurance demand, for this reason the dynamicity of the surveys is even more stressed with respect to Fier and Liebenberg (2013):

> A common thread among prior empirical analyses of life insurance demand is that researchers have employed a static (rather than dynamic) framework. Such an approach is reasonable given that household survey data are largely cross sectional and generally do not follow the same households over several time periods. While cross sectional data are useful in explaining current levels of insurance holdings in terms of current household characteristics (like marital status, employment status, and number of children), cross sectional data do not allow for an analysis of the impact of life events (e.g., getting married, changing jobs, or having a child) on life insurance demand.
> Limitations of data used in previous life insurance demand research are important because the life cycle literature suggests that life insurance purchases are likely to follow various "life events" such as getting married, having a child, purchasing a home, and getting a new job, for which cross-sectional data are not well-suited. Similarly, termination of life insurance is likely to follow other life events such as getting divorced, having a spouse die, becoming unemployed, and retiring. While previous literature has provided a strong theoretical foundation as well as empirical evidence on the determinants of life insurance demand based on cross-sectional data (within a static framework), the literature has not employed panel data (within a dynamic framework) to examine the various life events hypothesized to relate to life insurance demand.

The SCF surveys, with respect to the HRS surveys, do not ask explicitly whether the policyholder voluntarily lapsed the policy; In the 2010 paper they checked life insurance holdings in 1983 and then in 1989, in this way the analysis is not able to tell if the lapse was due to the will of the policyholder. For example they found that households who were more likely to buy a term life insurance were those that started a new job, therefore the term life could be a benefit which comes with the new employment. Households in which either spouse became unemployed were more than twice as likely as other households to surrender their whole life insurance, this finding may be increased evidence for the EFH or could be due to a lost benefit from the previous employment.

They divided households in four categories with respect to changes in life insurance holdings: new term life owners, new whole life owners, households dropping term insurance, and households dropping whole life insurance.

Before going further in their analysis, it is important to notice that, since term life insurance provide a benefit only if the death of the insured happens within a certain time and therefore the benefit is not granted, the insurance companies are not inclined to allow for a surrender option as this may cause adverse selection: policyholder in a good health status may ask for the surrender value if they do not believe it is likely they are going to die before the established date. This type of behavior holds for any contract where the benefit is not certain, for example the same reasoning can be applied in the opposite case: a pure endowment, which pays only if the insured is alive at a certain date, may push the policyholder in a bad health status to ask for an anticipated withdrawal of the cash value. Moreover if a saving component is not included in the term life insurance, these contracts do not have a surrender value as they only provide a pure risk coverage (Eling and Kiesenbauer, 2011).

Their paper (Liebenberg et al., 2010) models two variable: the change in life insurance ownership status (newly dropped or purchased insurance) and the quantity, or extent, of insurance coverage dropped or purchased. To do so they used Cragg's two part model (Cragg, 1971). This model was used to separately analyze the determinants of the change in ownership status (binary variable) and volume of coverage dropped or purchased (continuous variable observable upon the condition that a change in ownership status has taken place).

Tobit regression has been used in previous literature to model lapse behavior4, but the assumption made on the response variable would lead to an improper modeling of ownership status and extent if the determinants of these two variables were different. The Tobit model is used whenever it is believed that the process that govern the "participation" decision (the binary variable) and the "amount decision" (the continuous variable) have the same determinants. Liebenberg et al. (2010) rejects this assumption.

[4]Cox and Lin (2006) in their paper 'Annuity Lapse Rate Modeling: Tobit or not Tobit' found it suitable for modeling lapse rates with aggregate level data.

The Cragg's regression, in order to separately model the the change in ownership status (take as an example the lapse of a policy binary variable) and the change in the amount of insurance coverage (the amount of coverage the policyholder is giving up subsequent to decision of lapsing the policy), uses two models respectively. The first model is a probit regression (examining the determinants of dropped policy status) estimated on the full sample. The second model is a truncated regression (examining the determinants of dropped policy size) estimated on the sub sample of households that dropped policies.

As this thesis is concerned with lapse determinants I will only review the results referring to dropped (and not purchased) term and whole life insurance: table 4 panel B for the analysis in the univariate setting, the Cragg regression coefficients in table 7 and the relative interpretation for the probit regression in table 8.

Univariate difference in mean of the determinants for households that dropped a life insurance.

The determinants whose change in mean have the highest statistical significance (less than 1 % probability that they equal 0) for the drop of the term life insurance are: the loss of a spouse, retirement and the purchase of a whole life insurance. With regard to retirement this may be due to the fact that often a term life insurance is a benefit tied to employment.

The determinant whose change in mean has the highest statistical significance for the drop of the whole life insurance is the purchase of a term life insurance. A negative variation in worth and the loss of a job have a 10 % significance.

Probability interpretation of the Cragg's coefficients for dropped term life insurance (status)

Recently widowed policyholders are 184 % more likely to drop than those who were not.
Recently retired policyholders are 53 % more likely to drop.
Policyholders who purchased a whole life insurance are 275 % more likely to drop.
The determinants whose coefficients have the highest statistical significance are the loss of the spouse and the purchase of a whole life insurance, while recent retirement has a 5 % level significance.

Probability interpretation of the Cragg's coefficients for dropped whole life insurance (status)

Recently divorced policyholders are 84 % more likely to drop than those who were not.
Recently unemployed policyholders are 110 % more likely to drop.
Policyholders who purchased a term life insurance are 263 % more likely to drop.
The determinant whose coefficients have the highest statistical significance is the purchase of a term life insurance, while recent divorce and recent loss of a job have a 10 % level significance.

Cragg regression coefficients for dropped term life insurance (amount).

The determinant whose coefficient has the highest statistical significance is the amount of whole life insurance coverage bought.

Cragg regression coefficients for dropped whole life insurance (amount).

The determinant whose coefficient has the highest statistical significance is the retirement.

3.7 Comparison of the literature.

In this paragraph I will name and comment those explanatory variable analyzed by more than one paper and make a comparison on the results.

In table 5 of their paper, Eling and Kiesenbauer (2011) compare their results to the ones of other papers; Their summary is to be integrated with the one that follows:

Age.

Surrender become fewer as older the age at the contract date for Kagraoka (2005). Eling and Kiesenbauer (2010) find that lapse rates are constantly increasing until 25 years of age, then they remain stable until 40, then increasing again until 60 before they decrease. According to Fier and Liebenberg (2013) the first three youngest quartiles (46-76) were found to be more likely to lapse with respect to the elderly households which was picked as base line in the logistic regression in the second model specification, while the positive and negative coefficients for the first model specification (age and age squared) suggest an increasing and then decreasing probability of a lapse with respect to age.

The view of Eling and Kiesenbauer (2011) and Fier and Liebenberg (2013) in the first model specification are consistent with each other as the latter considers a sample of the population where the minimum age is 46. In the second model specification, where age has been divided in categories (46-62; 63-68; 69-76; >76), the coefficients of the regression only captures a sharp decrease from the third to the older category as Eling and Kiesenbauer (2011) describe; but it does not capture the increase from 40 to 60 as to this change in lapse activity corresponds only one category (46-62). Altogether the views are consistent. Kagrakoa (2005) instead considers age at the contract date, therefore we cannot make a comparison.

Gender.

Both Kagrakoa (2005) and Eling and Kiesenbauer (2011) find that women are less likely to lapse; while in Milhaud (2010) gender coefficients do not seem statistically significant.

Unemployment.

At a macro level the correlation between unemployment and surrender rates experienced in Korea from 1997 to 2002 is increased evidence for the EFH.

Kagrakoa (2005) consider the change in unemployment rate since the underwriting year and finds that it positively explains surrenders moreover it is found that surrender is not retarded to the change of unemployment rates: policyholders surrender their insurance contracts before their economic conditions get worse.

Fier and Liebenberg (2013) found no relation with the loss of a job but this may be due to the fact that only voluntarily lapse of the contracts were considered, while instead Liebenberg et al. (2010) found that households in which either spouse became unemployed were 110% more likely as other households to surrender their whole life insurance (10 % statistical significance), this finding however could be due to a lost benefit from the previous employment.

Loss of a spouse.

Fier and Liebenberg (2013) found that to the loss of the spouse corresponds a positive and significant increase in lapse probability only for households between 69 and 76.
Recently widowed policyholders are 184 % more likely to drop a term life insurance (highest statistical significance) (Liebenberg et al., 2010).

Retirement.

Fier and Liebenberg (2013) found that retirement is significant and positively related to the lapse decision only for younger households (46-62).
Recently retired policyholders are 53 % more likely to drop a term life insurance (5 % statistical significance) (Liebenberg et al., 2010).
The determinant whose coefficient, for the amount decision of dropped whole life insurance, has the highest statistical significance is the retirement (Liebenberg et al., 2010).

Divorce.

Fier and Liebenberg (2013) found that divorce is significant and positively related to the lapse decision only for younger households (46-62).
Recently divorced policyholders are 84 % more likely to drop whole-life insurance than

those who were not (10 % statistical significance) (Liebenberg et al., 2010).

Contract age.

Cerchiara et al. (2009) found that a duration in force of the policy of two and five years positively influences lapses, while it is very unlikely a lapsation of the policy within a year from the underwriting and at the tenth year; the very low lapse rate experienced thereafter is interpreted as a form of automatic policyholder selection for which those who have not surrendered within ten years are not likely to surrender anymore.

In the data set of Eling and Kiesenbauer (2011) total lapses are the highest for policies at the beginning of their life and then they decreases. Their interpretation is that the policyholder quickly figures out if premium amount and type of insurance fit his needs. A similar result is found by Kagraoka (2005): the peak in lapse activity is experienced at the third year and then it gradually decreases.

Chapter 4

Project Diary

To build the model two free software has been used: the multi-agent programmable modeling environment NetLogo (https://ccl.northwestern.edu/NetLogo/) (Wilensky, 1999) and the software environment for statistical computing R (https://www.r-project.org/) (R Core Team, 2013). The project diary describes the model at different stages of development.

4.1 Version 1

Introduction

In this first version of the program we have our agents, the insured party, buy two products at the beginning of the simulation: a unit linked contract, which is similar to an investment fund whose assets are mostly allocated within the equity segment, and a policy, where the premiums are invested in a fund re-evaluated each year at a fixed rate.
In the second type of contract, as the insurance company must provide a minimum guaranteed rate, the assets are allocated in non risky investments, mostly government bonds.

In this simulation all the premiums paid for the unit linked contract are invested in shares, therefore any capital gain, if present, will be taxed accordingly to Italian regulation at a 26% rate, moreover the insured party is allowed to surrender the contract at any time and withdraw the market value of his investments without paying a penalty.

The capital gain realized through the policy will be considered as derived from government bonds, currently taxed at 12.5% rate.
While for the first contract the insured party bears the risk, here the the risk is on the insurance company which has to guarantee the fixed rate no matter the returns exploited in the market, therefore it reserves to itself the right to have the insured party pay a penalty in case of a premature withdraw of the cash value of his contract.

In both cases, unless the contract is surrendered earlier, the benefit is set to be received at the expected retirement age. In Italy retirement occurs either after 42 years of contribution requirement or at the age of 67.

In the literature there are many theories regarding the drivers of surrender rates.
In this program we consider the emergency fund hypothesis. According to this theory the most influential variable for explaining policyholder surrender behavior is the income: the insured party starts using his investment as an “emergency fund” due to the

lack of liquidity caused by a drop in his income. A strong relation between income shocks and surrender rates was shown in the recent work of Fier and Liebenberg (2013).

In phase one our agents, when facing a liquidity need, they find themselves in a certain "state of the world" described by a multiplicity of variables: the economy state, which has an impact on the value of the unit linked contract through the price of the share in which the company has invested in and on the likelihood of receiving a further cut in their salaries or being hired, their age and the time they have to wait to receive their full benefit, the amount of their yearly premiums which they would stop paying in case of surrender, their consumption, the fixed rate at which the policy contract is re-evaluated each year, their wealth and current income, whether they have a degree or not, which has an effect on their starting income in case are hired again.
The agents randomly react surrendering the unit linked contract or the policy or none.
At the end of their life, we record the difference between their initial and final wealth. This measure will be used to tell whether a certain action given a certain state of the world has proven itself a good or bad strategy.

The final result of this first version of phase one of my project is a data set having in each row the data relative to a single agent who faced a lack of liquidity: the inputs of the data set are the states of the world in which the agent found himself when dealing with this problem and the action he took, the output is the difference between initial and final wealth.

In phase two we will train a neural net on this data and have the agents act based on the strategy which, according to the prediction of the neural net, maximizes the difference between final wealth and initial wealth at the time when the agent has not been able to cover the liquidity need.

Code

From the interface tab of the NetLogo model we are able to change the transition probabilities from one economy state to another.
The probabilities shown in the input boxes, as you open the file, are derived from the time series of yearly Italian GDP at market prices from 2000 to 2014 from Istat website, the Italian National Institute of Statistics (http://www.istat.it/en/).
Once you have obtained the GDP series, we consider a good (g) year if the GDP grows by more than 3%, a normal (n) year if by less and a recession (r) if the GDP decreases.
The succession we obtained is the following: g, g, g, n, g, g, n, r, n, n, r, r, n.
For such a short series it is easy to compute the transition probabilities by hand.
To calculate the maximum likelihood probability to pass from a 'n' state to a 'g' state we use the following formula:

$$\hat{P}(g{:}n) = \frac{\text{number of times the economy passed from an 'n' state to a 'g' state}}{\text{number of times the economy was in a 'n' state}}$$

Notice that at the denominator we do not consider the last state as we do not know what

state will follow. The result is 1/4.
For a large series I suggest to use the R ‘markovchain’ package.

From the interface tab we can modify the income tax rate, the corresponding income bracket and the taxation on realized capital gain from the two contracts.

We can set different ages at which the agent has the right to a state pension and the contribution requirement in years; these variables are important since we assume that the insured party will stipulate in the contract to receive the benefit when he retires.

Finally we have two procedures we can call from the interface tab: ‘setup’ and ‘go’, but before starting the model we need to have the mortality tables to simulate the deaths of the agents.
You can download the mortality table as two excel files, for male and female agents, from the Istat website.
What we need is the column denominated ‘qx’, it is an estimator of the probability of an individual of age ‘x’ to die before age x+1.
Copy and paste only this column in another sheet, then convert the excel file in a csv one.
Put the ‘mortalityTableMale.csv’ and ‘mortalityTableFemale.csv’ files, that you obtained from this procedure, at the same location where you have the NetLogo model.
It is important you call them this way before running the code otherwise the ‘create-mortalityTable’ procedure will give an error. The procedure stores the values of the csv file in a list for practical purposes.

```
to create-mortalityTable
  file-open "mortalityTableMale.csv"
  set mortalityTableMale []
  while [file-at-end? = false]
  [
    set mortalityTableMale lput file-read
mortalityTableMale
    ]
  file-close
  (Similarly for female agents)
end
```

In the first part of the code we have global variables and the own variables of the agents.
Global variables are common for all the agents in the ‘world’.
‘Turles-own’ variables are specific for each agent.

The latter includes: gender, age, whether the agent is graduated or not, whether he is employed, the amount of the yearly premium for both unit linked and policy contract, whether he is currently holding one or the other contract (‘unit’ and ‘policy’), whether he has faced a liquidity need (‘acted’), whether he is retired, his income, financial wealth, the current value of their contracts and the future value of the endowment contract (it’s the benefit he will receive if he pays all the premiums relative to that contract), the probability of being hired, fired, getting a raise and a cut, his perceived life expectancy, not based on the mortality tables, his yearly consumption and the random action he will make when facing liquidity need: whether to surrender one contract or the other or to keep them active (‘surrenderUnit’ and ‘surrenderPolicy’).

In phase one of this project, among the agents’ variables, we include: the state of the economy, the share price, the fixed rate to re-evaluate the policy contract and the penalty. We do so in order to train our agents to different sequences of the state of the economy and share price and to different types of policy contracts, with different penalties and minimum guaranteed rates as this has a major impact on the choice of whether to surrender this or that contract. This concept can be easily understood with an example: Imagine that, the neural net our agent will use in phase two to choose which contract to surrender, has been trained on a data set where the simulation that generated it has experienced a general negative trend of share prices. All the agents that randomly decided in phase one to keep the unit linked contract will most likely have a negative output (wealth - initial wealth) in their ‘own’ record, vice versa for those who surrendered it, therefore each agent in the data set will teach the neural network that unit linked contracts are a bad choice.

We have two solutions to properly train our agents.
The first is to keep those variables as ‘turles-own’. This will allow us to create the data set in only one simulation. This option is viable because the agents do not interact with each other, otherwise the pattern emerged would have been flawed and the network would not have been properly trained.
The second option is to run many simulations and then unify all information into one single data set.
In this version we are using the first one.

The ‘setup’ procedure starts deleting the input and output files from the previous simulation.
‘delete-csv-files’ checks whether the files are present in the NetLogo working directory and deletes them.

Male, graduate and unemployment percentage global variables are used to define the probability with which the agents present those characteristics.
We assumed all the agents are employed at the beginning of the simulation therefore, we see them colored green in the graphic user interface (GUI).

In the ‘setup-turtles’ procedure the turtles-own variable are defined.
We used two ways to generate a random variable with a discrete probability.
If we wish to generate this variable with a uniform distribution, where each possible value of the random variable is equally likely, we use either the ‘random’ function for integer variables or the ‘one-of’ function to pick one value from the specified ones.
If instead we wish to generate a value with a certain probability, we generate a uniformly distributed ‘random-float’ number between 0 and 1 and set the variable of interest accordingly to the desired probability; for example if we wish to have an agent graduated with a 40% probability we will set the ‘graduate’ variable to 1 upon the

condition that the random-float generated is less than 0.4.
The retirement age variable represents the age at which the agent believes he is going to retire at the beginning of the simulation, when he underwrites the two contracts with the insurance company. The agent chooses the age to receive the benefits from the insurer.
The agents in my simulation, according to Italian laws, retire after they either accumulated 42 years of contribution or at the age of 67.
In this procedure we already set which is the action the agents are going to perform in case they are not able to satisfy their consumption: with equal probabilities they will either surrender the policy, the unit linked contract, or do nothing.

The ‘create-mortalityTable’ stores in two lists the estimated probabilities of male and female agents to die within the year.

The ‘forever’ box in the ‘go’ button is checked and the simulation will call this procedure until each agent has died.

The agents first pay their income taxes, if their income exceeds the relative bracket.

```
to pay-taxes ;turtle procedure (tp)
  if (employed = 1 and retired = 0 and income >
incomeBracket)
  [let tax taxRate * (income - incomeBracket)
  set netIncome income - tax]
end
```

In the first step of the simulation the agents buy the two contracts and set the yearly premium as a percentage of the net income

```
to buy-unit
  if employed = 1 and unit = 0 and acted = 0
  [set unitPremium netIncome * 0.1
  set unit 1]
end

to buy-policy
  if employed = 1 and policy = 0 and acted = 0
  [set policyPremium netIncome * 0.1
    set futurePolicyFundValue policyPremium *
    (1 + fr) * ((1 + fr)^(retirementAge - age) - 1) / fr
    set policy 1]
end
```

and set their variable ‘futurePolicyFundValue’ at the value of the benefit they would get provided that all premiums will be paid; The final value (FV) of an annuity due of n years capitalized at an interest rate r is:

$$FV = (1 + r)\frac{(1 + r)^n - 1}{r}$$

This variable is important to define the consumption of the agent; we will see how in the ‘set-smoothedConsumption’ procedure.

The agents then, if they have not surrendered the policy, they did not retire yet and they have enough wealth, they pay the premiums.
The number of shares bought for the unit contract through the relative premium depends on the current price of a share.
The premium for the policy contract goes in the corresponding fund.

```
to pay-premiums
  ;unit
  if (unit = 1 and wealth > unitPremium)
  [set wealth wealth - unitPremium
  let n unitPremium / sharePrice
  set nShare nShare + n]
  ;policy
  if (policy = 1 and wealth > policyPremium)
  [set wealth wealth - policyPremium
   set policyFundValue policyFundValue + policyPremium ]
end
```

In the ‘set-smoothedConsumption’ procedure the agent defines, at the beginning of the simulation, the level of consumption he will keep during his life.
Let I be the agent’s yearly income, let R be his remaining working life, let W be the initial wealth, let E be the endowment he will receive at retirement age and let T be his expected remaining life time; we calculate his level of consumption C with the following formula:

$$C = \frac{I * R + W + E}{T}$$

The hypothesis behind it is that people try to keep the the same level of consumption during their life time without regarding temporary income variations (Friedman, 1957).

```
to set-smoothedConsumption ;tp
  if chosenConsumption = 0
  [
    set smoothedConsumption (
      (retirementAge - age) * (netIncome - policyPremium)
      + wealth  + futurePolicyFundValue
      )  / (lifeExpectancy - age)
  if smoothedConsumption > (retirementAge - age) *
(netIncome - policyPremium)
  [set smoothedConsumption (retirementAge - age) *
(netIncome - policyPremium)]
  set chosenConsumption 1
  ]
end
```

The second ‘if’ clause is to assure that this level of consumption is not too high, because the agent must wait until retirement for the policy.

It follows an unused procedure ‘set-consumption’ which simply set ‘consumption’ equal to ‘smoothedConsumption’. The initial purpose of this procedure was to increase ‘smoothedConsumption’ in case the agents where to deal with particular large expenses like a wedding, the first installment of the mortgage, a car or similar, but I found no literature linking surrender rates to these kind of expenses.

The ‘update-economy’ procedure has the economy pass from one state to the other based on the transition probabilities defined in the interface.

The ‘update-sharePrice’ procedure controls the dynamic of the share price depending on the state of the economy:
In a good state of the economy the share price will increase by a percentage between 3 and 6.
In a normal state of the economy the share price will increase by a percentage between 0 and 3.
In a recession state of the economy the share price will decrease by a percentage between 0 and 2.

At each time step the surrender values of the two contracts are calculated.
For the unit linked contract is subtracted only the eventual capital gain taxation.
For the policy contract it is also necessary to subtract the penalty which depends on the years left to the natural maturity of the contract.

```
to update-unitFundValue ;tp
  set unitFundValue nShare * sharePrice
  ifelse unitFundValue > 100
  [set unitSurrenderValue unitFundValue - (unitFundValue -
100) * nonGovBondTax]
  [set unitSurrenderValue unitFundValue]
end

to accrue-interests   ;tp
  set policyFundValue policyFundValue  * (1 + fr)
  let grossPolicySurrenderValue policyFundValue / (1 +
penalty) ^ (retirementAge - age)
  set policySurrenderValue grossPolicySurrenderValue * (1 -
govBondTax)
end
```

The agent then consumes upon the condition he has enough wealth to satisfy his consumption, if he does not, he makes a random choice based on the random variables defined in the ‘setup-turtles’ procedure: ‘surrenderUnit’, ‘surrenderPolicy’.
Depending on the second ‘if’ clause in the ‘consume-surrender-doNothing’ procedure the agent will either be asked to ‘surrender-unit’ , ‘surrender-policy’ or do nothing.
No matter what the action is, the agent has acted (the variable ‘acted’ is set to 1) in response to a liquidity need, therefore he will be part of our data set for the training.
To include the agents who ‘acted’ in our data set, we ask them to write on the input.csv

file, through the ‘file-write’ command, their variables together with their identification number.
There are two types of input variables: the state of the world (‘economyState’, ‘age’, ‘unitPremium’, ‘wealth’, ‘fr’ the fixed rate for the policy, ‘unitSurrenderValue’, ‘policySurrenderValue’, ‘consumption’, ‘graduate’ ,’retirementAge’, ‘netIncome’) and actions (‘surrenderUnit’, ‘surrenderPolicy’).

```
to consume-surrender-doNothing ;tp
   ifelse wealth > consumption
   [set wealth wealth - consumption]
   [if acted = 0 and retired = 0 [ifelse surrenderUnit = 1
[surrender-unit][ ifelse surrenderPolicy = 1 [surrender-
policy] [set label “N”] ]
       set wealthInit wealth
       file-open “input.csv”
       ;inputs are made of
       ;---> state of the world
       file-write who file-type “;”
       file-write economyState file-type “;”
       (...)
       file-write retirementAge file-type “;”
;relevant if agents retire at different ages
       file-write netIncome file-type “;”
       ;file-write employed file-type “;”
;redundant if netIncome is known
       ;file-write policyPremium file-type “;”   ;the
premium are equal
       ;file-write penalty file-type “;”
;irrelevant provided “policySurrenderValue”
       ;---> actions
       file-write surrenderUnit file-type “;”
       file-write surrenderPolicy file-type “\n”
       set acted 1
       ]
   file-close
   ]
end
```

Notice we do not feed the ‘penalty’ variable to the machine learning algorithm as we already record in the input file the surrender value for the policy. The same holds for the ‘employed’ variable which is redundant given the ‘netIncome’ variable.

The procedure used to surrender the contracts asks the agents to show in their built-in ‘label’ variable through the GUI whether they surrendered the unit linked contract (U), the policy (P) or none (N). This command helps the programmer in spotting possible mistakes in the code.

```
to surrender-unit ;tp
  set wealth wealth + unitSurrenderValue
  set unit 0 set label “U”
```

```
;the unit (U) has been surrendered
end

to surrender-policy ;tp
  set wealth wealth + policySurrenderValue
  set policy 0 set label “P”
;the policy (P) has been surrendered
end
```

When the agent reaches retirement age he stops receiving the yearly income and he collects the benefit through the surrender procedure, but in this case without paying any penalty since the withdrawal occurs at the established date.

We still need to have the agents show whether they surrendered the contracts because they simply received the benefit at maturity or if it is due to a lack of liquidity during their lives, therefore those agent that ‘acted’ will show this in the GUI trough the label ‘A’ after they retired, plus all agents that retired will set their color yellow.

```
to check-retire ;tp
  if age = retirementAge [
   if unit = 1 [surrender-unit] if policy = 1 [surrender-
policy]
   set color yellow set label “ “
   if acted = 1 [set label “A”]
  ]
end
```

Then in the code we have a procedure which updates the probability of getting a raise or a cut in an agents’ salary and the probability of being hired or fired depending on the state of the economy.
Through the following two procedures in the code the agents’ salary and employment status will change accordingly. The agents who lose their jobs are highlighted in red.

The simulation of the deaths is done through the use of the mortality tables lists created in the ‘setup’ procedure.
These lists have the ordered estimates of death probabilities; for example in position 0 we have the estimated probability of an agent, at birth, to die before celebrating the first birthday; in the last position (120) we have the estimated probability of a 120 years old agent to die within the year, which equals one.
When the agent dies, if he has faced a need of liquidity during his life, he will be asked to write in a csv file called ‘output’ the difference between his final wealth and his wealth at the time he faced the need of cash, together with his identification number.

```
to check-death ;tp
  ifelse gender = “M”
  [set pDeath item age mortalityTableMale]
  [set pDeath item age mortalityTableFemale]
  if random-float 1 < pDeath [
    if acted = 1
```

```
    [
      set output wealth - wealthInit
      file-open "output.csv"
      file-write who file-type ";" file-write output file-
type "\n"
      ]
    file-close
    die]
end
```

The phase one in NetLogo is concluded. We have the data we need in the same folder where the NetLogo file is. Now we set as working directory of R the same folder. We read the two data sets in R and order the rows based on the identification number of the agents.

```
inputData <- read.csv("input.csv", sep=";", header = FALSE)
inputData[,c(5,7,8,9,12)]=round(inputData[,c(5,7,8,9,12)],2
)
names(inputData) <-
c("who2","economyState","age","unitPremium",
"wealth","tmg","unitSurrenderValue","policySurrenderValue",
"consumption",
"graduated","retirementAge","netIncome",
"surrenderUnit","surrenderPolicy")

outputData <- read.csv("output.csv", sep=";", header =
FALSE)
names(outputData) <- c("who","outputVar")
outputData=round(outputData,2)

inputData=inputData[order(inputData$who2),]
outputData=outputData[order(outputData$who),]
cbind(inputData$who2,outputData$who) #check if they are in
order
```

The next step is to work with the inputs, modify them and set them up to train a neural net, but conceptually the data sets are now ready to train a machine learning algorithm at predicting the difference between initial and final wealth given the inputs.

Training a neural net on the data provided by the model specified in the first version did not produce a consistent (not necessarily rational) behavior for the policyholders. The first version of the project in phase one should be revisited and have the agents act coherently with the literature shown in paragraph 5 and then train a neural network or another machine learning algorithm on the data provided by this other model specification where agents do not choose to surrender the unit or the policy contract randomly but upon rules defined by the programmer consistently with the insights brought up by the literature reviewed. Having the agents act relying on the prediction made by the neural net could reveal itself to be too big of a challenge and out of the scope of an insurance company, moreover it will likely not be able to predict the choices of the policyholder. The machine learning algorithm trained could instead be used as a tool to advise policyholders on the outcome a certain choice might lead to and possibly provide a feasible alternative for the policyholders, such as an option to suspend or reduce the premium, to influence their behavior and reduce variability in their choices which have such a considerable impact on the liabilities of the insurance company.

4.2 Version 2.

In this second version of the model we are going to make an analysis on the effect that the surrender decision has on the liquidity of the insurance company.

As the paper of the Geneva Association (2012) points out, although the insurance business is not as liquidity dependent as the banking system, regulators currently discuss scenarios of liquidity problems in the insurance sector, for example in the life underwriting risk module within the 'Solvency II' framework it is required an amount of capital to cover losses due to a mass lapse event.

The paper distinguishes the term solvency by liquidity:

> While solvency refers to a company's overall ability to have sufficient assets to cover its liabilities, liquidity refers to its capability to cover current liabilities with current assets. Accordingly, liquidity management as part of insurers' asset liability management is the management of the cash inflows and cash outflows. Liquidity risk is a measure of the insufficiency of an insurer's cash resources in meeting its current or future cash needs. It is also the measure of the need that assets will have to be liquidated at a discount or that refinancing is only possible at a higher interest rate.

Introduction.

In this second model specification we have our agents buy a pure endowment contract at the beginning of the simulation through a lump sum payment.

The pure endowment is a contract for which the insured party receives a capital if he survives until a certain age specified in the contract.

We will run the model with two specifications. First we will not give the insured party the surrender option, therefore he will have to wait and survive up to the established date in the contract to receive the capital. This specification will allow us to see the insurance mutuality principle in action and we will focus on the number of policyholders needed to apply the principle.

In the second model specification we will allow the policyholder to surrender. In such model we will have the agents surrender with a certain probability defined by the user and apply a penalty user defined too.

The insurance company invests a fraction of the the premiums in zero coupon bonds with maturities equal to the maturities of its liabilities and therefore exploiting the maximum interest rates for its policyholders.

Before going further it is useful to discuss why insurance companies charge a penalty for surrendering the contract early. The insurance companies, as all rationale investors in a market, try to obtain the highest returns from its investments. The pure endowment, as most part of traditional insurance product, must be invested at a 'risk free' and currently very low interest rate. In order to provide the best returns for its clients and itself the insurance company invests its assets at longer duration usually providing higher interest rates than short term bonds.

If the insurance company is forced by early surrenders to liquidate or reinvest its assets at a lower interest rate, it should charge the cost on those very contracts.

In this model specification we do not consider the option of the insurance company to sell its assets. We assume that after a pure endowment with a certain maturity has been signed, the insurance company invests a fraction of the premium at the correspondent maturity and will receive the principal only at the established date, but it will still be forced to provide the value of the mathematical reserve diminished by the penalty in case of early surrender.

Theoretical Framework.

In this paragraph I will already refer to the NetLogo model variables by using the quotation marks.

The pure endowment binds the insurer to pay a benefit 'C' in case the policyholder of age 'x' survives another 'n' years until age x + n. Given these three variables the insurer calculates the premium in a lump sum 'U' paid when the policyholder is of age x as:

$$U = C \cdot P(T_x \geq n) \cdot (1 + techRate)^{-n} = C \cdot \ {}_nE_x$$

Where $P(T_x \geq n)$is the probability of a person of age x to survive another n years and the 'techRate' is the interest rate provided by the insurance company to the insured. The random variable T_x is the remaining lifetime at age 'x'. In actuarial math we refer to this probability with the following notation: ${}_n\mathrm{p}_\mathrm{x}$.

The product of the first two factors in the right hand side of the formula gives the expected value of the payment which multiplied by the third factor, is discounted of n years. The premium is therefore the actuarial present value of the benefit provided by the insurer.

Insurance companies uses favorable (higher for pure endowments and all contracts which pays in case of survival of the insured) survival probabilities and favorable (lower) interest rates to charge a higher premium. If favorable to the insurer, these two variables are called technical basis of the first order (demographic and financial). The premium ‘U’ so obtained contains an implicit safety margin.

As for now we are interested in the mutuality principle of insurance companies and on the effects of lapse rates on liquidity risk, in my simulation the user will be provided the value of the fair technical interest rate and the the actual death probabilities. The fair technical basis are called basis of the second order.

In order to find the survival probability we use the following:

$$P(T_x \geq n) = P(T_x \geq 1) \cdot P(T_{x+1} \geq 1) \ldots P(T_{x+n-1} \geq 1)$$

We multiply the probability of surviving each year for n years where the one year survival probability is obtained as:

$$P(T_x \geq 1) = 1 - q_x$$

Where $q_x$is the the probability of dying within the year: $P(T_x < 1)$. Our input for the model will be directly the the mortality table of q_x for $x = \{0,1,\ldots,120\}$ where 120 is the maximum age, therefore we will have that $q_{120} = 1$.

Although we will start from the mortality rates it is useful to explain how these are estimated:

$$q_x = \frac{l_x - l_{x+1}}{l_x}$$

Where $l_x$are the living at age x. In our case we used the most recent life tables available in the ISTAT website.

At this point, in order to provide the surrender option to the policyholder, we must have a formula to calculate the value of the contract ‘Vt’ for each time step *t* in the simulation from 0 to ‘n’. As the premium is paid by the policyholder at the beginning, when the insured is of age *x*, the value of the contract will represent a debt of the insurance company towards the policyholder and is calculated as:

$$V_t = C \cdot P(T_{x+t} \geq n - t) \cdot (1 + techRate)^{-(n-t)}$$

V_t is referred to as the mathematical reserve and is the actuarial present value of the difference between benefits and premiums. If $V_t$is positive it represents a debt for the insured company, and this is our case as the premium has been paid at 0 < *t*.

In theory the mathematical reserve represents the exit value of the contract that the insurer must pay at *t*, but in practice insurance companies charge a penalty for contracts surrendered early. If the insurance company did not provide a surrender option, the maturity of the contract would be certain. The choice of the policyholder to exit the contract before the maturity exposes the insurance company to a liquidity risk therefore the latter must consider how to invest the premiums, duration and amount, and the amount of the penalty to charge.

Usually insurance companies do not allow for a surrender of the contract within the first three years, and thereafter they charge a ‘penalty’ ranging between 2 and 4 per cent. The ‘surrenderValue’ is calculated as follows:

$$SurrenderValue_t = \frac{V_t}{(1 + penalty)^{-(n-t)}}$$

Code

First the user must connect the model to R. From the R command line we have to install the package Rserve (Urbanek, 2013) then type “library(Rserve) ; Rserve(args=‘--vanilla’)”, this command will work both on Linux and Windows OS. The author of the package recommends you do not use it on a Windows machine. I have been running my model on the 14.04 Long Term Support Ubuntu OS.

You can download the NetLogo Rserve extension at http://rserve-ext.sourceforge.net/. The folder must be unzipped in the ‘extensions’ folder in the NetLogo directory. The folder also provides some useful tutorials.

In order to be able to use the functions of the extensions you must name it in the code tab in this way “extensions[rserve]”.

At this point we can start our analysis from the interface tab. Press the button ‘Open R’ to open the connection.

In this tab we have buttons, which call for procedures present in the code tab, and sliders and inputs to define global variables.

The ‘minCapital’ and ‘maxCapital’ variables defines the minimum and maximum amount a policyholder can insure, which is the amount they will receive if they survive up to the established age.

The ‘setup’ button will create the policyholders, their relative contract terms, premium excluded, and one insurer.

Based on policyholders’ characteristics, age and gender, and contract terms, duration and amount insured, the ‘calc-fairTechnicalRate’ will return the value of an average fair technical rate. The user must then type the ‘techRate’ variable in the relative input box. To charge an implicit safety margin insurance companies uses a technical rate lower than the actual interest rate they can invest assets at: this spread builds the profit of the financial component of the contract.

Then the ‘define-premium’ procedure sets the amount of the premium that the policyholders have to pay at the beginning of the simulation.

We define the other variables. The probability that each policyholder will lapse his policy at each time step of the simulation is set through the variable ‘lapseRate’. We define the fraction of premiums invested at the relative yield based on the maturity of the contract, and finally the ‘penalty’ charged in case of surrender.

We will now discuss the code tab.

The ‘Setup’ button will call the relative procedure in the code tab which creates two mortality tables for gender. This table will be used to simulate the process of the deaths of the agents. We will also use those probability to price the premiums and to calculate the mathematical reserves. Notice we are using the same probabilities to simulate the deaths and charge the premiums, therefore we cannot obtain profit through the demographic component. In this model we focus on the liquidity of the firm and we do not want a high implicit margin to hide the otherwise too low penalty for surrendering the contract.

As explained in the theoretical framework to price the premiums we need the survival probabilities.

We will use the intuitive R language to make an otherwise complicated calculation. We have in a list the death probabilities and we need in another list the survival probabilities for each age, and we now that for every x in the list:

$$p_x = P(T_x \geq 1) = 1 - P(T_x < 1) = q_x$$

In the code tab, the ‘create-lifeTable’ procedure is written as

```
rserve:put “mortalityTableMale” mortalityTableMale
rserve:eval “lifeTableMale <- (1 - mortalityTableMale)”
set lifeTableMale rserve:get “lifeTableMale”
```

The colon notation is used to call functions of the extensions. The ‘put’ function will create a variable in R named as the first argument and having as value the second argument. We are creating the mortality rates list within R. The second function makes an evaluation within R. In this case we are making the calculation on the whole list with one command. Through the ‘get’ function we store the list back in the NetLogo environment. We will proceed analogously for the female life tables.

The following procedure ‘create-TermStructure’ will store in a list the interest rates at which an institution can lend money for different maturities. Term Structures varies across countries and across institutions. The rates used in my model refer to the interest rates at which an Italian Insurance company can lend its money.

We set as list two global variables which will be used to calculate the fair technical interest rate ‘techRate’: ‘maturityListTemp’ and ‘interestRatesTemp’.

Then we create and define the variables of the policyholders and the insurer. Notice that at the beginning of the code we defined two types of agents, insurers and policyholders, through the ‘breed’ keyword.

We create a number of policyholders equal to the global variable ‘num-people’ defined in the slider in the interface tab.

```
create-policyholders num-people
 [
   ifelse (random-float 1 < 0.5)
   [set gender "M"] [set gender "F"]
   set age 35 + random(31)
   set x age
   set maxBenefitAge 65
   set benefitAge x + 1 + random (maxBenefitAge - x)
   set n benefitAge - x
set C minCapital + random-float (maxCapital - minCapital)
   ]
```

The above code creates either male or female policyholders with a 50% probability. Policyholders underwriting the contract have an age between 35 and 65 years.

As explained in the theoretical framework ‘x’ (age at contract underwriting year), ‘n’ (duration of the contract) and ‘C’ (capital insured), are terms of the contract needed to price the premium.

The ‘x’ and ‘n’ variable are created randomly in such a way to prevent having a policyholder receiving the capital after 65th years of age as this seem an unlikely choice for this type of contract.

The capital insured ‘C’ is a random number between a minimum and maximum defined by the user in the interface tab.

We create one insurer and define the ‘contracts’ variable which will keep track of the number of contracts held by the insurer thought the simulation. We set as lists three insurers-own variables needed for the simulation:

```
to setup-insurers

  create-insurers 1
  [
    set contracts num-people
    set premiumList []
    set maturityList []
    set interestRates []
    ]
end
```

We now have the data of the policyholders necessary to calculate an average fair technical rate:

```
to calc-fairTechRate
  ask policyholders
  [
    set maturityListTemp lput n maturityListTemp
    set interestRatesTemp lput item (n - 1) termStructure
interestRatesTemp
    ]
    rserve:put "interestRatesTemp" interestRatesTemp
    rserve:put "maturityListTemp" maturityListTemp
    show rserve:get "weighted.mean(interestRatesTemp,
maturityListTemp)"
end
```

With this procedure we ask the policyholders to put in two lists the maturities and the relative interest rates earned by the insurance company on the relative fraction of premium invested. The ‘lput’ function adds the first argument to the second argument, which is a list. Notice that the ‘ask’ function works as a “for cycle” for each agent in the agent set ‘policyholders’, therefore the two lists will be ordered properly: in position ‘i’ of the ‘interestRatesTemp’ list we will have the interest rate earned by the insurance company for the i-th element (maturity) of the ‘maturityListTemp’. The interest rates are extracted from the yield curve of the insurance company: taking into account that the first element of a list is the ‘item’ 0, we are adding to the interest rates list the ‘n’-th interest rate, which is the return obtained by the company for investing money for a period of ‘n’ years.

Finally to calculate the average fair technical rate, which represents the rate at which the insurance company invests its money, we calculate an average of the interest rates obtained, based on the maturities of the contract, weighted by the maturities.

Later, when we will run the simulation, by applying the fair technical rate to calculate the premiums and investing all premiums without giving the surrender option to the contracts, the user will notice how, for a high number of policyholders and a low capital insured, the fund of the company will tend to zero at the end of the simulation (when all contracts have either reached maturity or the death of the insured happened before the maturity).

The user must now set in the input box ‘techRate’ the chosen technical rate. I would suggest the fair one for the purposes of our analysis.

Now we have all the variables needed to calculate the premiums: ‘x’, ‘n’, ‘C’, ‘gender’ and ‘techRate’. We will do so by calling the ‘define-premium’ from the relative button in the interface tab.

```
to define-premium
  ask policyholders
```

```
    [
      rserve:put "x" x
      rserve:put "n" n
      ifelse gender = "M"
      [rserve:eval "probs <- lifeTableMale[(x+1):(x+n+1)]"]
      [rserve:eval "probs <- lifeTableFemale[(x+1):(x+n+1)]"]
      let npx rserve:get "prod(probs)"
      set U C * (1 + techRate)^(- n) * npx
    ]
  end
```

We use R to calculate the probability of a person of age ‘x’ to survive another ‘n’ years. Recall that:

$$P(T_x \geq n) = {}_np_x = P(T_x \geq 1) \cdot P(T_{x+1} \geq 1) \dots P(T_{x+n-1} \geq 1)$$
$$= {}_1p_x \cdot {}_1p_{x+1} \cdots {}_1p_{x+n-1}$$

The variables needed by R to calculate this probability are: the age of the insured, the duration of the contract, his/her gender and the one year survival probabilities. The latter is already in the memory of R. We export the others to R through the ‘put’ function. We calculate the probability by making the product of the relevant one year survival probabilities and import it back in the NetLogo memory to make the calculation of the premium.

We are differentiating the price based on the gender of the insured. As of 21st December 2012 the European Court of Justice forced the insurance companies of the member States to treat male and female policyholders equally in terms of benefits and premiums. In this version of the model we want to observe the effect of lapses on an actuarially fair portfolio, therefore we will discriminate the prices based on the gender of the insured.

The button ‘go’ will call the relative procedure iteratively until the insurance company will have no more contracts in his portfolio.

The ‘ticks’ variable represents time expressed as years.

At year zero, the policyholders underwrite the contract and pay the insurance company the premium.

The insurance company stores in a list the amount of the premiums and in another list the relative maturities. The company will need this values to invest the premiums in assets having the same maturity as the relative liability.

```
to go
  ask policyholders with [paid = 0]
  [
    let uu U let nn n
    if nn > 30 [set nn 30]
    ask insurer num-people
    [
```

```
        set fund fund + uu
        set premiumList lput uu premiumList
        set maturityList lput nn maturityList
        set interestRates lput item (nn - 1) termStructure
interestRates ]
    set paid 1
    ]
```

In the code we have the agents run this commands upon the condition they have not paid the premium yet.
The policyholders store the premium and the maturity of their contract in a temporary variable which is then returned to the insurer.

The insurer increases his ‘fund’ correspondingly and updates the two lists together with a third list containing the interest rates gained in the market through the investment of the relative premium.

In the first year the insurer has to invest the premiums. We will run the following lines upon the condition the insurer has not invested the premiums yet:

```
ask insurer num-people
  [ if invested = 0 [
    rserve:put "premiumList" premiumList
    rserve:put "investedProportion" investedProportion
    rserve:put "maturityList" maturityList
    rserve:put "interestRates" interestRates
```

We exported the variables in R.

```
  rserve:eval "investmentList <- premiumList *
investedProportion"
 rserve:eval "investment <- sum(investmentList)"
 set investment rserve:get "investment"
 set fund fund – investment
```

We create the ‘investment’ variable in R, import it in NetLogo and subtract it from the fund of the insurer.

```
      rserve:eval "finalvalueList <- investmentList * (1 +
interestRates) ^ maturityList"
```

‘finalvalueList’ represents the list of principals returned to the insurance company at the corresponding maturity. Notice that the above variables are all vectors. In this case the mathematical operators make the calculation elementwise.

```
 rserve:eval "ds <- data.frame(cbind(maturityList,
finalvalueList))"
 rserve:eval "tempVec <- c(rep(0,30))"
 rserve:eval "for(i in 1:30){tempVec[i] <- i}"
 rserve:eval "dsIntegration <-
```

```
data.frame(cbind(tempVec,c(rep(0,30))))"
      rserve:eval "names(dsIntegration) <- names(ds) "
      rserve:eval "ds <- rbind(ds,dsIntegration)"
      rserve:eval "ds <- aggregate(. ~ maturityList, data=ds,
sum)"
```

The dataset ‘ds’, stored in R memory, will make the calculation much easier throughout the simulation. This dataset has 30 rows, as many as the number of years the simulation will last, and two columns, one with the maturities and the other with the corresponding amount earned in that year. Therefore in the second column, for row j, we have the sum of the amounts the insurer will earn on the investments made with a maturity of j years.

Notice we have more investments with the same maturity. The ‘aggregate’ function sum the “final value” of the investments having the same maturity.

```
set invested 1
    ] ]
```

Finally we set the variable ‘invested’ to one so that this command will run only at the beginning of the simulation.

```
  ask policyholders
  [
    if age < x + n and random-float 1 < lapseRate and
lapsed = 0
     [
      rserve:put "age" age
      rserve:put "time" (benefitAge - age)
      ifelse gender = "M"
      [rserve:eval "probs <-
lifeTableMale[(age+1):(age+time+1)]"]
      [rserve:eval "probs <-
lifeTableFemale[(age+1):(age+time+1)]"]
      let time_p_age rserve:get "prod(probs)"
      set Vt C * (1 + techRate)^(- (benefitAge - age)) *
time_p_age
      set surrenderValue Vt / (1 + penalty)^(benefitAge -
age)
      let sv surrenderValue
      ask insurer num-people [set fund fund - sv set
contracts contracts - 1]
      set lapsed 1
      ]
```

In the lines of code above we ask the policyholders to surrender their contracts with a probability equal to the ‘lapseRate’ variable, upon the condition they have not extinguished the policy already, either by surrendering it or surviving and receiving the capital insured.

If the conditions are met, and the policyholder “chooses” to surrender the policy, we have to calculate the mathematical reserve and subtract the penalty. The value obtained will be returned to the policyholder and the fund of the company will decrease correspondingly.

The policyholder is asking the capital back, before the natural maturity ‘n’. The insurance company has invested his premium at a the longer maturity ‘n’, but it is forced now to pay back the policyholder and this generate a liquidity risk for the company which has two way of protecting itself: apply a penalty and investing only a fraction of the premiums.

```
ifelse gender = "M"
[set pDeath item age mortalityTableMale]
[set pDeath item age mortalityTableFemale]
if random-float 1 < pDeath [if lapsed = 0 and benefit = 0
   [ask insurer num-people [set contracts contracts - 1]]
  die]
```

We have the policyholders die each year based on the death probabilities q_x. The death of the insured party extinguishes the contract without the payment of the insurance company.

```
    if age = x + n and lapsed = 0
    [
      let cc C
      ask insurer num-people [set fund fund - cc set
contracts contracts - 1]
      set benefit 1
      ]
    set age age + 1
  ]
```

If the policyholder has not surrendered the contract earlier and has survived the established date, he receives the insured capital.

```
ask insurer num-people
[
  rserve:put "ticks" ticks
  if ticks < 30
  [
    let add rserve:get " as.numeric(ds[ticks+1,2])"
    set fund fund + add
    ]
  ]
```

We use the dataset ‘ds’ created at the beginning of the simulation to increase the fund of the company by the corresponding amount each year. The sum of the “final values” for maturity ‘ticks’ + 1 are stored in the second column. Notice that at the end of year zero the insurance company receives the principal for the investments with maturity one year.

Results.

First we are going to test the mutuality principle of insurance companies by sharing the risk among a large number of policyholders. Then the scope of this first analysis will be to find a number below which such principle does not apply anymore.

Model Specification 1. Mutuality Principle

10 thousand policyholders each insuring a capital of 5,000, with fair technical basis, without allowing for the surrender option and investing 100% of the premiums.

The expected value of the final value of the fund is zero, as we used fair technical basis.

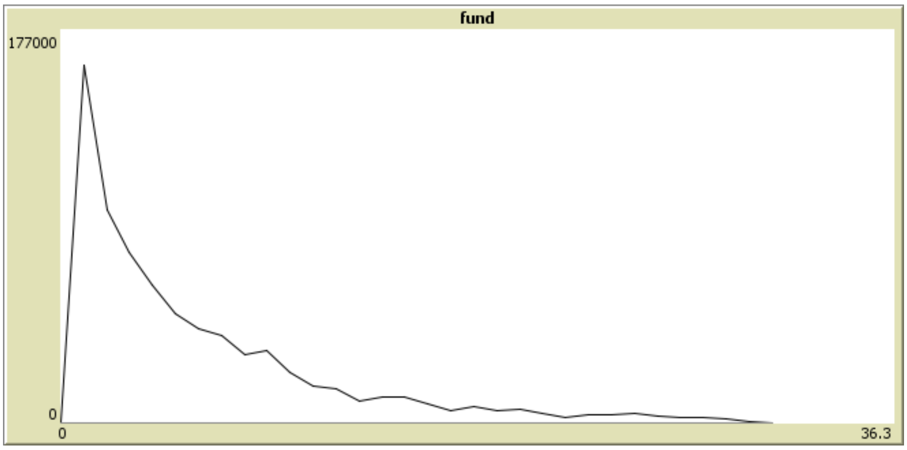


Figure 1

Notice that the fund of the insurance company is the highest at the beginning of the simulation, because we have most of the insured parties underwrite a pure endowment with short duration of the contract, therefore the company invests most of its resources at short maturities.

In the first run, at time step 30, we have a final value of the fund of -62,946.

We run the model under this specification 500 times and record in a spreadsheet the realizations of following statistic:

$$\hat{X}_{10,000} = \frac{final\ 'fund'value}{total\ sum\ insured} \cdot 100$$

We proportionate the final value of the fund by the total sum insured by the company.

Let $\underline{X}_{10,000}$ be the realizations of the statistic defined, then we use as estimate of its variance, the sample variance $\mathrm{VAR}(\underline{\mathrm{X}}_{10,000})$.

Similarly $\underline{X}_{1,000}$ is a random sample of size 500 of the statistic $\hat{X}_{1.000}$ generated by a model with same specifications as *Model Specification 1* except for a smaller number of insured parties: 1,000.

As we proportionated the final value of the fund by the total amount insured, we expect the variance of the statistic generated by the model with a higher number of insured parties ($\hat{X}_{10,000}$) to be smaller than the variance of $\underline{X}_{1,000}$. Nevertheless we cannot appreciate the difference without a model:

$$\frac{\mathrm{VAR}\left(\underline{\mathrm{X}}_{10,000}\right)}{\mathrm{VAR}\left(\underline{\mathrm{X}}_{1,000}\right)} = 0.09$$

To better capture the mutuality principle of insurance companies, we run the model for different numbers of insured parties.

For the graph on the left we run the model for the number of insured defined in the abscissa, 500 times.

The denominator in the formula above ($\mathrm{VAR}\left(\underline{\mathrm{X}}_{1,000}\right)$) corresponds to the point in the upper left part of the graph, while the numerator, to the point in the bottom right part of the graph.

With the specifications of our model, increasing the number of insured above 3000 insured parties does not substantially decrease the variance of the statistic $\hat{X}$. This implies that to achieve mutuality an insurance company would need at least 3000 insured.

For the graph on the right we run the model with a smaller pace of number of insured, up to 3000 insured parties.

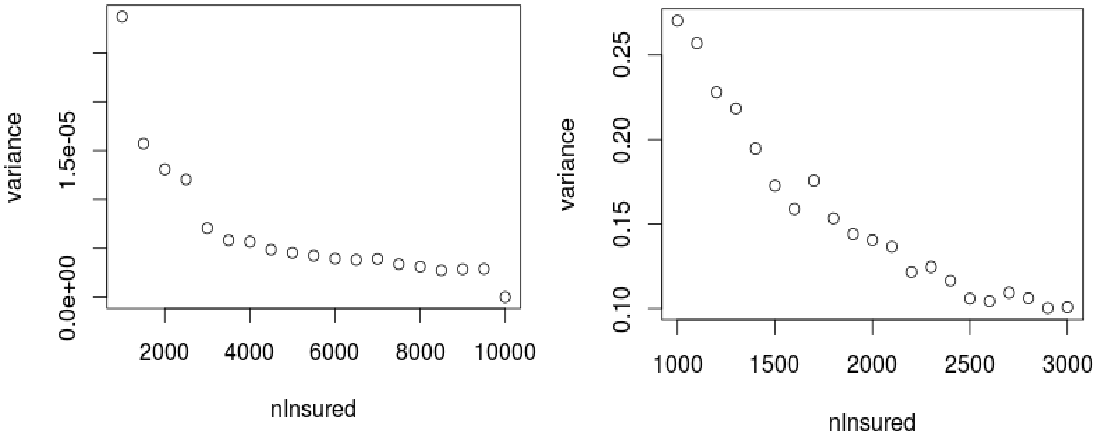


Figure 2

So far premiums were fixed for all insured parties. We are now going to test if variability in the insured capital has an effect on the variability of the final value of the fund.

We run the model 500 times, under the first model specifications but allowing policyholders to insure an amount between 1,000 and 9,000, and we refer to the same statistic, under this model specification as $\hat{X}_{variable\ C}$, and we compare it to $\hat{X}_{10.000}$ which we now refer to as $\hat{X}_{fixed\ C}$ as we had all policyholders insure an amount of 5,000.

On average we expect the insurance company to insure the same amount for both model specifications as we specified a range for the capital insured uniformly distributed between 1,000 and 9,000 with expected value 5,000.

The ratio of the variances of the two random samples is:

$$\frac{\mathrm{VAR}\left(\underline{\mathrm{X}}_{fixed\ C}\right)}{\mathrm{VAR}\left(\underline{\mathrm{X}}_{\mathrm{variable\ C}}\right)} = 0.76$$

To test whether the variability of the final value of the fund is lower for insurance companies allowing for a smaller variability of the insured capital we need to test the null hypothesis that the two variances are equal.

The ratio of variances of two random variables normally distributed is distributed as a Fisher random variable.

We use the Shapiro-Wilk statistic to test the null hypothesis that the two random variables are normally distributed.

We set the random samples $\underline{X}_{fixed\ C}$ and $\underline{X}_{\text{variable C}}$, as two vectors in R named ‘Xf’ and ‘Xv’ respectively, then:

```
> shapiro.test(Xf)

   Shapiro-Wilk normality test

data:  Xf
W = 0.99714, p-value = 0.5354
```

```
> shapiro.test(Xv)

   Shapiro-Wilk normality test

data:  Xv
W = 0.99698, p-value = 0.4846
```

We cannot reject the null hypothesis that the random variables are normally distributed.

We compare the sample quintiles to the theoretical quintiles of the normal distribution.

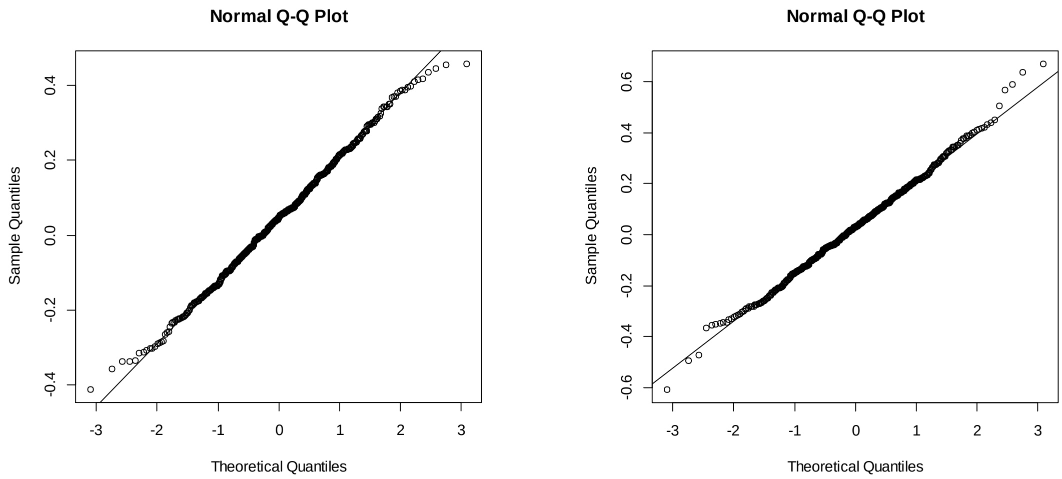


Figure 3

We plot the histogram of the frequency on the normal distribution.

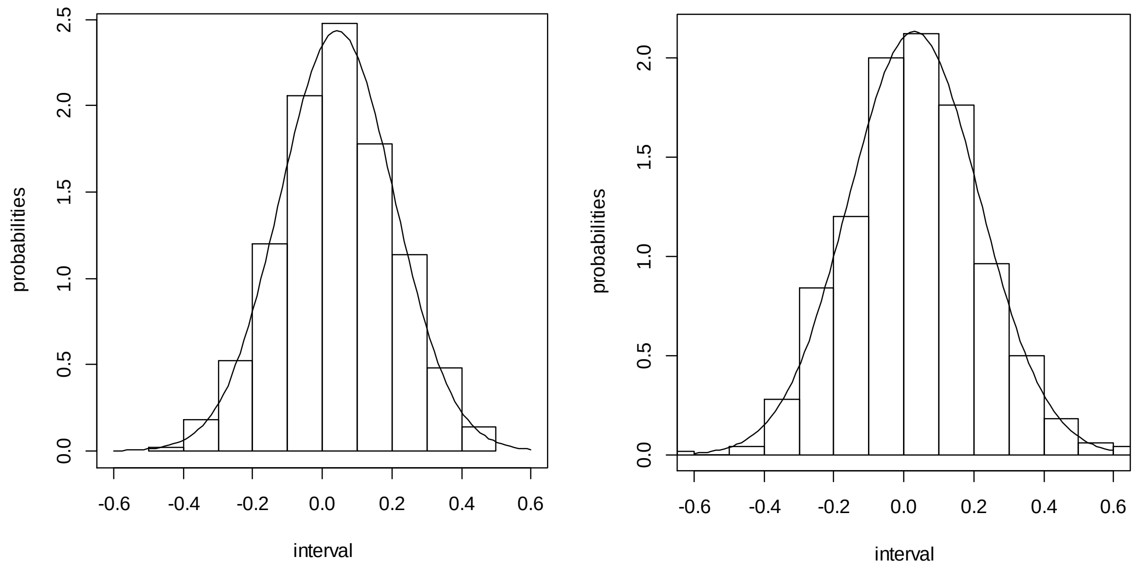


Figure 4

Assuming the statistics are normally distributed it follows that the ratio of their variances has a Fisher distribution.

We conduct a variance ratio test on the two random samples.

```
> var.test(Xf,Xv)

        F test to compare two variances

data:  Xf and Xv
F = 0.76597, num df = 499, denom df = 499, p-value = 0.002959
alternative hypothesis: true ratio of variances is not equal to
1
95 percent confidence interval:
 0.6425630 0.9130825
sample estimates:
ratio of variances
         0.7659719
```

We reject the hypothesis that the variances are equal.

If our assumptions are correct we can say that the variability of the final value of the fund is lower for insurance companies allowing for a lower variability of the insured capital among contracts.

Model Specification 2. Analysis of Surrenders.

10 thousand policyholders surrendering their policy at each step with a 3% probability, each insuring a capital of 5,000, with fair technical basis, investing 100% of the premiums.

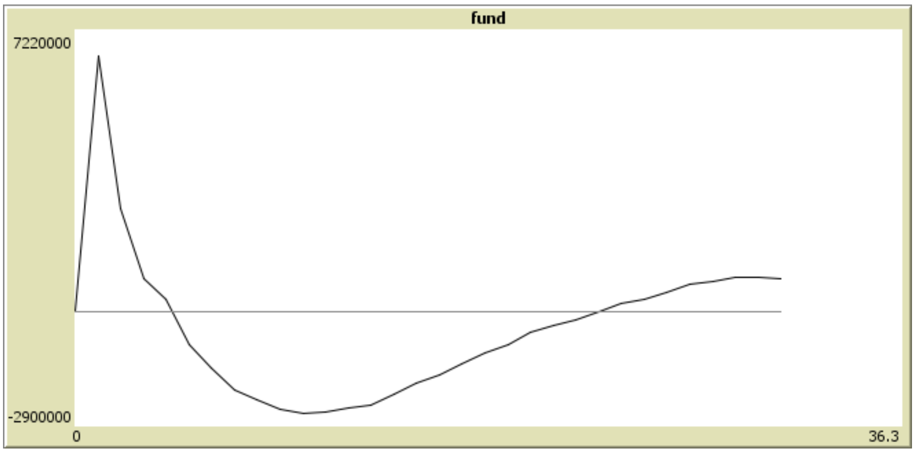


Figure 5

As policyholders surrender their policy, the insurance company cannot repay them since it has invested at longer maturities.

The policyholder, by surrendering the policy, gives up the interests embedded in his contract in favor of the insurance company, which ends up with a positive final value of the fund.

Liquidity risk due to lapse rates is dealt by insurance companies with the application of a penalty.

We run the simulation again, applying a penalty of 3.5%

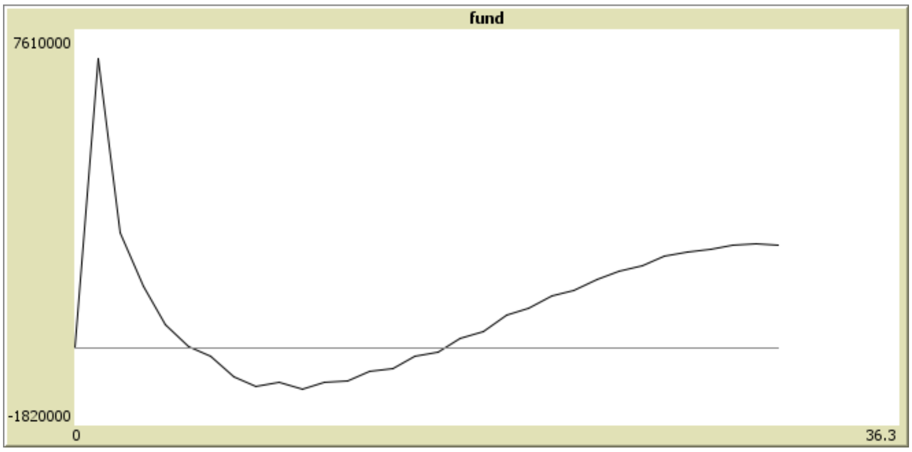
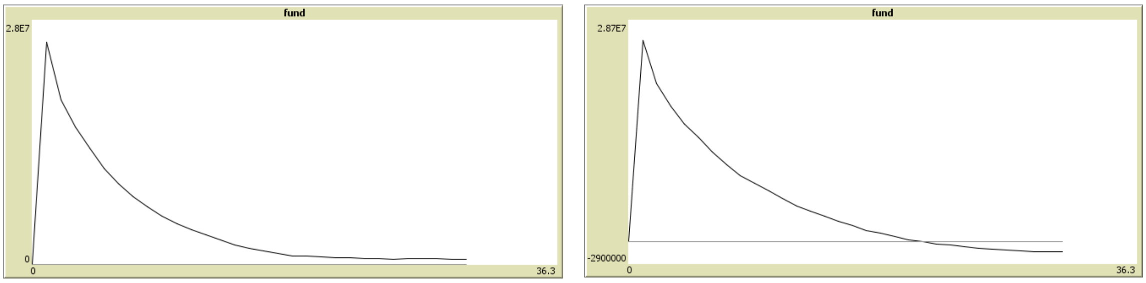


Figure 6

Applying a 3.5% penalty reduces the liquidity shortage, but it is not enough to face all the anticipated requests of capital. If the insurance company were to raise the penalty, it would lose market share on future contracts.

Another solution is to invest a fraction of the premiums, and keep the rest to pay back policyholders asking for the surrender value of their contract.

We run the model again, having the insurance company conservatively saving 50% of its premiums to face early surrenders.



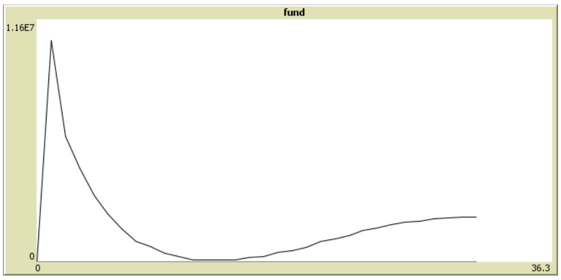
Lapse rate 3% Lapse rate 0.5%

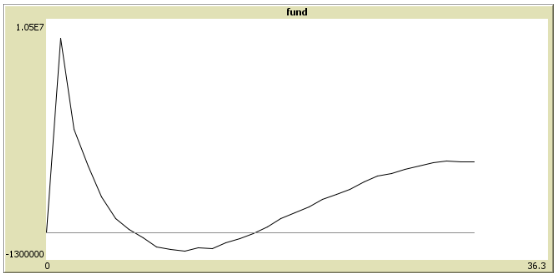
Table 1

This solution works only in case the insurance company correctly estimates future lapse rates. If the estimate happens to be too high and the effective lapse rate is 0.5% instead of 3%, the insurance company cannot meet its obligations.

If more policyholders than expected keep their contract and ask for their interests, investing only half of the premiums reveal itself to be a poor choice.

On the other side, if the company were to invest 90% of the premiums:





Lapse rate 3%

Lapse rate 5%

Table 2

If the estimate happens to be too low, the insurance company is still not prepared to pay back policyholders.

4.3 Version 3.

Introduction

This version of the model is an extension of the previous model. We are studying the liquidity of the insurance company while adding other two insurance products to the portfolio: a term life insurance and an endowment.

Before going further we need to explain a correction we applied in this version with respect to the order of the actions made by the agents in the previous version. Before we had the agents act in the following order:

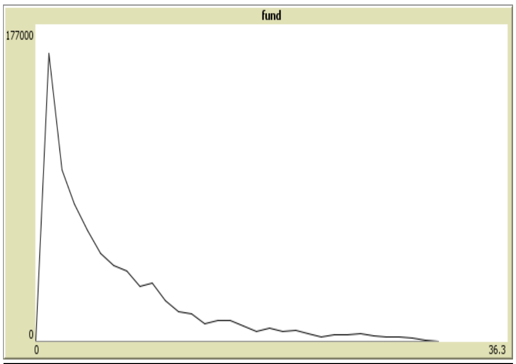
- The policyholders receive the benefit for the pure endowment component upon the following condition: `age = x + n`
- The policyholders age: `set age age + 1`
- The company receives the returns, at maturity, of the invested premiums: `set fund fund rserve:get " ds[ticks+1,2]"`
- The clock is updated: `tick`

The purpose of the investment is to cover the expected value of the benefit at maturity, therefore the maturity of the investment and the maturity of the benefit should occur at the same time step, but with the model specification above, we the insurance company first receives the return on the investment, and then the following year the policyholder receives the corresponding benefit, if alive.

The consequence is that the company delays any liquidity need at the end of the simulation, indeed we saw in the results of the previous version that, in the case the insurance company does not allow for surrender option, it does not experience liquidity issues until the extinction of all contracts.

Due to the actuarially fair financial and demographic variable, the expected value of the fund should be zero at each time step, while with the previous specification we have a decreasing fund and the expected value of the fund is zero only at the end of the simulation:

Previous model



Current Model

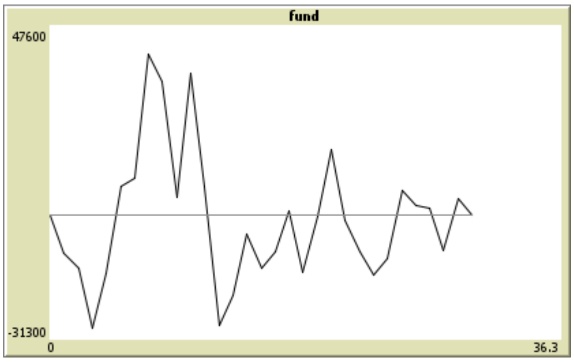


Table 3

In the current model the order of the actions is:

- We update the value of the fund with the investments of the three types of premiums
- The policyholders receive the benefit for the insurance component upon the condition: `random-float 1 < pDeath and age < x + n`
- The policyholders age: `set age age + 1`
- The policyholders receive the benefit for the pure endowment component upon the condition: `age = x + n`
- The clock is updated

Notice we set the maturity of the expected benefit of the pure endowment after the update of the age. This will anticipate the cash flow out due to the benefit one year earlier, in correspondence with the maturity of the relative premium.

The second step is due to the integration in the model of a term life insurance product which pays upon the condition that the insured party dies within 'n' years.

Theoretical Framework

We already explained how to calculate the equivalence premium for the pure endowment contract in the previous model.

The term life insurance contract binds the insurer to pay a benefit 'C' to the designated beneficiary upon the condition that the insured party, of age 'x', dies within 'n' years, before age x + n.

The uncertainty in the pure endowment contract is relative only to the realization of the event that the insurer will survive until age x + n. Such event has a probability equal to ${}_np_x = \frac{l_{x+n}}{l_x}$. If the event occurs, the payment is due at age x + n. There is no uncertainty on the timing of the benefit; therefore the insurer invests the premium at a maturity of ‘n’ years. Since the interest rate at ‘n’ years, i_n, has been used for the pricing of the contract, it will also be used to calculate the mathematical reserve in case the policyholder exit the contract prematurely.

The term life insurance contract pays when the insured party dies, if the event occurs within ‘n’ years. We have two levels of uncertainties, relative to the ‘if’ and the ‘when’ of the payment benefit. This has an effect on the maturity of the investment of the contract. In order to simplify calculation, in actuarial mathematics, we assume that the payment will be due at the end of the year of death.

If an insured party of age x = 70 underwrites a term life insurance of n = 2 years, over a capital insured of C = 100, the insurance company must provide for the expected value of the benefit for the first and the second year.

If the insured party dies within the first year with a probability of $1 - {}_1p_{50} = \frac{l_{50}-l_{51}}{l_{50}} = 0.01845$, the company will pay the benefit, 100, at the end of the first year, therefore it must invest at time 0 for a maturity of 1 year, with relative interest rate $i_1 = 0.0023$, in order to have the expected value of the benefit at the end of the first year: 1.845. The insurance company invests now at a 1 year maturity the amount $\frac{1.845}{1+0.0023} = 1.84$

If the insured party dies in the second year, we assumed that the benefit will be due at the end of the second year. The probability that the insured party will die the second year, evaluated at time 0, equal $\frac{l_{51}-l_{52}}{l_{50}} = 0.0202$. The insurance company must invest now at a two year maturity, earning the relative interest rate $i_2 = 0.0025$, such an amount to be able to cover the expected value of the benefit, 2.02, at the end of the second year. The insurance company invests now at a 2 year maturity the amount $\frac{2.02}{(1+0.0025)^2} = 2.01$

Notice that probability is evaluated at time 0: the conditioning on the event is that the insured party is alive at time 0, not at 1. We do not know in 0 if he will die within the first year. Let me stress that this probability is different from the values we have in the ‘mortalityTable’ which correspond to $\frac{l_x-l_{x+1}}{l_x}$.

Finally we found that the equivalence premium, which is the premium that make the expected profit of the insurer null, equals $1.84 + 2.01 = 3.85$.

The general formula for the equivalence premium is:

$$\sum_{h=0}^{n-1} \frac{C \cdot \dfrac{l_{x+h} - l_{x+h+1}}{l_{x+h}}}{(1 + i_{h+1})^{h+1}} = C \cdot {}_nA_x$$

Notice we are not using only one technical rate. We are discounting the expected value of the benefit for each maturity at the relative interest. Nonetheless we are required to provide a technical rate for the calculation of the reserve.

We need to find that value *x* such that:

$$\sum_{h=0}^{n-1} \frac{C \cdot \frac{l_{x+h} - l_{x+h+1}}{l_{x+h}}}{(1+i_{h+1})^{h+1}} = \sum_{h=0}^{n-1} \frac{C \cdot \frac{l_{x+h} - l_{x+h+1}}{l_{x+h}}}{(1+x)^{h+1}}$$

In the example before we have that $3.85 = 1.84\, v + 2.02\, v^{\,2}$, where we transformed the unknown in the discount factor. We take as technical rate the positive solution 0.002437

For higher order equation we have at maximum two real solutions, as in this case. Notice that a polynomial can cross the *x* axis more than two times, but for how the problem is structured the equation has at maximum two real roots, one positive and one negative and we set the technical rate equal to the positive solution :

$$\sum_{h=0}^{n-1} \frac{l_{x+h} - l_{x+h+1}}{l_{x+h}} v^{h+1} - {}_{n}A_{x} = 0$$

To calculate the solution we use the function *polyroot* of the base package of R. The function takes as argument the coefficients of the equation in increasing order.

For an individual of age x = 18, insuring the capital for n = 30 years we implement the code in the following way:

```
x<-18
n<-30
```

We define the relative contract features

```
h1qx<-(l[x:(x+n-1)]-l[(x+1):(x+n)])/l[x]
factors<-(1/(1+termStructure[1:n]))^(1:n)
nAx<-sum(factors*h1qx)
```

We calculate the equivalence premium which is the first coefficient of the equation

```
roots<-polyroot(c(-nAx,h1qx))
```

The vector h1qx is a list of the deferred death probabilities used for the calculation of the premium which is also the list of coefficients of the equation, recall that:

$${}_{h/1}q_{x} = \frac{l_{x+h} - l_{x+h+1}}{l_{x+h}}$$

```
realRoots<-Re(roots)[abs(Im(roots)) < 1e-6]
```

We store in this variable the real roots

```
techFactor<-realRoots[which(realRoots>0)]
```

We take as solution only the positive one

```
techRate<-1/techFactor-1
> techRate
[1] 0.01570015
```

We will use this rate to discount the future expected benefit of the insurer at the time the policyholder surrender the contracts. This value represents the exit price from the contract for the insurance company and is calculated as the premium but for a shorter coverage equal to the time left to maturity.

In order to avoid adverse selection life insurers offer the surrender option only if the contract provide a benefit with certainty (Perna and Sibillo, 2007), this is why we are including in this version of the model a third product called Endowment.

An endowment contract binds the insurer to pay the policyholder the capital when he dies, if the death occurs within a certain age, if not the insurer is still bound to pay the capital when the policyholder reaches the established age.

In this model we are giving the surrender option only to policyholder which purchased the endowment. Notice that such contract is the combination of a term life insurance and a pure endowment with equal maturity and capital insured.

The coefficient to calculate the premium for this contract equals the sum of the coefficients of the two contracts:

$$Endowment\ coefficient = \sum_{h=0}^{n-1} \frac{\frac{l_{x+h} - l_{x+h+1}}{l_{x+h}}}{(1+i)^{h+1}} + \frac{\frac{l_{x+n}}{l_x}}{(1+i)^n} = {}_nA_x + {}_nE_x$$

Notice that for $i=0$ the coefficient equals one, therefore the premium would equal the capital insured since the benefit is certain.

The Code

In the interface tab we added inputs to change the relevant variables for each type of contract. These variables are:

- The minimum and maximum capital insured accepted
- Minimum and maximum age of the policyholder
- Minimum and maximum duration of the contract

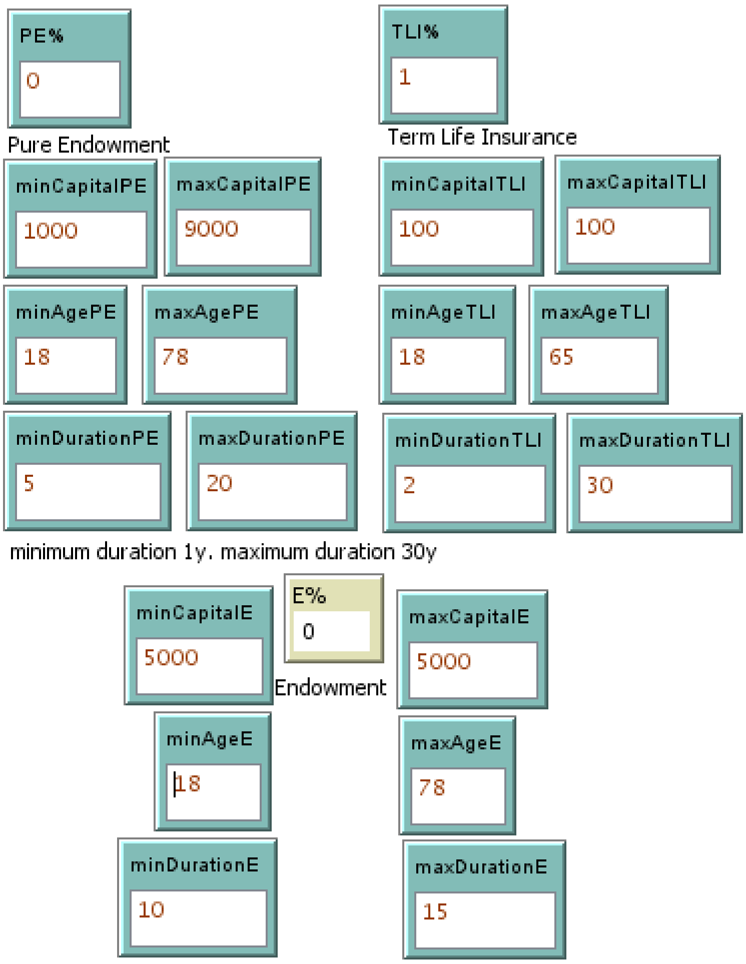


Figure 7

We can choose different ranges for each type of contract and we can choose the percentage of policyholders having the three types of contracts.

In this model the 'investedProportion' slider, refers only to the premium of the Endowment contract, as only for this contract we need to cover for the risk of an early withdrawal.

We now move to the code tab. As we go through the code, I will point out lines of code written to check the consistency of the model.

As in the previous model we store at the beginning of the simulation the "termStructure", "mortalityTable" and "lifeTable" both in the NetLogo and R memory.

In order to double check our results we introduce another list: the living for each age of the life table. We call this list as "l", which stands for living at age *x*.

We create the policyholders. We highlight that part of code defining the relevant contract features:

```
    let j random-float 1
    ifelse j < PE% [set guarantee "PE"] [ifelse j < (PE% + TLI%)
[set guarantee "TLI"][set guarantee "E"]]

    if guarantee = "PE"
    [
      set age minAgePE + random(1 + maxAgePE - minAgePE)
      set x age
      set n minDurationPE + random (1 + maxDurationPE -
minDurationPE)
      set benefitAge x + n
      set C minCapitalPE + random (1 + maxCapitalPE - minCapitalPE)
    ]
```

Similarly for the other two types of guarantees.

We define the premiums through the “define-premium” policyholder procedure.

```
  ask policyholders
  [
    rserve:put "x" x
    rserve:put "n" n
    rserve:put "C" C
    set npx rserve:get "l[x+n]/l[x]"
    set techRate_PE item (n - 1) termStructure
    set nEx (1 + techRate_PE)^(- n) * npx
    ;set finalValue C * npx ;check

    if guarantee = "PE"[set U_PE C * nEx]
```

We calculate the premium for the pure endowment as specified in the previous model. If the policyholder holds this type of contract, the premium is defined as the capital multiplied by the relative coefficient.

To double check the capitalization of the premium of the insurance company, we store in the “finalValue” variable the expected value of the benefit the insurer must provide for at maturity. This check works only if invested proportion equals 1, otherwise the “finalValue” variable will be higher.

The following two blocks of code are used to define the premium for the term life insurance contract

```
    ;rserve:eval "h1qx2 <- c(rep(0,n))"
    ;rserve:eval "for (h in 0:(n-1)){h1qx2[(h+1)]<-(l[x+h]-
l[x+h+1]) / l[x]}" ;check
    rserve:eval "h1qx<-(l[x:(x+n-1)]-l[(x+1):(x+n)])/l[x]"
    set h1qx rserve:get "h1qx" ;check
```

We calculate the deferred death probabilities. Although it is not necessary to import the “h1qx” variable in NetLogo we store it in order to check that for each policyholder the comand `sum(h1qx) + npx` returns one. We can also check that the first element in the h1qx list correspond to the mortality rate used to simulate the deaths.

```
rserve:eval "factors<-1/((1+termStructure[1:n])^(1:n))"
set nAx rserve:get "sum(factors*h1qx)"
```

We calculate the coefficient ${}_nA_x$.

```
;rserve:put "nAx" nAx
;rserve:eval "coeffs <- c(-nAx,h1qx)"
;rserve:eval "roots<-polyroot(c(-nAx,h1qx))"
;rserve:eval "realRoots<-Re(roots)[abs(Im(roots)) < 1e-6]"
;rserve:eval "techFactor<-realRoots[which(realRoots>0)]"
```

The above code should be used to compute the corresponding technical rate as we explained in the theoretical framework. Unfortunately the “polyroot” function does not work through the “rserve” library. By running that line of code, NetLogo returns this error: `java.lang.NullPointerException`.

I was not able to find the cause of the error, but I found a loophole to avoid the problem. The procedure to avoid the problem is called “write-data_TLI” and is implemented after the “define-premium” procedure. We will view it now not to lose the thread of the argument and we will view the end of the “define-premium” procedure after.

```
to write-data_TLI
  if file-exists? "tRateData.csv" = true
  [file-delete "tRateData.csv"]

  ask policyholders
  [
    file-open "tRateData.csv"
    file-write who
    file-write nAx
    file-write x
    file-write n file-type "\n"
    file-close
    ]
end
```

The above procedure write in a csv file the necessary data to have R directly compute the equivalent technical rate since NetLogo is not able to do so through the rserve library.

At the end of the “setup” procedure we need to run the following R script to compute the technical rate from the `"tRateData.csv"` created by the `"write-data_TLI"`. Be sure to set as working directory, through the setwd() function, the same folder where you have the NetLogo model where the previous procedure has created the csv file.

```
data_tli <- read.csv("tRateData.csv",header=F, sep="")
```

We read the dataset created by NetLogo.

```
data_tli<-cbind(data_tli,rep(0,nrow(data_tli)))
data_tli<-cbind(data_tli,rep(0,nrow(data_tli)))
data_tli<-cbind(data_tli,rep(0,nrow(data_tli)))
names(data_tli)<-
c("who","nAx","x","n","nAxCheck","nAxCheck2","techRate_TLI")
```

The following is how the data will look like at this point

```
> data_tli[1:5,]
  who         nAx  x  n nAxCheck nAxCheck2 techRate_TLI
1 128 0.146071279 64 10        0         0            0
2  69 0.455133738 60 26        0         0            0
3 207 0.070035202 56 10        0         0            0
...
```

We will compute again in R the coefficient using directly the interest rates of the term structure for the first check.
We check again the value using the equivalent technical rate computed through the polyroot function. We store the technical rate in the last column.

```
termStructure <- read.csv("termStructure.csv",header=F)
termStructure<-termStructure$V1
l <- read.csv("l.csv",header=F)
l <- l$V1
```

We also need the term structure and the living at age x. Then we can run the following for cycle:

```
for (i in 1:nrow(data_tli)){

data_tli[i,5]<-
sum((1/(1+termStructure[1:data_tli[i,4]])^(1:data_tli[i,4]))*
((l[data_tli[i,3]:(data_tli[i,3]+data_tli[i,4]-1)]-
l[(data_tli[i,3]+1):(data_tli[i,3]+data_tli[i,4])])/l[data_tli[i,3]
]))
h1qx<-((l[data_tli[i,3]:(data_tli[i,3]+data_tli[i,4]-1)]-
l[(data_tli[i,3]+1):(data_tli[i,3]+data_tli[i,4])])/l[data_tli[i,3]
])
roots<-polyroot(c(-data_tli[i,2],h1qx))
realRoots<-Re(roots)[abs(Im(roots)) < 1e-6]
techFactor<-realRoots[which(realRoots>0)]
techRate<-1/techFactor-1
data_tli[i,7]<-techRate
data_tli[i,6]<-sum((1/(1+data_tli[i,7])^(1:data_tli[i,4]))*
((l[data_tli[i,3]:(data_tli[i,3]+data_tli[i,4]-1)]-
l[(data_tli[i,3]+1):(data_tli[i,3]+data_tli[i,4])])/l[data_tli[i,3]
]))

}
```

At the end of the cycle we check we have computed correctly the technical rate:

```
> data_tli[1:5,]
  who         nAx  x  n    nAxCheck   nAxCheck2 techRate_TLI
1 128 0.146071279 64 10 0.146071279 0.146071279  0.006557198
2  69 0.455133738 60 26 0.455133738 0.455133738  0.014203705
3 207 0.070035202 56 10 0.070035202 0.070035202  0.006585670
4 299 0.016568825 18 29 0.016568825 0.016568825  0.015205805
5 197 0.008430648 26 14 0.008430648 0.008430648  0.008858495
```

We write the data in a table which will be located in the same folder of the NetLogo model with the name `"tRateIN_TLI.csv"`.

```
write.table(data_tli[,c(1,7)], file =
"tRateIN_TLI.csv",row.names=FALSE,col.names=FALSE, sep=" ")
```

We go back to the NetLogo model and press the “read-data_TLI” button which calls the following procedure

```
to read-data_TLI

  file-open "tRateIN_TLI.csv"
  while [file-at-end? = false]
  [
    let ww file-read
    ask policyholder ww [set techRate_TLI file-read]
    ]
  file-close

end
```

In the above procedure we ask each policyholder to set as techRate_TLI variable the value found in the table in correspondence of the relative “who” variable.

We can now go back to the “define-premium” procedure:

```
    set investmentByMaturity rserve:get "C*(factors*h1qx)"
    ;set finalValueByMaturity rserve:get "C*h1qx" ;check

    if guarantee = "TLI"[set U_TLI C * nAx]

    if guarantee = "E"[set U_PE C * nEx set U_TLI C * nAx]

      ]
end
```

Similarly for how we checked the pure endowment we are directly storing the final value of the investments made without capitalizing the relative premium. We stress again that this check can be used only if the insurer invests the whole premium.

Finally, depending on the type of guarantee, we set the premium. To make the computation more efficient we divide the premium of the endowment in the two relative components.

In the setup procedure we are also creating the following three datasets in R

```
 rserve:put "num_people" num-people
 rserve:eval "temp_TLI<-
matrix(0,nrow=30,ncol=num_people)"
 rserve:eval "temp_E_TLI<-
matrix(0,nrow=30,ncol=num_people)"
 rserve:eval "tempFinalValue_TLI<-
matrix(0,nrow=30,ncol=num_people)"; check
```

We will use these temporary arrays later in the code to calculate the investment of the term life insurance and the term life insurance component in the endowment contracts. We will put in each column the amount of the investment for each policyholder for every maturity (the rows) and later we will sum the rows to obtain the amount invested for each maturity.

The following part of code is a policyholder procedure run at the beginning of the simulation. We are using the policyholder own variable ‘done’ to have the policyholder make these actions only at the beginning of the code, therefore this procedure will end asking the policyholders to set this variable equal to 1. This procedure is used to set up the necessary variable for R and NetLogo to calculate the following output: ‘ds_PE’, ‘finalValueList_TLI’, ‘ds_E_PE’ and ‘finalValueList_E_TLI’.

They are lists, of 30 elements, as many as the maximum duration we can set for a contract, of the final value of the investments made from the capitalization of the premiums obtained from the Pure endowment, the term Life insurance and the endowment (the pure endowment component and the insurance component) respectively.

At each step we are increasing the fund of the element in the corresponding position; for example by issuing pure endowments contract with a minimum duration of 2 years and a maximum duration of 10 years, ds_PE, where ds stands for data set given the structure we used to build it, has the following shape:

maturity	amount
1	0
2	637.223
…	…
10	636.248
11	0
…	0
30	0

Table 4

Since the company has nothing invested with a maturity of one year, at the first time step we are not increasing the fund. The same holds for the years between the 11th and the 30th .

We had to split the investments because when we are changing the invested proportion of the premiums to face the early surrenders, we are applying it only to the premiums of the endowment since the surrender option is given only to this type of contract.

```
ask policyholders with [done = 0]
  [
    if guarantee = "PE"
    [
      let uu_pe U_PE let nn n
      ;let fv finalvalue ;check
      ask insurer num-people
      [
        set premiumList_PE lput uu_pe premiumList_PE
        set maturityList_PE lput nn maturityList_PE
        set interestRates_PE lput item (nn - 1) termStructure
interestRates_PE
        ;set finalvalueList_PE lput fv finalvalueList_PE ;check
        ]
       ]
```

We are asking each policyholder to add the the 'finalValue' variable we defined previously, to check later if it corresponds to the capitalization of the premium.

```
    if guarantee = "TLI"
    [
      let uu_tli U_TLI
      let ibm investmentByMaturity
      let w who + 1
      let fvbm finalValueByMaturity ;check
      ask insurer num-people
      [
        set premiumList_TLI lput uu_tli premiumList_TLI
        rserve:put "investmentByMaturity_TLI" ibm
        rserve:put "w" w ask insurer num-people[rserve:eval
"temp_TLI[1:length(investmentByMaturity_TLI),w]<-
t(investmentByMaturity_TLI)"]
        ;rserve:put "finalValueByMaturity_TLI" fvbm
        ;rserve:put "w" w ask insurer num-people[rserve:eval
"tempFinalValue_TLI[1:length(finalValueByMaturity_TLI),w]<-
t(finalValueByMaturity_TLI)"] ;check
        ]
      ]
```

We added in the temporary dataset created at the beginning of the code the amount of the investments for each maturity which will be later capitalized. The "tempFinalValue" dataset contains directly the final value of the investments and by summing its rows we obtain what we defined as 'finalValueList_TLI'.

```
    if guarantee = "E"
    [
      let uu_pe U_PE let nn n
      ask insurer num-people
      [
        set premiumList_E_PE lput uu_pe premiumList_E_PE
        set maturityList_E_PE lput nn maturityList_E_PE
        set interestRates_E_PE lput item (nn - 1) termStructure
interestRates_E_PE
        ]

      let uu_tli U_TLI
      let ibm investmentByMaturity
      let w who + 1
      ask insurer num-people
      [
        set premiumList_E_TLI lput uu_tli premiumList_E_TLI
        rserve:put "investmentByMaturity_E_TLI" ibm
        rserve:put "w" w ask insurer num-people[rserve:eval
"temp_E_TLI[1:length(investmentByMaturity_E_TLI),w]<-
t(investmentByMaturity_E_TLI)"]
        ]
      ]
    set done 1
  ]
```

Finally we repeat the same code but for the endowment contract.

```
      if paid = 0 [
        ask insurer num-people
        [set fund fund + sum [U_TLI] of policyholders + sum [U_PE]
      of policyholders]
        set paid 1
        ]
```

We have the policyholders pay the premiums.

```
    ask insurers with [invested = 0]
    [
      ;;;; Pure Endowment
      rserve:put "premiumList_PE" premiumList_PE
      rserve:put "maturityList_PE" maturityList_PE
      rserve:put "interestRates_PE" interestRates_PE
      rserve:eval "investmentList_PE <- premiumList_PE"
      rserve:eval "investment_PE <- sum(investmentList_PE)"
      set fund fund - rserve:get "investment_PE"
      rserve:eval "finalvalueList_PE <- investmentList_PE * (1 +
interestRates_PE) ^ maturityList_PE"
      rserve:eval "ds_PE <- data.frame(cbind(maturityList_PE,
finalvalueList_PE))"
      rserve:eval "tempVec <- c(rep(0,30))"
      rserve:eval "for(i in 1:30){tempVec[i] <- i}"
      rserve:eval "dsIntegration <-
data.frame(cbind(tempVec,c(rep(0,30))))"
```

```
      rserve:eval "names(dsIntegration) <- names(ds_PE) "
      rserve:eval "ds_PE <- rbind(ds_PE,dsIntegration)"
      rserve:eval "ds_PE <- aggregate(. ~ maturityList_PE,
data=ds_PE, sum)"
```

With respect to the previous code notice we are investing the whole premiums for the endowment contract.

```
      ;;;; Term Life Insurance
      rserve:put "premiumList_TLI" premiumList_TLI
      rserve:eval "investmentByMaturity_TLI<-rowSums(temp_TLI)"
      rserve:eval "investment_TLI<-sum(premiumList_TLI)"
      set fund fund - rserve:get "investment_TLI"
      rserve:eval "finalValueList_TLI<-
investmentByMaturity_TLI*(1+termStructure)^(1:30)"
      ;rserve:eval "finalValueList_TLI2<-
rowSums(tempFinalValue_TLI)" ;check
```

We capitalize the investments for the relative maturity and check if it equals the final value calculated directly. As we do not give the option to exit the contract for the term life insurance, we are investing the whole premium.

```
      ;;;;;;;; Endowment
      rserve:put "premiumList_E_PE" premiumList_E_PE
      rserve:put "investedProportion" investedProportion
      rserve:put "maturityList_E_PE" maturityList_E_PE
      rserve:put "interestRates_E_PE" interestRates_E_PE
      rserve:eval "investmentList_E_PE <- premiumList_E_PE *
investedProportion"
      rserve:eval "investment_E_PE <- sum(investmentList_E_PE)"
      set fund fund - rserve:get "investment_E_PE"
      rserve:eval "finalvalueList_E_PE <- investmentList_E_PE * (1
+ interestRates_E_PE) ^ maturityList_E_PE"
      rserve:eval "ds_E_PE <- data.frame(cbind(maturityList_E_PE,
finalvalueList_E_PE))"
      rserve:eval "tempVec <- c(rep(0,30))"
      rserve:eval "for(i in 1:30){tempVec[i] <- i}"
      rserve:eval "dsIntegration <-
data.frame(cbind(tempVec,c(rep(0,30))))"
      rserve:eval "names(dsIntegration) <- names(ds_E_PE) "
      rserve:eval "ds_E_PE <- rbind(ds_E_PE,dsIntegration)"
      rserve:eval "ds_E_PE <- aggregate(. ~ maturityList_E_PE,
data=ds_E_PE, sum)"

      rserve:put "premiumList_E_TLI" premiumList_E_TLI
      rserve:eval "investmentByMaturity_E_TLI<-
rowSums(temp_E_TLI)*investedProportion"
      rserve:eval "investment_E_TLI<-
sum(premiumList_E_TLI)*investedProportion"
      set fund fund - rserve:get "investment_E_TLI"
      rserve:eval "finalValueList_E_TLI<-
investmentByMaturity_E_TLI*(1+termStructure)^(1:30)"

      set invested 1
```

```
      ]
```

The endowment is divided in an analogous way but we are multiplying the investment by the 'investedProportion' variable.

The following code will repeat at each step until the end of the simulation.

```
  ask insurer num-people
  [
    rserve:put "ticks" ticks
    if ticks < 30
    [
      set fund fund + rserve:get "as.numeric(ds_PE[ticks+1,2])"
     ; show (word "From PE premiums: + " rserve:get
"as.numeric(ds[ticks+1,2])") ;check
      set fund fund + rserve:get "finalValueList_TLI[ticks+1]"
     ; show (word "From TLI premiums: + " rserve:get
"finalValueList_TLI[ticks+1]");check
      set fund fund + rserve:get "as.numeric(ds_E_PE[ticks+1,2])"
     ; show (word "From E_PE premiums: + " rserve:get
"as.numeric(ds_E_TLI[ticks+1,2])") ;check
      set fund fund + rserve:get "finalValueList_E_TLI[ticks+1]"
     ; show (word "From E_TLI premiums: + " rserve:get
"finalValueList_E_TLI[ticks+1]");check
      ]
    ]
```

The value of the fund is updated as the investments reaches their maturities. We can show the increments by type of premium invested in the interface tab during the simulation.

```
 ask policyholders
  [
    if age < x + n and random-float 1 < lapseRate and lapsed = 0
and guarantee = "E"
    [
      rserve:put "age" age
      rserve:put "time" (benefitAge - age)
      rserve:eval "probs <- lifeTable[(age):(age+time-1)]"
      set time_p_age rserve:get "prod(probs)"
      let n-tEx+t (1 + techRate_PE)^(- (benefitAge - age)) *
time_p_age
      set Vt_PE C * n-tEx+t

      ;rserve:put "benefitAge" benefitAge
      ;rserve:put "age" age
      ;set time_p_age2 rserve:get "l[benefitAge]/l[age]" ;check

      rserve:put "x" x
      rserve:put "n" n
      rserve:put "t" age - x
      rserve:eval "h1qage<-(l[age:(x+n-1)]-
```

```
l[(age+1):(x+n)])/l[age]"
      rserve:put "techRate_TLI" techRate_TLI
      let n-tAx+t rserve:get "sum( (1/(1+techRate_TLI)^(1:(n-t))) *
h1qage)"
      set Vt_TLI C * n-tAx+t

      set surrenderValue (Vt_PE + Vt_TLI) / (1 +
penalty)^(benefitAge - age)
      let sv surrenderValue
      ask insurer num-people [set fund fund - sv]
      set lapsed 1
      ]
```

We allow for the surrender option only for the endowment contract. The policyholder withdraws the value of the mathematical reserve of the two contracts discounted by the penalty.
The probability that the insured party survives until maturity at the evaluation date is both calculated as the product of the one year survival probabilities and as the ratio of the living at x + n and x.

```
to check-death ;tp
  set pDeath item (age - 1) mortalityTable
  if random-float 1 < pDeath [if guarantee != "PE" and age <
x + n and lapsed = 0
    [
      let cc C
      ask insurer num-people [set fund fund - cc]
      ;show  (word"Benefits  due  to  premature  deaths:  -  "  C)
;check
      ]
    die]
end
```

For those policyholders that insured the capital for the death event (term life insurance and endowment) we have the insurer reduce the fund value if the policyholder dies before x + n.

```
set age age + 1

if age = x + n and lapsed = 0 and guarantee != "TLI"
[
  let cc C
  ;show (word"Benefits due to survivors: - " C) ;check
  ask insurer num-people [set fund fund - cc]
  ]
```

If the policyholder does not die, we increase by one its age, and if he reaches the benefit age x + n the insurer is bound to pay the capital.

Results

In this model we introduced a new and fairly complicated theoretical framework for the integration of the term life insurance. Therefore the first analysis aims at finding possible mistakes in the code by looking at the output of the model.

We run the model 500 times and record the final value of the fund for a portfolio composed of term life insurance products. We plot a histogram of the results.

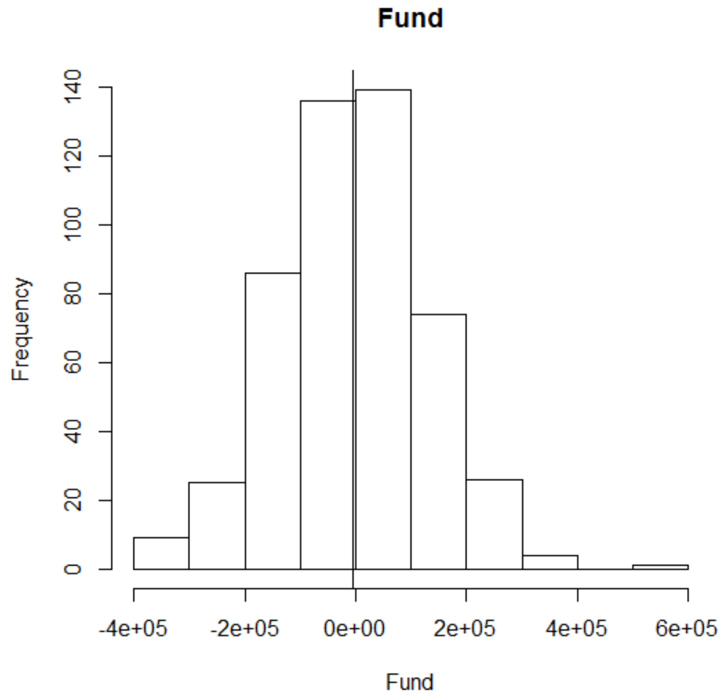


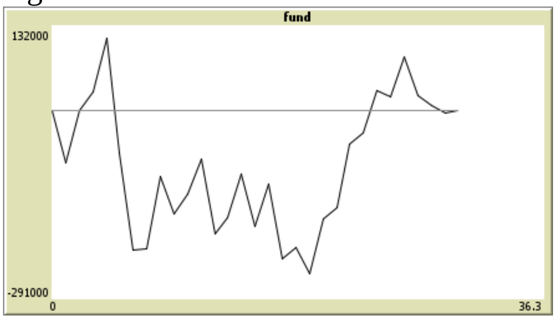
Figure 8

As expected we find that the frequencies of the final value of the fund are centered in zero and we plot a vertical line in correspondence of the mean of the random sample.

Model Specification 1.

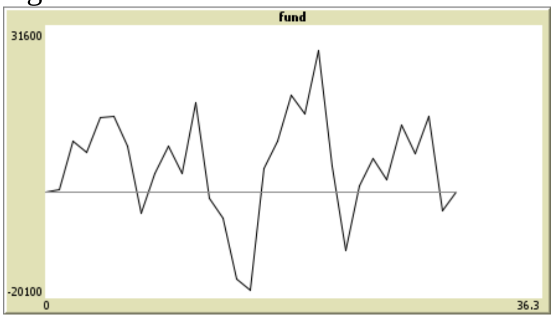
# Policyholder	Penalty	Lapse probability	Invested proportion
10.000		0	100%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			60-80 and 20-30
Duration			2-30

Age: 60-80



Minimum fund value: -252 thousand
Maximum fund value: 112 thousand

Age: 20-30



Minimum fund value: -18 thousand
Maximum fund value: 26 thousand

Table 5

For a portfolio composed of only Endowments, where the payment is certain, as we provide the policyholder the same interest the insurer earns, the final value of the fund will be zero, therefore we are now more interested in the fluctuations of the fund before each contract has reached its maturity.

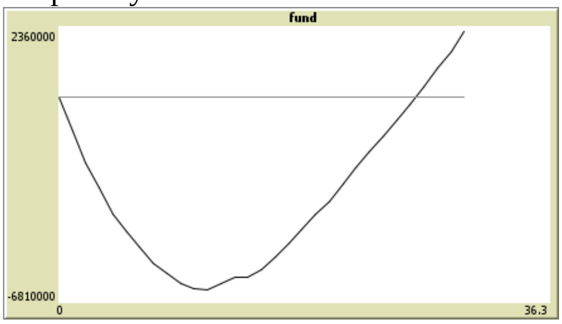
By insuring a portfolio of individuals between 60 and 80 years old, we notice a fluctuation of the fund between 112.000 and 252.000 from 0. Consider we are insuring a total of $10.000 \cdot 5.000 = 50\, million$ and the maximum deviation from the expected value, 0, corresponds to the 0.50% of the total sum insured.

With a younger portfolio, between 20 and 30 years of age, the variability diminishes from a 18.000 low to a 26.000 high.

Model Specification 2

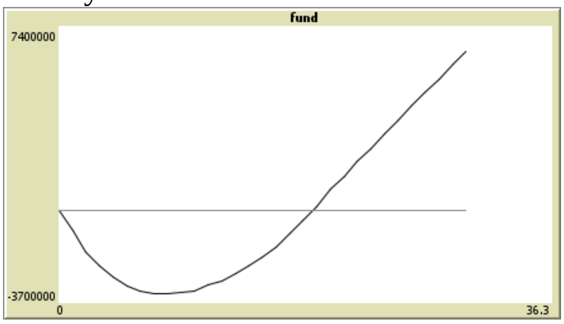
# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	0 and 3.5%	**3%**	100%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			18-80
Duration			2-30

No penalty



Minimum fund value: -6.383 thousand
Maximum fund value: 2.210 thousand

Penalty: 3.5%



Minimum fund value: -3.363 thousand
Maximum fund value: 6.381 thousand

Table 6

The value of the fund becomes negative at the beginning of the simulation due to the high number of early surrenders and the longer maturity of the investment of the company. Then, as those investment reach maturity, the value of the fund becomes positive since the policyholders who lapsed gave up part of the interest they would have earned by holding the contract until maturity.

With a portfolio completely exposed to lapse risk (100% of endowment contracts), and a 3% probability of lapsing the policy, the insurance company faces a liquidity shortage of about 6 million at the 11th year. Although we can see a clear path in the value of the fund, due to the high number of lapses (at the end of the simulation around 20% of the policyholders have lapsed the policy), the variability of the minimum value of the fund is high but we cannot appreciate it by watching the graph since the minimum and maximum value displayed in the graph differ of about 8.5 million.

To have an idea, in 50 simulations, the minimum value of the fund averaged -6.38 million, and it reached a minimum, among all simulations, of -6.72 million

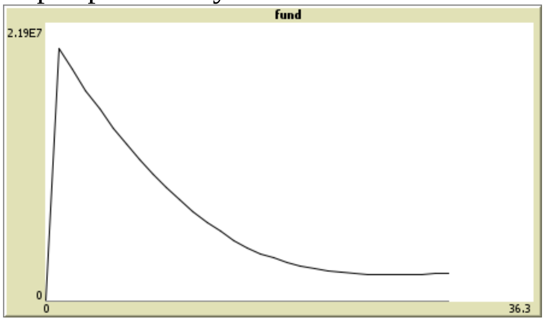
By applying a penalty of 3.5% the liquidity shortage lowers down to 3.4 million at the 9th year.

Model Specification 3

Similarly for what we did in the previous version of the model, we deal with the early surrenders by investing half of the premiums of the endowment contract.

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	3% and 0.5%	**50%**
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			18-80
Duration			2-30

Lapse probability: 3%



Lapse probability: 0.5%

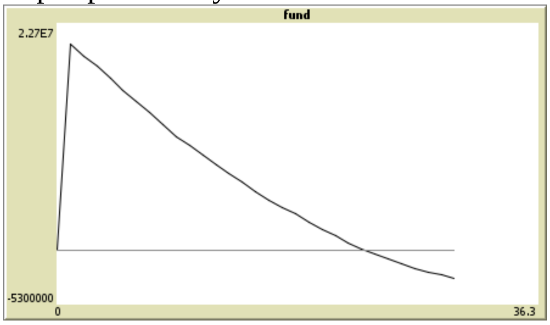


Table 7

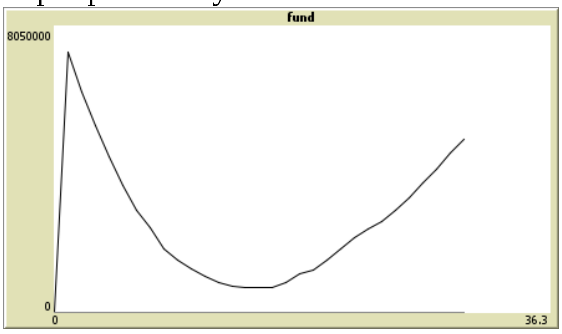
If our estimate on the lapse activity of the policyholders is correct, this strategy works. If our estimate on the lapse rate is too high, investing half of the premiums is not enough to meet the obligations of those contracts we did not expect they would have reached maturity.

Model Specification 4

We test the strategy of saving only 20% of the premium under the usual 3% lapse rate assumption and under a stressed scenario.

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	3% and 5%	**80%**
Product	Pure endowment	Term life insurance	Endowment
Percentage	0	0	100%
Sum insured			5.000
Age			18-80
Duration			2-30

Lapse probability: 3%



Lapse probability: 5%

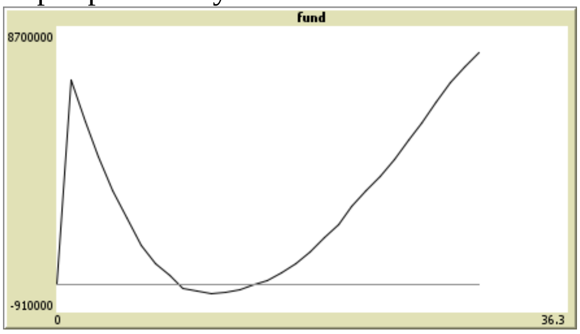


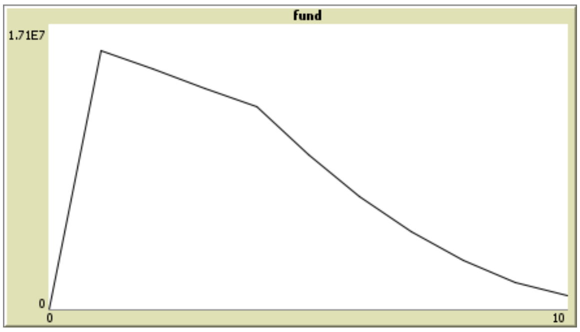
Table 8

If our assumption is correct, the strategy works, but if the estimate is too low, the amount saved is not enough. With the 5% lapse probability the minimum value of the fund averaged a – 412 thousand in 50 simulations.

Model Specification 5

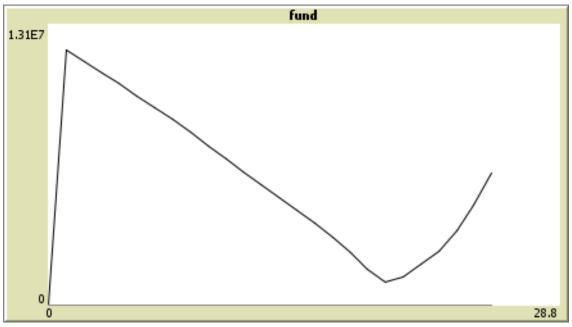
# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	3%	65%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			18-30
Duration			**5-10** and **20-25**

Duration: 5-10



Number of early surrenders: 2.016
Final fund value: 825 thousand

Duration: 20-25



Number of early surrenders: 4.893
Final fund value: 6.158 thousand
Minimum fund value: 1.053 thousand

Table 9

In these two simulations the lapse rates is the same, but in the first we are giving the surrender option to contracts lasting between 5 and 10 years, therefore most of this contract will reach maturity, while in the second simulation nearly half of the portfolio has surrendered the contract. Moreover the penalty paid to surrender the contract will be higher on average as the time to maturity is longer.

Model Specification 6

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	5%	80%
Product	Pure endowment	Term life insurance	Endowment
Percentage	40%	40%	**20%**
Sum insured	5.000	5.000	5.000
Age	18-80	18-80	18-80
Duration	2-30	2-30	2-30

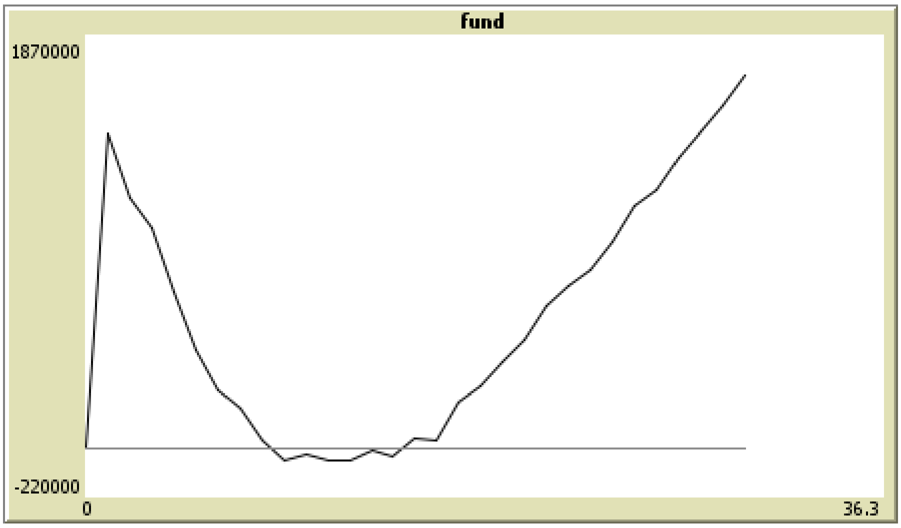


Table 10

In this model specification the insurance company is exposed to lapse risk only on 20% of the portfolio. With respect to the 4th model specification we record, on average, a minimum value of the fund of –133 thousand.

Model Specification 7

Similarly for what we did with the pure endowment under the previous version of the model, we test if the variability of the final value of the fund of an insurance company with a portfolio of term life insurances is higher if it allows for variability in the sum insured of the contracts.

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000			100%
Product		Term life insurance	
Percentage		100%	
Sum insured		**5K and 1-9K**	
Age		30-70	
Duration		2-30	

Table 11

We run the model for both specifications defined above (one where all policyholders insure a fixed sum of 5.000, another allowing to insure a sum between 1.000 and 9.000) 500 times.

We record in a spreadsheet the final value of the fund.

The sample variance of the random sample obtained from the model specification where we allowed for variability in the sum insured among policyholders is greater than the variance of the other sample; their ratio equals 0.79

We use the Shapiro-Wilk statistic to test the null hypothesis that the final value of the fund under the two model specifications is normally distributed and we cannot reject the hypothesis in either case (p-values of 0.82 and 0.67, in the first and second specification respectively).

Assuming normality of the distributions we use a variance ratio test of the equality of two variances; with a p-value of 0.0088 we reject the null hypothesis for the alternative: the true ratio of the variances does not equal 1.

4.4 Version 4.

With regard to the error returned by the 'polyroot' function, my supervisor and I have been able to find a better loophole to the problem.

The run of the R script from the R console after the setup procedure will not be necessary anymore. In the setup procedure we are having NetLogo write that script through the 'file-write' built in function of NetLogo, and we are also asking NetLogo to run it through the 'source' R function within the Rserve extension. This avoids calling the 'polyroot' function directly from the Rserve extension and the program does not return an error.

In conclusion we are able to run the model by calling the "setup" and "go" procedures only.

Introduction

This version of the model is an extension of the previous model where policyholders surrender their policy with a probability obtained from the coefficients of the generalized linear model of Milhaud et al. (2010).

In the previous model all policyholders lapsed their policy with the same probability defined by the user. In this model each policyholder has his own probability of lapsing the contract which will depend on the variable taken into account in the regression made by Milhaud et al. (2010).

The model used in their paper is a binomial GLM with logit link function.

The variable considered are: contract age, premium frequency, underwriting age, face amount, saving premium, risk premium, contract type and gender.

We report the coefficients from their paper:

Coef. (var. type)	modality : correspondance	coefficient estimate	std error	p-value	effect
β_0 (continuous)		10.63398	1.48281	7.42e-13	>0
$\beta_{duration}$ (categorical)	1 : [0,12] (in month)	0 (reference)			nul
	2 :]12,18]	-1.31804	0.15450	$<2e-16$	<0
	3 :]18,24]	-2.66856	0.14016	$<2e-16$	<0
	4 :]24,30]	-2.75744	0.14799	$<2e-16$	<0
	5 :]30,36]	-3.09368	0.14294	$<2e-16$	<0
	6 :]36,42]	-3.54961	0.15080	$<2e-16$	<0
	7 :]42,48]	-3.72161	0.14980	$<2e-16$	<0
	8 :]48,54]	-4.10431	0.15772	$<2e-16$	<0
	9 : >54	-5.49307	0.14037	$<2e-16$	<0
$\beta_{premium\ frequency}$ (categorical) (in month)	Monthly	0 (reference)			nul
	Bi-monthly	0.92656	0.62071	0.135504	>0
	Quarterly	-0.03284	0.10270	0.749148	<0
	Half-yearly	-0.22055	0.16681	0.186128	<0
	Annual	0.43613	0.10690	4.51e-05	>0
	Single	-0.28494	0.38155	0.455177	<0
$\beta_{underwriting\ age}$ (categorical)	1 : [0,20[(years old)	0 (reference)			nul
	2 : [20,30[	0.28378	0.13912	0.041376	>0
	3 : [30,40[	-0.01146	0.13663	0.933163	<0
	4 : [40,50[	-0.26266	0.14077	0.062054	<0
	5 : [50,60[	-0.42098	0.15136	0.005416	<0
	6 : [60,70[	-0.66396	0.19531	0.000675	<0
	7 : >70	-0.75323	0.23417	0.001297	<0
$\beta_{face\ amount}$ (categorical)	1* :	0 (reference)			nul
	2* :	-5.79014	1.46592	7.82e-05	<0
	3* :	-7.14918	1.46631	1.08e-06	<0
$\beta_{risk\ premium}$ (categorical)	1* :	0 (reference)			nul
	2* :	0.36060	0.11719	0.002091	>0
	3* :	0.26300	0.14041	0.061068	>0
$\beta_{saving\ premium}$ (categorical)	1* :	0 (reference)			nul
	2* :	0.93642	0.13099	8.74e-13	>0
	3* :	1.32983	0.14955	$<2e-16$	>0
$\beta_{contract\ type}$ (categorical)	PP con PB	0 (reference)			nul
	PP sin PB	-16.79213	114.05786	0.882955	<0
	PU con PB	-7.48389	1.51757	8.16e-07	<0
	PU sin PB	-12.43284	1.08499	$<2e-16$	<0
β_{gender}	Female	0 (reference)			nul
	Male	-0.08543	0.04854	0.078401	<0

* Note : for confidentiality reasons, the real ranges of the face amount, the risk premium and saving premium are omitted.

Table 12

Notice they define contract age as duration when instead they do not refer to the overall duration of the contract but at the contract age. We notice that the probability of lapsing the policy decreases as the contract approaches maturity.

The linear predictor in GLM estimates a function of the variable of interest, which in this case is a probability between 0 and 1.

The logistic regression is a regression of the logarithm of the odd ratio:

$$logit(p) = log\left(\frac{p}{1-p}\right) = a + bx + \mathrm{cy} + \cdots$$

If x is a categorical variable like contract age and all the variable of this paper, then its coefficient, b, has value -5.49 for contract after the 9th year of age. We can compare it to the reference level, 0, (for contract in the first year of their lives) keeping all other variables at their reference level too except for "contract type" which we set to -12.43 as all the contracts in our portfolio have a unique premium (PU) and there is no profit benefit participation scheme (sin PB):

$$log\left(\frac{p*}{1-p*}\right) = \mathrm{a} - 5.49 - 12.43 \text{ and } log\left(\frac{p}{1-p}\right) = \mathrm{a} - 12.43$$

Where a is the intercept, $p*$ is the estimate of the probability of lapsing the contract for a policyholder with all variables at the reference level except for the contract age variable which is greater than 9, and p is the estimate of the probability of lapsing the contract for a policyholder with all variables at the reference level.

$$\frac{\frac{p*}{1-p*}}{\frac{p}{1-p}} = \exp(-5.49) = 0.0041$$

By exponentiation of the coefficient we obtained the ratio of the odds ratio: the probability of lapsing the policy is much lower for contract at the final stages of their life.

$$p* = \mathrm{logit}^{-1}(\mathrm{a}+\mathrm{b}+\mathrm{c}) = \frac{1}{1+\exp-(10.63-5.49-12.43)} = 0.06\%$$
$$p = \mathrm{logit}^{-1}(\mathrm{a}+\mathrm{b}) = \frac{1}{1+\exp-(10.63-12.43)} = 14\%$$

The above example provides a tool to the reader to interpret the coefficients.

Since in our model the time unit is the year, we will make an average of the coefficients provided in their paper

For the variable “premium frequency”, as for the variable “contract type”, we are setting the coefficient to -0.28 which corresponds to the payment of the premium in a lump sum, as in our model.

For the variables “face amount”, “saving premium” and “risk premium” we use the reference level since the author does not provide real ranges for these variable for confidentiality reasons.

With regard to the variable “gender” the p-value associated with it does not seem statistically relevant, as the author points out.

We are using the variable “underwriting age” which corresponds to the variable ‘x’ in our model. The probability of lapsing the policy is the highest for policyholder between 20 and 30 years of age and then it gradually decreases.

The code

We set the intercept as a global variable common to all policyholder.

```
set intercept 10.63 - 0.28 - 12.43
```

The first element correspond to the real intercept of the model, to which we add the coefficient -0.28 as all policyholders pay the premium in a lump sum and the coefficient -12.43 since the type of contracts modeled to not include a profit participation scheme for the policyholder.

Then we add as policyholder own variable the following: ‘pLapse’, ‘coeffContractAge’ and ‘coeffUwAge’.

The first is the lapse probability, the second and the third are the coefficients relative to “contract age” and “underwriting age” which will depends on the NetLogo variable transformations ‘age – x’ and ‘x’ respectively.

The variables modeled in the paper are categorical ones, therefore we will need a for cycle to set the coefficients to the proper level:

```
ask policyholders
[
  set contractAge age - x
  ifelse contractAge = 0 [set coeffContractAge 0]
  [
    ifelse contractAge = 1 [set coeffContractAge -1.985]
    [
      ifelse contractAge = 2 [set coeffContractAge -2.92]
      [
        ifelse contractAge = 3 [set coeffContractAge -3.63]
        [
          ifelse contractAge = 4 [set coeffContractAge -4.1]

          [set coeffContractAge -5.49]
          ]
        ]
      ]
```

```
      ]

    ifelse x >= 0 and x < 20 [set coeffUwAge 0]
    [
      ifelse x >= 20 and x < 30 [set coeffUwAge 0.28]
      [
        ifelse x >= 30 and x < 40 [set coeffUwAge -0.011]
        [
          ifelse x >= 40 and x < 50 [set coeffUwAge -0.26]
          [
            ifelse x >= 50 and x < 60 [set coeffUwAge -0.42]
            [
              ifelse x >= 60 and x < 70 [set coeffUwAge -0.66]
              [set coeffUwAge -0.75]
              ]
            ]
          ]
        ]
      ]

    set pLapse 1 / (1 + exp(- (intercept + coeffContractAge +
coeffUwAge )))
    if age < x + n and random-float 1 < pLapse and lapsed = 0 and
guarantee = "E"
    [
      ...
```

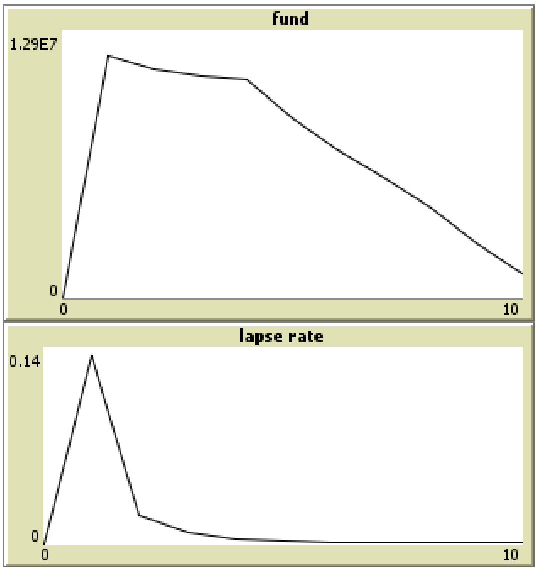
Finally we obtain the lapse probability by mean of the inverse of the logit function and have the policyholder lapse with such probability.

Results

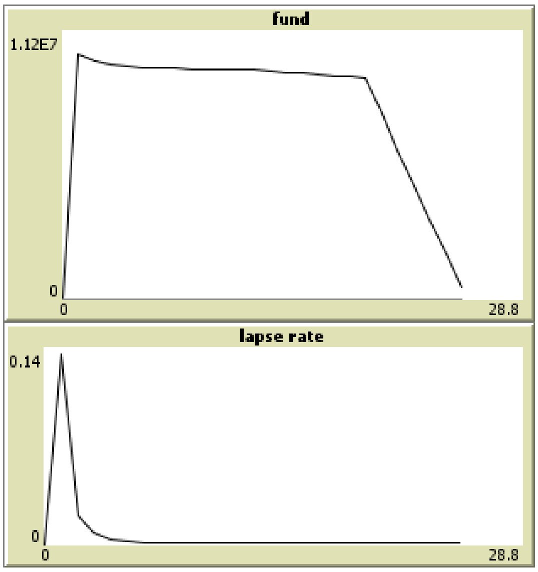
Model Specification 1

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	*Policyholder own*	65%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			18-30
Duration			**5-10** and **20-25**

Duration: 5-10



Duration: 20-25



Number of early surrenders: 1653
Final fund value: 1.224 thousand

Number of early surrenders: 1.695
Final fund value: 518 thousand

Table 13

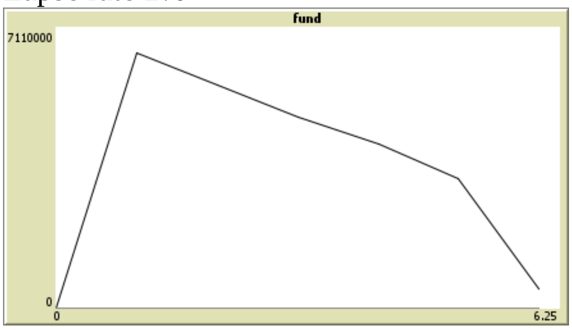
The first model specification (on the left) gave rise to an output similar to the model specification 5, in the previous version of the model: lapse rate is not as high to justify such a low proportion of invested premiums; indeed, as contracts reach maturity, the insurance company has not collected enough from its investment to meet obligations and its fund reaches a dangerous minimum of 1.224 thousand.

In the previous model (specification 5 of previous version), in correspondence of the image on the right, we had a completely different situation: although lapse rate had the same intensity as the model with lower maturities, the effect was a high number of surrenders due to the long maturity of the investments; instead in this new version, although the maturity has the same length (20 – 25 years), considering the lower lapse probability for contracts in the final stages of their life, surrenders are not as numerous as the previous version of the model (1.695 against 4.893). In the end the company has to meet the obligations for more contracts than expected and with the highest interests rates because of the long term nature of these contracts, but the proceeds obtained from the small fraction of invested premiums is barely sufficient to meet those obligations and it jeopardizes the liquidity of the firm.

Model Specification 2

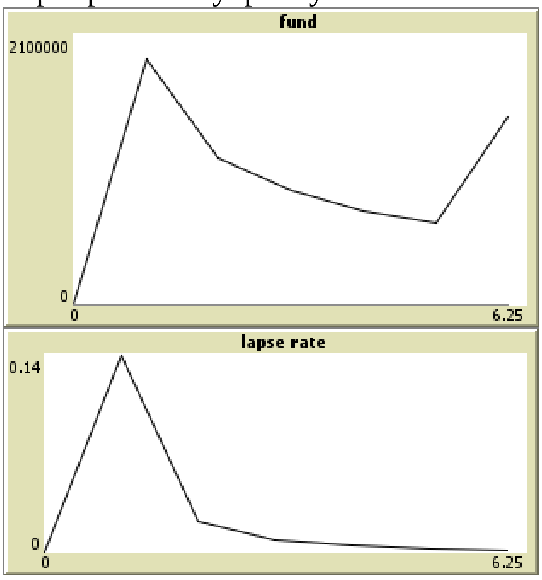
# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	**2% and P.h.-own**	85%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			20-30
Duration			6

Lapse rate 2%



Number of early surrenders: 1.089
Final fund value: 491 thousand

Lapse probability: policyholder-own



Surrenders: 1.672 or 16,72%
Final fund value: 1.454 thousand
Minimum fund value: 636 thousand

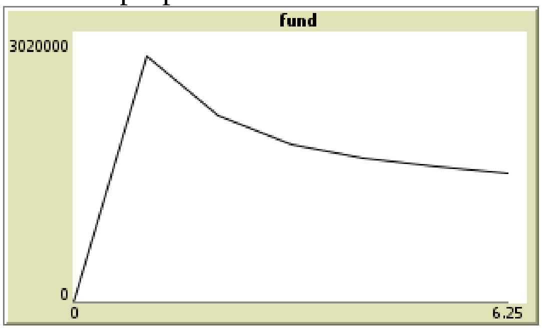
Table 14

In this case we have the opposite situation. With the previous model specification, the 2% lapse rate hypothesis is too low for investing only 85% of the premiums, while with the new model specification, the younger age of the policyholders and the short term maturity of the investment yields in a higher number of surrenders which decreases the value of the fund in the first years, jeopardizing liquidity, and in the end the company receives the proceeds of most part of the investments without having to repay the policyholders who surrendered earlier increasing the final value of the fund.

Consider the randomness of these simulations and the riskiness of investing a too high fraction of the premiums: see for example a different outcome of the simulation (figure on the right) for a smaller number of insured parties (1.000) and therefore a higher variability in lapses.

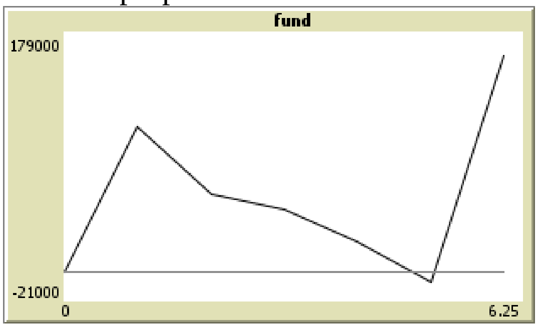
Assuming the correct estimation of the coefficients of Milhaud et al. (2010), a less risky action by the management would be to invest a lower fraction of the premiums (83% figure on the left).

Invested proportion: 83%



Number of early surrenders: 1.684
Final fund value: 1.447 thousand

Invested proportion: 85%



Surrenders: 18% of all portfolio

Table 15

Chapter 5

Comparative Analysis

5.1 Comparative analysis without the surrender option

For this type of analysis the reader must bear in mind that the invested proportion of the premiums is 100% as there is no lapse risk.

Mutuality principle.

The second version of the model (paragraph 4.2) has the following specifications:

Contract type: Pure endowment
Sum insured: 5.000
Underwriting age: 35-65
Maximum benefit age: 65

We recorded the variance of the following statistics from a sample of size 500 (500 simulations):

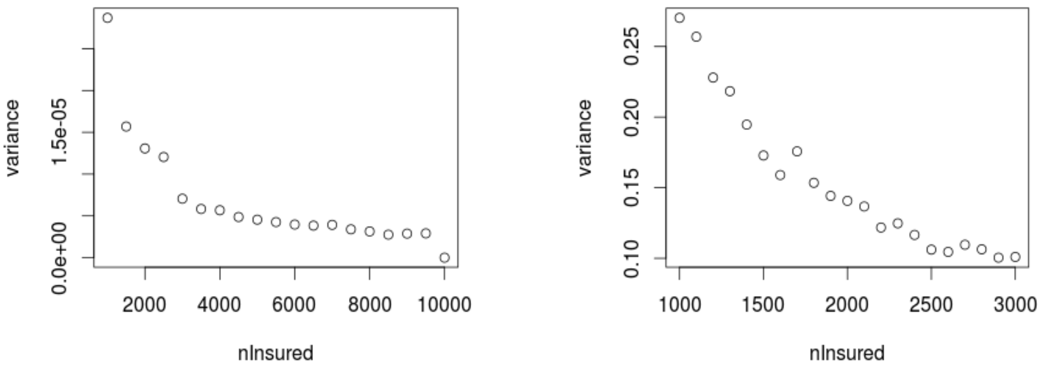
$$\hat{X}_i = \frac{final\ 'fund'value}{total\ sum\ insured} \qquad \text{with } i = 1.000, 1.500, \dots, 10.000$$

Where *i* is the number of policyholders in each simulation.

The result is a set of 19 random samples of size 500:

$$\underline{X_i} \text{ with } i = 1.000, 1.500, \dots, 10.000$$

In figure 2 we reported the variances of these random samples



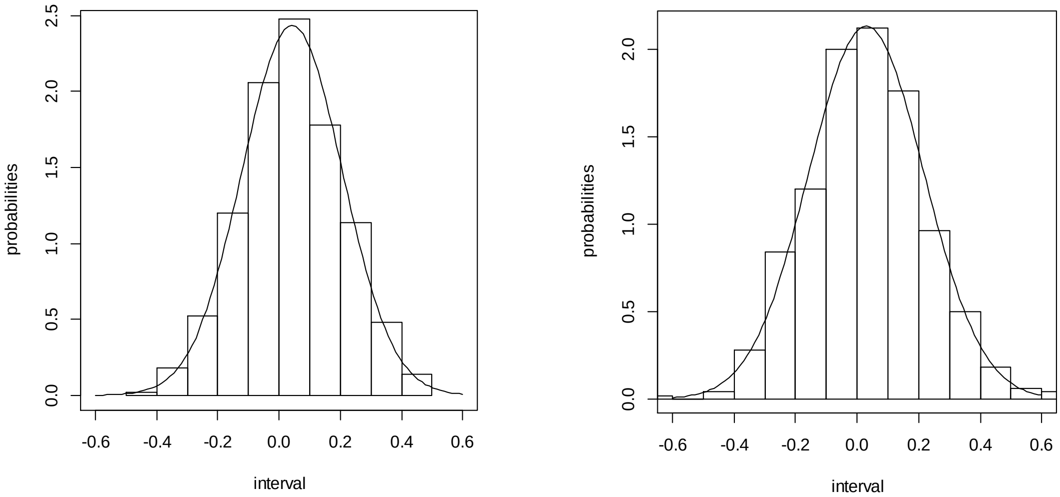
With the this model specifications we notice that above 3.000 policyholders the variance of the statistics decreases at a slower pace.

Fixed and variable sum insured.

We run the model under the previous model specification, but with a sum insured varying between 1.000 and 9.000, 500 times and we record in each simulation the value of the statistic defined before. We call the random sample $\underline{\mathrm{X}}_{\mathrm{variable\ C}}$ and we compare its variance to the sample called $\underline{\mathrm{X}}_{10.000}$ which we will now refer to as $\underline{\mathrm{X}}_{\mathrm{fixed\ C}}$:

$$\frac{\mathrm{VAR}\left(\underline{\mathrm{X}}_{fixed\ C}\right)}{\mathrm{VAR}\left(\underline{\mathrm{X}}_{\mathrm{variable\ C}}\right)} = 0.76$$

The sample variance of the random sample at the numerator is lower. After testing the normality of the two random samples with the Shapiro-Wilk statistic, we assume that the above statistics are normally distributed:



We can therefore assume that the ratio of the two variances has a Fisher distribution and we can test the null hypothesis that the two variances are equal. A variance ratio test on the two random samples returns a p-value of 0.03%, we therefore reject the null hypothesis in favor of the alternative: the true ratio of the population variance does not equal 1 and the variance of the final value of the fund is higher by allowing for variability in the sum insured.

Under model specification 7 in paragraph 4.3, we conducted a similar analysis for term life insurance policies:

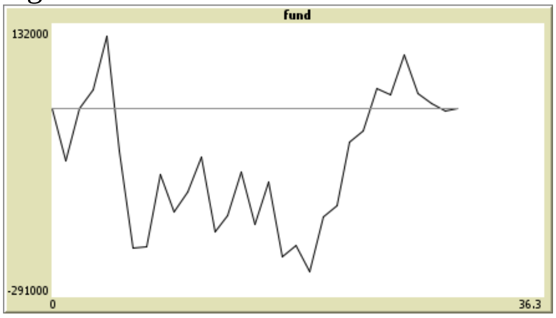
Contract type: Term Life insurance
Age: 30 - 70
Duration: 2 - 30
Sum insured: 5K and 1 - 9 K

In this case the statistic used is directly the final value of the fund; Also for term life insurance products the variance of the sample obtained from the model where we allowed for variability in the sum insured has greater variance with a variance ratio of 0.79; by first assuming normality of the final value of the fund, a variance ratio test on the two samples returned a p-value of 0.0088.

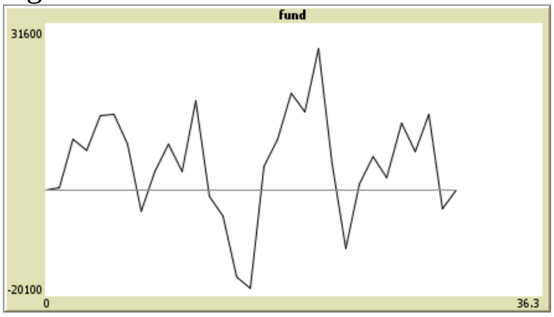
Variability in the value of the fund due to age.

Under model specification 1 in paragraph 4.3 where we have a portfolio of endowment policies, we looked at the value of the fund through all the steps in two scenarios: one where we insure older policyholders (60-80) and one where we insure younger ones (20-30):

Age: 60-80



Age: 20-30



Minimum fund value: -252 thousand
Maximum fund value: 112 thousand

Minimum fund value: -18 thousand
Maximum fund value: 26 thousand

The variability in the first case is much higher.
By running the model many times one could record the number of times the fund crosses a certain threshold and estimate a probability.

5.2 Comparative analysis with the surrender option

The effect of the penalty.

Model Specification 2. Paragraph 4.3.

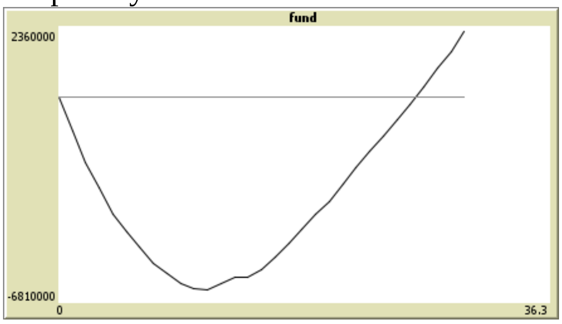
In this model, given a 3% year constant lapse rate, we first allow the policyholder to withdraw the whole mathematical reserve without a penalty and in the second run we apply a penalty of 3.5% specified as:

$$Surrender\ Value_t = \frac{Mathematical\ Reserve_t}{(1 + penalty)^{-(n-t)}}$$

In both run we invest 100% of the premiums with a portfolio of only Endowment contracts:

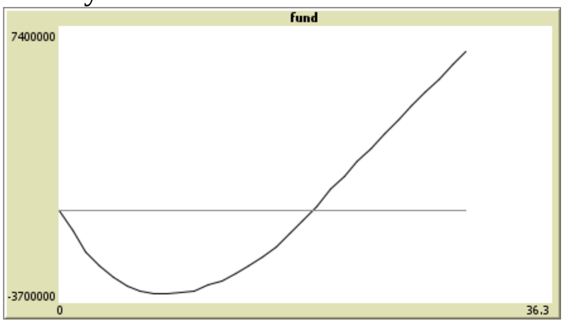
Sum insured: 5.000
Age: 18-80
Duration: 2-30

No penalty



Minimum fund value: -6.383 thousand
Maximum fund value: 2.210 thousand

Penalty: 3.5%



Minimum fund value: -3.363 thousand
Maximum fund value: 6.381 thousand

Without the penalty the value of the fund is negative at the beginning of the simulation due to early surrenders, then, as the corresponding investment reach maturity, the value of the fund becomes positive since the policyholders who lapsed gave up part of the interest they would have earned by holding the contract until maturity.

Notice that, except for pure endowment contracts, when I refer to “maturity” I am also referring to the possibility that the insurer dies and receives the whole sum insured. A policyholder dying one year after underwriting the contract receives the whole sum insured, while a policyholder, by surrendering the contract after one year receives the mathematical reserve discounted by the time to maturity, which is a much smaller value than the sum insured.

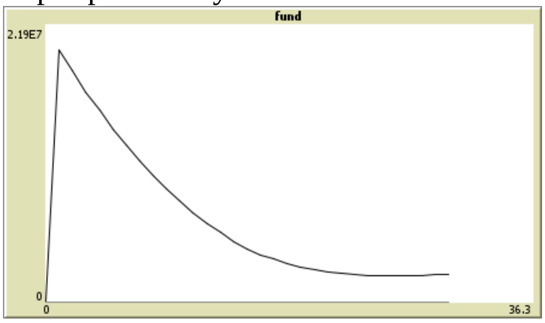
By applying a penalty of 3.5% the liquidity shortage lowers down to 3.4 million and the final value of the fund is higher due to the combination of the penalty and financial discount of the reserve.

Invested proportion.

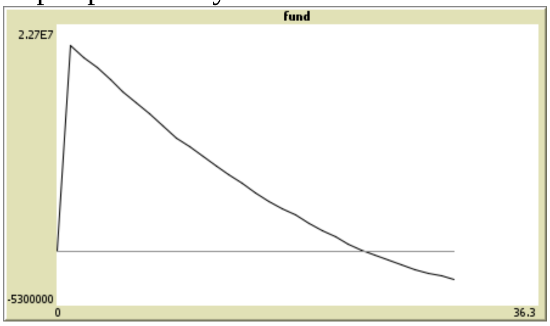
Model Specification 3. Paragraph 4.3

We have the insurer invest half of the premiums to cover for early surrenders.
We compare the outcome of a model where our estimate of the lapse rate is correct at a 3% annual rate with another scenario of lower lapse rates.

Lapse probability: 3%



Lapse probability: 0.5%

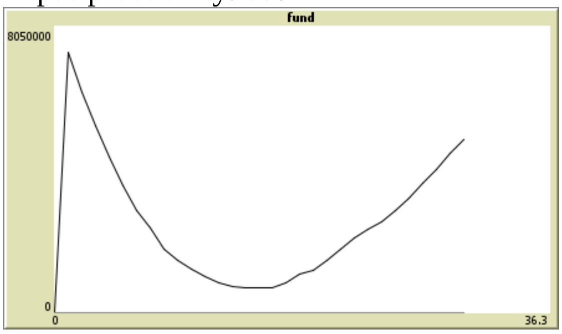


If our estimate on the lapse activity of the policyholders is correct the strategy works.
For the stressed scenario investing half of the premiums is not enough to meet the obligations of those contracts we did not expect they would have reached maturity.
By reaching maturity the policyholder asks for the whole sum insured which the company is not able to recover from its low investment.

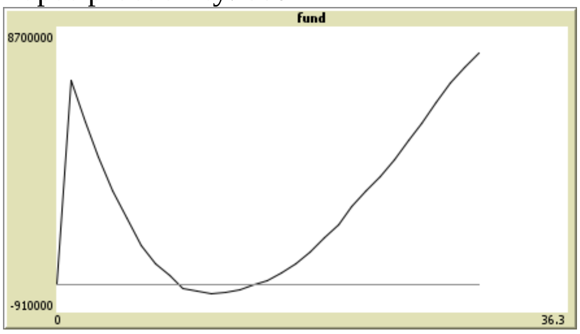
Model Specification 4. Paragraph 4.3

We test the strategy of saving only 20% of the premium under the usual 3% lapse rate assumption and under a scenario of increased lapse rates.

Lapse probability: 3%



Lapse probability: 5%



Under the stressed scenario the amount saved is not enough to meet the requests of early surrenders. With the 5% lapse probability the minimum value of the fund averaged a – 412 thousand in 50 simulations.

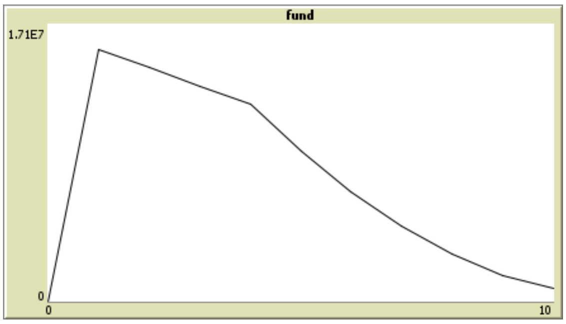
The effect of the duration of the contract on actual lapse activity.

Model Specification 5. Paragraph 4.3

In both the following simulations we set a constant 3% lapse rate, but in the first the policyholders insure the capital for 5-10 years while in the second the policyholder asks for a longer coverage.

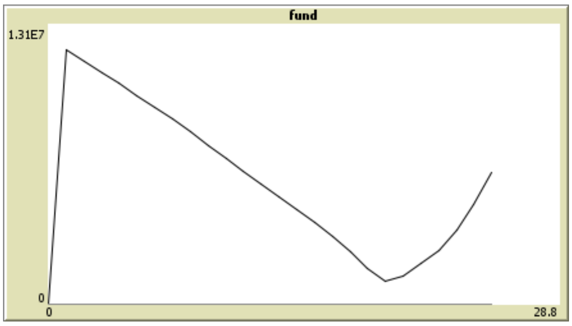
In both model specifications the company invests 65% of the premiums.

Duration: 5-10



Number of early surrenders: 2.016
Final fund value: 825 thousand

Duration: 20-25



Number of early surrenders: 4.893
Final fund value: 6.158 thousand
Minimum fund value: 1.053 thousand

In the first specification due to the shorter duration most of the contracts reach maturity asking for the whole sum insured and the 65% invested proportion of the premiums is not enough to meet obligations, while in the second simulation nearly half of the portfolio has surrendered the contract and this brings the fund to a dangerous 1.053 thousand minimum due to the too high invested proportion, but it then increases its value due to two reasons:

- On average a higher penalty is paid to surrender the contract because of the longer maturity and therefore the longer time to maturity when the policyholder chooses to surrender the contract.

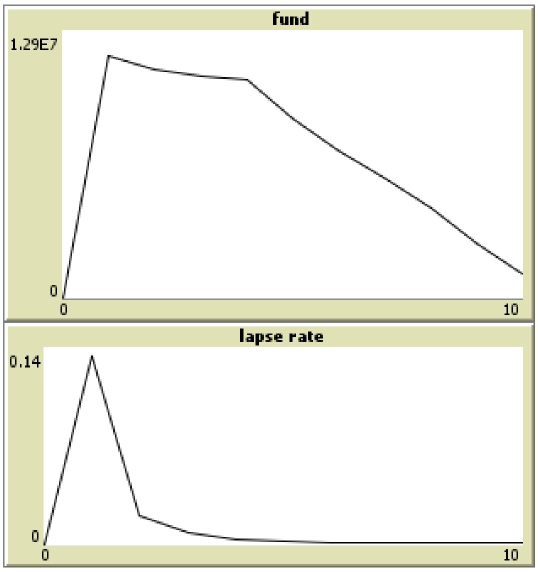
- The actual number of surrenders is higher due to the longer maturity and therefore a small number of policyholders will ask the whole sum insured.

Policyholder specific lapse probability.

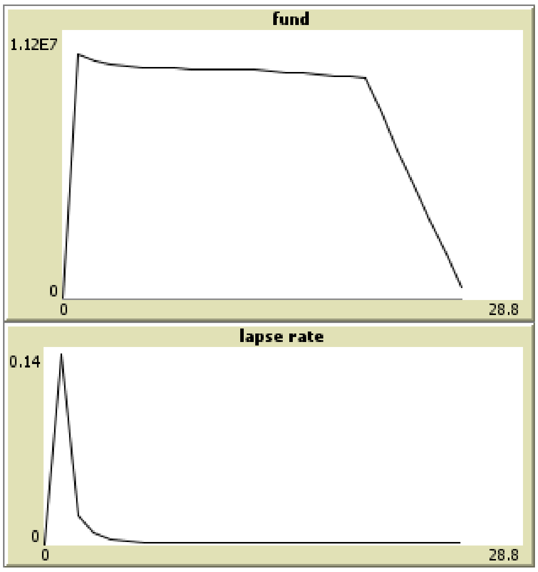
We make the same comparison, but with the last version of the model (4^{th}) where we have policyholders lapse with a probability based on the coefficient of the logit regression of Milhaud et al. (2010).

Model Specification 1. Paragraph 4.4.

Duration: 5-10



Duration: 20-25



Number of early surrenders: 1653
Final fund value: 1.224 thousand

Number of early surrenders: 1.695
Final fund value: 518 thousand

For the image on the left we have a similar situation of the previous model specification: about 2 thousand policyholders surrender the contract with a similar outcome.

Since we have policyholders surrender the contract with the probability defined by Milhaud et al. (2010), total surrenders do not increase as much due to the decreasing coefficient for the variable ‘contract age’ (see table 12 in the paragraph 4.4).

The final outcome is an under investment of the premiums which becomes evident when survivors start asking for the sum insured at year 25. The drop in the value of the fund is even more evident due to the high interest rates because of the long term nature of these contracts.

Comparison between a flat unique lapse rate and a dynamic policyholder specific lapse probability.

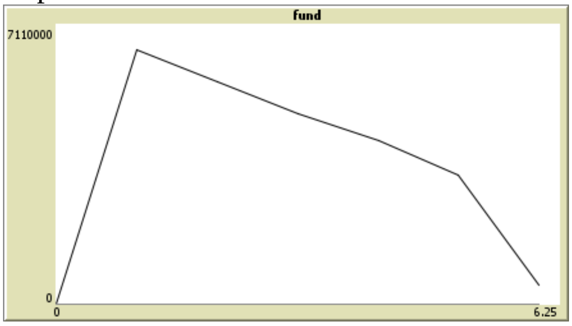
Both the following model specification have:

- An age at underwriting between 20 and 30 years.
- A contract duration of 6 years.
- An invested proportion of 85%.

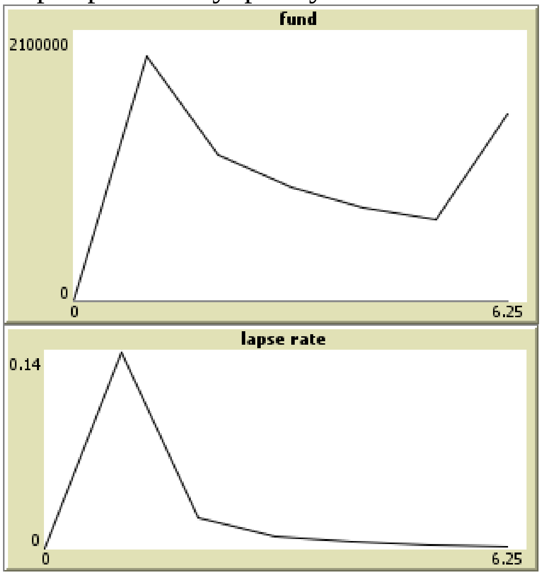
On the left we have the development of the fund assuming a 2% constant lapse rate; on the right we use the coefficients of the paper by Milhaud et al. (2010).

Model Specification 2. Paragraph 4.4.

Lapse rate 2%



Lapse probability: policyholder-own



Number of early surrenders: 1.089
Final fund value: 491 thousand

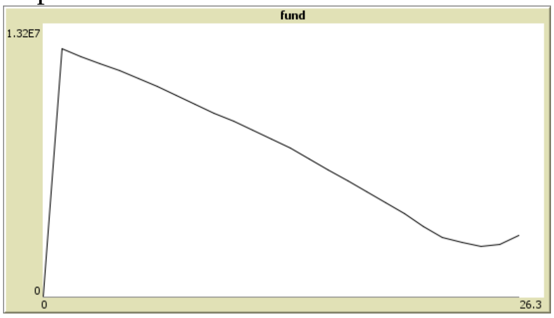
Surrenders: 1.672 or 16,72%
Final fund value: 1.454 thousand
Minimum fund value: 636 thousand

With a short term maturity (6 years to which it corresponds a high lapse probability based on the coefficients in table 12) and the young underwriting age of the contract (see table 12) total surrenders increases from 1.089 to 1.454 whereas by switching to the new specific policyholder lapse probability with the long maturities,

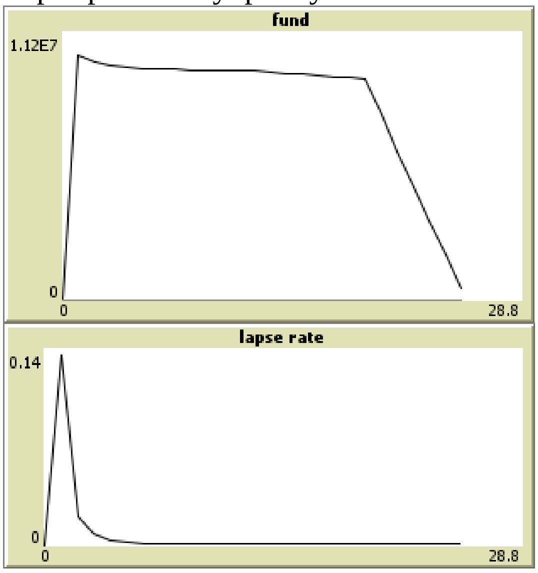
# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	*2% and P.h.-own*	65%
Product	Pure endowment	Term life insurance	Endowment
Percentage			100%
Sum insured			5.000
Age			18-30
Duration			**20-25**

we have the following transition,

Lapse rate 2%



Lapse probability: policyholder-own



Number of early surrenders: 3.625
Minimum fund value: 2.456 thousand
Final fund value: 2.970 thousand

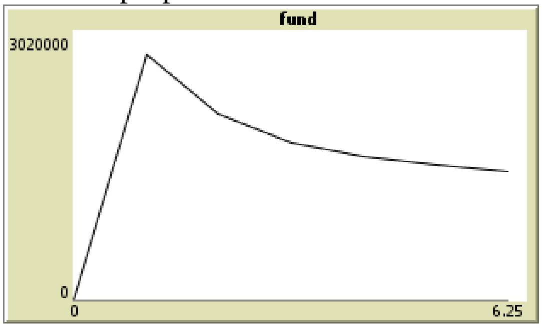
Number of early surrenders: 1.695
Final fund value: 518 thousand

which describes the opposite result. While with the short term maturities the insurer experienced an underinvestment of the premiums with the 2% flat rate, he should instead increase the amount invested for longer term maturities if he believes in the new framework set by Milhaud et al. (2010).

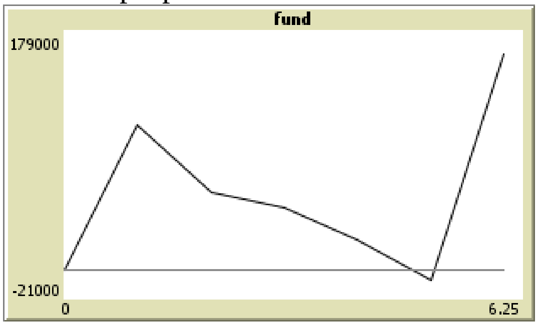
Consider the randomness of these simulations and the riskiness of investing a too high fraction of the premiums: In the next simulation we see what the outcome may have been with a smaller number of insured parties and therefore a greater variability for the short term maturity setting (figure on the right).

Assuming the correct estimation of the coefficients of Milhaud et al. (2010), a less risky action by the management would be to invest a lower fraction of the premiums (83% figure on the left).

Invested proportion: 83%



Invested proportion: 85%



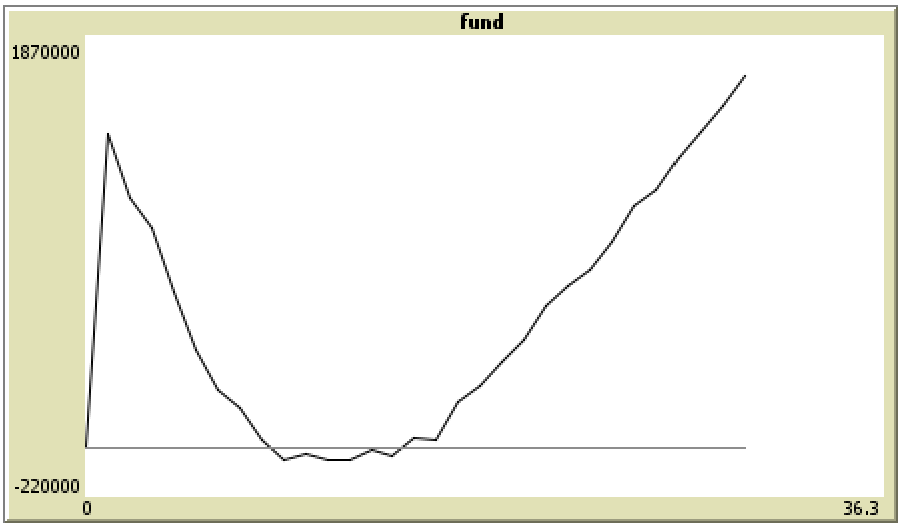
Number of early surrenders: 1.684
Final fund value: 1.447 thousand

Surrenders: 18% of all portfolio

A diversified portfolio.

Model Specification 6. Paragraph 4.3

# Policyholder	Penalty	Lapse probability	Invested proportion
10.000	3.5%	5%	80%
Product	Pure endowment	Term life insurance	Endowment
Percentage	40%	40%	**20%**
Sum insured	5.000	5.000	5.000
Age	18-80	18-80	18-80
Duration	2-30	2-30	2-30



In this model specification the insurance company is exposed to lapse risk only on 20% of the portfolio. With respect to the 4th model specification, where with the same stressed scenario of 5% lapse probability the minimum value of the fund averaged a –412 thousand in 50 simulations, we record, on average, a minimum value of the fund of –133 thousand.

Conclusions

We started with the main theories behind lapse determinants: in paragraph 1.2 we introduced the interest rate hypothesis and the emergency fund hypothesis.

According to the first, the insured party surrenders the contract if he can find greater returns in the market. The emergency fund hypothesis links the lapse decision with a worsened economic condition for the policyholder.

Both theories have been widely investigated with respect to their relation with macro variables like interest and unemployment rate respectively, but little or no differentiation on micro variables has been considered by the industry and the academic community.

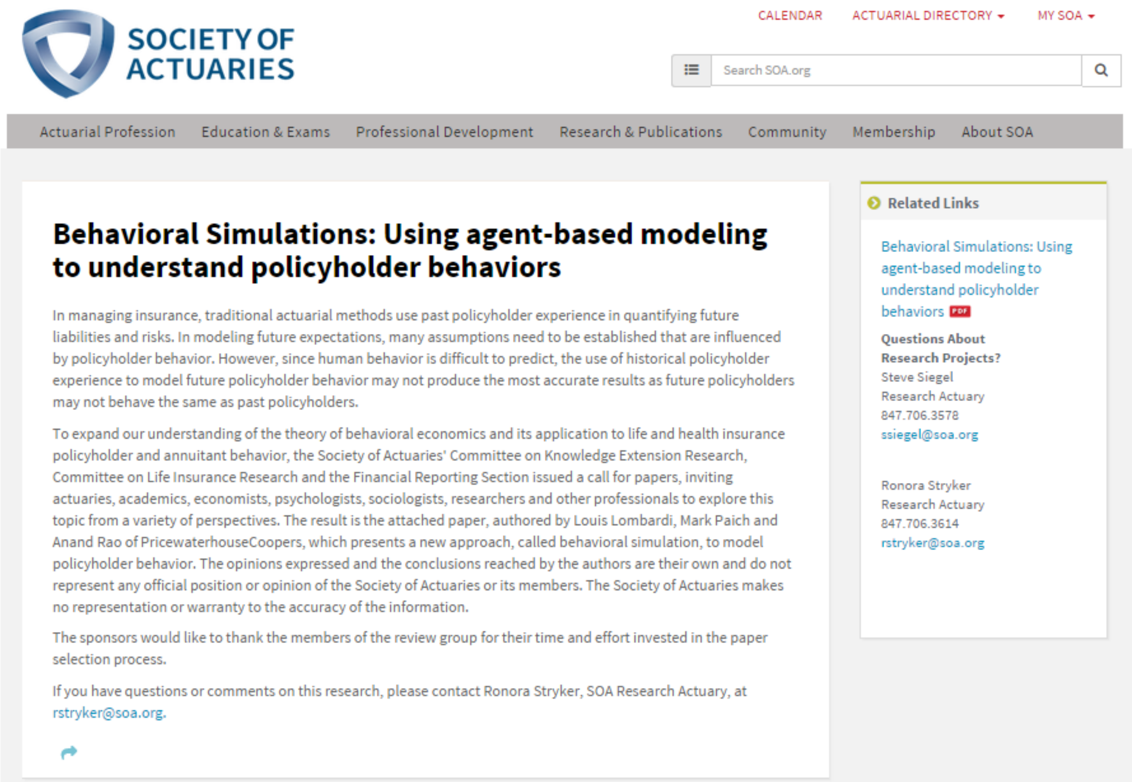
Only in the last 5 years research has been done on empirical data to model lapse risk at a micro level. All the papers on this topic used generalized linear models in their attempt. A review of these papers can be found in chapter 3. The literature is divided based on the methodology used. Particular attention has been posed on the type of variable modeled, whether the lapse probability, the number of lapses within a pool of contracts or the amount withdrawn.

In paragraph 3.7 we conceptually describe the most important explanatory variables used and the interpretations of the authors in how they affect the lapse decision.

In paragraph 2.1 we briefly explain the importance of micro level data to model a complex system. Other than the quantitative analyses reviewed in chapter 3, we comment in this paragraph the work by Campbell et al. (2014) where the concepts of behavioral economics are dealt with in an insurance perspective to explain policyholder behaviors.

Some applications of ABM in life and non-life insurance are listed in paragraph 2.2.

On this topic in 2012 the American Society of Actuaries (SoA) issued a call for research on policyholder's behavior. The SoA invited actuaries and academics to explore the field of behavioral economics applied to insurance with a post in their site accompanied by a paper called *Behavioral Simulations: Using agent-based modeling to understand policyholder behavior* (Lombardi et al., 2012)



SOCIETY OF ACTUARIES

CALENDAR ACTUARIAL DIRECTORY MY SOA

Search SOA.org

Actuarial Profession | Education & Exams | Professional Development | Research & Publications | Community | Membership | About SOA

Behavioral Simulations: Using agent-based modeling to understand policyholder behaviors

In managing insurance, traditional actuarial methods use past policyholder experience in quantifying future liabilities and risks. In modeling future expectations, many assumptions need to be established that are influenced by policyholder behavior. However, since human behavior is difficult to predict, the use of historical policyholder experience to model future policyholder behavior may not produce the most accurate results as future policyholders may not behave the same as past policyholders.

To expand our understanding of the theory of behavioral economics and its application to life and health insurance policyholder and annuitant behavior, the Society of Actuaries' Committee on Knowledge Extension Research, Committee on Life Insurance Research and the Financial Reporting Section issued a call for papers, inviting actuaries, academics, economists, psychologists, sociologists, researchers and other professionals to explore this topic from a variety of perspectives. The result is the attached paper, authored by Louis Lombardi, Mark Paich and Anand Rao of PricewaterhouseCoopers, which presents a new approach, called behavioral simulation, to model policyholder behavior. The opinions expressed and the conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information.

The sponsors would like to thank the members of the review group for their time and effort invested in the paper selection process.

If you have questions or comments on this research, please contact Ronora Stryker, SOA Research Actuary, at rstryker@soa.org.

Related Links

Behavioral Simulations: Using agent-based modeling to understand policyholder behaviors PDF

Questions About Research Projects?
Steve Siegel
Research Actuary
847.706.3578
ssiegel@soa.org

Ronora Stryker
Research Actuary
847.706.3614
rstryker@soa.org

Their paper describes in detail how the economy state and policyholder characteristics affect its decision through his life.

This paper inspired the first version of the model in paragraph 4.1 where we used their intuition to build a tool that would advise policyholders before deciding whether to lapse or not the policy given the characteristics of the insured party such as age, type of job, income and the state of the economy. With a model we can observe if a certain decision yields a positive or negative outcome and by advising the policyholders constructing a tool that would influence their choices and stabilize cash outflows.

This first model has not been completed; it is an input to inspire further works. Before using an ABM to influence policyholder decision it is necessary to develop a preliminary model to explain the effects that the surrender decision has on the asset and liability management of an insurance company. I will return on the first model in the last paragraph ‘further developments’.

From paragraph 4.2 we model lapse risk relying on the characteristic of the single policyholder, the type of contract and its terms. The ABM describes the interactions between insurer and insured parties.

We consider the effect of an early surrender of the policy on the liquidity of the insurance company.

In my model the single characteristics of the policyholder affect the firm in two ways:

- The maturity, age and duration of the contract affects the investments of the firm.
- The lapse probability depends on policyholder and contract characteristics.

After observing the outcome of the model, insurance companies, in order to deal with early surrenders, can either:

- Change its portfolio composition for new production.
- Modify contract features.
- Apply a different penalty.
- Change its investments.

My model is a prototype implemented in the agent-based model environment *NetLogo* to be further developed to meet the needs of a complex business as life insurance.
In its current version it can offer insights on what would be the appropriate management actions to deal with early surrenders, given a specific liability composition, in terms of asset allocation.

We assume the insurance company has a portfolio distributed between three types of contract: a pure endowment, a term life insurance and an endowment.

The pure endowment pays the insured party at maturity, if alive.

The term life insurance pays the insured party when he dies, if the death occurs before maturity.

The endowment is a combination of the two, with same maturity and insured capital: it either pays when the insured party dies or at maturity if he survives.

The insurer allows to surrender the contract only to policyholders of the endowment as the payment is certain and no adverse selection is possible.

The insurer invests the premium in order to have at each year the expected value of the benefit. To do so the insurance company invests the premiums at the corresponding maturity of the expected benefit with its particular interest rate given by the current term structure.

With an increasing term structure, in order to guarantee the highest possible return to its policyholder and itself, the insurance company invests the premium at the corresponding maturities of the expected benefits; therefore any early surrender of the contract jeopardizes the liquidity of the firm. To protect itself an insurance company can either apply a penalty for early surrenders or invest only a fraction of its premiums.

I build two main model specifications, one where the user can define an estimate of the probability of surrendering the contract which will remain constant throughout the simulation for all policyholders (paragraph 4.3) and another where each policyholder surrenders the contract with a probability based on the coefficients of the logit regression by Milhaud et al. (2010) (paragraph 4.4).

The version developed in paragraph 4.3 with respect to the one in paragraph 4.2 includes three types of contract instead of one and it allows the user to define the ranges of application of the relevant features for each contract: underwriting age, duration and sum insured.

The results of both models are summarized in paragraph 5.2.

In *model specification 2* (paragraph 4.3 – third version) we analyze the effect of the penalty alone in reducing liquidity risk, then in *model specification 3* and *4* (4.3) we introduced the possibility for the insurance company to save a fraction of the premiums to face early surrenders. The two model specification uses a different fraction of invested premiums (lower and higher respectively) and they are both studied with respect to the expected lapse rate (3%) and a stressed scenario (a scenario of low lapse rates and high lapse rate respectively). The model is able to give a quantitative result on how likely it is in the stressed scenario (lower/higher) for *model specification 3/4* to incur in liquidity risk due to a too low/high invested fraction of premiums.

In *model specification 5* (4.3) we estimated the effect of the duration of the contracts of the portfolio on actual lapse activity. Clearly to a longer duration of the contracts corresponds a higher number of lapses for the same annual lapse rate. This result is to be seen in comparison with *model specification 1* (4.4 – fourth version). The higher lapse activity is mitigated by the fact that lapse probability lowers as the policyholder nears longer maturities according to Milhaud at al. (2010) (see table 12). While for the third version of the model we had an overinvestment of the premiums, for the fourth version we have an under investment as the actual lapse activity is lower.

The situation is inverted for the comparison made in model specification 2 (4.4). Due to the lower maturity and the young age of the policyholders the total lapse activity is higher for the fourth model specification for which we have an overinvestment.

By lowering the number of insured parties to 1.000 and therefore increasing variability we can see a realization of liquidity risk. The model suggests a slightly lower fraction of invest premiums.

In *model specification 6* (4.3) we can see a more practical application by allowing the insurance company to diversify lapse risk with a portfolio less exposed to policyholder behavior.

The model can also be used to make an analysis on the effect that various characteristics of the liabilities, such as the age of the insured party and the size of the portfolio, have on the liquidity of the firm, without considering the surrender option.

The results of this type of analysis are summarized in paragraph 5.1.

On the mutuality principle of life insurance companies we highlight how for the portfolio specified in *model specification 1* (4.2) there is not substantial benefit in terms of reduced variability in the final value of the fund by increasing the number of policyholders above 3.000.

In the same paragraph we collected statistical evidence that variability of the final value of the fund is increasing with respect to variability in the sum insured.

In *model specification 1* (4.3) we noticed that variability in the liquidity of the firm through the simulation is increasing with respect to underwriting age.

The model is capable of giving a quantitative answer to all of these problems. Using Monte Carlo simulation we can calculate a probability for the occurrence of an event given certain assumptions on portfolio composition and the other user defined features.

However the model has been used in this thesis to give a conceptual explanation of how the surrender decision affects the liquidity of the firm. A quantitative use of the model as an operational tool for life insurance companies would be possible by extending the model to include a variety of contracts and other characteristics which will be discussed in the following paragraph.

Further developments

An operational tool for life insurance companies

A first order of improvement necessary to model lapse risk regards the composition of the liabilities of the firm.

The main improvement that the model needs is the inclusion of with profit participation contracts in the simulation.

In my model we assume that financial interests equal the technical rate returned to the policyholder; this assumption was needed in a first implementation of the model to check for the actuarial equivalence of premiums and benefits.

A deviation of financial interests from technical interests and subsequent inclusion of a participation rate and a minimum guaranteed rate are a necessary step to refine the model.

With profit contracts are probably the most difficult contracts to model due to the interaction between policyholder behaviors and management actions.

A high lapse rate can force the management to sell assets to meet policyholder's demands which can trigger the realization of gains or losses and impact the benefit returned to the insured parties with subsequent reaction on their part.

The complexity in modeling this type of contract is well known to the regulator (EIOPA, 2013):

> In life insurance, the nature and complexity of the risks would for example be impacted by the financial options and guarantees embedded into the contracts (such as surrender or other take-up options), particularly those with profit participation features.

As for unit linked contracts, which have also been excluded from my simulation, I do not believe it would be difficult to include them since the early surrendered policy would cause a symmetrical cash outflow in both assets and liabilities. As a matter of fact insurance companies do not charge a penalty for surrendering this type of contract.

A second order of refinement of the model would be to include in the simulation the possibility for the firm to influence policyholder choices whenever they are forced to lapse the contract due to a drop in their income or face any expense which may occur in their lives.

An agent-based model is the most suitable instrument to explain the interactions among the following three factors:

- Policyholder choices.

- Insurer attempt to retain the customer and stabilize cash outflows.
- The state of the economy or the environment in general and policyholder characteristics considered as a whole.

How each of these factors interacts with the others is described in paragraph 4.1 where I laid the foundations for this model.

With regard to the second element, management actions influence in a variety of ways the interest returned to policyholders. We should include in the model the possibility for the insurance company to influence policyholders' decision depending on the minimum return guaranteed and the return obtained by the company in the market. If the policyholder becomes a cost due to the persistent state of low interest rate, we can include in the ABM the push of the management towards the agency network to influence policyholders in surrendering the contract.

Another common practice in life insurance is to realize losses to lower the interest rate returned to policyholders without having their contract going "in the money", below the minimum guaranteed. We should include in the model a non-discretionary practice, like a target return, for these types of actions.

References

Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). Classification and Regression Trees. Taylor & Francis.

Cameron, A., Trivedi, P. (2013). Regression Analysis of Count Data, 2nd edition, Econometric Society Monograph No.53, Cambridge University Press, 1998.

Campbell, J., Chan, M., Li, K., Lombardi, L., Lombardi, L., Purushotham and M., Rao, A. (2014). Modeling of Policyholder Behavior for Life Insurance and Annuity Products. pp. 6-48. https://www.soa.org/Research/Research-Projects/Life-Insurance/research-2014-modeling-policy.aspx

Casti, J., Jones, R. (2000). Insurance World: A process for understanding risk flows due to catastrophes in the Insurance/Reinsurance Industry. http://s3.amazonaws.com/zanran_storage/www.iiasa.ac.at/ContentPages/8489563.pdf

CEIOPS (2009). CEIOPS' Advice for Level 2 Implementing Measures on Solvency II. https://eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Final-Advice-on-Standard-Formula-Life-underwriting-risk.pdf

Cerchiara, R., Edwards, M. and Gambini, A. (2009). Generalized Linear Models in Life Insurance: Decrements and Risk factor analysis under Solvency II. *Working Paper. AFIR Colloquium Rome. http://www.actuaries.org/AFIR/Colloquia/Rome2/Cerchiara_Edwards_Gambini.pdf*

Clark, D., Morgan, E., Kent, J.(2013). Dynamic Policyholder Behavior and Management Actions Survey Report. http://it.milliman.com/insight/2013/Dynamic-Policyholder-Behaviour-and-Management-Actions-Survey-Report/

Cox, S., Lin, Y. (2006). Annuity Lapse Rate Modeling: Tobit or not Tobit. *Society of Actuaries.* https://cba.unl.edu/research/articles/1287/download.pdf

Cragg, J. (1971). Some Statistical Models for Limited Dependent Variables With Application to the Demand for Durable Goods. *Econometrica*, 39(5). pp. 829-844.

EIOPA (2013). Technical Specification on the Long Term Guarantee Assessment. https://eiopa.europa.eu/Publications/QIS/A_-_Technical_Specification_on_the_Long_Term_Guarantee_Assessment__Part_I_.pdf

Eling, M. and Kochanski, M. (2012). Research on Lapse in Life Insurance – What has been done and what needs to be done?. *Working Papers on Risk Management and Insurance*, No. 126. http://www.ivw.unisg.ch/~/media/internet/content/dateien/instituteundcenters/ivw/wps/wp126.pdf

Eling, M. and Kiesenbauer, D. (2011). What Policy Features Determine Life Insurance

Lapse? An Analysis of the German Market. *Working Papers on Risk Management and Insurance*, No. 95. http://www.ivw.unisg.ch/~/media/internet/content/dateien/instituteundcenters/ivw/wps/wp95.pdf

Epstein, J. (2008). Why Model?. *Journal of Artificial Societies and Social Simulation*, 11(412). http://jasss.soc.surrey.ac.uk/11/4/12.html

Epstein, J. and Axtell, R. (1996). Growing artificial societies: social science from the bottom up, Brookings Institution Press.

Fier, G., Liebenberg, A. (2013). Life Insurance Lapse Behavior. *North American Actuarial Journal*, 17(2). http://www.aria.org/meetings/2012%20Meetings/2E-Life%20Insurance%20Lapse%20Behavior.pdf

Friedman, M. (1957). The Permanent Income Hypothesis. In M. Friedman, ed., *A Theory of the Consumption Function*. Princeton University Press, pp. 20-37.

Geneva Association (2012). Surrenders in the Life Insurance Industry and their Impact on Liquidity. https://www.genevaassociation.org/media/99452/ga2012-surrenders_in_the_life_insurance_industry.pdf

Haefeli, D. and Ruprecht, W. (2012). Surrenders in the Life Insurance Industry and their Impact on Liquidity. *The Geneva Association.* https://www.genevaassociation.org/media/99452/ga2012-surrenders_in_the_life_insurance_industry.pdf

Humphrey, B. (2013). Dealing With Zeros in Economic Data. *University of Alberta, Department of Economics. https://www.ualberta.ca/~bhumphre/class/zeros_v1.pdf*

Kagrakoa, Y., (2005). Modeling insurance surrenders by the negative binomial model. *Working paper*. http://www.musashi.jp/~kagraoka/research/NBM_017.pdf

Kiesenbauer, D. (2012). Main determinants of lapse in the German life insurance industry. *North American Actuarial Journal*, 16(1).

Kuo, W., Chenghsien, T. and Wei-Kuang, C., (2003). An Empirical Study on the Lapse Rate: The Cointegration Approach. *Journal of Risk and Insurance*, 70(3). pp. 489-508.

Lombardi, L., Paich, M., Rao, A. (2012). Behavioral Simulations: Using agent-based modeling to understand policyholder behaviors. https://www.soa.org/Research/Research-Projects/Risk-Management/Behavioral-Simulations/

Liebenberg, A., Carson, J., Dumm, E. (2010). A Dynamic Analysis of the Demand for Life Insurance. *Risk Theory Society Annual Seminar. http://www.aria.org/rts/proceedings/2010/default.htm*

LIMRA (2011). Buyer/Nonbuyer Study Middle Market sample. Accessed from https://www.soa.org/files/pd/annual-mtg/2011-chicago-annual-mtg-145.pdf

Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2010). Surrender triggers in life insurance: classification and risk predictions. *Working paper. http://isfa.univ-lyon1.fr/vmaume/sites/default/files/documents/lapse.pdf*

Mills, A. (2010). Complexity Science: An introduction (and invitation) for actuaries. https://www.soa.org/Research/Research-Projects/Risk-Management/Behavioral-Simulations/

Moudiki, T., Planchet, F. (2014). Economics Scenarios Generation for Insurance. *8th Rmetrics Workshop on R in Finance and Insurance*, Paris, June 27, 2014. http://www.ressources-actuarielles.net/EXT/ISFA/fp-isfa.nsf/a5ee867f10dc2d79c1256f130065f05f/39b54166464089afc12572b0003d88c2/$FILE/TM_ESG.pdf

Perna, C., Sibillo, M. (2007). Mathematical and Statistical Methods for Insurance and Finance. Springer

Pinquet, J., Guillen, M. and Ayuso, M. (2011). Commitment and Lapse Behavior in Long-Term Insurance: A Case Study. *Journal of Risk and Insurance*, 78 (4). pp. 983-1002.

Outreville, J. (1990). Whole-life insurance lapse rates and the emergency fund hypothesis. *Insurance: Mathematics and Economics*, 9 (4). pp. 249-255.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Renshaw, A. and Haberman, S. (1986). Statistical Analysis of Life Assurance Lapses . *Journal of the Institute of Actuaries*, 113 (3). pp. 459-497.

Urbanek, S. (2013). Rserve: Binary R server. R package version 1.7-3. http://CRAN.R-project.org/package=Rserve

Wang, S. and Mango, D. (2003). Research outside the actuarial comfort zone. Actuarial Review, 30 (1).

Weindorfer, B. (2012). QIS5: A review of the results for EEA Member States, Austria and Germany. http://www.fh-vie.ac.at/var/em_plain_site/storage/original/application/9ddc7c5ebb9c47c024125a04c6632db3.pdf

Wilensky, U. (1998). NetLogo Wealth Distribution model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

http://ccl.northwestern.edu/NetLogo/models/WealthDistribution.

Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/NetLogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

Yaari, M. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies*, 32 (2), pp. 137-50.