PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

César da Rocha Neves

Essay on Dynamic Modeling in Life Insurance and Private Pension: Longevity, Surrender and Embedded Options

TESE DE DOUTORADO

Thesis presented to the Programa de Pós-graduação em Engenharia Elétrica of the Departamento de Engenharia Elétrica, PUC-Rio, as partial fulfillment of the requirements for de degree of Doctor em Engenharia Elétrica.

Advisor: Prof. Cristiano Augusto Coelho Fernades
Co-Advisor: Prof. Álvaro de Lima Veiga Filho

Rio de Janeiro
November 2015

César da Rocha Neves

Essay on Dynamic Modeling in Life Insurance and Private Pension: Longevity, Surrender and Embedded Options

TESE DE DOUTORADO

Thesis presented to the Programa de Pós-graduação em Engenharia Elétrica of the Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio, as partial fulfillment of the requeriments for de degree of Doctor.

Prof. Cristiano Augusto Coelho Fernandes
Advisor
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Álvaro de Lima Veiga Filho
Co-Advisor
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Antonio Carlos Figueiredo Pinto
Departamento de Administração – PUC-Rio

Prof. Marcelo Cunha Medeiros
Departamento de Economia – PUC-Rio

Prof. Nikolai Valtchev Kolev
USP

Prof. Adrian Heringer Pizzinga
UFF

Dr. Rodrigo Simões Atherino
JGP Global de Recursos Ltda

Prof. José Eugenio Leal
Coordinator of the Centro Técnico
Científico da PUC-Rio

Rio de Janeiro, November 27th, 2015.

César da Rocha Neves

César da Rocha Neves graduated in 1999 in Actuarial Science at Rio de Janeiro Federal University (UFRJ). In 2004, he got his master's degree in Production Engineering - Operation Research at COPPE/UFRJ. César is analyst at the Brazilian Insurance Supervisor (SUSEP) and Professor in Actuarial Science at the Department of Statistics and Actuarial Sciences of the Rio de Janeiro State University (UERJ) and the Brazilian College of Insurance (ESNS). His research focus on actuarial sciences, modeling and risk management, applications of time series to insurance and pension plan.

Ficha Catalográfica

To my daughters Laura and Alice.

Acknowledgments

To my wife Viviane for her love and understanding.

To my beautiful and beloved little girls Laura and Alice. I hope they are proud of this work. Laura participated actively throughout the doctorate and Alice was born during the course, bringing me more joy.

To my parents, who always encouraged me to study. To all my family and friends for their help and motivation.

To Professor Álvaro Veiga, who believed in my project and accepted me to the doctorate.

To my advisor Professor Cristiano Fernandes, for his partnership and help with the articles and the lectures presented in conferences.

To SUSEP and UERJ, for granting me the opportunity to attend the doctorate.

To FUNENSEG and my friend Claudio Contador, for the support and motivation.

To the members of the doctoral examination board, whose suggestions were very important for the improvement of the work.

Abstract

Neves, César da Rocha Neves; Fernandes, Cristiano Augusto Coelho (Advisor); Veiga Filho, Álvaro de Lima (Co-advisor). **Essay on Dynamic Modeling in Life Insurance and Private Pension: Longevity, Surrender and Embedded Options**. Rio de Janeiro, 2015. 163p. PhD Thesis - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

In this thesis we propose four dynamic models to help life insurers and pension plans to measure and manage their risk factors and annuity plans. In the first two essays, we propose models to forecast longevity gains of a population, which is an important risk factor. In the first paper, a multivariate time series model using the seemingly unrelated time series equation (SUTSE) framework is proposed to forecast longevity gains and mortality rates. In the second paper, a multivariate structural time series model with common stochastic trends is proposed to forecast longevity gains of a population with a short time series of observed mortality rates, using the information of a related population for which longer mortality time series exist. In the third paper, another important risk factor is modeled – surrender rates. We propose a multi-stage stochastic model to forecast them using Monte Carlo simulation after a sequence of GLM, ARMA-GARCH and multivariate copula fitting is executed. Assuming the importance of the embedded options valuation to maintain the solvency of annuity plans, in the fourth paper we propose a model for evaluating the value of embedded options in the Brazilian unit-linked plans.

Keywords

Dynamic modeling; longevity; mortality rates; surrender rates; embedded options; SUTSE model; life insurance; private pension.

Resumo

Neves, César da Rocha Neves; Fernandes, Cristiano Augusto Coelho (Orientador); Veiga Filho, Álvaro de Lima (Co-orientador). **Ensaios de Modelagem Dinâmica Aplicada a Seguro de Vida e Previdência: Longevidade, Resgate e Opções Embutidas**. Rio de Janeiro, 2015. 163p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Nesta tese, propomos quatro modelos dinâmicos para ajudar as seguradoras e fundos de pensão a medir e gerencias seus fatores de risco e seus planos de anuidade. Nos primeiros dois ensaios, propomos modelos de previsão de ganhos de longevidade de uma população, que é um importante fator de risco. No primeiro artigo, um modelo de séries temporais multivariado usando a abordagem SUTSE (seemingly unrelated time series equation) é proposto para prever ganhos de longevidade e taxas de mortalidade. No segundo artigo, um modelo estrutural multivariado com tendências estocásticas comuns é proposto para prever os ganhos de longevidade de uma população com uma curta série temporal de taxas de mortalidade, usando as informações de uma população relacionada, para qual uma longa série temporal de taxas de mortalidade é disponível. No terceiro artigo, outro importante fator de risco é modelado – taxas de cancelamento. Apresentamos um modelo estocástico multiestágio para previsão das taxas de cancelamento usando simulação de Monte Carlo depois de uma sequência de ajustes GLM, ARMA-GARCH e cópula multivariada ser executada. No quarto artigo, assumindo a necessidade de se avaliar as opções embutidas para manter a solvência dos planos de anuidade, propomos um modelo para mensuração das opções embutidas nos planos *unit-linkeds* brasileiros.

Palavras-chave

Modelagem dinâmica; longevidade; taxas de mortalidade; taxas de cancelamento; opções embutidas; seguro de vida, previdência complementar.

Contents

List of Figures

List of Tables

1
Introduction

The analysis of the solvency in the insurance and private pension markets is a concern to insurers, supervisors, policyholders and stakeholders. In recent years an increasing number of international regulations and guidelines on this matter have been issued. Among these are Solvency II[1] and the core principles and guidelines published by the International Association of Insurance Supervisors (IAIS)[2]. Mention should also be made of IFRS 4, which presents accounting standards for insurance operations published by the International Accounting Standards Board (IASB).

Thus, the actuarial, statistical and financial literature contains many papers and theses about risk modeling and management, capital and liabilities measurement and embedded options pricing. Considering this international scenario, in this thesis we present four papers applying dynamic modeling to estimate two important risk factors in life insurance and pension plans – mortality risk and surrender risks – and to evaluate embedded options in annuity contracts.

Nowadays, the actuarial risk associated with the longevity gain of a population is of great concern to insurers, and as such attracts great interest from actuaries and other academics. Longevity risk, unlike most other actuarial risks, is not totally diversifiable and is independent of the size of the covered population. The observed trend of declining mortality rates requires an increase on provisions and capital to assure meeting the future commitments of policyholders and beneficiaries. Thus, adequate modeling of the future longevity gains is crucial for insurers and pension funds.

In consequence, in the first two papers we propose dynamic models to forecast the mortality improvements of a population and show how they can be used to manage the longevity risk. In the first paper, we assume that the time series of mortality rates of different age groups of a population are not directly related to each other, but are subject to similar influences. These similar

---

[1] Solvency II is a European Union (EU) Directive that codifies and harmonizes the EU insurance regulation.
[2] IAIS is an organization of international insurance supervisors and regulators.

influences are captured by assuming that both observable mortality rates and unobservable components of different age groups are contemporaneously correlated.

To implement this model we adopt the seemingly unrelated time series equation (SUTSE) framework. This class of models takes the form of linear Gaussian state space models, allowing the use of the Kalman filter to estimate the unobservable components and the parameters of interest. Additionally, we discuss how the proposed model can be used by insurers and pension funds to manage the risk of declining mortality rates. The results indicate that the proposed SUTSE model outperforms the benchmark models.

In the second paper, we propose a multivariate model with stochastic trends to forecast the longevity gains of a population with short time series of observed mortality rates. This state space model also makes use of the SUTSE model to jointly model time series of mortality rates associated with particular age groups. But, in this approach a more parsimonious dependence is derived from the SUTSE model, by testing for restrictions on the covariance matrix associated with the trend noise. Under such restrictions, a SUTSE model collapses to a common trends model, one in which a reduced number of trends is able to satisfactorily explain the multivariate time series. Due to ease of treating missing data in state space models, this framework can be used to estimate longevity rates for populations with short time series. The strategy rests on the joint modeling of such a set of series with the corresponding mortality rate series obtained from a population with a long time series of data (the related population) that has similar mortality characteristics to the population with the shorter time series. Thus, we assume that the time series of both populations are cointegrated.

So, we apply the proposed model to Brazilian male and female populations using different related populations to find the best goodness of fit. Additionally, we estimate the distribution of Brazilian policyholders' future mortality rates through the model. Furthermore, the complete life expectancies of current Brazilian policyholders are estimated as well as their temporal evolutions. Then, we analyze the forecasting of the life expectancies over time, considering the longevity gain forecasted by the model. To exemplify how the model can be used to measure underwriting risk, using Monte Carlo simulation, we obtain distributions of present values of cash flows generated by company expenses from

hypothetical beneficiary groups receiving life annuities. Through this distribution, we calculate the best estimate of the liabilities and the capital based on underwriting risk. By means of Monte Carlo simulation, the idiosyncratic risk effect on the process of calculating an amount of underwriting capital is also shown, since we simulate using beneficiary groups with different sizes.

In the third paper, we propose a multi-stage stochastic model to forecast surrender rates from life insurance using multivariate elliptical copulas and financial variables by means of Monte Carlo simulation after a sequence of GLM, ARMA-GARCH and multivariate copula fitting is executed. This approach is quite relevant nowadays, since adequate surrender rates are an essential factor that must be considered in the realistic valuation of life insurance liabilities and actuarial risks, which are now required by solvency and accounting standards. For instance, the forecast of surrender rates can be used for valuation of embedded options to estimate the policyholder behavior with respect to exercising contractual options. Moreover, companies must forecast surrender rates in order to manage risks that arise due to mismatches between assets and liabilities (ALM).

In our proposed dynamic model, we study the dependence among the surrender rate time series through a multivariate elliptical copula framework. During recent economic crises, highly uncommon surrender rates were observed in the insurance industry. So, to really explain the dependence structure, we incorporate in our copula model a proxy for the stock market return index as one of the marginal distributions. Our model can also be used to simulate future surrender rates given a specific financial scenario, which can be chosen in a stress test context to analyze policyholder behavior when faced with a financial crisis. To simulate this scenario, we propose a specific algorithm for simulation of multivariate elliptical copulas conditioned on a marginal distribution, which is the stock market return residual distribution in our application.

In the fourth paper, we present a model to obtain the best estimate of the Brazilian unit-linked embedded options under real-world measure. Particularly in Brazil, these plans have several embedded options, such as guaranteed annuity option, option to defer annuitization, surrender option, option to switch the type of annuity at the moment of exercising the guaranteed annuity option, option to shut down, option to transfer the fund to other insurer and option to increase the coverage by paying of premiums. Therefore, insurers have to value these options

to calculate liabilities and the need for capital and hence to maintain their solvency. The objective of the fourth paper is to discuss the embedded options in Brazilian unit-linked plans, which are the most important Brazilian annuity products, and propose a model to obtain their best estimate. As in Solvency II, we assume that the best estimate corresponds to the probability weighted average of future cash flows, taking account of the time value of money by using the relevant risk-free interest rate term structure. Like other insurance markets, the Brazilian annuity market is incomplete, but this market is also not arbitrage-free, as we demonstrate in this thesis by showing the arbitrage opportunities. So, in our approach to evaluate the best estimate of Brazilian embedded options we have to assume real-world probabilities, unlike the majority of other studies.

Our approach does not assume an optimal policyholder behavior. Instead of this assumption, considering the Brazilian annuity market's characteristics, we model the surrender rates before the retirement date through a model proposed in the third paper that allows rational and irrational surrender. This model uses the mortality rates forecasted by the SUTSE model presented in the second paper in the simulation process. In addition, we model the annuity decision considering that policyholders have the right to change the kind of annuity at the moment of retirement, the option to defer the retirement date, and the possibility of choosing a self-annuitization strategy, which is modeled by a jump process. On the other hand, we also consider that policyholders can increase the value of the embedded options if they continue to pay regular premiums, pay additional premiums or transfer their funds from other plans or other insurers to their unit-linked plans. These movements are also modeled by means of jump processes. We then will apply the model using Monte Carlo simulation, and will report some sensibility analyses of the parameters used in the model.

In the fourth paper, presented in chapter 5, we use the model presented in chapter 3 to forecast mortality rates and the model presented in chapter 4 to forecast surrender rates.

The first paper presented in chapter 2 was published in the Journal of Forecasting 34(8), 661-674, 2015. The second paper was accepted in the North American Actuarial Journal and it will be published in 2016 (in process). The third paper was published in the North American Actuarial Journal, 18(2), 343-362, 2014. Finally, the fourth paper presented in chapter 5 is the result of the

research project supported by FundaciÃ³n Mapfre, and it was published in the "Notebooks of the Foundation" collection – Cuardenos de la FundaciÃ³n C/211 (ISBN: 978-84-9844-573-2).

2

Paper 1: Forecasting longevity gains using a seemingly unrelated time series model

In this paper a multivariate time series model using the seemingly unrelated time series equation (SUTSE) framework is proposed to forecast longevity gains. The proposed model is represented in state space form and uses Kalman filtering to estimate the unobservable components and fixed parameters. We apply the model both to male mortality rates of Portugal and the US. Our results compare favorably, in terms of MAPE, in sample and out-of-sample, to those obtained by the Lee-Carter method and some of its extensions.

2.1
Introduction

The actuarial risk associated with the longevity gain of a population is of great concern to managers of insurance companies and pension funds (both public and private), and as such attracts great interest from actuaries and other academics. The longevity risk, unlike most other actuarial risks, is not totally diversifiable and is independent of the size of the covered population.

The observed trend of declining mortality rates requires an increase on provisions and capital by insurance companies, in order to assure meeting the future commitments of policyholders and beneficiaries. Thus, adequate modeling of the future mortality improvement is crucial for insurers and pension funds. Because of this, the International Association of Insurance Supervisors (IAIS), through its Insurance Core Principles [3], recommends that the obligations of insurance companies should be evaluated on consistent bases, by an economic valuation that reflects the prospective future cash flows. In Europe, Solvency II [4] determines that the best estimate of the provision for future commitments must be

measured based on current information and realistic predictions, using adequate, applicable and relevant actuarial and statistical methods.

Numerous models have been proposed in the literature to forecast mortality rates, and consequently to manage longevity risk. Among these, the most widely used is that of Lee and Carter (1992). In this method, the unobservable component that measures the evolution of the trend over time is unique for all age groups. Factors are estimated corresponding to the relative weights of the unobservable component for each age group. Estimation of the method involves performing two steps: singular value decomposition (SVD) and ordinary least squares estimation. To forecast the future rates, ARIMA models for time series are used. In particular, for the population studied in Lee and Carter (1992), the authors used a random walk with drift.

De Jong and Tickle (2006) suggested other methods to estimate the parameters of the Lee-Carter model. Instead of using SVD, model estimation is accomplished through maximum likelihood by applying the Kalman filter, considering more than one unobservable component to model the evolution of the trend in time. At the end of their paper they suggested the use of a stochastic drift, which corresponds to the multivariate local linear trend model (Harvey, 1989). Hári et al. (2008) used a state space approach with Kalman filtering to estimate and predict future mortality rates. They proposed an extension of the Lee-Carter method as reformulated by Girosi and King (2005). In their approach a multivariate state space model with stochastic drift with few latent factors is proposed to capture the common movements among all age groups.

Gao and Hu (2009) also proposed models based on Kalman filter to forecast mortality rates using a dynamic mortality factor model considering the conditional heteroscedasticity of mortality. Some other recent works have been proposed to predict future mortality structures and manage the longevity risk, such as Plat (2009), Haberman and Renshaw (2009, 2011, 2012), Cairns (2011), D'Amato et al. (2012) and French and O'Hare (2013).

In this paper, unlike the models that use Kalman filter mentioned previously, we consider that time series of mortality rates of different age groups of a population are not directly related to each other, but are subject to similar influences in the sense of Fernández and Harvey (1990). These similar influences on the populations are exemplified by better eating habits, improvements in life

quality and medical care, all of which can significantly reduce future mortality rates. Formally they are captured in the proposed model by assuming that both observable mortality rates and unobservable components of different age groups are contemporaneously correlated.

Lee and Carter (1992) concluded that the correlations between actual and fitted rates often are substantial and persist over very wide age gaps, but they avoid incorporating this into their stochastic model. On the other hand, Barrieu et al. (2012) concluded that dependence between ages is an important component in modeling mortality. In D'Amato et al. (2012) the dependence structure of neighboring observations in the population is captured to improve forecasting mortality through the Lee-Carter sieve bootstrap method. In our time series model, by construction, the dependence structure of all mortality rates time series and of all stochastic components that compose the trends of the time series is duly accounted for. This strategy captures the similar influences in the process of mortality improvement of the all age groups. Thus, the proposed procedure estimates the stochastic trend for each age group and the correlations among them to obtain a more realistic mortality model.

To implement the model we adopt a framework analogous to the seemingly unrelated regression equation (SURE) called the seemingly unrelated time series equation (SUTSE), as described by Harvey (1989) and Harvey and Koopman (1997). This class of models, which Harvey (1989) called multivariate structural SUTSE models, takes the form of linear Gaussian state space models, allowing the use of the Kalman filter to estimate the unobservable components and the fixed parameters of the model.

We apply our model to the mortality rates of male populations of Portugal and the US and to test model adequacy and accuracy, both in sample and out-of-sample, we compare our results to some benchmarks in the literature, namely: the Lee and Carter (1992) method and its state space form adaptations, as well to its variants with a cohort effect, the model by D'Amato et al. (2012) and the model by French and O'Hare (2013). Additionally, we discuss how the proposed model can be used by insurers and pension funds to manage the risk of declining

mortality rates. We implemented the proposed SUTSE model in the STAMP 8.3 program[5].

The remainder of the paper is organized as follows. In the section 2.2 we present the model proposed to estimate the longevity gains. In the section 2.3 we apply the model to the Portuguese and American populations. The section 2.4 concludes.

2.2
SUTSE model to predict the longevity gains

Before presenting the proposed model for forecasting longevity gain of a population, we first introduce the state space model and the Kalman filter in general form.

2.2.1
State space approach

In this paper, we write the model in linear Gaussian state space following the general form, as defined in Durbin and Koopman (2012). Let y_t be a time series vector.

The state space form is defined as follows:

- observation equation: $y_t = Z_t \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, H_t)$

- state equation: $\qquad \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim N(0, Q_t)$ $\qquad\qquad$ (2-1)

where $t = 1, \ldots, T$, y_t is a $N \times 1$ vector of observations, α_t is an unobserved $m \times 1$ vector, the system matrices have theses dimensions: $Z_t \in \Re^{N \times m}$, $T_t \in \Re^{m \times m}$, $R_t \in \Re^{m \times r}$, $H_t \in \Re^{N \times N}$, $Q_t \in \Re^{r \times r}$, N is the number of time series and m is the number of components of the state vector α_t. The initial state vector $\alpha_1 \sim N(a_1, P_1)$ is independent of $\varepsilon_1, \cdots, \varepsilon_T$ and of $\eta_1, \cdots, \eta_T$. ε_t and η_t are serially and mutually independent for all t.

---

[5]STAMP is a statistical/econometric program to model time series with unobservable and irregular components such as trend, seasonality and cycle. http://stamp-software.com/.

2.2.2
Kalman filter

The Kalman filter is a recursive procedure to compute the estimator for the state space vector at time *t*, based on the information available up to that time (Harvey, 1989). When the shocks have Gaussian distribution, the estimator found will be optimal in terms of mean squared error. The recursive equations of the Kalman filter are:

$$E(\alpha_{t+1} \mid Y_t) = a_{t+1} = T_t a_t + K_t \nu_t;$$
$$V(\alpha_{t+1} \mid Y_t) = P_{t+1} = T_t P_t L_t' + R_t Q_t R_t';$$
$$\nu_t = y_t - E(y_t \mid Y_{t-1}) = y_t - Z_t a_t \text{ is the innovations vector;} \qquad (2\text{-}2)$$
$$K_t = T_t P_t Z_t' F_t^{-1} \text{ is the Kalman gain;}$$
$$F_t = Z_t P_t Z_t' + H_t \text{ is the covariance matrix of the innovations; and}$$
$$L_t = T_t - K_t Z_t.$$

where $Y_t = \{y_1, \cdots, y_t\}$ and $\nu_t \sim N(0, F_t)$.

2.2.3
Structural SUTSE model

In the structural models of Harvey (1989), a time series is decomposed into components of interest, such as trend, seasonality and cycle. Due to the characteristics of the data, in the case of mortality rates we only specify the trend component. Since these series evolve probabilistically over time, we work with stochastic trends. In this case, the trends are called local components.

In Lee and Carter (1992) mortality rates age-dependent and time-dependent terms are predicted ignoring the existing correlation structure between the different age groups. Renshaw and Haberman (2006) incorporated a cohort effect in such model. Currie (2006) simplified that model through the Age-Period-Cohort (APC) model. Cairns et al. (2009) stated that there is a significant cohort effect in mortality improvements. Plat (2009) also used a cohort effect and described some problems with modeling cohort effect. On the other hand, D'Amato et al. (2012) by use of the Lee-Carter sieve bootstrap method considered

the dependence in the error terms and captured the dependence structure of neighboring observations in the studied population.

In our proposed SUTSE model we estimate a trend for each age group of a population and also fully specify the dependence structure among these age groups. In section 2.3.2, we will show that by incorporating such realistic features we can improve forecasting when compared to other competing models. In our model the dependence structure amongst the mortality rates time series are captured by a full covariance matrix of the shocks on the observation equation. This covariance is the only connection between the mortality rate series for different age groups. Additionally, each age group has its own unobservable components to explain the evolution of the mortality rate time series, and these components are also interconnected through full disturbance covariance matrices. These covariance matrices can capture existing similar influences in mortality improvements at different age groups, associated with better eating habits, improvement in life quality and progress in medical care.

In the next section, we present a SUTSE model considering that each trend is composed of a level and a slope, both of which are stochastic. Such flexibility is valuable to pick up possible existing changes in mortality trends which can contribute to improve forecasting mortality rates. This model is known as the local linear trend model. In section 2.2.3.2, we propose a particular case of such model, that which has a deterministic slope, as in Lee and Carter (1992).

In our proposed model, the observable series y_t are given by the logarithm of the central mortality rates for each age x. As is known, log transformation can improve residuals diagnostics in many situations and guarantees positive mortality rates. Bowers et al. (1997) defined the central mortality rate as follows:

$$m_{x,t} = \frac{\int_0^1 l_{x+s,t}\mu_{x+s,t}ds}{\int_0^1 l_{x+s,t}ds} = \frac{d_{x,t}}{L_{x,t}} \approx \frac{q_{x,t}}{1-\frac{q_{x,t}}{2}} \qquad (2\text{-}3)$$

where $m_{x,t}$ is the central mortality rate for age x at time t, $\mu_{x,t}$ is the force of mortality (hazard rate) for age x at time t, $l_{x,t}$ is the number of survivors with age x at the start of the year at time t, $d_{x,t}$ is the number of deaths between ages x and $x+1$ at time t, $L_{x,t}$ is the number of years lived by the population between

ages x and $x+1$ at time t, representing the number of people exposed to risk, and $q_{x,t}$ is the probability of dying between the ages of x and $x+1$ at time t. To simplify the notation, in this paper we will drop the index x of the mortality rate $m_{x,t}$.

2.2.3.1
Local linear trend model (LLT)

We first introduce a univariate structural model composed of only the stochastic trend component μ_t. This model is represented in the following form:

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N\!\left(0,\sigma_\varepsilon^2\right) \\
\mu_{t+1} &= \mu_t + \beta_t + \eta_t & \eta_t &\sim N\!\left(0,\sigma_\eta^2\right) \\
\beta_{t+1} &= \beta_t + \varsigma_t & \varsigma_t &\sim N\!\left(0,\sigma_\varsigma^2\right) \quad \text{for} \quad t=1,\dots,T
\end{aligned}
\tag{2-4}
$$

where y_t is the datum observed at time t, μ_t is the level at time t, β_t is the slope at time t, and the observation, level and slope disturbances at time $t\left(\varepsilon_t, \eta_t \text{ and } \varsigma_t\right)$ are serially and mutually independent. When σ_η^2 and σ_ς^2 are both different than zero, then we have a local linear trend model (LLT). If only σ_ς^2 is equal to zero, we have a model with deterministic slope (i.e., the level follows a random walk with constant drift). Finally, if only σ_η^2 is equal to zero, the model is said to have a smooth trend. However, if both variances are equal to zero, it can be shown that the model collapses to a deterministic linear trend.

In this section we propose a multivariate SUTSE model where the log of the mortality rate, for each age group x, follows a stochastic trend μ_t, given by the following equations:

$$
\begin{aligned}
\log(m_t) &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N\!\left[0, \Sigma_\varepsilon\right] \\
\mu_{t+1} &= \mu_t + \beta_t + \eta_t & & \\
\beta_{t+1} &= \beta_t + \varsigma_t & \begin{bmatrix}\eta_t \\ \varsigma_t\end{bmatrix} &\sim N\!\left[\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\Sigma_\eta & 0 \\ 0 & \Sigma_\varsigma\end{bmatrix}\right] \\
E\!\left(\varepsilon_{t,k}\varepsilon_{t,s}\right), E\!\left(\eta_{t,k}\eta_{t,s}\right), & E\!\left(\varsigma_{t,k}\varsigma_{t,s}\right) \neq 0 \quad \forall k,s; & t &= 1,\dots,T.
\end{aligned}
\tag{2-5}
$$

where $\log(m_t)$ is a $N \times 1$ vector of logarithms of the central mortality rates of the age groups at t, N is the number of time series, i.e., the number of age groups, $\alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}$, μ_t is the $N \times 1$ level vector at t, β_t is the $N \times 1$ slope vector at t and ε_t, η_t and ς_t are $N \times 1$ disturbances vectors. We define $\Sigma_\varepsilon, \Sigma_\eta$ and Σ_ς as time invariant full matrices, which is characteristic of SUTSE models. Then local linear trend model can be written in state space form as follows[6]:

$$\log(m_t)_{N\times 1} = [I_N \quad 0]_{N\times 2N} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}_{2N\times 1} + \varepsilon_t, \qquad \varepsilon_t \sim N[0, \Sigma_\varepsilon] \qquad (2-6)$$

$$\begin{pmatrix} \mu_{t+1} \\ \beta_{t+1} \end{pmatrix}_{2N\times 1} = \begin{bmatrix} I_N & I_N \\ 0 & I_N \end{bmatrix}_{2N\times 2N} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}_{2N\times 1} + \begin{bmatrix} I_N & 0 \\ 0 & I_N \end{bmatrix}_{2N\times 2N} \begin{bmatrix} \eta_t \\ \varsigma_t \end{bmatrix}_{2N\times 1}, \begin{bmatrix} \eta_t \\ \varsigma_t \end{bmatrix} \sim N \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \Sigma_\eta & 0 \\ 0 & \Sigma_\varsigma \end{bmatrix} \end{bmatrix}$$

$$E(\varepsilon_{t,k}\varepsilon_{t,s}), E(\eta_{t,k}\eta_{t,s}), E(\varsigma_{t,k}\varsigma_{t,s}) \neq 0 \quad \forall k,s; \qquad t = 1,\ldots,T.$$

where I_N is the identity matrix of size N.

In equations (2-5) and (2-6), we see that the logs of the central mortality rates of the different age groups are not directly related to each other. They are interconnected by the covariance matrices $\Sigma_\eta, \Sigma_\varsigma$ and Σ_ε, which play the role of capturing the similar influences on the temporal development of the levels, slopes and of the observable variable itself in the different age groups. This is a main characteristic of SUTSE.

2.2.3.2
Model with deterministic slope

A particular case of the local linear trend model described before is to consider a model with deterministic slope. This adaptation makes sense from the empirical point of view, because it is very common for time series not to have sufficient variance in the trend to justify a variant slope over time. For this purpose, in this section we set $\Sigma_\varsigma = 0$ so that $\beta_{t+1} = \beta_t = \beta$. With this, the trend

---

[6] Equation (2-6) is equivalent to the expression found in Harvey (1989, chapter 8, p.432-433), which uses Kronecker product.

μ_{t+1} remains stochastic, but with fixed slope. Therefore, the equation of the level μ_{t+1} becomes a random walk plus a constant drift, as follows:

$$
\begin{aligned}
\log(m_t) &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N[0, \Sigma_\varepsilon] \\
\mu_{t+1} &= \mu_t + \beta + \eta_t & \eta_t &\sim N[0, \Sigma_\eta] \\
E(\varepsilon_{t,k}\varepsilon_{t,s}), E(\eta_{t,k}\eta_{t,s}) &\neq 0 \quad \forall k,s & t &= 1,\dots,T
\end{aligned}
\tag{2-7}
$$

where $\beta \in \Re^{N \times 1}$ is the drift for all t.

This model with deterministic slope can be written in state space form as follows:

$$
\log(m_t)_{N \times 1} = \begin{bmatrix} I_N & 0 \end{bmatrix}_{N \times 2N} \begin{bmatrix} \mu_t \\ \beta \end{bmatrix}_{2N \times 1} + \varepsilon_t, \qquad \varepsilon_t \sim N[0, \Sigma_\varepsilon] \tag{2-8}
$$

$$
\begin{pmatrix} \mu_{t+1} \\ \beta \end{pmatrix}_{2N \times 1} = \begin{bmatrix} I_N & I_N \\ 0 & I_N \end{bmatrix}_{2N \times 2N} \begin{bmatrix} \mu_t \\ \beta \end{bmatrix}_{2N \times 1} + \begin{bmatrix} I_N \\ 0 \end{bmatrix}_{2N \times N} [\eta_t]_{N \times 1}, \qquad \eta_t \sim N[0, \Sigma_\eta]
$$

2.2.3.3
Estimation of model parameters

We assume that the initial state space vector has distribution $N(a_1, P_1)$. To estimate the fixed and unknown parameters (Ψ) of the models, we use the likelihood function which following Durbin and Koopman (2012), is given by:

$$
L(\Psi) = f(y_1, \cdots, y_T; \Psi) = f(y_1) \prod_{t=2}^{T} f(y_t \mid Y_{t-1}; \Psi) \tag{2-9}
$$

Then:

$$
\log L(\Psi) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \left(\log|F_t| + v_t' F_t^{-1} v_t \right) \tag{2-10}
$$

where $\Psi = \{\Sigma_\varepsilon, \Sigma_\eta, \Sigma_\varsigma\}$ in the model of section 2.2.3.1; $\Psi = \{\Sigma_\varepsilon, \Sigma_\eta, \beta\}$ in the model of section 2.2.3.2; $y_t = \log(m_t)$, $y_t \mid Y_{t-1} \sim N(Z_t a_t, F_t)$ and $Z_t = \begin{bmatrix} I_N & 0 \end{bmatrix} \in \Re^{N \times 2N}$ *in the model of section 2.2.3.1; and* $Z_t = \begin{bmatrix} I_N \end{bmatrix} \in \Re^{N \times N}$ in the model of section 2.2.3.2, with a_t and F_t being calculated by Kalman filter (eq.(2-2)).

To estimate Ψ, we used the STAMP 8.3 program, which employs concepts of diffuse priors for exact estimation of the model's unknown parameters. Details of the approach used by this software can be found in the appendix of chapter 9 of

Koopman et al. (2007) and concepts associated with the use of the diffuse likelihood distribution are discussed in chapter 7 of Durbin and Koopman (2012), whose expression is given by Koopman (1997) as follows:

$$\log L_d(\Psi) = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{d}\omega_t - \frac{1}{2}\sum_{t=1}^{T}\left(\log|F_t| + v_t' F_t^{-1} v_t\right) \tag{2-11}$$

where

$$\omega_t = \begin{cases} \log|F_{\infty,t}|, & \text{if } F_{\infty,t} \text{ is positive definite,} \\ \log|F_{*,t}| + v_t^{(0)'} F_{*,t} v_t^{(0)}, & \text{if } F_{\infty,t} = 0 \end{cases}$$

and the elements that comprise ω_t are well explained in Durbin and Koopman (2012, chapter 5).

2.2.3.4
Forecast s steps ahead

The forecasting functions obtained by extrapolating the models s steps ahead are presented next:

$$E\left(\log(m_{T+s})\mid Y_T\right) = E(\mu_T\mid Y_T) + s\,E(\beta_T\mid Y_T), \text{ in the local linear trend model,}$$

and

$$E\left(\log(m_{T+s})\mid Y_T\right) = E(\mu_T\mid Y_T) + s\,\hat{\beta}, \text{ in the model with deterministic slope.}$$

In turn, the variance is given by:

$$V\left(\log(m_{T+s})\mid Y_T\right) = Z\,V(\alpha_{T+s}\mid Y_T)Z' + \Sigma_\varepsilon, \tag{2-12}$$

with $V\left(\alpha_{T+s}\mid Y_T\right) = T^s V\left(\alpha_T\mid Y_T\right)T'^s + \sum_{i=0}^{s-1} T^i.Q.T'^i,$

where $s = 1, 2, \ldots$ is the forecast horizon, $E(\mu_T\mid Y_T)$ is the vector of the means of the smoothed levels at time T, $E(\beta_T\mid Y_T)$ is the vector of the means of the smoothed slopes at time T; $\hat{\beta}$ is the estimated deterministic slope, $V(\alpha_{T+s}\mid Y_T)$ is the conditional covariance matrix of the state vector at time $T+s$ given the information available up to time T, and $\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ in local linear trend model and $\alpha_t = \mu_t$ in the model with deterministic slope,

$$Z = \begin{bmatrix} I_N & 0 \end{bmatrix}_{Nx2N}, \quad T = \begin{bmatrix} I_N & I_N \\ 0 & I_N \end{bmatrix} \text{ e } Q = \begin{bmatrix} \Sigma_\eta & 0 \\ 0 & \Sigma_\varsigma \end{bmatrix}_{2Nx2N}, \text{ for the LLT, and}$$

$$Z = \begin{bmatrix} I_N \end{bmatrix}_{NxN}, \quad T = \begin{bmatrix} I_N \end{bmatrix} \text{ e } Q = \begin{bmatrix} \Sigma_\eta \end{bmatrix}_{NxN}, \text{ for the model with deterministic slope.}$$

Hence, the shape of the forecasting function of the logarithms of the mortality rates is a straight line for all age groups x. The conditional variance, and consequently the confidence interval, increases as the forecast horizon grows. Under Gaussianity, we have the following distribution for the log of the future mortality rates given the information available up to time T:

$$\log(m_{x,T+s}) \sim N(E(\log(m_{x,T+s})|Y_T), V(\log(m_{x,T+s})|Y_T)), \; x = 1,\ldots,N \text{ e } s = 1, 2,\ldots$$

(2-13)

where x is the age group and s is the forecast horizon.

From equation (2-13), it follows that the predicted central mortality rates have lognormal distribution given by:

$$m_{x,T+s} \sim \log normal(E(\log(m_{x,T+s})|Y_T), V(\log(m_{x,T+s})|Y_T)) \quad (2-14)$$

Therefore, the forecasting function of the central mortality rates is given by:

$$E(m_{x,T+s}|Y_T) = \exp(E(\log(m_{x,T+s})|Y_T) + 0.5\, V(\log(m_{x,T+s})|Y_T)),$$

with variance given by:

$$V(m_{x,T+s}|Y_T) = \exp[2E(\log(m_{x,T+s})|Y_T) + V(\log(m_{x,T+s})|Y_T)][\exp(V(\log(m_{x,T+s})|Y_T))-1]$$

(2-15)

The model predicts a natural damping of the reduction of the expected mortality rates as s grows, because $E(\log(m_{x,T+s})|Y_T) < 0$, $V(\log(m_{x,T+s})|Y_T) > 0$ and the values of the variances of the logarithms of the future rates increase with time. We believe this damping is a benefit of the model because it is reasonable to assume that the future longevity gain will not continue at the same pace as in recent decades.

However, the increase in variance starting at a certain future year $T+s$ may generate a forecast for growth of the mortality rates for some age group. This can be interpreted as an inflection point (ip). From this point on, the model would not predict the decline of the mortality rates. Hence, we can use the assumption that $E(m_{x,T+s}|Y_T)$ remains constant as of $s = ip$. Alternatively, we can also use as an estimator the median of the distribution of $m_{x,T+s}$, which is the estimator that

minimizes the absolute errors, being a more realistic estimator for asymmetric distributions. This latter alternative is more conservative with respect to solvency, but is unable to pick up the observed damping on the mortality rates.

The distribution of the factor of longevity gain s steps ahead beginning at time T is also lognormal and is given by:

$$G_{x,T+s} = \frac{m_{x,T+s}}{m_{x,T}}$$
(2-16)

Then: $G_{x,T+s} \sim \log normal\left(E\left(\log\left(m_{x,T+s}\right)\mid Y_T\right) - \log\left(m_{x,T}\right), V\left(\log\left(m_{x,T+s}\right)\mid Y_T\right)\right)$

where $m_{x,T}$ is the central mortality rate observed for age x at time T, and $G_{x,T+s}$ is the factor of longevity gain at age x between time T and $T+s$ given the information available up to time T. Therefore, the forecasting function of the longevity gain is given by:

$$E(G_{x,T+s}\mid Y_T) = \exp\left(E\left(\log\left(m_{x,T+s}\right)\mid Y_T\right) - \log\left(m_{x,T}\right) + 0.5\,V\left(\log\left(m_{x,T+s}\right)\mid Y_T\right)\right),$$

with variance given by:

$$V(G_{x,T+s}\mid Y_T) = \exp\left[2E\left(\log\left(m_{x,T+s}\right)\mid Y_T\right) + V\left(\log\left(m_{x,T+s}\right)\mid Y_T\right)\right]\left[\exp\left(V\left(\log\left(m_{x,T+s}\right)\mid Y_T\right)\right) - 1\right]$$
(2-17)

We can forecast the distribution of $m_{x,T+s}$ for a determined subset of the population with known $m_{x,T}$ using the longevity gain distribution of the population. With this, we can approximate the distribution of the mortality rates by the following expression:

$$m_{x,T+s} \approx m_{x,T} \times G_{x,T+s},$$
(2-18)

with the following moments:

$$E(m_{x,T+s}\mid Y_T) = m_{x,T}\,E(G_{x,T+s}\mid Y_T)$$

$$V(m_{x,T+s}\mid Y_T) = m_{x,T}{}^2\,V(G_{x,T+s}\mid Y_T)$$

Insurers and pension funds that do not have sufficient historic data to apply the proposed models, but manage to estimate $m_{x,T}$ at some point, can utilize equation (2-18) to predict their mortality rates. Nevertheless, they should use the distribution of the longevity gain of a population that has the same demographic characteristics as their policyholders or beneficiaries.

2.3

Application of the SUTSE model

Firstly, we apply the models proposed in section 2.2.3 to the Portuguese male population to check the fit and highlight the main prediction characteristics of the proposed SUTSE models. The mortality rates of the Portuguese male population were obtained from the Human Mortality Database[7], covering the period from 1940 to 2009, containing a total of 70 observations for each age. We have left out the last 5 years of data for out-of-sample validation.

Models were implemented using the program STAMP 8.3. Since there are not enough observations for the mortality rates time series, we had to reduce the number of parameters to be estimated in order to ensure likelihood convergence. In order to attain such objective we decided to work with ten homogeneous age groups, namely: <1 year, 1-9 years, 10-19 years, 20-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, 70-79 years and ≥ 80 years. For each of these groups central mortality rates were calculated using the following formula:

$$m_{x,t} = \frac{\sum_{i \in x} d_{i,t}}{\sum_{i \in x} L_{i,t}} \tag{2-19}$$

where $m_{x,t}$ is the central mortality rate for age group x at time t, $d_{i,t}$ is the number of deaths of people with age i at time t, and $L_{i,t}$ is the number of people exposed to risk at age i at time t.

Furthermore, to ensure convergence of our models, we reduce the number of parameters of the covariance matrices to be estimated by adopting Cholesky decomposition for such matrices. More specifically we used that $\Sigma = \Theta D \Theta'$ (see Koopman et al., 2007, chapter 9) for all covariance matrices in the estimation of the model's parameters, where Θ is a lower-triangular matrix with unity values on the leading diagonal and D is a non-negative diagonal matrix. Then, to reduce the number of parameters to be estimated we assume that all members of the D matrices are proportional, namely:

---

[7] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on July 16, 2012).

$$D_\varepsilon = hD, \quad D_\eta = w_1 D, \quad D_\varsigma = w_2 D, \tag{2-20}$$

where

$D = diag(d, \ldots, d) \in \Re^{N \times N}$, and

h, w_1 and w_2 are a non-negative scalars.

In addition, in the estimation process, we opt to set $h = 1$, $0 < w_1 \leq 1$ and $0 < w_2 \leq 1$. Optimal values for w_1 and w_2 were found through a grid search in (0,1], choosing the pair that maximized the likelihood. The results are presented in Table 2.1.

Table 2.1. Values of w_1 and w_2 used in the SUTSE models.

Model	w_1	w_2
LLT	0.1	0.05
model with determ. slope	1	_

In sections 2.3.1 and 2.3.2, we will see that the strategy to reduce the number of parameters in the SUTSE model here adopted will not negatively affect the goodness of fit and the predictive accuracy of the SUTSE model.

2.3.1
Model fitting

Diagnostic checking is performed using the standardized innovations associated with the models, which are tested for normality, homoscedasticity and serial uncorrelatedness. For this purpose, we use the tests of Bowman-Shenton, the heteroscedasticity test - H(h) (Durbin and Koopman, 2012) and Box-Ljung, respectively. To test the goodness of fit of the logs of the mortality rates modeled by equations (2-5) and (2-7), we use an easily interpreted discrepancy measure the mean absolute percentage error (MAPE), whose expression is as follows:

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_{t|t-1}}{y_t} \right| \tag{2-21}$$

where n is the number of observations considered, y_t is the observed time series value and $\hat{y}_{t|t-1}$ is the one step ahead value predicted by the model.

After fitting the local linear trend model (eq. (2-5)) and the model with deterministic slope (eq. (2-7)) to the logged series of central mortality rates, we carried out diagnosis of the standardized innovations. The results, presented in Tables 2.2 and 2.3 below, indicate that this series is well fitted by the proposed models.

Table 2.2. LLT model: diagnostics of the standardized innovations and fit of the model within the sample for each age group (p-values in parenthesis).

Age group (in years)	Normality	Heterosced.	Autocorrel.	MAPE (in logs)
<1	9.863 (0.007)	1.204 (0.337)	5.094 (0.747)	1.850%
1-9	1.859 (0.395)	0.291 (0.997)	3.536 (0.896)	1.369%
10-19	2.049 (0.359)	0.775 (0.718)	9.722 (0.285)	1.009%
20-29	2.247 (0.325)	0.429 (0.971)	1.461 (0.993)	1.136%
30-39	0.040 (0.980)	0.402 (0.979)	8.809 (0.359)	0.928%
40-49	0.189 (0.910)	0.317 (0.994)	15.232 (0.055)	0.989%
50-59	1.340 (0.512)	0.220 (0.999)	11.947 (0.154)	0.979%
60-69	0.647 (0.724)	0.273 (0.998)	7.464 (0.488)	1.096%
70-79	1.337 (0.512)	0.165 (0.999)	12.105 (0.147)	1.534%
≥ 80	10.061 (0.007)	0.191 (0.999)	13.661 (0.886)	1.931%

Table 2.3. SUTSE model with deterministic slope: diagnostics of the standardized innovations and fit of the model within the sample for each age group (p-values in parenthesis).

Age group (in years)	Normality	Heterosced.	Autocorrel.	MAPE (in logs)
<1	8.209 (0.016)	1.200 (0.340)	4.614 (0.798)	1.943%
1-9	2.941 (0.230)	0.332 (0.993)	2.823 (0.950)	1.323%
10-19	1.135 (0.567)	1.012 (0.489)	50.573 (0 .000)	1.089%
20-29	7.022 (0.030)	0.611 (0.866)	17.705 (0.0234)	1.113%
30-39	3.018 (0.221)	0.769 (0.724)	13.703 (0.090)	1.022%
40-49	0.626 (0.7310)	0.382 (0.984)	12.321 (0.141)	0.991%
50-59	0.137 (0.933)	0.268 (0.998)	7.696 (0.463)	0.951%
60-69	0.230 (0.891)	0.335 (0.992)	7.007 (0.536)	1.108%
70-79	3.091 (0.213)	0.162 (0.999)	13.029 (0.111)	1.552%
≥ 80	8.574 (0.014)	0.185 (0.999)	15.337 (0.053)	1.952%

Since we adopted a trend for each age group, the fit within the sample is very good, as can be seen in the last column of Tables 2.2 and 2.3. Furthermore, Figure 2.1 shows the goodness of fit of the models described in equations (2-5) and (2-7), from the plot of $\log(m_{x,t}) \times \log(\hat{m}_{x,t|t-1})$, where $\log(\hat{m}_{x,t|t-1}) = E(\log(m_{x,t})|Y_{t-1})$, for two age groups.
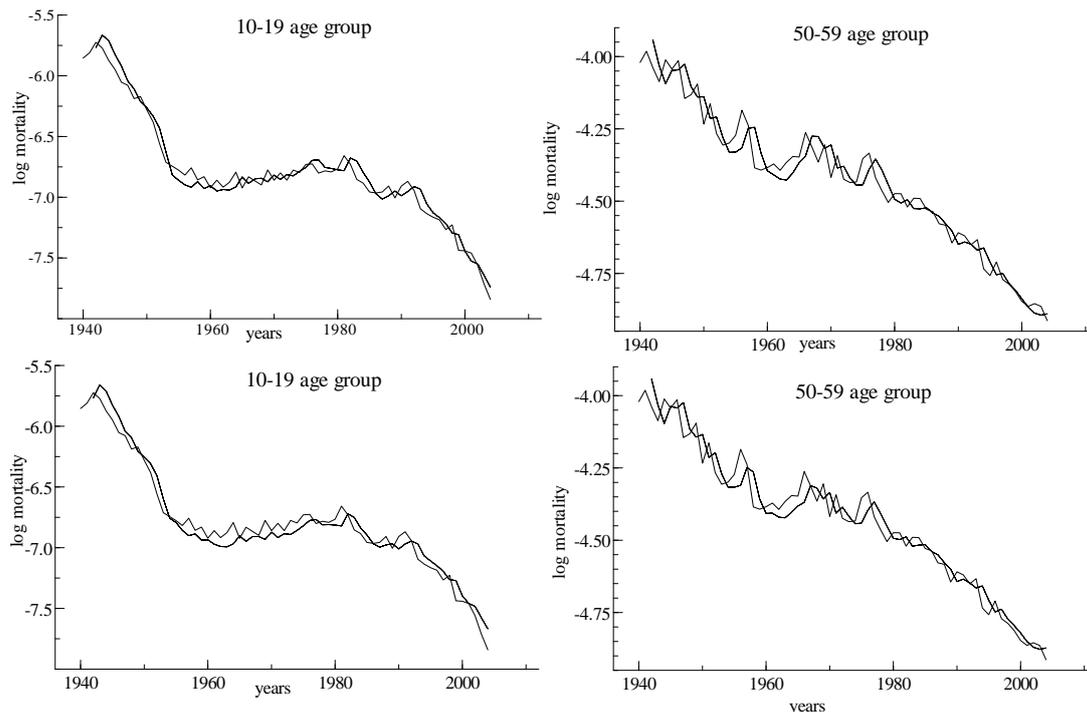
Figure 2.1. Observed log mortality, in solid lines, versus estimated log mortality, in dashed lines, for the 10-19 and 50-59 age groups. The two top graphs are by local linear trend model and the two bottom graphs by the model with deterministic slope.

Even for the age groups in which the mortality rates were stable between 1960 and 1980, such as for the 10-19 year-old age group, the models managed to capture this movement without the need of interventions.

From the estimated disturbance correlation matrix presented in Table 2.4 one can see that the correlations of the log-mortality rates are higher between consecutive age groups and also among older age groups.

Table 2.4. Disturbance correlation matrix of the log-mortality using the LLT model.

Age group	<1	1-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	≥ 80
<1	1	-	-	-	-	-	-	-	-	-
1-9	0.73	1	-	-	-	-	-	-	-	-
10-19	0.20	0.27	1	-	-	-	-	-	-	-
20-29	0.20	0.27	0.35	1	-	-	-	-	-	-
30-39	0.07	0.16	0.29	0.47	1	-	-	-	-	-
40-49	0.05	0.04	0.25	0.28	0.65	1	-	-	-	-
50-59	0.06	0.11	0.12	0.12	0.52	0.68	1	-	-	-
60-69	0.09	0.11	0.11	0.11	0.45	0.49	0.71	1	-	-
70-79	0.05	0.07	0.10	0.10	0.36	0.36	0.58	0.80	1	-
≥ 80	0.02	0.02	0.03	0.03	0.25	0.28	0.48	0.72	0.80	1

2.3.2
Models comparison

In this section we compare the predictive accuracy of the models proposed in section 2.2.3 with the benchmark procedure adopted to forecast mortality rates, the Lee and Carter model (1992), and its adaptations. We also compare the SUTSE models with other relevant models in the mortality rate literature, such as Renshaw and Haberman (2006) and Currie (2006), which capture a cohort effect, and also with the models by D'Amato et al. (2012) and by French and O'Hare (2013).

To summarize, the models used in our model comparison exercise are:

(1)   Lee and Carter (1992).

(2)   Lee and Carter in state space form, estimated by the Kalman filter as in De Jong and Tickle (2006):

$$y_t = \log(m_t) - a = b\,k_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, \Sigma_\varepsilon) \qquad \text{(2-22)}$$
$$k_{t+1} = \beta + k_t + \eta_t \qquad\qquad \eta_t \sim N(0, \sigma_\eta^2)$$

where $y_t \in \Re^{N \times 1}$, $\varepsilon_t \in \Re^{N \times 1}$, $k_t \in \Re$, β is the deterministic drift $\in \Re$,

$$a = \frac{\sum_{t=1}^{T} \log(m_t)}{T} \in \Re^{N \times 1}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N = 1 - \sum_{x=1}^{N-1} b_x \end{bmatrix} \in \Re^{N \times 1}, \eta_t \in \Re, \sigma_\eta^2 \in \Re,$$

$\Sigma_\varepsilon = diag\left(\sigma_\varepsilon^2, \sigma_\varepsilon^2, \cdots \sigma_\varepsilon^2\right) \in \Re^{N \times N}$ and N is the number of time series (age groups).

This model has only one unobservable component (k_t), which measures the evolution of the trend in time for all age groups. The vector b corresponds to the weights associated with each age group.

(3)   Extension of model (2) with addition of a stochastic drift instead of the fixed β of eq. (2-22), modeled by means of a random walk, equivalent to a local linear trend model with only one component (k_t), linking all the series, as suggested by De Jong and Tickle (2006).

(4)   Extension of model (2) with addition of a stochastic drift, modeled by means of an AR(1), equivalent to a damped trend model.

(5)   Renshaw and Haberman (2006)

(6)   Currie (2006)

(7)   D'Amato et al. (2012)

(8)   French and O'Hare (2013): the authors, following Forni et al. (2005), adopted a dynamic factor model to forecast mortality rates. They used a very simple form of a common trend model, with only one long run component driving the log mortality rates for all ages. One has to bear in mind that their

application goodness of fit is reported for the log mortality rates, while our results are for mortality rates. Also they consider age groups (from 10 to 89 years old) different than those we have worked with.

Models (2) to (4) were estimated by Kalman filter using Ssfpack/S-Plus[8] (Koopman et al., 2008). It should be noticed that they are not proper SUTSE models since there is no direct correlation between the log-mortality series. Instead, the link between the series is established by the underlying trends. Models (5) and (6) were estimated as described by Cairns et al. (2009). Model (8) was implemented by means of a Matlab program available in Forni´s website[9].

We adopt the MAPE criteria to compare the forecasting power among the models[10]. Table 2.5 shows the MAPE values, both in sample and out-of-sample, for the different models. In a second exercise we also fitted the SUTSE models to the mortality rates of the US male population, where data was also obtained from the Human Mortality Database, in the period from 1933 to 2009. As in the case of Portugal, the last five years of were used for out-of-sample testing.

---

[8] SsfPack for S-Plus is available in the S+FinMetrics module.
[9] http://morgana.unimore.it/forni_mario/matlab.htm.
[10] We do not use AIC or BIC metrics in addition to MAPE because the competing models use different dependent variables.

Table 2.5. Comparison of the models by the MAPE method – Portugal and USA.

MODEL	MAPE			
	Portugal		United States	
	in sample	out of sample	in sample	out of sample
SUTSE - LLT	**5.10%**	**7.55%**	**2.30%**	**2.72%**
SUTSE with determ. slope	**5.17%**	**11.64%**	**2.73%**	**4.83%**
(1) Lee- Carter (1992)	12.80%	19.26%	8.69%	17.35%
(2) Lee-Carter in state space (SS)	11.57%	17.98%	6.48%	10.44%
(3) Lee-Carter in SS – LLT	13.29%	14.34%	6.60%	11.80%
(4) Lee-Carter in SS – damped trend	15.40%	14.98%	6.59%	11.90%
(5) Renshaw and Haberman (2006)	6.10%	19.77%	2.51%	7.59%
(6) Currie (2006)	22.18%	46.21%	10.35%	36.79%
(7) D'Amato et al. (2012)	12.84%	19.74%	9.82%	19.78%
(8) French and O'Hare (2013)	39.01%	61.66%	39.20%	42.94%

One can see that, both in sample and out-of-sample, our proposed SUTSE models are the ones that best fit the data for both populations. For US male population the cohort effect is valuable to forecast the future mortality rates, as one can see by comparing the MAPE value obtained by Renshaw and Haberman (2006) with that by Lee and Carter (1992). Nevertheless, as one can see from Table 2.5, both our SUTSE models produced similar in sample, and significantly better out-of-sample MAPE when compared to that model.

Furthermore, we compare the forecasting performance amongst the competing models applying the Diebold-Mariano (DM) test (Diebold and

Mariano, 2002). We test the null hypothesis of no difference in the accuracy of two competing forecasts, using MAPE as the loss function. Our SUTSE models are compared among themselves and also compared with the best competing model for each population (see Table 2.5).

Table 2.6 shows the p-values for the null hypothesis of no difference in the accuracy of two competing forecasts. Under DM test, we can conclude that the local linear trend model produces the best forecasting power among all models, and that the SUTSE model with deterministic slope is better than the other competing models, for both populations.

Table 2.6. DM test: p-values for the null hypothesis of no difference in the accuracy of two competing forecasts.

Comparison	Portugal	United States
SUTSE - LLT vs. SUTSE with determ. slope	0.015	0.000
SUTSE - LLT vs. best competing model	0.000	0.000
SUTSE with determ. slope vs. best competing model	0.000	0.000

Therefore, for the studied populations, our results show that the proposed SUTSE models outperformed the competing models even when the cohort effects are explicitly modeled. The main advantage of SUTSE models is that they are able to capture stochastic trends on mortality rate for each age group and also to account for dependence between these trends. The other advantage is that one can easily derive confidence intervals for future mortality rates (eq. (2-14)) and also for longevity gain factors (eq. (2-16)), allowing the proper incorporation of uncertainty in such projections.

2.3.3
Prediction

When fitted to a log of a series, SUTSE models produce forecasting variances that growth with the time horizon. Given the relationship between the mean of a log normally distributed variable and the variance of a normally distributed variable, this produces a log normally distributed variable with a changing mean. When such a model was fitted to the mortality rate of Portugal´s male population, one can notice, in the local linear trend model, after a certain time period, the occurrence of a steady growth on the mortality rates projections for the majority of the age groups. This does not seem justified on the light of the observed improvements on the overall life quality of the Portuguese population in the last decades or so. More specifically, this inflection point depends on the values of $V\left(\log\left(m_{x,T+s}\right)|Y_T\right)$ and $E\left(\beta_T|Y_T\right)$ for each series, as can be seen by combining equations (2-12) and (2-15). Such undesirable behavior can be circumvented by considering two alternative estimators for the future mortality rates in the local linear trend model, namely:

i. $AE_{x,T+s}=$ minimum of $\left(E\left(m_{x,T+s}|Y_T\right),AE_{x,T+s-1}\right)$, where $AE_{x,T}=E\left(m_{x,T}|Y_T\right)$; and

ii.   Median of $\left(m_{x,T+s}|Y_T\right)$.

The second alternative is more conservative and its use is recommended when the model is used to analyze the solvency of companies. However, it does not contemplate the damping of future mortality rates as does the first alternative. Nevertheless, when considering a back test this alternative estimator produces a MAPE value (7.42%) very similar to that found when the mean of future mortality rates is used as an estimator. So, we can assume that this alternative estimator can produce realistic future mortality rates.

Figure 2.2 depicts the predicted mortality rates for the 10-19 age group using the aforementioned predictors produced by local linear trend model. Note that the inflection point for this age group occurs in 2045. As it can be seen the confidence interval increases with time, a typical feature presented by models with stochastic trends.
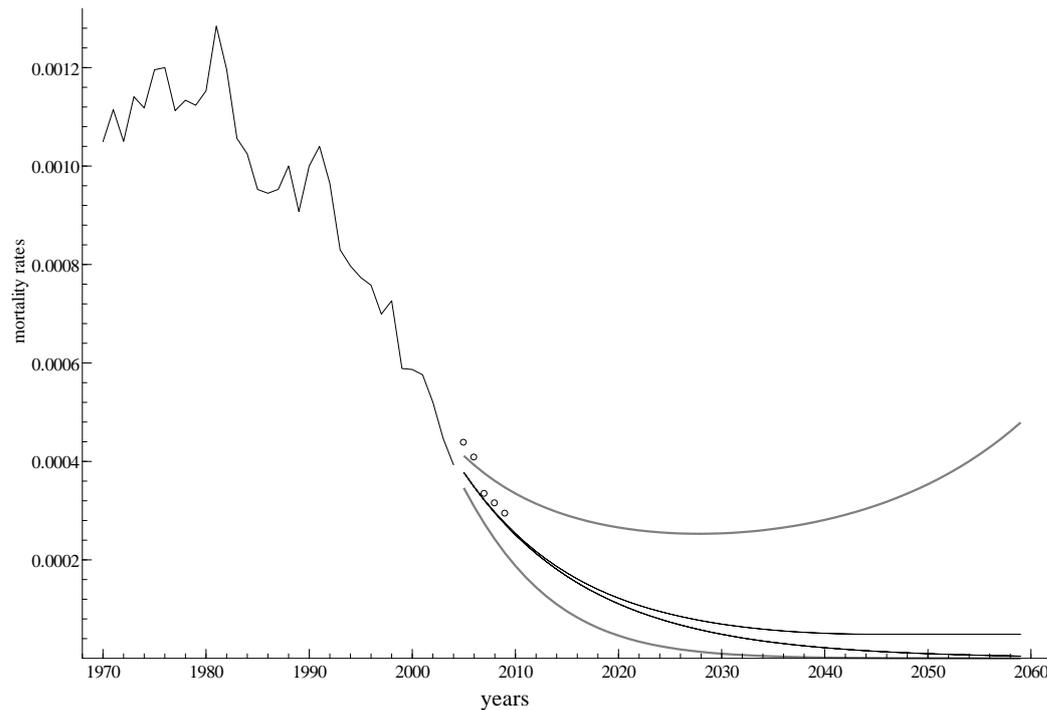
Figure 2.2. Predicted mortality rates for the 10-19 age group by the LLT. The solid line represents actual mortality rates, in-sample; the circles actual mortality rates, out-of-sample; the long dashed line is the $AE_{x,T+s}$; the short dashed line is the median of $\left(m_{x,T+s} \mid Y_T\right)$; the dotted lines the upper and lower limits of the 95% confidence interval for the predicted mortality rates.

One can notice that for the model with deterministic slope, the predicted mortality rate values always decline over time given that the forecast variances are smaller than those obtained via local linear trend model. In this case, we choose to use the conditional expected value as an estimator for the future values (Figure 2.3). As shown in Tables 2.5 and 2.6, the out-of-sample forecasts of such model are not as well adjusted as the local linear trend model but are superior to those produced by the competing models. This conclusion is only valid for the five-year period considered in our study (2005 to 2009). Nevertheless, the deterministic slope model produces a much smaller confidence interval than local linear trend model. For this reason, we assume that this model produce better predictions when considering longer time horizons, especially when working with simulation.
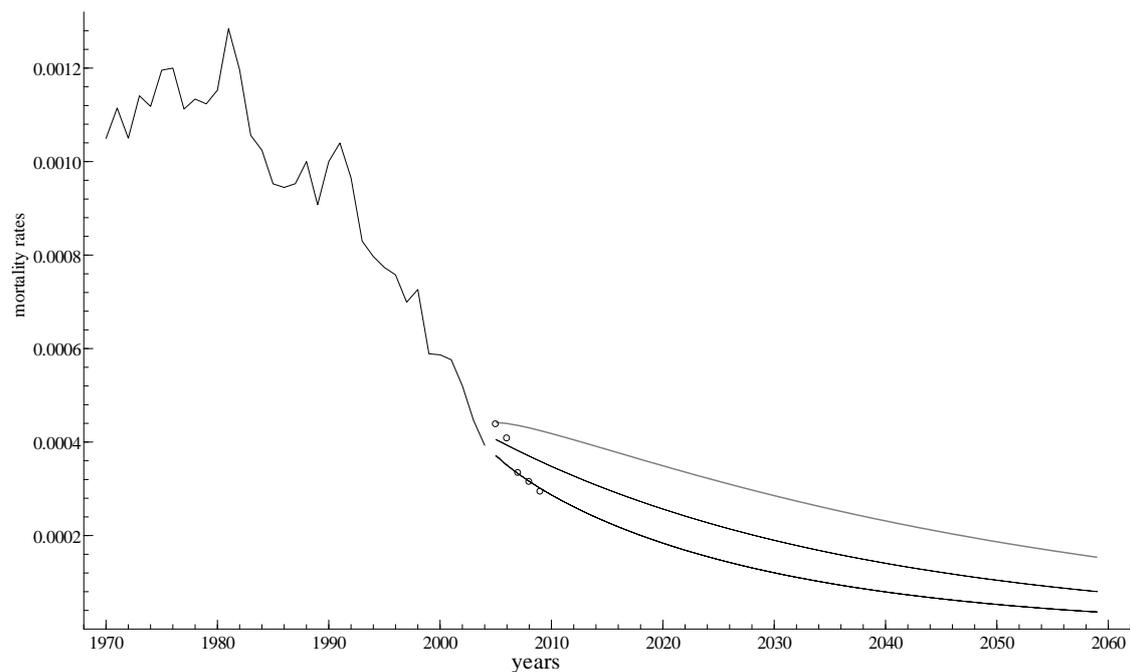
Figure 2.3. Predicted mortality rates for the 10-19 age group by the SUTSE model with deterministic slope. The solid line represents actual mortality rates, in-sample; the circles actual mortality rates, out-of-sample; the dashed line is the $E\!\left(m_{x,T+s} \mid Y_T\right)$; the dotted lines the upper and lower limits of the 95% confidence interval for the predicted mortality rates.

In Figure 2.4 we highlight the differences between the expected long-term longevity gains between our proposed models and that of Lee and Carter (1992), considering the last four age groups. For the local linear trend model, we present the two alternative estimators:

i. $LAE_{x,T+s} = \text{minimum of}\left(E\!\left(G_{x,T+s} \mid Y_T\right), LAE_{x,T+s-1}\right)$, where $LAE_{x,T} = 1$; and

ii. Median of $\left(G_{x,T+s} \mid Y_T\right)$.

When we use the median, there is a sharp reduction in the forecast mortality rates because they reflect the last projected slope for each age group without damping. We stress that the expected mortality improvement in the deterministic slope model is smaller than that of Lee and Carter (1992) because of the natural inclusion in this model of the damping effect in the formula for the mean. This is a naturally adequate model to the Portuguese population since it

projects a damping trend that captures the observed reduction of mortality rates over time. The assumption that the decrease in mortality rates observed in recent decades will not continue in the future at the same pace appears highly plausible to us.

We can see that, for example, for the 80 and over age group, over a forecast horizon of 55 years, the SUTSE model with fixed slope projects an average factor of longevity gain of 84.55% against 69.98% produced by Lee-Carter model. These gains represent a reduction of the mortality rates of about 15% and 30%, respectively, based on the last year of the sample (2004).
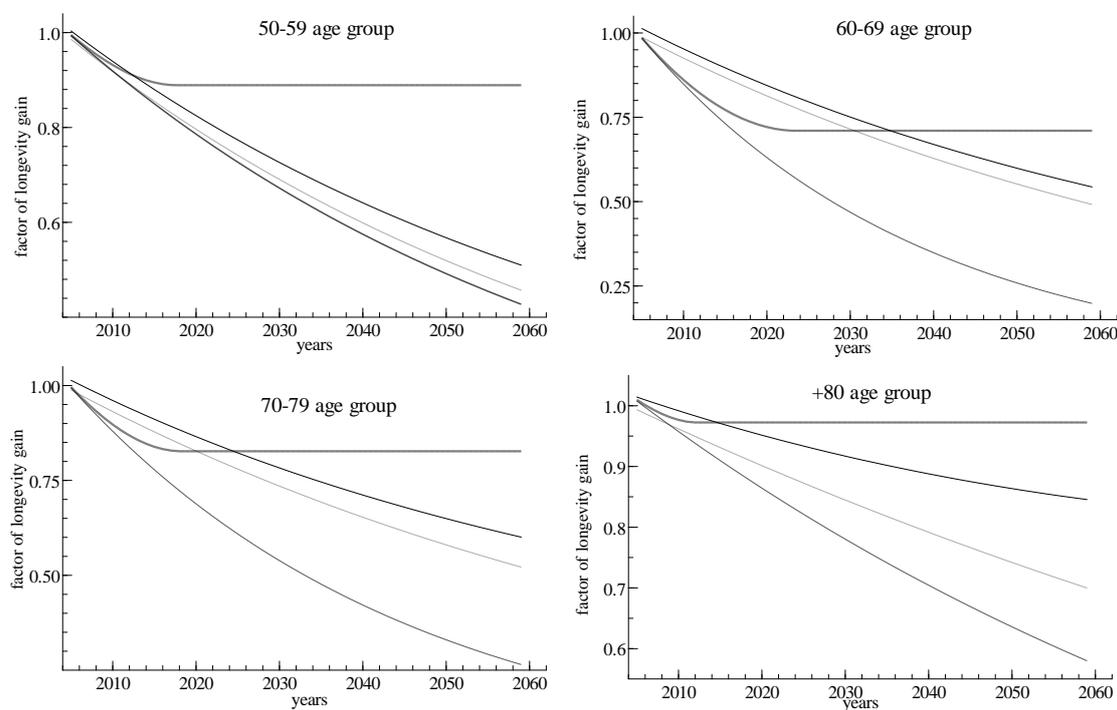
Figure 2.4. Forecast of the expected factors of longevity gain for the last four age groups: 50-59 years, 60-69 years, 70-79 years and 80+ years. SUTSE LLT: $LAE_{x,T+s}$ is in dotted lines and median of $\left(G_{x,T+s} \mid Y_T\right)$ in long dashed lines; SUTSE model with deterministic slope, in solid line; Lee and Carter (1992) in short dashed lines.

2.4
Conclusions

In this paper a multivariate SUTSE framework was proposed to forecast longevity gains for different age groups of a population. Models were fitted, independently, to male mortality rates of Portuguese and US populations. In both cases SUTSE models outperformed, in sample and out-of-sample, Lee-Carter like

models, even when the cohort effects are observed in the studied population. More specifically, for the tested populations, the local linear trend model presented the lowest MAPE for short term forecast, as it can be seen in Table 2.5, where we display results for five years of out-of-sample forecasting. However, for long term forecast, a SUTSE model with deterministic slope is more appropriate since its forecasted variances are smaller than those of the local linear trend model. One can note that our model with deterministic slope also performed much better in back-testing than Lee and Carter model (1992) and other competing models.

As was shown, in SUTSE models dependence between different time series is captured by full disturbance covariance matrices associated with disturbances that drive the observable log mortality series and also the unobserved stochastic trend. The SUTSE models characteristics allow that similar influences in the mortality rate trends of age groups are captured resulting in a model of a good forecasting power. Our results also showed that the disturbance correlations of the log-mortality rates are higher for consecutive age groups and also among older age groups. In practice the use of SUTSE models by insurers and pension funds will be dictated by the availability of time series whose length will suffice to ensure likelihood convergence.

3

Paper 2: Forecasting longevity gains for a population with short time series using a structural SUTSE model: an application to Brazilian annuity plans

In this paper, a multivariate structural time series model with common stochastic trends is proposed to forecast longevity gains of a population with a short time series of observed mortality rates, using the information of a related population for which longer mortality time series exist. The state space model proposed here makes use of the seemingly unrelated time series equation (SUTSE) and applies the concepts of related series and common trends to construct a proper model to predict the future mortality rates of a population with little available information. This common trends approach works with by assuming the two populations' mortality rates are affected by common factors. Further, we show how this model can be used by insurers and pension funds to forecast mortality rates of policyholders and beneficiaries. We apply the proposed model to Brazilian annuity plans where life expectancies and their temporal evolution are predicted using the forecasted longevity gains. Finally, to demonstrate how the model can be used in actuarial practice, the best estimate of the liabilities and the capital based on underwriting risk are estimated by means of Monte Carlo simulation. The idiosyncratic risk effect in the process of calculating an amount of underwriting capital is also illustrated using that simulation

3.1
Introduction

In order to maintain the solvency of insurers and pension funds, longevity gains have to be realistically measured by actuaries and risk managers. Longevity risk, unlike most other actuarial risks, does not have a perfect hedge and is independent of the size of the covered population. The continuing downward trend in mortality rates observed in most developed and developing countries requires increasing provision and capital to meet future commitments to

policyholders and beneficiaries. In this paper, to evaluate the commitments of Brazilian insurers and pension funds that sell plans with annuity payments, we propose a multivariate model with stochastic trends to forecast the longevity gains of a population with short time series of observed mortality rates.

Relevant scholarly literature proposes numerous models for forecasting mortality rates and managing longevity risk. Among these, the most widely used by practitioners is probably that of Lee and Carter (1992). In this model, the unobservable component that measures the evolution of the trend over time is common to all age groups. Factors are estimated corresponding to the relative weights of the unobservable component for each age group. Its estimation is accomplished in two steps: singular value decomposition (SVD) and ordinary least squares (OLS). To forecast future mortality rates, ARIMA models for time series are then used.

In the sequel we briefly review those extensions of the Lee-Carter method that have made use of the state space framework and the Kalman filter, since this is the approach used in our proposed model. In De Jong and Tickle (2006), the authors extend the Lee-Carter model by introducing an extra unobservable component in order to capture better the evolution of the trend in time. Since the model is cast into the state space form, the Kalman filter is used for parameter estimation. Hári et al. (2008) used a multivariate state space model with few latent factors to capture the common movements among mortality time series associated with different age groups. To forecast mortality rates, Gao and Hu (2009) also proposed models based on a Kalman filter, which contains a dynamic mortality factor model that assumes conditional heteroscedasticity on the mortality rate time series. Lazar and Denuit (2009) also used state space representation to frame known multivariate time series models to forecast future death rates of a population: dynamic factor analysis and vector-error correction (VEC), where concepts of common trends and cointegration are exploited.

To model mortality rates for a population with limited data, Li et al. (2004) proposed a modified Lee-Carter model and applied it to data in which there are few observations at uneven intervals. Additionally, some important studies are available which tackle joint modeling of the mortality rates of two populations. For example, Li and Lee (2005) extended the Lee-Carter method assuming, at the outset, that mortality trends of a group of populations are driven by a common

factor. If the difference between the group trend and that of an individual population is systematic and significant, an idiosyncratic factor for this population is specified. To model the mortality evolution of small populations, Jarner and Kryger (2011) proposed a methodology based on the existence of a larger reference population, used to estimate the underlying long-term trend, with use of the small population to estimate the deviation from this trend by employing a multivariate stationary time series model. Li and Hardy (2011) examined basis risk in index longevity hedges considering four extensions of the Lee-Carter model, in which dependence between the populations may be captured by the following structures: both populations are jointly driven by the same single time-varying index; the two populations are cointegrated; the populations depend on a common age factor; and there is an augmented common factor model in which a population-specific time-varying index is added to the common factor model with the property that it will converge towards a certain constant level over time. To model the joint development over time of mortality rates in a pair of related populations, Cairns et al. (2011) proposed an age-period-cohort model that incorporates a mean-reverting stochastic spread that allows for different trends in mortality improvement rates in the short run and parallel improvements in mortality in the long run. Dowd et al. (2011) proposed a "gravity" model to estimate mortality rates for two related populations with different sizes. They modeled the larger population independently and the smaller population in terms of spreads relative to the evolution of the former, but the spreads and cohort effects between the populations depend on gravity or spread reversion parameters for the two effects. We should also mention the following relevant articles on the subject of prediction of future mortality structures, such as Delwarde et al. (2007), Plat (2009) and Haberman and Renshaw (2009, 2011 and 2012).

The state space model proposed here makes use of the seemingly unrelated time series equation (SUTSE) (see Harvey, 1989, Chapter 8) to model the time series of mortality rates associated with particular age groups. In such a framework, initially a general multivariate trend model is specified, in which each time series has its own independent stochastic trend component. Dependence between the individual trend components is made possible by assuming that such trends are driven by noises with a multivariate normal distribution. Such structure enables one to introduce similar influences on the mortality trends (Fernández and

Harvey, 1990), such as better eating habits and improvements in life quality and medical care, all of which can significantly reduce future mortality rates. A more parsimonious dependence structure may be derived from the aforementioned SUTSE model by testing for restrictions on the covariance matrix associated with the trend noise. Under such restrictions, a SUTSE model collapses to a common trends model, one in which a reduced number of trends is able to satisfactorily explain the multivariate time series. Due to the ease of treating missing data in state space models, this framework can be used to estimate longevity rates for populations with short time series, which is typical of most developing countries. The strategy rests on the joint modeling of such a set of series with the corresponding mortality rate series obtained from a population with a long time series of data (the related population) that has similar mortality characteristics to the population with the shorter time series.

In this paper, we apply our proposed model to the Brazilian male and female populations using American and Portuguese populations as related populations. Additionally, we also show how this model can be used by insurers and pension funds to manage the risk of declining mortality rates. We estimate the distribution of Brazilian policyholders' future mortality rates through the model. Furthermore, the complete life expectancies of current Brazilian policyholders are estimated as well as their temporal evolutions. Then, we analyze the forecasting of the life expectancies over time, considering the longevity gain forecasted by the model.

To exemplify how the model can be used to measure underwriting risk, using Monte Carlo simulation, we obtain distributions of present values of cash flows generated by company expenses from hypothetical beneficiary groups receiving life annuities. Through this distribution, we calculate the best estimate of the liabilities and the capital based on underwriting risk, the latter derived from the tail distribution of present values of the cash flows. By means of Monte Carlo simulation, the idiosyncratic risk effect on the process of calculating an amount of underwriting capital is also shown, since we simulate using beneficiary groups with different sizes.

The remainder of the paper is organized as follows. In section 3.2, we provide a description of the proposed model. In section 3.3, we apply the model to the Brazilian population. In section 3.4, we apply the model to forecast mortality

rates of Brazilian policyholders and estimate life expectancies and their temporal evolutions. In section 3.5, we show a valuation of the best estimate of the liabilities and the capital based on underwriting risk. The last section concludes.

3.2
The SUTSE model

Before presenting the structural SUTSE model, we give a formal introduction of state space models and the Kalman filter.

3.2.1
State space and Kalman filter

A model represented in linear Gaussian state space (Durbin and Koopman, 2012) is defined as follows:

- observation equation: $y_t = Z_t \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, H_t)$ (3-1)

- state equation: $\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim N(0, Q_t)$

where $t = 1, \ldots, T$, y_t is a $N \times 1$ vector of observations, α_t is an unobserved $m \times 1$ vector, the system matrices have theses dimensions: $Z_t \in \Re^{N \times m}$, $T_t \in \Re^{m \times m}$, $R_t \in \Re^{m \times r}$, $H_t \in \Re^{N \times N}$, $Q_t \in \Re^{r \times r}$, N is the number of time series and m is the number of components of the state vector α_t. The initial state vector $\alpha_1 \sim N(a_1, P_1)$ is independent of $\varepsilon_1, \cdots, \varepsilon_T$ and of $\eta_1, \cdots, \eta_T$. ε_t and η_t are serially and mutually independent for all t.

In state space models, parameters and unobservable component estimates are estimated by use of the Kalman filter. This filter is a recursive procedure to compute the estimator for the state space vector at time t, based on the information available up to that time (Harvey, 1989). The recursive equations of the Kalman filter are given below:

$$E(\alpha_{t+1} | Y_t) = a_{t+1} = T_t a_t + K_t v_t;$$

$$V\left(\alpha_{t+1} \mid Y_t\right) = P_{t+1} = T_t P_t L_t^{'} + R_t Q_t R_t^{'};$$

$$v_t = y_t - E(y_t \mid Y_{t-1}) = y_t - Z_t a_t \text{ is the innovations vector;}$$ (3-2)

$$K_t = T_t P_t Z_t^{'} F_t^{-1} \text{ is the Kalman gain;}$$

$$F_t = Z_t P_t Z_t^{'} + H_t \text{ is the covariance matrix of the innovations; and}$$

$$L_t = T_t - K_t Z_t.$$

where $Y_t = \left\{y_1, \cdots, y_t\right\}$ and $v_t \sim N\left(0, F_t\right)$.

3.2.2
The proposed model

In this structural model (Harvey, 1989) the only component of interest to be estimated is the trend. We adopt a framework analogous to the seemingly unrelated regression equation (SURE), called the seemingly unrelated time series equation (SUTSE) by Harvey (1989). The main characteristic of a SUTSE model is that the dependence structure among the time series is captured by a full covariance matrix of the shocks of the observation equation (see eq. (3-1)). Additionally, each time series has its own trend components. These stochastic components are interconnected through a full disturbance covariance matrix.

In practice, using such a framework to estimate trends for a short time series of mortality rates can lead to imprecise trend estimates, and consequently unreliable longevity forecasting. One way around this problem is to adopt the concepts of related series and common trends for the SUTSE model (Harvey, 1989) and then to use them to construct a proper model to predict the future mortality rates of a population with little available information. This common trends approach works by assuming the two populations' mortality rates are affected by common factors, as adopted by Li and Lee (2005), Jarner and Kryger (2011), Li and Hardy (2011), Cairns et al. (2011) and Dowd et al. (2011).

Central to our model is the assumption that there is a population whose time series of mortality rates are related to those of the population of interest. The chosen reference population must have a long time series, sufficient to estimate its mortality trends. The model estimates a trend for every age group of each population and fully specifies the dependence structure among these age groups.

Nevertheless, our approach assumes that the disturbances of the unobserved components of the two populations are perfectly correlated. Therefore, each unobservable component of each age group of the populations used in the model is related by means of a linear equation. Complete data exist for population 1, which is the related population, while the data on population 2, the population of interest, are incomplete.

More specifically, in our approach we assume N common trends in a structural SUTSE model that corresponds to the N time series of population 1. Therefore, as with Li and Hardy (2011), we assume that the time series are cointegrated, but here this condition is imposed by setting up a common trend for the SUTSE model approach. To impose this condition, the model adopts the premise that each level and slope of the N series of population 2 is a linear combination of the same component of the respective age group of the related population 1. With this, we force rank equal to N for the covariance matrices with size $2N$ x $2N$ of the unobservable components' disturbances.

As stated in Dowd et al. (2011), a "biologically reasonable" mortality model should allow for the interdependence of the mortality rates of both populations that are subject to common influences. This characteristic is presented in the proposed model, since the full covariance matrix of the shocks of both populations captures the similar influences in the process of mortality improvements, such as better eating habits, life quality and medical care. Besides this, the diagonal matrices link the levels and slopes of population 2 with the levels and slopes of population 1. Additionally, the dependence structure among the mortality rates at different ages is also captured by a full covariance matrix of the shocks of the observation equation. Another advantage of the proposed model is that one can easily derive confidence intervals for both future mortality rates and longevity gain factors, allowing proper incorporation of uncertainty in such projections.

Our proposed model is a particular case of the local linear trend model. The related time series approach is accomplished by associating the series of each age group by means of the diagonal matrices, and putting the variables of each age group in a cluster, as set out in Harvey (1989). The observable time series are

given by the logarithm of the central mortality rates for each age x [11]. The vector $\log(m_t)$ is partitioned into two sub-vectors: the first associated with the age groups of the population with complete data and the second associated with the age groups of the population with incomplete data. Analogously, the other variables included in the model are also partitioned. The expressions of the model are as follows:

$$\log(m_t) = (\theta^+)^{-1} \mu_t + \varepsilon_t$$
$$\mu_{t+1} = \mu_t + \theta^+ (\chi^+)^{-1} \beta_t + \eta_t \qquad (3\text{-}3)$$
$$\beta_{t+1} = \beta_t + \varsigma_t$$

where

$t = 1, \ldots, T$;

$\varepsilon_t \sim N(0, \Sigma_\varepsilon)$;

$$\eta_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\eta,1} & 0 \\ 0 & 0 \end{bmatrix} \right) ;$$

$$\varsigma_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\varsigma,1} & 0 \\ 0 & 0 \end{bmatrix} \right) ;$$

$$\log(m_t) = \begin{bmatrix} \log(m_{1,t}) \\ \log(m_{2,t}) \end{bmatrix}; \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}; \quad \beta_t = \chi^+ \begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \end{bmatrix} = \begin{bmatrix} \beta_{1,t} \\ \beta \end{bmatrix};$$

$$\varsigma_t = \chi^+ \begin{bmatrix} \varsigma_{1,t} \\ \varsigma_{2,t} \end{bmatrix}; \quad \chi^+ = \begin{bmatrix} I_N & 0 \\ -\chi & I_N \end{bmatrix}; \quad \theta^+ = \begin{bmatrix} I_N & 0 \\ -\theta & I_N \end{bmatrix};$$

$$\mu_t = \theta^+ \begin{bmatrix} \mu_{1,t} \\ \mu_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_{1,t} \\ \mu_t^+ \end{bmatrix}; \quad \eta_t = \theta^+ \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \end{bmatrix};$$

$\mu_{t+1}^+ = \mu_t^+ + (\chi - \theta)\beta_{1,t} + \beta$, where $\mu_t^+ = \mu_0 + t.\beta + (\chi - \theta)\sum_{i=1}^{t-1} \beta_{1,i}$ and $\mu_1^+ = \mu_0 + \beta$;

$\chi = diag(\chi_1, \ldots, \chi_N)$ is the $N \times N$ loading matrix of the slope;

$\theta = diag(\theta_1, \ldots, \theta_N)$ is the $N \times N$ loading matrix of the level;

Σ_ε is a $2N \times 2N$ full matrix and $\Sigma_{1,\eta}, \Sigma_{1,\varsigma}$ are $N \times N$ full matrices;

---

[11] Since log transformation can improve residuals diagnostics in many situations and guarantees positive mortality rates, we chose to work with logs.

$$\Sigma_{\varepsilon} = \begin{bmatrix} \Sigma_{\varepsilon,1} & \Sigma_{\varepsilon,12} \\ \Sigma_{\varepsilon,21} & \Sigma_{\varepsilon,2} \end{bmatrix} ; \quad \Sigma_{\eta} = \begin{bmatrix} \Sigma_{\eta,1} & 0 \\ 0 & 0 \end{bmatrix} ; \quad \Sigma_{\varsigma} = \begin{bmatrix} \Sigma_{\varsigma,1} & 0 \\ 0 & 0 \end{bmatrix} ; \quad \text{(all covariance}$$

matrices are time invariants); $\log(m_{a,t})$ is a $N \times 1$ vector of logarithms of the central mortality rates of population a at time t, a is 1 or 2; N is the number of age groups in each population; $\varepsilon_{a,t}$ is a $N \times 1$ disturbance vector of population a at t ; $\beta_{a,t}$ is the $N \times 1$ slope vector of population a at t; $\mu_{a,t}$ is the $N \times 1$ level vector of population a at t; and μ_0 and β are $N \times 1$ constant vectors respectively of the level and slope. A particular case of the model occurs when $\Sigma_{\varsigma,1} = 0$ (i.e., the drifts are fixed). With this, we only have the common levels.

In addition, because of the common trends, the series $\log(m_{1,t})$ and $\log(m_{2,t})$ are cointegrated of order (2, 1). Then, we can recompose the unobserved data of the time series of population 2 based on the relation between the populations, which is given by this formula:

$$\log(m_{2,t}) = \theta \log(m_{1,t}) + \mu_t^+ + \varepsilon_{2,t} - \theta \varepsilon_{1,t} \tag{3-4}$$

To estimate the fixed and unknown parameters (Ψ) of the model, we use the STAMP 8.3 program, which employs concepts of diffuse priors for exact estimation of the model's unknown parameters. Details of the approach used by this software can be found in the appendix of chapter 9 of Koopman *et al.* (2007) and concepts associated with the use of the diffuse likelihood distribution are discussed in chapter 7 of Durbin and Koopman (2012), whose expression is given by Koopman (1997) as follows:

$$\log L_d(\Psi) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{d} \omega_t - \frac{1}{2} \sum_{t=d+1}^{T} \left(\log|F_t| + v_t' F_t^{-1} v_t \right) \tag{3-5}$$

where

$$\omega_t = \begin{cases} \log|F_{\infty,t}|, & \text{if } F_{\infty,t} \text{ is positive definite,} \\ \log|F_{*,t}| + v_t^{(0)'} F_{*,t} v_t^{(0)}, & \text{if } F_{\infty,t} = 0 \end{cases}$$

and the elements that comprise ω_t are well explained in Durbin and Koopman (2012, chapter 5).

Insurers and pension funds with little available information can apply the model presented in this section. They just have to find a related population with their insured groups and apply the longevity gain model. Because of the model's assumption, the chosen population should have mortality rates correlated with the mortality rates of policyholders and beneficiaries. In practice, companies can test the model using several related populations and select one that produces the best fit.

3.2.3
Forecast s steps ahead

The forecasting functions obtained by extrapolating the models s steps ahead are given by theses expressions:

$$E\big(\log(m_{T+s})|Y_T\big)=\left(\hat{\theta}^{+}\right)^{-1}E(\mu_T|Y_T)+s\left(\hat{\chi}^{+}\right)^{-1}E(\beta_T|Y_T)$$

$$E\big(\log(m_{1,T+s})|Y_T\big)=E(\mu_{1,T}|Y_T)+s\,E(\beta_{1,T}|Y_T) \qquad (3\text{-}6)$$

$$E\big(\log(m_{2,T+s})|Y_T\big)=\hat{\theta}\,E(\mu_{1,T}|Y_T)+E(\mu_T^{+}|Y_T)+s\left(\hat{\chi}\,E(\beta_{1,T}|Y_T)+\hat{\beta}\right)$$

In turn, the variance matrix is given by:

$$V\big(\log(m_{T+s})|Y_T\big)=Z\,V(\alpha_{T+s}|Y_T)Z'+\sum\nolimits_{\varepsilon}, \qquad (3\text{-}7)$$

where $s=1,2,\ldots$ is the forecast horizon, $E(\mu_T|Y_T)$ is the vector of the means of the smoothed levels at time T, $E(\beta_T|Y_T)$ is the vector of the means of the smoothed slopes at time T, and $V(\alpha_{T+s}|Y_T)=T^{s}V(\alpha_T|Y_T)T'^{s}+\sum_{i=0}^{s-1}T^{i}.Q.T'^{i}$ is the conditional covariance matrix of the state vector at time $T+s$, given the information available up to time T,

$$\alpha_t=\begin{bmatrix}\mu_t\\ \beta_t\end{bmatrix}_{4N\times1},\; Z=\big[(\theta^{+})^{-1}\;\;0\big]_{2N\times4N},\; T=\begin{bmatrix}I_{2N} & \theta^{+}(\chi^{+})^{-1}\\ 0 & I_{2N}\end{bmatrix}_{4N\times4N}\;\text{and}$$

$$Q=\begin{bmatrix}\Sigma_{\eta} & 0\\ 0 & \Sigma_{\varsigma}\end{bmatrix}_{4N\times4N}.$$

Hence, the shape of the prediction function for the logarithms of the mortality rates is a straight line for all age groups x. The conditional variance and

consequently the confidence interval increase as the forecast horizon grows. Under Gaussianity, we have the following distribution for the log of the future mortality rates:

$$\log(m_{a,x,T+s}) \mid Y_T \sim N\big(E(\log(m_{a,x,T+s}) \mid Y_T), V(\log(m_{a,x,T+s}) \mid Y_T)\big), \tag{3-8}$$

where $a = 1$ ou 2, $x = 1, \ldots, N$, $s = 1, 2, \ldots$, a is the population, x is the age group and s is the forecast horizon.

From equation (3-8), it follows that the predicted central mortality rates have lognormal distribution, given by this expression:

$$m_{a,x,T+s} \mid Y_T \sim \log normal\big(E(\log(m_{a,x,T+s}) \mid Y_T), V(\log(m_{a,x,T+s}) \mid Y_T)\big) \tag{3-9}$$

Therefore, the prediction function of the central mortality rates is given by:

$$E(m_{a,x,T+s} \mid Y_T) = \exp\big(E(\log(m_{a,x,T+s}) \mid Y_T) + 0.5\, V(\log(m_{a,x,T+s}) \mid Y_T)\big),$$

with variance given by:

$$V(m_{a,x,T+s} \mid Y_T) = \exp\big[2E(\log(m_{a,x,T+s}) \mid Y_T) + V(\log(m_{a,x,T+s}) \mid Y_T)\big]\big[\exp(V(\log(m_{a,x,T+s}) \mid Y_T)) - 1\big]$$

$$\tag{3-10}$$

The model predicts a natural damping of the reduction of the expected mortality rates as s grows, because $E(\log(m_{a,x,T+s}) \mid Y_n) < 0$ and $V(\log(m_{a,x,T+s}) \mid Y_n) > 0$ and the values of the variances of the logarithms of the future rates increase with time. The distribution of the factor of longevity gain s steps ahead beginning at time T is also lognormal and is given by:

$$G_{a,x,T+s} = \frac{m_{a,x,T+s}}{m_{a,x,T}} \tag{3-11}$$

Then: $G_{a,x,T+s} \sim \log normal\big(E(\log(m_{a,x,T+s}) \mid Y_T) - \log(m_{a,x,T}), V(\log(m_{a,x,T+s}) \mid Y_T)\big)$

where $m_{a,x,T}$ is the central mortality rate observed for age x of population a at time T, and $G_{a,x,T+s}$ is the factor of longevity gain for age x of population a between time T and $T+s$, given Y_T. Therefore, the function to forecast the factor of longevity gain is given by:

$$E(G_{a,x,T+s} \mid Y_T) = \exp\big(E(\log(m_{a,x,T+s}) \mid Y_T) - \log(m_{a,x,T}) + 0.5\, V(\log(m_{a,x,T+s}) \mid Y_T)\big),$$

with variance given by:

$$V(G_{a,x,T+s} \mid Y_T) = \exp\big[2E(\log(m_{a,x,T+s}) \mid Y_T) + V(\log(m_{a,x,T+s}) \mid Y_T)\big]\big[\exp(V(\log(m_{a,x,T+s}) \mid Y_T)) - 1\big]$$

$$\tag{3-12}$$

3.3

Application to the Brazilian population

In this section, the SUTSE model is applied to forecast the Brazilian longevity gains, for both genders. The Brazilian mortality rates come from the census bureau, the Brazilian Institute of Geography and Statistics[12] (IBGE), but there are only 13 years of data available (from 1998 to 2010). Therefore, for the purpose of our approach we must use the information of a related population for which longer mortality time series exist.

Due to a lack of historical information about Brazilian mortality rates, Silva (2010) identified the country which is the most similar to Brazil concerning relevant socioeconomic variables to predict the evolution of the mortality rates applying the matching technique (propensity score). The author applied this technique on 21 Organization for Economic Co-operation and Development (OECD) sample countries and, from the results, Portugal was chosen as the basis for projections of Brazilian mortality and acquisition of factors of improvement.

Taking into account those results, Portugal as a related population was tested. The Portuguese mortality rates were obtained from the Human Mortality Database[13]. The time series start in 1940 and continue to 2009. To test our approach we also selected another population from the Human Mortality Database. In this database, there are 37 countries, 3 of which are from the American continent (Chile, Canada and US). Among these, the country most similar to Brazil is the US, considering that their populations are the largest of the continent, formed by transplanted populations (immigrants and black slaves) (Ruiz, 2005), presenting the same ethnic groups. Furthermore, both countries have high degrees of urbanization and are large countries in area. Thus, we also suppose the US as a related population to test the goodness of fit of the proposed model. The US mortality rates time series start in 1933 and continue to 2009.

In the modeling we used the Brazilian data up to 2009 (12 years of observations) because that year coincides with the last year of information for the chosen countries. We also used the Brazilian mortality rate for 2010 to test the

---

[12] Available at www.ibge.gov.br. The data were downloaded on July 26, 2012.
[13] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on July 16, 2012.

model's out-of-sample prediction. We implemented the model with the STAMP 8.3 program. Since there are insufficient observations for the mortality rate time series, we had to reduce the number of parameters to be estimated in order to ensure the likelihood convergence. To do this, we decided to work with ten homogeneous age groups, namely: <1 year, 1-9 years, 10-19 years, 20-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, 70-79 years and ≥ 80 years. For each of these groups, central mortality rates were calculated using the following formula:

$$m_{a,x,t} = \frac{\sum_{i \in x} d_{a,i,t}}{\sum_{i \in x} L_{a,i,t}} \tag{3-13}$$

where $m_{a,x,t}$ is the central mortality rate for age group x of population a at time t, $d_{a,i,t}$ is the number of deaths of people with age i of population a at time t, and $L_{a,i,t}$ is the number of people exposed to risk at age i of population a at time t.

We analyze the evolution of the log mortality rates and life expectancy of the chosen populations. Between 1998 and 2009, there is a mortality improvement for the majority of the age groups of the populations. Figure 3.1 presents the observed log mortality rates of two representative age groups for both genders. We can see that the Brazilian and US log mortality rates have smoother trends than the Portuguese populations.
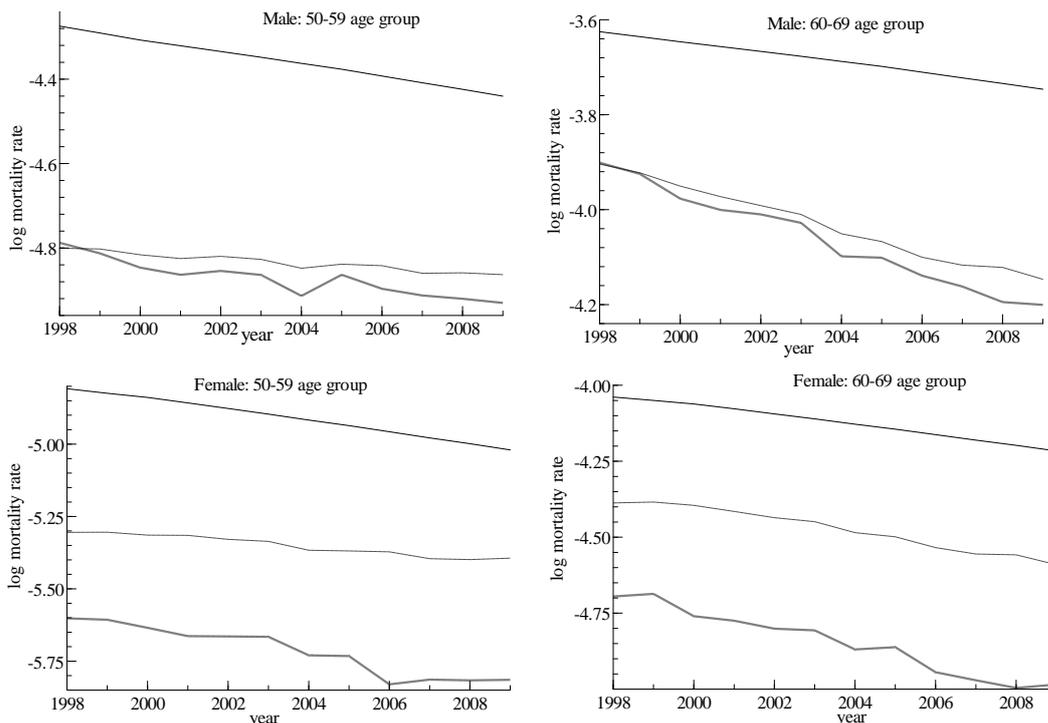
Figure 3.1. Brazilian log mortality rates, in solid lines, US log mortality rates, in dashed lines, and Portuguese log mortality rates, in dotted lines, for the 50-59 and 60-59 age groups.
Note: Two top graphs are for male populations, and two bottom graphs for female populations.

Since Brazil is a developing country, its life expectancy is lower than the other selected countries. Nonetheless, like Portugal and US, there is a trend of increase in the life expectancy for both genders, since Brazil has seen a substantial reduction in inequality over the past two decades. As an example, Figure 3.2 shows the evolution of the life expectancy for 60-year-olds. Table 3.1 presents the percentage increase in the life expectancy for three representative ages over twelve years.

Table 3.1. Percentage increase in the life expectancies between 1998 and 2009: Brazil, Portugal and US

Ages	male			female		
	Brazil	Portugal	US	Brazil	Portugal	US
40	5.19%	6.57%	5.80%	5.49%	5.63%	3.53%
50	5.59%	8.10%	7.47%	6.32%	3.62%	4.72%
60	5.44%	11.13%	10.84%	7.39%	7.39%	6.32%
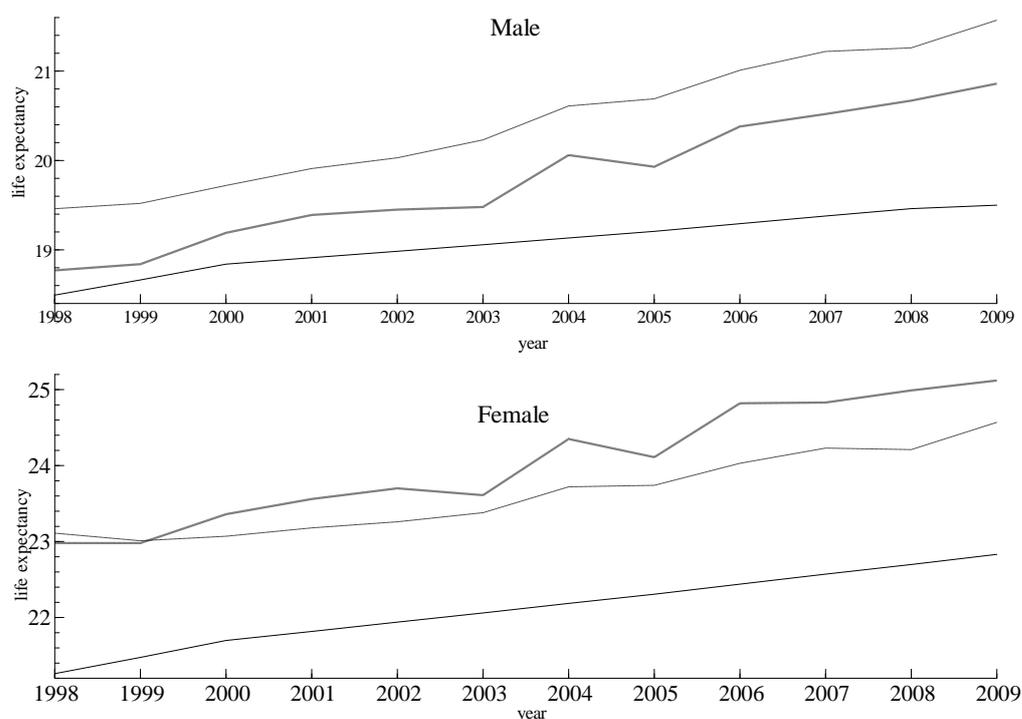
Figure 3.2. Life expectancies for 60-year-old. Brazilian life expectancy, in solid lines, US life expectancy, in dashed lines, and Portuguese life expectancy, in dotted lines.
Note: Top graph are for male populations, and bottom graph for female populations.

Except for the comparison between the Brazilian and American female populations, Table 3.1 shows that the Brazilian population attained a lower percentage increase of life expectancy. Furthermore, the Brazilian life expectancy corresponds to the life expectancy presented by related populations some years ago. So, since the proposed model can also be applied assuming there is a lag

between the data of the related population and those of the studied population, we test it considering lags of 5, 10, 15 and 20 years to try to obtain the best fit. According to maximum likelihood and AIC criteria, for both genders, 5 years lag is a better fit than others lags. Consequently, their results are compared to those of the approach without lag.

Additionally, to ensure convergence for such a parameter-rich model, we reduce the number of parameters of the covariance matrices to be estimated by adopting Cholesky decomposition for the matrices. More specifically, we use that $\Sigma = \Theta D \Theta'$ (see Koopman et al., 2007, chapter 9), where Θ is a lower-triangular matrix with unity values on the leading diagonal and D is a non-negative diagonal matrix. This entails assuming that all members of the D matrices are proportional, namely:

$$D_\varepsilon = hD, \quad D_{\eta,1} = w_1 D^*, \quad D_{\varsigma,1} = w_2 D^*, \tag{3-14}$$

where

$$D = diag(d,\ldots,d) \in \Re^{2N \times 2N}$$

$$D^* = diag(d,\ldots,d) \in \Re^{N \times N}$$

and h, w_1 and w_2 are a non-negative scalars.

In addition, in the estimation process, we opt to set $h = 1$, $0 < w_1 \le 1$ and $0 < w_2 \le 1$. Optimal values for w_1 and w_2 are found through a grid search in $(0,1]$, choosing the pair that maximizes the likelihood. The results are presented in Table 3.2.

Table 3.2. Values of w_1 and w_2 utilized in the Kalman filter

Related population	w_1	w_2
Portugal – male – without lag	1	0.05
Portugal – male – lag of 5 years	1	0.05
USA – male – without lag	1	1
USA – male – lag of 5 years	1	1
Portugal – female – without lag	0.75	0.05
Portugal – female – lag of 5 years	1	0.10
USA – female – without lag	1	0.50
USA – female – lag of 5 years	1	0.50

We use MAPE (mean absolute percentage error) criteria to select between the related population (US and Portugal) the one that best fits the Brazilian population. From Table 3.3 it can be seen that, analyzing in sample and out of sample results, for the male population, the best model is that in which the related population is the American male population with a 5-year lag. For the female populations depicted in Table 3.3, a good compromise is the model that considers the American female population without lag as the related population. For comparison with our results, it is important to note that the Lee-Carter model fitted to the same data gives the following out of sample MAPE: 0.70% for the Brazilian male population and 1.14% for the female population.

Table 3.3. MAPE for SUTSE models applied to different related populations by male and female.

Related Population	Male		Female	
	in sample	out of sample	in sample	out of sample
Portugal - without lag	0.20%	0.59%	0.17%	0.78%
Portugal - lag of 5 years	0.18%	0.57%	0.14%	0.80%
US - without lag	0.18%	0.66%	**0.15%**	**0.58%**
US – lag of 5 years	**0.19%**	**0.47%**	0.16%	0.74%

Note: In sample: between 1998 and 2009, and out of sample: 2010.

For the best-fitted models, diagnostic checking is performed using the standardized innovations associated with the models, which are tested for normality, homoscedasticity and serial uncorrelatedness. For this purpose, we use the tests of Bowman-Shenton, Box-Ljung and the heteroscedasticity test - H(h) (Durbin and Koopman, 2012), respectively. Table 3.4 presents the results for those tests, with the overall message that the proposed model is well specified for both mortality series, male and female.

Table 3.4. Diagnostics of the standardized innovations of the SUTSE models applied to Brazilian mortality rates.

Age group (years)	Brazilian Male			Brazilian Female		
	Normal.	Heterosced.	Autocor.	Normal.	Heterosced.	Autocor.
<1	1.2285 (0.541)	0.0349 (0.989)	7.8674 (0.164)	0.47614 (0.788)	0.0333 (0.990)	4.5093 (0.479)
1-9	0.8290 (0.661)	0.8408 (0.555)	13.961 (0.896)	0.4077 (0.816)	0.4313 (0.746)	9.8010 (0.081)
10-19	2.8421 (0.241)	0.0470 (0.984)	6.6232 (0.250)	2.1555 (0.340)	0.5014 (0.708)	5.319 (0.378)
20-29	2.9846 (0.225)	0.0291 (0.992)	17.3330 (0.004)	0.7912 (0.673)	3.7836 (0.152)	5.5493 (0.353)
30-39	0.28676 (0.866)	0.0375 (0.988)	11.1490 (0.049)	2.1794 (0.336)	4.2847 (0.132)	11.5030 (0.042)
40-49	2.5065 (0.286)	0.0350 (0.990)	13.9472 (0.016)	4.3799 (0.112)	1.2531 (0.429)	3.4536 (0.630)
50-59	2.8003 (0.247)	0.8408 (0.555)	12.1720 (0.033)	2.2346 (0.327)	9.6120 (0.048)	7.0747 (0.215)
60-69	7.1530 (0.028)	80.928 (0.002)	13.988 (0.016)	4.0033 (0.135)	5.6424 (0.095)	3.2173 (0.667)
70-79	1.1631 (0.559)	0.2222 (0.876)	7.0427 (0.218)	7.5993 (0.022)	0.3233 (0.956)	3.4725 (0.628)
≥80	6.4945 (0.039)	0.6469 (0.635)	3.6412 (0.602)	11.2720 (0.004)	0.5503 (0.682)	1.9880 (0.6851)

Note: Period of analysis: 1998 to 2009.

3.3.1
Forecasting longevity gains

We forecasted longevity gains 50 steps ahead for Brazilian male and female populations, from 2010 to 2059, using both our proposed model and the Lee-Carter method. Our SUTSE model produces, for both genders, lower predicted mortality improvements when compared to those obtained from the Lee-

Carter method. This, in turn, means that the projected mortality rates of the SUTSE model are higher than those according to the Lee-Carter method. As examples, Figures 3.3 and 3.4 show the forecast of the factors of longevity gain for both Brazilian populations in the 50-59 and 60-69 age groups.
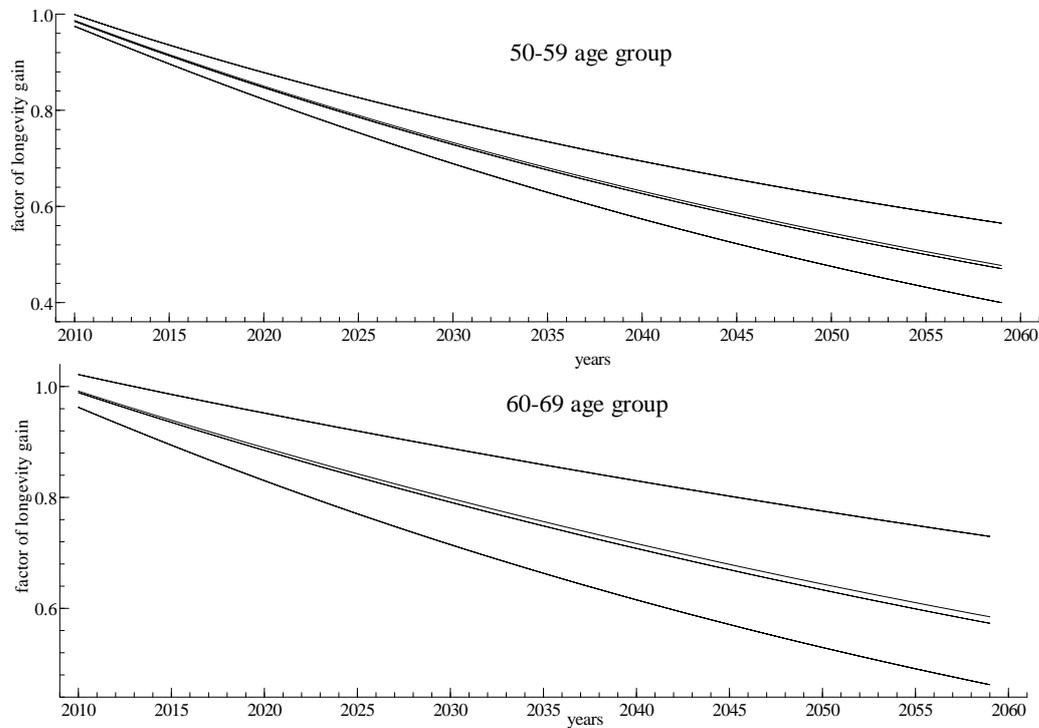
Figure 3.3. Forecast of the factors of longevity gain for the Brazilian male population in the 50-59 and 60-69 age groups.
Note: SUTSE model: solid line represents the expected factors, and short dashed lines are upper and lower limits of the 95% confidence interval. Lee-Carter method: long dashed line is predicted factors.
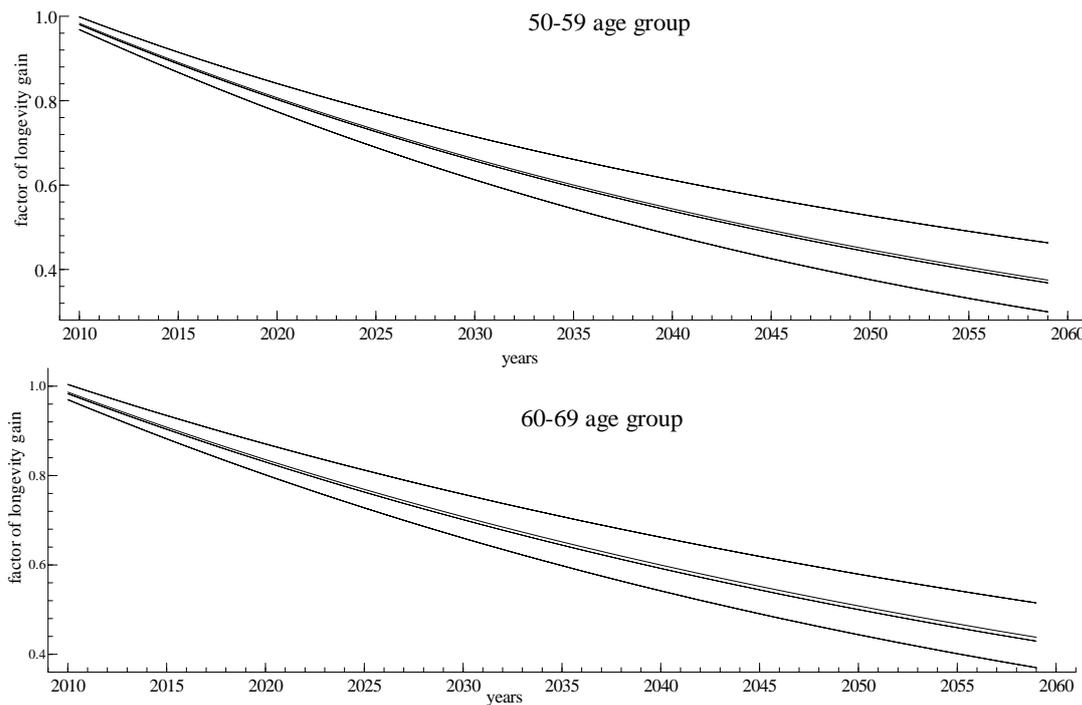
Figure 3.4. Forecast of the factors of longevity gain for the Brazilian female population in the 50-59 and 60-69 age groups.
Note: SUTSE model: solid line represents the expected factors, and short dashed lines are upper and lower limits of the 95% confidence interval. Lee-Carter method:long dashed line is the predicted factors.

We now explain the difference in magnitude of the predicted factor of longevity gain produced by SUTSE and Lee-Carter. The latter model is based on a fixed drift that is projected over time, and in our particular application, it has been estimated using only 12 years of data. In contrast, the SUTSE model is able to rely on trend estimates using data from a related population. Besides that, our SUTSE model also incorporates a natural damping effect in calculating these predictions, due to the anti-log transformation, thereby performing better. For example, for the 80 and over male age group, the SUTSE model projects an average reduction of 13% for the mortality rates over 50 years while this figure is 25% in the Lee-Carter model (both cases using 2009 as the base year).

We note that our model predicted higher mortality improvements for the Brazilian female population than for the male population. This result implies that in Brazil, differences of life expectancies between genders will increase over the years. Appendix 7.1 (Tables 7.1 and 7.2) displays predicted values of factors of longevity gain obtained by the SUTSE model for all male and female age groups.

A quick look at these figures shows, as expected, that predicted mortality improvements are lower for older age groups.

3.4
Our model in practice: its use by insurers and pension funds

Insurers and pension funds must adequately evaluate the future mortality rates in order to calculate correctly their provisions and capital based on underwriting risk. To model longevity risk, one needs a time series of mortality rates, even if it has short length. Nevertheless, some companies do not have such historical data. In practice, companies can obtain mortality rates of the insured group in time T through a graduation model, while others may associate their present mortality experience to a known mortality table.

Using the SUTSE model, insurers and pension funds can forecast the $m_{x,T+s}$ distributions for a group with known $m_{x,T}$ at time T using the longevity gain distribution of a determined population. Companies must assume that policyholders and beneficiaries have the same longevity gain distribution as a benchmark population, despite having different mortality rates at time T. This allows obtaining a proxy for the s steps ahead prediction of the central mortality rate of the covered group by the following expression:

$$m_{x,T+s} \approx m_{x,T} \times G_{a,x,T+s} \tag{3-15}$$

where, from equations (3-9) and (3-11), it can be easily seen that these distributions are also lognormal. We now obtain the mean and variance for such variables conditional on observations up to time T:

$$E(m_{x,T+s} \mid Y_T) = m_{x,T} \, E(G_{a,x,T+s} \mid Y_T)$$

$$V(m_{x,T+s} \mid Y_T) = m_{x,T}^{\,2} \, V(G_{a,x,T+s} \mid Y_T)$$

where $G_{a,x,T+s}$ is the longevity gain for age x of population a between time T and $T+s$ (see equation (3-11)) and $m_{x,T}$ is the central mortality rate for the covered group x at time T.

By means of equation (3-15), we can obtain the distribution of the future central mortality rate of policyholders and beneficiaries of Brazilian annuity plans. We assume that the mortality rates of a Brazilian policyholder and

beneficiaries on the date of valuation ($m_{x,T}$ in equation (3-15)) correspond to those found in the BR-EMS 2010 mortality tables for survival coverage, for males and females (Oliveira et al., 2012)[14]. Finally, we also assume that the longevity gain distribution of the insured group is equal to that of the Brazilian population (see section 3.3).

Figures 3.5 and 3.6 depict, for males and females, the central mortality rates forecasted by the SUTSE model for 50- and 60-year-olds.
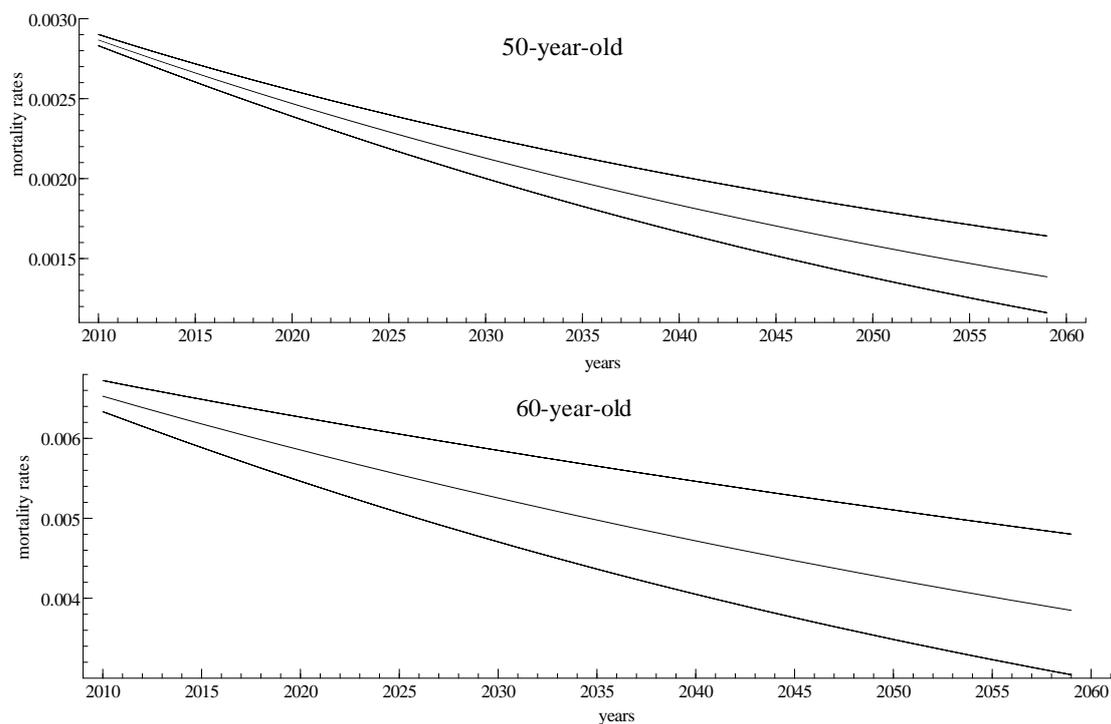
Figure 3.5. Forecasted central mortality rates for 50- and 60-year-old men.
Note: Solid line represents expected factors mortality rates, and dashed lines are upper and lower limits of the 95% confidence interval.

---

[14] The BR-EMS 2010 table was built using Brazilian insurance market experience.
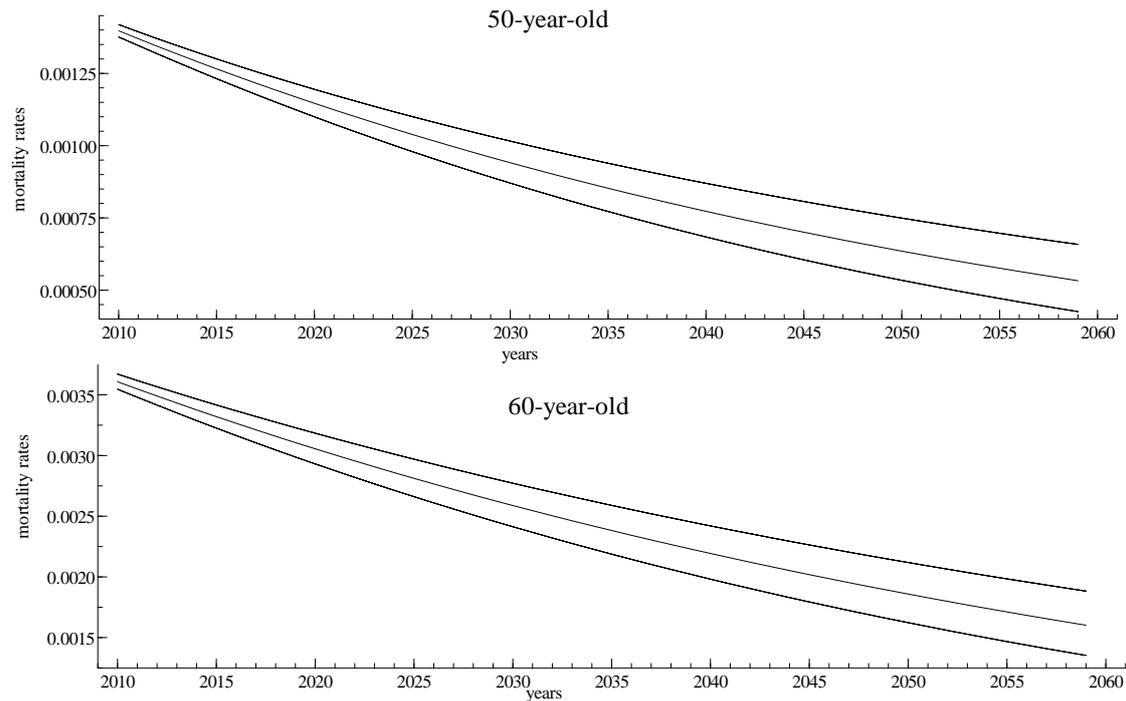
Figure 3.6. Forecasted central mortality rates for 50- and 60-year-old women.
Note: Solid line represents expected factors mortality rates, and dashed lines are upper and lower limits of the 95% confidence interval.

The importance of having a proper estimate for the longevity risk can be concretely illustrated by estimating the complete life expectancy with and without mortality improvements. The expectation of future lifetime for a person with age x at time T can be expressed (Bowers et al., 1997) by:

$$e_{x,T}^{o} = E\big(T(x)\,|\,Y_T\big) = \int_{0}^{\infty} {}_t p_x \, dt \tag{3-16}$$

where $e_{x,T}^{o}$ is complete life expectancy of age x at T, $T(x)$ is the random variable future lifetime of age x at T, and ${}_t p_x$ is the probability that an age x person in T will attain age $x + t$, given the information up to T.

Using the forecasted longevity gains and executing a discretization of $e_{x,T}^{o}$, life expectancies are calculated as follows:

$$e_{x,T}^{o} = \sum_{t=1}^{\infty} {}_t p_x + 0.5 \tag{3-17}$$

where ${}_t p_x = \prod_{s=1}^{t} p_{x+s-1,T+s}$, $p_{x,T+s} = 1 - q_{x,T+s}$ and $q_{x,T+s} \approx \dfrac{E\big(m_{x,T+s}\,|\,Y_T\big)}{1 + \dfrac{E\big(m_{x,T+s}\,|\,Y_T\big)}{2}}.$

Table 3.5 shows estimated life expectancies at time *T* for some selected ages, both with the longevity gain assumption and without it (from BR-EMS 2010). To compare the longevity gain assumptions, it is illustrated the life expectancies calculated using the mortality improvements forecasted by the SUTSE model and the Lee-Carter method.

Table 3.5. Estimated complete life expectancies at valuation time for ages 50 to 90.

Ages	Male			Female		
	(1)	(2)	(3)	(1)	(2)	(3)
50	34.23	35.64	36.15	38.38	40.56	41.10
60	25.48	26.36	26.68	29.11	30.57	30.91
70	17.59	18.08	18.25	20.52	21.36	21.55
80	10.99	11.24	11.31	12.81	13.22	13.31
90	6.05	6.16	6.18	6.61	6.68	6.70

Note: (1) do not consider longevity gains, (2) consider longevity gains forecasted by the SUTSE model and (3) consider longevity gains forecasted by the Lee-Carter model.

From Table 3.5, one can easily see that the SUTSE model produces higher life expectancies than those obtained from straight use of the BR-EMS 2010 table. From it one can note that for a 60-year-old, the differences between life expectancies is around one year (for both genders), which means one more year of cash outflows for an insurer. If we compare genders, female life expectancy is longer than male, a well-documented fact in actuarial literature. Moreover, since the forecasted mortality improvements are larger using Lee-Carter method, we can see that the life expectancies derived from it are larger than those found by the SUTSE model. This can be understood by the fact that Lee-Carter's projections are obtained from a shorter mortality rates time series, that from the Brazilian population (only 12 annual observations). These projections, by construction, put too much weight in the most recent observations, and also are not very reliable, given the small number of observations used for its estimation. On the other hand, the proposed SUTSE model produces mortality rates for the Brazilian population by combining the original Brazilian series with a longer time series of mortality rates from a related population. Therefore, from a statistical point of view, the mortality projections are based on more observations. Consequently, the life

expectancies obtained by our proposed model are more reliable than those of the Lee-Carter method.

Since mortality rates are stochastic in nature and we assume a mortality improvement, the temporal evolution of life expectancies can be captured by the following expression:

$$e^{o}_{x,T+k} = E\left(T(x)^{(T+k)} \mid Y_T\right) = \int_{0}^{\infty} {}_t p_x^{(T+k)} dt \qquad (3\text{-}18)$$

where $e^{o}_{x,T+k}$ is a complete life expectancy of age x at time $T+k$, $T(x)^{(T+k)}$ is a random variable that expresses future lifetime of age x at time $T+k$ and ${}_t p_x^{(T+k)}$ is the probability that a person of age x at time $T+k$ will attain age $x+t$, given the information up to time T.

$$e^{o}_{x,T+k} = \sum_{t=1}^{\infty} {}_t p_x^{(T+k)} + 0.5 \qquad (3\text{-}19)$$

where $\qquad {}_t p_x^{(T+k)} = \prod_{s=1}^{t} p_{x+s-1,T+k+s} \qquad , \qquad p_{x,T+k+s} = 1 - q_{x,T+k+s} \qquad$ and

$$q_{x,T+k+s} \approx \frac{E\left(m_{x,T+k+s} \mid Y_T\right)}{1 + \dfrac{E\left(m_{x,T+k+s} \mid Y_T\right)}{2}}.$$

As an example, in Figure 3.7 we can look at the 30-years-ahead life expectancy prediction for male and female 60-year-olds. A careful look at such figure shows life expectancy increments of 1.19 years for males and 1.68 for females. The higher increment of the life expectancy for the female population are also found in the Brazilian original data as one can see in Table 3.1.
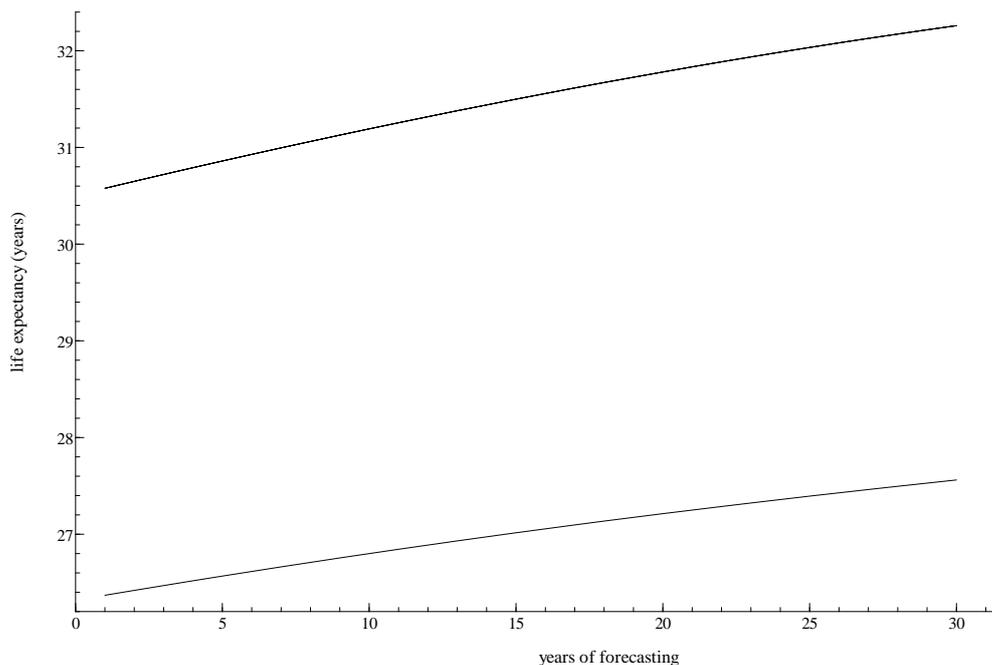
Figure 3.7. Temporal evolution of life expectancy for male and female 60-year-olds. Solid line for males and dashed line for females.

3.5

Valuation of technical provision and capital based on underwriting risk

As stated by IAIS (2012), the technical provision for solvency purposes is a significant component in the valuation of solvency and it should be measured on a consistent basis. Capital for solvency purposes is another important aspect of solvency valuation. When solvency capital is added to provision, the resulting value should guarantee, within a certain confidence level required by the insurance supervisor, that insurance commitments will be paid. In this section, the best estimate of the liabilities is defined as the present value of expected cash flows of insurance contract commitments. Moreover, the capital based on underwriting risk is a highlighted value in the equity that has the role of covering the risks from losses greater than those expected by an insurer. In order to ensure the solvency of a company, the capital should be measured using the tail of the distribution of losses through the extreme value approach. The value of capital is a measure of risk associated with the right tail of that distribution less the value of provision for solvency purposes.

In the calculation of underwriting risk, disregarding the interest rate risk (not measured in this paper), consideration must be given to the idiosyncratic risk[15], which is a function of the size of an insured group. The larger the group exposed to risk, the smaller the idiosyncratic risk will be, considering the law of large numbers. Because of this, small companies tend to have underwriting capital more weighty than that of big companies.

Through Monte Carlo simulation, using our longevity gain model (the SUTSE model) and assuming that the BR-EMS 2010 mortality table correctly reflects mortality rates of Brazilian beneficiaries, we estimate the present values of future cash flows of hypothetical beneficiary groups that are receiving lifetime annuities. To highlight the effect of the longevity risk, the differences between future real interest rates and interest rates fixed by insurers are not predicted. Thus, in the simulation, these interest rate values are equated.

The present values of future cash flows are simulated for groups of different sizes to emphasize the value of idiosyncratic risk. The algorithm used in stochastic simulation for each beneficiary k is synthesized below:

- set age and income of beneficiary *k* at the date of valuation;

- simulate *n* times for each beneficiary *k*, as follows:

    1. generate a random value for the longevity gain, by applying equation (3-11);

    2. apply equation (3-15);

    3. compute the death probability based on this approximation:

$$q_{x,t} \approx \frac{m_{x,t}}{1 + \dfrac{m_{x,t}}{2}}$$

    4. generate a uniform random variable (0, 1);

       . if this value is greater than the death probability, then add the present value of the annual income at time *t* with the amount of expenses and go to time *t+1*;

       . otherwise, end simulation *i* and go to simulation *i+1* in the total of *n* simulations;

- go to beneficiary k+1.

---

[15] International Actuarial Association classifies this risk as level risk (see in "A global framework for insurance solvency assessment", 2004 - www.actuarios.org)

From this simulation one generates n present values of cash flows for each beneficiary k, where the sum of these k values produces a distribution of size n of future expenses of the insurance company. The expected value of this distribution corresponds to the value of the best estimate and the tail of the distribution is used to measure the value of the underwriting capital.

We emphasize that if n tends to infinity, the value of the provision tends to the amount calculated by the mean of annuities, which is based on the average of predicted death probabilities, as below:

$$a_{x,T} = \sum_{t=1}^{\infty} {}_t p_x \times (i + i_t)^{-t}, \tag{3-20}$$

where $a_{x,T}$ is the annuity with income payment at the end of the year to age x at period T, i_t is the real interest rate corresponding to time t, ${}_t p_x = \prod_{s=1}^{t} p_{x+s-1,T+s}$,

$$p_{x,T+s} = 1 - q_{x,T+s} \quad \text{and} \quad q_{x,T+s} \approx \frac{E\big(m_{x,T+s} \mid Y_T\big)}{1 + \dfrac{E\big(m_{x,T+s} \mid Y_T\big)}{2}}.$$

To best illustrate the idiosyncratic risk estimate, we choose to use the value at risk (VaR) as a measure of risk. Tsay (2005) defines VaR as the maximum loss of a financial position during a period, given probability α. Thus, VaR can be represented as follows:

$$P\big(Z_T \geq VaR \mid Y_T\big) = \alpha \tag{3-21}$$

where Z_T is the present value of cash flows at T.

To show the computation of the best estimate of the provision and the underwriting capital, simulations are performed for four hypothetical groups of 50, 100, 500 and 1,000 beneficiaries. All beneficiaries are 60-year-old men with an annuity of \$36,000.00.

We measured the ratio between the best estimate of the liability assuming the longevity gains obtained by the SUTSE model and the Lee-Carter method. These ratios for different annual interest rates are shown in Table 3.6. Since the best estimate is calculated as the expected value of the distribution, the results depicted in Table 3.6 are equal for all beneficiary groups. As it is expected, the best estimates are larger using Lee-Carter. It can be seen that the observed gradient in the ratios are accompanied by an increase in interest rates. For

instance, assuming a policy with 4% annual interest rate and presuming this is equal to the real interest rate, the insurer would account for 2.6% less if it forecasts the mortality rates using our approach.

Table 3.6. Ratio of the best estimates at valuation time.

Real interest rates (annual)	Ratio
0%	0.955
2%	0.966
4%	0.974
6%	0.980

Note: Ratio between the best estimates assuming the longevity gains obtained by the SUTSE model and the Lee-Carter model

In Table 3.7, using our SUTSE model as longevity gain assumption, the values of capital based on underwriting risk for the four beneficiary groups are calculated, with interest rates fixed at 4% per year. These values are shown in the percentage of amounts of the best estimate. Considering the results of Table 3.7, it is clear that the smaller the group is, the greater the underwriting risk will be. It appears that the reduction of the value of capital is more pronounced between 50 and 500 beneficiaries than between 500 and 1,000. This means that when the group reaches a certain size, it tends to have only longevity risk, eliminating the idiosyncratic risk.

Table 3.7. Percentage of capital based on underwriting risk in relation to the best estimate at valuation time.

Groups	# beneficiaries	Percentage of capital
1	50	7.86%
2	100	5.60%
3	500	2.60%
4	1000	1.94%

Note: Percentage of capital based on underwriting risk, given $\alpha = 2.5\%$, in relation to the best estimate of the liabilities.

For example, an insurer with 50 beneficiaries would have capital of 7.86% of the best estimate, while for the group with 1,000 beneficiaries the capital is less

than one fourth of that value (1.94%). Therefore, the smaller insurer has a competitive disadvantage because a higher capital requirement reduces the return on investment. If the smaller insurer accounts for the same percentage of provision as the larger insurer, the shareholders and beneficiaries of the former would not be protected against bankruptcy. Under this hypothesis, the probability that the present value of expenses of this insurer would be higher than capital plus the best estimate is around 32%. Thus, it is clear that management of underwriting risk should consider idiosyncratic risk to maintain the company's solvency.

3.6
Conclusion

In this paper we introduced a multivariate structural SUTSE model to forecast longevity gains for populations with a short time series of observed mortality rates, using a related population for which there exists long time series of mortality rates. Such a class of models enables one to model jointly the different mortality rate time series through a common trends framework. This guarantees interdependence between the mortality rate time series of both populations, assuming the two populations' mortality rates are affected by common factors.

The proposed model was applied to Brazilian male and female mortality rates, considering the Portuguese and the American populations as related populations, one at a time. Using MAPE as a measure of forecasting performance, we chose the five-year lagged mortality rates time series for the American population as the companion time series to help estimate the trend for the Brazilian male mortality rates. For the Brazilian female population, the related time series were the contemporaneous mortality rates of the American female population. As can it be seen in section 3.3.1, for the Brazilian population, the predicted mortality rates from our model are higher than those found by the traditional Lee-Carter method. As a result our proposed model will produce shorter life expectancies and lower values for both the liabilities and capital based on underwriting risk, when compared to those obtained from Lee-Carter method. Given the better statistical properties of the estimates obtained from the SUTSE

model when compared to those obtained from Lee-Carter method for a short time series of observed mortality rates, the former should be preferred.

From the results of this practical exercise, we can derive a number of findings on the behavior of Brazilian policyholders also shared by the populations of other countries. First, as is observed in many countries, in Brazil women have longer life expectancies than men. In addition, the increment in life expectancy is higher for females than for males. This implies that the gender gap in life expectancy can increase with time. Furthermore, the model produces lower mortality improvements for older age groups.

As a final exercise, using Monte Carlo simulation, we obtained the best estimate of the liabilities for four hypothetical beneficiary groups, considering the longevity gain model. Additionally, the capital levels based on underwriting risk were obtained for these beneficiary groups in relation to the values of the best estimates. Having in mind the idiosyncratic risk, we can conclude that the larger the beneficiary group is, the lower the capital value in relation to the value of best estimate will be. Therefore, in order to ensure the solvency of an undertaking, one also has to consider the idiosyncratic risk in the process of capital valuation.

4

Paper 3: Forecasting Surrender Rates using Elliptical Copulas and Financial Variables

A multistage stochastic model to forecast surrender rates for life insurance and pension plans is proposed. Surrender rates are forecasted by means of Monte Carlo simulation after a sequence of Generalized Linear Model (GLM), ARMA-GARCH, and multivariate copula fitting is executed. To produce a conditional forecast on a stock market index, in our application we used the residuals of an ARMA-GARCH model fitted to the Brazilian stock market index returns, which generates one of the marginal distributions used in the dependence modeling through copulas. This strategy is adopted to explain the high and uncommon surrender rates observed during the recent economic crisis. After applying known simulation methods for multivariate elliptical copulas, we proceeded backwards to obtain the forecasted distributions of surrender rates by application, in the sequel, of ARMA-GARCH and GLM models. Additionally, our approach produced an algorithm able to simulate multivariate elliptical copulas conditioned on a marginal distribution. Using this algorithm, surrender rates can be simulated conditioned on stock index residuals, which allows insurers and pension funds to simulate future surrender rates assuming a financial stress scenario with no need to predict the stock market index.

4.1
Introduction

In this paper a multi-stage stochastic model is proposed to forecast surrender rates from life insurance and pension plans using multivariate elliptical copulas and financial variables. This approach is quite relevant nowadays, since adequate surrender rates play an essential role in the realistic valuation of life insurance and pension plan liabilities or actuarial risks, which are now required by solvency and accounting standards. For instance, the forecast of surrender rates can be used in valuation of embedded options in order to estimate the

policyholder behavior with respect to exercise contractual options. Solvency II presents some rules about policyholder behavior, such as that it should be appropriately based on statistical and empirical evidence and that one should not assume it to be independent of financial markets. De Giovanni (2010) models a surrender option embedded using a rational expectation approach allowing rational and irrational surrender rates. Kling et al. (2014) analyze the impact of policyholder behavior on pricing, hedging and hedge efficiency of variable annuities, presenting five different assumptions regarding surrender rates. Furthermore, companies must forecast surrender rates in order to manage risks that arise due to mismatches between assets and liabilities (ALM).

There are few published work dealing with the modeling of surrender and lapse rates. One of the most cited is Kim (2005), where a surrender and lapse rate model is proposed making using of economic variables as explanatory variables, such as market interest, credit, unemployment and economy growth rates. Since the surrender and lapse rates had been badly affected by the financial crisis from December 1997 to December 1998, the author decided to use a dummy variable to capture such unexpected behavior. Cox and Lin (2006) used a similar approach, but they modeled annuity lapse rates via Tobit models.

Tsai et al. (2002) applied a cointegrated vector autoregression approach to estimate an empirical relation between the lapse rates and interest rates, deriving a long-term relation between these variables. Kuo et al. (2003) also used a cointegration approach to study lapse rates, but considered unemployment rates in addition to lapse rates and interest rates as variables in the time-series vector.

Considering that surrender rates were also affected during the economic crises, Loisel and Milhaud (2011) proposed a stochastic model of the surrender rates to be applied in place of the classical S-shaped deterministic curve. Additionally, some papers in the literature have been published to evaluate surrender options. For instance, Tsai (2012) calculated the impact of surrender options to reserve durations using the empirical VAR model proposed by Tsai et al. (2002). Knoller et al. (2015) focused on policyholder surrender behavior in the variable annuity products. One of the conclusions was that surrender depends on the value of the embedded option and hence indirectly on the development of financial markets.

As pointed out by Loisel and Milhaud (2011), surrender risk is extremely complex to model since it depends, simultaneously, on factors from different sources, such as policyholders' characteristics, personal desires and needs, contract features and time elapsed and also the economic and financial context. In order to capture some of those factors in the process of forecasting surrender rates in plans with annuity payments, we propose a model to simulate future rates using elliptical copulas and financial variables.

Our database consists of monthly male and female surrender rates of a large Brazilian life insurer from January 2006 to December 2011, where the last twelve months were reserved for model validation. At the outset, cluster analysis is used to group surrender rates in fewer age groups, for both males and females. In the second stage, the Generalized Linear Model (GLM) approach as used by Kim (2005) is applied to our dataset. In the third stage of our modeling process, to study the dependence among the surrender rate time series through a copula framework, the GLM residual time series of each age/gender group are modeled by an ARMA-GARCH process to make them independent and identically distributed. Finally, to capture dependence among these i.i.d. time series, elliptical copulas are fitted. It should be noted that in Brazil during the economic crisis of 2008, highly uncommon surrender rates were observed in the insurance industry. Therefore, to really explain the dependence structure we decided to incorporate in our copula modeling the residuals of returns of the Brazilian stock market index (Ibovespa) as one of the marginal distributions. The proposed approach is then evaluated through an out-of-sample back test.

Our approach can also be used to simulate future surrender rates given a specific financial scenario, which can be chosen in a stress test context to analyze policyholder behavior when faced by a financial crisis. To simulate this scenario, we present a specific algorithm for simulation of elliptical copulas conditioned on a marginal distribution, which is the Ibovespa residual distribution in our application.

The rest of the paper is organized as follows. In the section 4.2, the proposed approach is presented, while in the section 4.3 we apply the model to the data from a Brazilian life insurer. In the section 4.4, we illustrate a simulation of surrender rates assuming a financial crisis scenario. The section 4.5 concludes.

4.2
The model

The proposed model is a stochastic model to simulate future surrender rates by means of Monte Carlo simulation. As will be seen, the construction of the full procedure rests on judicious applications of cluster analysis, GLM, ARMA-GARCH processes and copulas.

We chose to work with multivariate Gaussian copulas (MGC) and multivariate Student's *t* copula (MTC). The advantage of these copulas is that one can specify different levels of correlation between the marginal distributions, since they are characterized by a range of parameters and can be fitted flexibly to the data. A brief review of multivariate copula theories, which are used in our model, is presented in Appendix 7.2.1 for the sake of completeness.

We start our analysis by applying cluster analysis to the full set of surrender rate time series, in order to attain data reduction. The known partitioning method called k-means[16] is applied (Jain et al, 1999). As a result, we end up with *k* time series of monthly surrender rates of observations to be modeled. In the second stage, GLM is used to explain the surrender rate time series. We opted to apply such framework because it is commonly used by actuaries to model rates and our approach is an extension of Kim (2005). In this context, some explanatory variables must be used to explain the characteristics of the groups, such as age, gender, type of contract and policy age since issue. Due to the quality of information available, we decided to include only the variables age and gender.

It is a known fact in the insurance literature the effect of interest rates on surrender rates. Furthermore, short-term interest rates forecast can be directly obtained from the current term structure. These make interest rates a natural explanatory variable in the GLM framework. We could also have tried other relevant macroeconomic variables, such as unemployment and economy growth rates. Nevertheless, such a decision would make it necessary to forecast these variables ex ante to predict the surrender rates, but this could increase model risk.

---

[16] In the k-means method, each observation is classified as belonging to one of *k* age/gender groups through calculation of the values of the centroid for each group.

In order to keep the modeling strategy self-contained, we opted not to use these potential predictors.

Within the GLM framework we chose a binomial distribution for the response variable, the surrender rates. Differently from Kim (2005), we test three types of link functions: logit, probit and complementary log-log, which have the following forms:

Logit function: $\log\left(\dfrac{w_{x,t}}{1 - w_{x,t}}\right) = \beta_0 + \beta_1 X_{1,t} + \ldots + \beta_j X_{j,t}$ (4-1)

probit function: $\Phi^{-1}(w_{x,t}) = \beta_0 + \beta_1 X_{1,t} + \ldots + \beta_j X_{j,t}$

complementary log-log function: $\log(-\log(1 - w_{x,t})) = \beta_0 + \beta_1 X_{1,t} + \ldots + \beta_j X_{j,t}$

where $w_{x,t}$ is the surrender rate for age/gender group x at time t, $X_{i,t}$ is the explanatory variables i at time t and β_i is the coefficient related to $X_{i,t}$, $i = 1, 2, \ldots, j$, $t = 1, 2, \ldots, q$, q is the estimating period, $x = 1, 2, \ldots, k$ and k is the number of age/gender groups.

Goodness fit for the GLM is tested using the deviance, which is a measure of distance between the log-likelihood of the saturated model and the log-likelihood of the fitted model (De Jong and Heller, 2008). The smaller the deviance value, the better the model fit will be.

One should note that, in principle, GLM residuals should not be i.i.d., since such framework does not deal with the time series structure of the data. Thus, before modeling the dependence structure by means of copulas, the GLM residuals are modeled by ARMA-GARCH processes, so that, hopefully, i.i.d. residuals are obtained. The i.i.d. condition is tested by means of BDS test, which is a test for null hypothesis of i.i.d. time series (Brock et al., 1987)[17]. The ARMA-GARCH model applied to each time series of GLM residuals is given by:

$$w_{x,t} - \hat{w}_{x,t} = r_{x,t} = \phi_{0,x} + \sum_{i=1}^{p} \phi_{i,x} r_{x,t-i} - \sum_{i=1}^{r} \theta_{i,x} a_{x,t-i} + \sigma_{x,t} \varepsilon_{x,t}$$

(4-2)

$$\sigma_{x,t}^2 = \alpha_{0,x} + \sum_{i=1}^{m} \alpha_{i,x} a_{x,t-i}^2 + \sum_{i=1}^{s} \beta_{j,x} \sigma_{x,t-j}^2$$

---

[17] In the BDS test, the null hypothesis is that the time series is i.i.d. against the following set of alternatives: linear dependence, nonlinear dependence, non-stationarity and chaos. This test was applied using bootstrapped p-values.

where $w_{x,t}$ are the surrender rates observed for age/gender group x at time t, $\hat{w}_{x,t}$ is the surrender rates for age/gender group x at time t estimated by GLM, $r_{x,t}$ is the residual for age/gender group x at time t, $\varepsilon_{x,t}$ are a sequence of i.i.d. random variables with mean 0 and variance 1 for the age/gender group x,

$$a_{x,t} = r_{x,t} - \left(\phi_{0,x} + \sum_{i=1}^{p} \phi_{i,x} r_{x,t-i} - \sum_{i=1}^{r} \theta_{i,x} a_{x,t-i} \right) , \quad \alpha_{0,x} > 0 , \quad \alpha_{i,x} \geq 0 , \quad \beta_{j,x} \geq 0 ,$$

$$\sum_{i=1}^{\max(m,s)} \left(\alpha_{i,x} + \beta_{j,x} \right) < 1, \left| \theta_{i,x} \right| < 1, \left| \phi_{i,x} \right| < 1, t = 1,2,...,q \text{ and } x = 1,2,...,k, \text{ and } p, r, m \text{ and}$$

s are non-negative integers.

Assuming that the k time series of residuals from the fitted ARMA-GARCH processes are i.i.d., then the joint dependence of these residuals can be captured by means of copulas, whose marginals are obtained from the ARMA-GARCH processes.

Kim (2005) and Loisel and Milhaud (2011) pointed out in their works that surrender rates tend to increase during economic crises. This happened in Brazil in the aftermath of the 2008 global financial crisis, when the surrender rates were uncommonly high. In order to capture this policyholder behavior, the proposed model uses a financial variable, which we chose as the returns of the main stock market index in Brazil (Ibovespa). The monthly returns of this index are also modeled through ARMA-GARCH processes and the residuals, independent and identically distributed, are also used to construct one of the marginal distributions in the dependence modeling.

To finalize our modeling strategy, we apply both multivariate Gaussian copula and multivariate Student's t copula to the aforementioned residual series in order to capture the dependence structure among them. Then, by simulating these residuals and applying equations (4-2) and (4-1), the distributions of surrender rates s steps ahead are obtained. The modeling process is summarized in Figure 4.1.

Data: Time series of surrender rates from an insurer	**Data**: Time series of stock market index returns

Cluster

ARMA-GARCH

Surrender rates clustered in age/gender	i.i.d. residuals of the stock market index returns

GLM

k time series of surrender rate GLM residuals	

ARMA-GARCH

k time series of i.i.d. surrender rate residuals	**Dependence modeling: multivariate Gaussian copula and multivariate Student's t copula**

Figure 4.1.Modeling Process of the Residuals Dependence

The dependence of the n residual time series is modeled using multivariate elliptical copulas so that the surrender rates can be predicted. The advantage of this model is that one can simulate the surrender rates without needing to forecast the returns of the stock market index. In order to attain this, the return's residuals have to be simulated using a known univariate distribution, such as Gaussian or Student's t (for the marginal distributions), and the chosen multivariate copula (for the joint distribution).

The cumulative joint distribution of the residuals is written in terms of marginal distributions as in equation (7-2), where x_1 is the time series of the stock market index residuals and $x_2,\ldots,x_n$ are the residuals of the (*n-1*) time series of surrender rates (the $\varepsilon_x's$ in eq. (4-2)).

To estimate the parameters of the MGC and MTC, we apply the canonical maximum likelihood method (CML). Firstly, during the ARMA-GARCH processes, the marginal distributions of the i.i.d. residuals are obtained with the respective parameters. Then, the set of parameters of the multivariate copula must be obtained.

For MGC, the components of correlation matrix R are estimated by means of the algorithm for CML described in Cherubini et al. (2004, chapter 5, page 160). For MTC, the set of parameters is estimated through the three-stage KME-CML method, described in Fantazzini (2010, section 3). The estimation stages are:

- transform the ARMA-GARCH residuals into uniform variables using the empirical distribution function;

- collect all pairwise estimates of the sample Kendall's tau given in an empirical Kendall's tau matrix R^{τ}, and then construct the correlation matrix R using this relationship $\hat{R}_{i,j} = \sin\left(\frac{\pi}{2}\hat{R}_{i,j}^{\tau}\right)$; and

- estimate the degrees of freedom of the t distribution by maximizing the log-likelihood of eq. (7-8).

After estimating the parameters of the marginal distributions and of the copulas, we simulate samples from the joint distribution of the ARMA-GARCH residuals using the known algorithms for elliptical copula simulation (Cherubini et al., 2004, chapter 6). Then we proceed backwards to obtain the (*n-1*) distributions of surrender rates, for each month, first through application of eq. (4-2), and then of eq. (4-1).

Once both Gaussian and Student's t copulas are fitted and surrender rates simulated, to choose the best fitting copula we can use goodness of fit tests in the out-of-sample period (the last twelve months, from January 2011 to December 2011). For this, we use the mean absolute percentage error (MAPE), with formula given by:

$$MAPE = \frac{100\%}{k \times m} \sum_{x=1}^{k} \sum_{t=1}^{m} \left| \frac{w_{x,t} - \hat{w}_{x,t}}{w_{x,t}} \right| \tag{4-3}$$

where m is the number of out-of-sample months, k is the number of age/gender groups, $w_{x,t}$ is the observed value of the surrender rates for age/gender group x at

time t and $\hat{w}_{x,t}$ is the values of the surrender rates predicted by our model for age/gender group x at time t.

Additionally, the Kupiec test (Kupiec, 1995) can be computed to investigate the accuracy of probability prediction using such copulas. In this test, one checks whether the number of observed surrender rates out of the simulated confidence interval of 95% is consistent with this chosen confidence level. Using these two methods, one is able to pick the best elliptical copula to simulate the surrender rates, considering the dependence structure among the residuals. Figure 4.2 summarizes the full process developed for forecasting surrender rates.
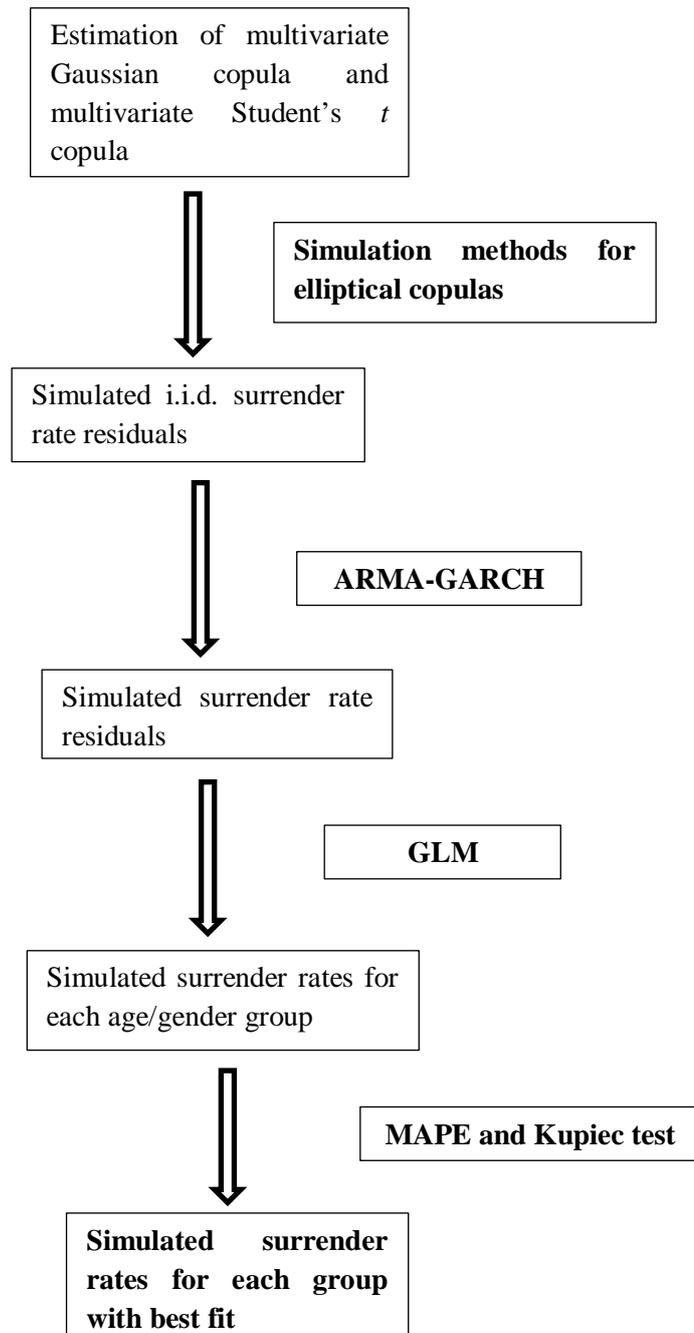
Figure 4.2. Forecasting the Surrender Rates through Simulation by Elliptical Copulas.

The proposed approach for surrender rate simulation can be used assuming a particular scenario for the financial crisis using copulas conditioned on a certain financial variable. The expression of the multivariate copula conditioned on one marginal distribution is given by eq. (7-10).

The model derived from such strategy is able to simulate the residuals of surrender rates considering a given residual of the stock market index return at a determined time, and consequently it will obtain the distributions of surrender rates given the stressful situation in the financial market using the dependence structure. This approach can be useful in financial stress scenario processes to evaluate liabilities and risks, being an excellent tool to manage the surrender options and the mismatches between assets and liabilities. In order to model it, we have to adopt a marginal distribution for the residuals of the stock market index returns and also to assume that a crisis can be characterized by the occurrence of extreme values in such time series. This can be done by choosing prespecified quantiles of the cumulative distribution function of that marginal distribution.

Thus, to simulate the surrender rate residuals, we propose a simulation method for elliptical copulas conditioned on one marginal distribution, which is an adaptation of the algorithms described in Cherubini et al. (2004, chapter 6). The proposed algorithms are described next.

For Multivariate Gaussian Copula:

- Find the Cholesky decomposition A of the correlation matrix R;

- Simulate $n\text{-}1$ independent random variables $z = (z_2, \ldots, z_n)^T$ from $N(0,1)$;

- Set $u_1 = F_1(t_1)$, where t_1 is the observed/given residual of the stock market index return;

- Set $x_1 = \Phi^{-1}(u_1)$, where Φ denotes the univariate standard normal distribution function;

- Set $z_1 = x_1 / A[1,1]$, where $A[1,1] = 1$;

- Set $z = (z_1, z_2, \ldots, z_n)^T$;

- Set $x = Az$;

- Set $u_i = \Phi(x_i)$, with $i = 2, \ldots, n$; and

- Obtain $(u_2, \ldots, u_n)^T = (F_2(t_2), \ldots, F_n(t_n))^T$, where $F_i(t_i)$ denotes the ith marginal, $i = 2, \ldots, n$.

For Multivariate Student's t Copula:

- Find the Cholesky decomposition A of the correlation matrix R;

- Simulate $n\text{-}1$ independent random variables $z = (z_2, \ldots, z_n)^T$ from $N(0,1)$;

- Simulate a random variable s from χ_υ^2 independent of z;

- Set $u_1 = F_1(t_1)$, where t_1 is the observed/given residual of the stock market index return;

- Set $x_1 = T_\upsilon^{-1}(u_1)$, where T_υ denotes the univariate Student's t distribution function;

- Set $y_1 = \dfrac{x_1}{\sqrt{\upsilon/s}}$;

- Set $z_1 = y_1 / A[1,1]$, where $A[1,1] = 1$;

- Set $z = (z_1, z_2, \ldots, z_n)^T$;

- Set $y = Az$;

- Set $x = \sqrt{\upsilon/s} \cdot y$;

- Set $u_i = T_\upsilon(x_i)$, with $i = 2, \ldots, n$; and

- Obtain $(u_2, \ldots, u_n)^T = (F_2(t_2), \ldots, F_n(t_n))^T$, where $F_i(t_i)$ denotes the ith marginal, $i = 2, \ldots, n$.

The residuals generated from this algorithm are used to obtain the conditional distribution of the surrender rates through application of eq. (4-2) and eq. (4-1). As a result of this process, one obtains the conditional distribution of surrender rates on the stock market index returns for each age/gender group. Goodness of fit in the out-of-sample period is also done by means of MAPE and Kupiec test. This full process is summarized in Figure 4.3.

```
┌──────────────────────────────────────────┐
│ Enter with values for residuals of the     │
│ stock market index returns or with the     │
│ quantiles of the cumulative distribution   │
│ function of these residuals                 │
│ (marginal distribution)                     │
└──────────────────────────────────────────┘
```

Simulation methods for elliptical copulas conditioned on one marginal distribution

Simulated i.i.d. surrender rate residuals

ARMA-GARCH

Simulated surrender rate residuals

GLM

Simulated surrender rates for each age/gender group

MAPE and Kupiec test

Simulated surrender rates for each group with best fit

Figure 4.3. Forecasting the Surrender Rates through Simulation by Conditional Elliptical Copulas.

4.3
Application

Our database consists of male and female monthly surrender rates of a large Brazilian life insurer from January 2006 to December 2011. More

specifically, the data were obtained from policyholders having plans with annuity payments, grouped by age and gender. It is important to stress that the majority of Brazilian policyholders with annuity plans belong to unit-linked plans. The data series contain monthly surrender rates for policyholders with age between 24 years and 80 years old. In order to test the out-of-sample model, we drop the last twelve months of data from the estimation period.

The surrender rates are grouped in five homogenous age groups for each gender, using cluster analysis. The resulting groups are 24-28 years, 29-53 years, 54-62 years, 63-71 years and 72-80 years. The last three age groups include policyholders in retirement age brackets.

In the GLM stage, the only macroeconomic variable used as explanatory variable is the monthly Brazilian real short-term interest rate from January 2006 to December 2010, which is taken from inflation-indexed bonds called "NTN-B notes", linked to the IPCA (consumer price index). Additionally, we include age and gender as dummy variables in the linear predictor. The model is also tested including 1-month and 2-month lagged interest rates, to investigate whether there are lagged effects of interest rates on monthly surrender rates.

The contributions of the lagged interest rates in the value of deviance using the three different link functions are shown in Table 4.1. Since the differences are not statistically significant, the lagged interest rate variables are not used in the model.

Table 4.1. Deviances using the Lagged Interest Rates as Explanatory Variables for Different Link Functions.

model	Link function		
	Logit	Probit	Compl. log-log
(1)	0.519	0.520	0.519
(2)	0.524	0.526	0.524
(3)	0.526	0.527	0.526

Note: Where (1) is using 1- and 2-month lagged interest rates as explanatory variables; (2) is using only 1-month lagged interest rates; and (3) is using no lagged interest rates.

Table 4.1 suggests that no discernible differences exist among the deviance functions of the three link functions. In fact, a Chi-square test for differences between the link functions indicates no statistical significance between

the deviances. Since the logit is well known for its case of interpretation, we chose to adopt it as link function in the GLM modeling of the surrender rates.

The percentage of change in the odds for the logit link function is given by the following formula:

$$\Delta\left(\frac{w_{x,t}}{1-w_{x,t}}\right)\% = \left(e^{\beta_i \Delta X_i} - 1\right).\,100 \qquad (4\text{-}4)$$

where $w_{x,t}$ is the surrender rate for age/gender group x at time t, β_i is the coefficient related to the explanatory variable X_i, $x=1,2,...,10$, $t=1,2,...,60$ and the remaining variables are kept fixed.

Our results show that the odds for females are 1.35% higher than those for males. Table 4.2 shows the odds variation of the surrender rates for each age group in relation to the oldest age group (72-80 years). It can be seen that the odds of the four first age groups are higher than that of the last age group, and that the younger the age group the higher the odds are. Table 4.3 depicts the sensitivity of the monthly surrender rate variation in relation to interest rate variation. When there is a positive variation in the interest rates, the monthly surrender rates also increase. This behavior can be explained by assuming that during this period there were other more attractive financial products in the market, considering that some annuity plans fix the guaranteed interest rates in the contract and unit-linked plans fix the interest rates in the guaranteed annuities.

Table 4.2. Percentage Variations in the Odds for Age Groups.

Age groups (years)	$\Delta\left(\frac{w_{x,t}}{1-w_{x,t}}\right)\%$
24-28	36.40%
29-53	19.87%
54-62	20.31%
63-71	9.19%

Table 4.3. Percentage Variations in the Odds Given the Interest Rate Variation.

Δ annual interest rates	$\Delta\left(\dfrac{w_{x,t}}{1-w_{x,t}}\right)\%$
2%	16.04%
1%	7.72%
0.75%	5.74%
0.50%	3.79%
0.25%	1.88%
-0.25%	-1.84%
-0.50%	-3.65%
-0.75%	-5.43%
-1%	-7.17%
-2%	-13.82%

We now apply the ARMA-GARCH model (see eq. (4-2)) to the GLM residuals and to the monthly returns of the Brazilian stock market index (Ibovespa), generating a set of potentially i.i.d. residuals. Particularly, to find the best ARMA-GARCH model for the Ibovespa returns, we use the data from January 2000 to December 2010. These returns are from the BM&FBOVESPA, the São Paulo Stock, Mercantile and Futures Exchange. In fact, application of the BDS test to those residuals did not reject the null hypothesis of i.i.d series for any of them. Table 4.4 shows, for each age/gender group, the values of the model's parameters, with their *p*-values, and the distributions of the i.i.d. random variables.

Table 4.4. ARMA-GARCH Fitting Results for both GLM Residuals and Ibovespa Residuals.

Residuals	ϕ_0	ϕ_1	θ_1	α_0	α_1	β_1	Distribution
24-28 / male	-	0.997 (0.000)	0.6217 (0.000)	0.00002 (0.542)	-	-	Student's t $(\upsilon=5)$
29-53 / male	-	0.951 (0.000)	-0.426 (0.000)	0.00001 (0.462)	-	-	Student's t $(\upsilon=5)$
54-62 / male	-	0.839 (0.000)	-0.604 (0.004)	0.00001 (0.054)	-	-	Student's t $(\upsilon=5)$
63-71/ male	-	0.777 (0.000)	-0.532 (0.003)	0.00001 (0.133)	-	-	Student's t $(\upsilon=5)$
72-80 / male	-	0.821 (0.000)	-0.544 (0.000)	0.00002 (0.384)	-	-	Student's t $(\upsilon=5)$
24-28 / female	-	0.896 (0.000)	-0.376 (0.037)	0.00001 (0.031)	0.312 (0.274)	-	Student's t $(\upsilon=5)$
29-53 / female	-	0.905 (0.000)	-0.353 (0.017)	0.00001 (0.519)	0.398 (0.579)	0.513 (0.296)	Student's t $(\upsilon=6)$
54-62 / female	-	0.828 (0.000)	-0.513 (0.002)	0.00001 (0.169)	-	-	Student's t $(\upsilon=5)$
63-71/ female	-	0.803 (0.000)	-0.438 (0.004)	0.00001 (0.387)	-	-	Student's t $(\upsilon=5)$
72-80 / female	-	0.852 (0.000)	-0.471 (0.001)	0.00001 (0.021)	-	-	Student's t $(\upsilon=5)$
Ibovespa returns	0.017 (0.017)	-	-	0.00076 (0.453)	0.171 (0.051)	0.710 (0.000)	Standard Gaussian

In order to study the dependence structure among the residuals, first we obtain the empirical estimation of different measures of dependence, such as Pearson's correlation, Kendall's tau, and Spearman's rho, which are presented in Appendix 7.2.2.

From Appendix 7.2.2, one can see that residuals of the male age groups show positive correlation with each other, being stronger among consecutive age groups. For males, one cannot reject, for the measures of concordance Kendall's tau and Spearman's rho, the null hypothesis of zero correlation between the 24-28 and 63-71 age groups and between the 24-28 and 72-80 groups, using a significance level of 0.05. For female age groups, similar behavior is found, but the null hypothesis of zero correlation cannot be rejected only between the 24-28 and 72-80 age groups. For Pearson's correlation, for both genders, the null hypothesis of zero correlation between two age groups of the same gender is always rejected.

Table 4.5 depicts the estimates for the values of the measures of dependence between the residuals of the Ibovespa returns and the residuals of

each age/gender group. Table 4.6 presents the *p*-values for the null hypothesis of no dependence for each of those measures.

Table 4.5. Measures of Dependence.

Age/gender groups	Pearson's correlation	Kendall's tau	Spearman's rho
24-28 / male	-0.078	-0.045	-0.058
29-53 / male	-0.159	-0.141	-0.187
54-62 / male	-0.405	-0.202	-0.274
63-71/ male	-0.506	-0.202	-0.299
72-80 / male	-0.514	-0.216	-0.307
24-28 / female	0.072	0.048	0.067
29-53 / female	0.042	0.054	0.080
54-62 / female	-0.104	-0.096	-0.155
63-71/ female	-0.141	-0.096	-0.131
72-80 / female	-0.135	-0.076	-0.124

Note: Measures of dependence between residuals of the Ibovespa returns and the residuals of each age/gender group.

Table 4.6.P-values of the Correlation Hypotheses Test.

Age/gender groups	Pearson's correlation	Kendall's tau	Spearman's rho
24-28 / male	0.551	0.610	0.654
29-53 / male	0.223	0.111	0.149
54-62 / male	0.001	0.022	0.034
63-71/ male	0.000	0.022	0.021
72-80 / male	0.000	0.014	0.018
24-28 / female	0.582	0.583	0.605
29-53 / female	0.744	0.540	0.535
54-62 / female	0.427	0.278	0.232
63-71/ female	0.282	0.278	0.313
72-80 / female	0.302	0.385	0.337

By economic reasoning, given the increase of surrender rates during the economic crisis, we expected negative dependence between the residuals of the Ibovespa returns and the residuals of each age/gender group. Nevertheless, the results from Tables 4.5 and 4.6 show that the negative correlations between the residuals of the Ibovespa return and the residuals of age/gender groups are statistically significant only for the following male age groups: 54-62, 63-71 and 72-80 years. This may suggest that Brazilian men near retirement ages are more prone to surrender their life insurance plans due to the perception of an economic crisis. For other male age groups and also for women near retirement age, the correlations are also negative, but we cannot reject that their values are zero.

The negative correlations may be understood by assuming that the majority of annuity policyholders (89% in December 2012[18]) belong to unit-linked plans and some of them are able to invest in stocks. Then, by observing the negative stock market index returns, policyholders decided to switch to less risky financial products, assuming that for those policyholders near retirement age there was less time available to try to recover from losses incurred during the financial crisis. Another possible explanation for such a condition is that risk aversion increases with age, and therefore during the crisis period older policyholders decided to withdraw their money and switch to less risky investments.

The dependence between the residuals of the Ibovespa returns and the residuals of a particular age/gender group can be analyzed by contour plot of an empirical copula. To produce such a plot, one needs samples drawn from the estimated cumulative distribution functions, as follows:

$$\left(u_i, v_i\right) = \left(F_{X_p}\left(x_p^{(i)}\right), F_{X_z}\left(x_z^{(i)}\right)\right) \tag{4-5}$$

where

$i = 1, 2, \ldots, q$; q is the size of the sample, i.e., 60 months;

$p = 1, 2, \ldots, n$;

$z = 1, 2, \ldots, n$ and $z \neq p$;

n is the number of marginal distributions used in the multivariate copulas, $n = 11$; and

x_1 is the time series of the stock market index residuals and $x_2, \ldots, x_n$ are the residuals of the $(n\text{-}1)$ time series of surrender rates ($\varepsilon_x's$ in eq. (4.2)).

Thus, the empirical copula $\hat{C}$ and its frequency $\hat{c}$ are defined at points $\left(\dfrac{i}{q}, \dfrac{j}{q}\right)$, as Deheuvels (1978), by the following expressions:

$$\hat{C}\left(\frac{i}{q}, \frac{j}{q}\right) = \frac{1}{q}\sum_{l=1}^{q} 1_{\{u_l \leq u_{(i)}, v_l \leq v_{(j)}\}} \tag{4-6}$$

$$\hat{c}\left(\frac{i}{q}, \frac{j}{q}\right) = \begin{cases} \dfrac{1}{q} & \text{if } \left(u_{(i)}, v_{(j)}\right) \text{is a element of the sample} \\ 0 & \text{otherwise} \end{cases}$$

---

[18] Data from SUSEP (Brazilian Insurance Supervisor), www.susep.gov.br.

where $u_{(1)} \leq u_{(2)} \leq \ldots \leq u_{(q)}$ and $v_{(1)} \leq v_{(2)} \leq \ldots \leq v_{(q)}$ are the order statistics of a univariate sample from copula $\hat{C}$.

To illustrate the difference between independent residuals and residuals with negative measure of dependence, Figure 4.4 shows two empirical copula densities. The first is built using the time series of residuals of the male 63-71 age group as $F^{-1}(v) = x_5$ and the residuals of the Ibovespa returns as $F^{-1}(u) = x_1$, and the other empirical copula uses the latter and the residuals of the female 24-28 age group as $F^{-1}(v) = x_7$. In the first graph of Figure 4.4, the negative dependence can be observed.
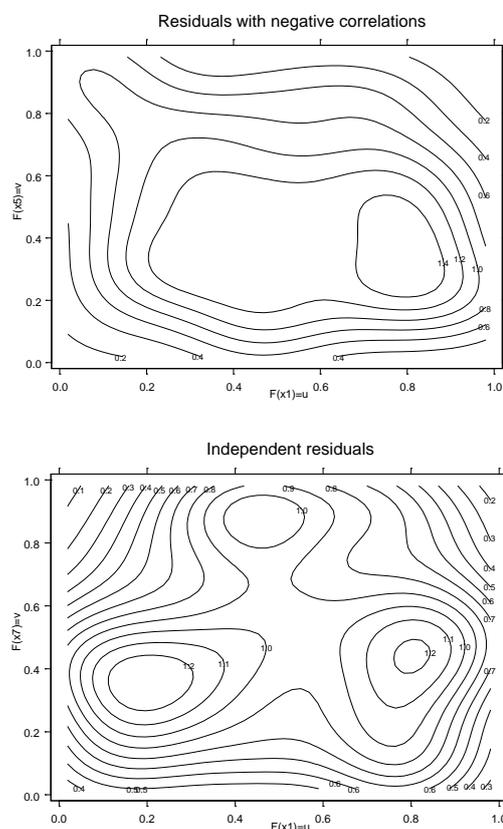
Figure 4.4. Plots Contours of the Empirical Copula Densities

Dependence among the eleven time series of residuals (from Ibovespa returns and ten age/gender groups) is modeled by multivariate Gaussian copula and multivariate Student's t copula. Applying the CML method, we obtain estimates for the MGC parameter (i.e., the correlation matrix R, from eq. (7-5) and eq. (7-6)) and by means of the three-stage KME-CML method, the MTC

parameters (correlation matrix R and υ from eq. (7-7) and eq. (7-8)). After maximizing the log-likelihood of equation (7-8), we found 26 degrees of freedom as the MTC parameter, which is a high number and can produce similar results to the Gaussian copula.

4.3.1
Comparison between elliptical copulas

We now investigate the predictive ability of our modeling procedure through an out-of-sample exercise. We reserved the monthly surrender rates of 2011 from the sample to test for goodness of fit using the MGC and MTC to estimate the residuals' dependence structure. Using the MGC and MTC simulation algorithms, we obtain the residuals' marginal distributions and by applying eq. (4-2) and eq. (4-1), we obtain the distribution of forecasted monthly surrender rates. Then, the predicted values of the surrender rates are compared with observed values for each age/gender group.

For the MGC, the MAPE value is 15.8%, while by MTC it is 15.7%. Assuming that the means of forecasted residuals are zero, the predicted values of surrender rates via copulas is similar to that obtained using only the GLM approach, but the copula step allows one to simulate from the distributions of surrender rates considering the dependence structure. Model prediction ability is tested through the Kupiec test, using a confidence level of 95%. For both MGC and MTC copulas we are unable to reject the null hypothesis of correct coverage, with p-values of 0.67 and 0.37, respectively. Since the number of degrees of freedom of the MTC is high (26), MTC and MGC produce similar results. Therefore, we opted to simulate surrender rates from the MGC. Figures 4.5 and 4.6 depict the observed surrender rates and predicted values for the months of 2011, including a confidence interval of 95%.
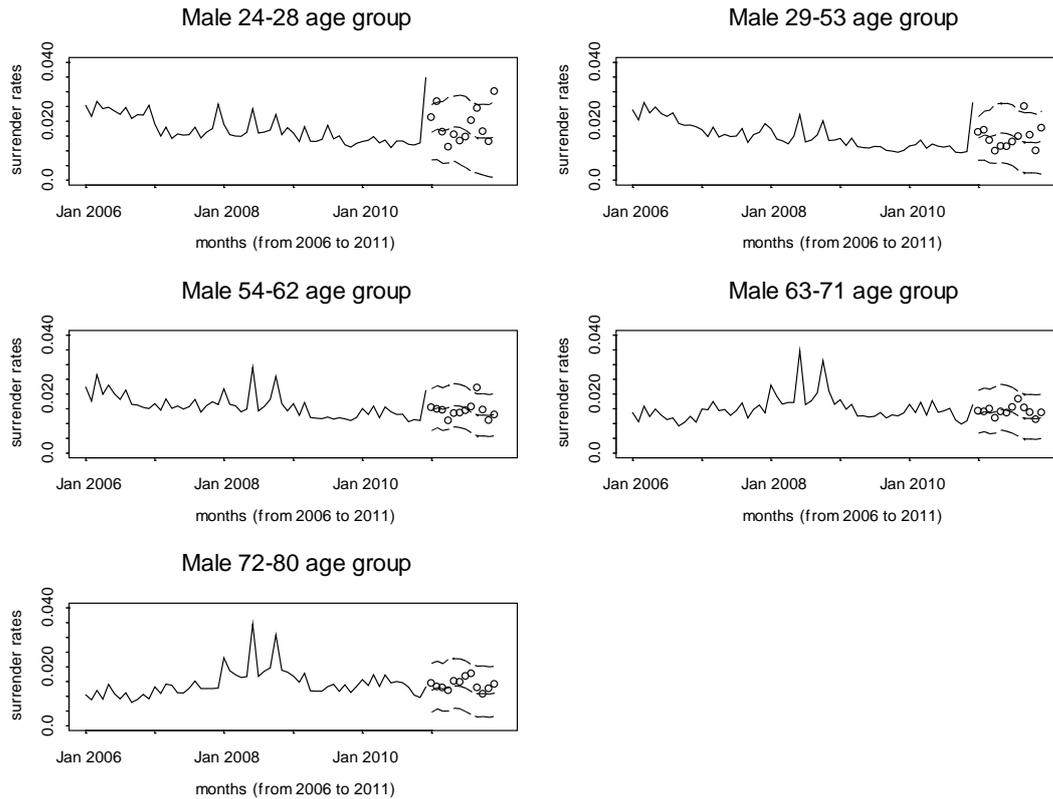
Figure 4.5.Observed and Forecasted Surrender Rates for Males.

Note: Observed monthly surrender rates up to 2010, for each age group, in solid lines; predicted value of surrender rates for 2011 and its confidence interval of 95% in dashed lines; and observed surrender rates in 2011 as circles.
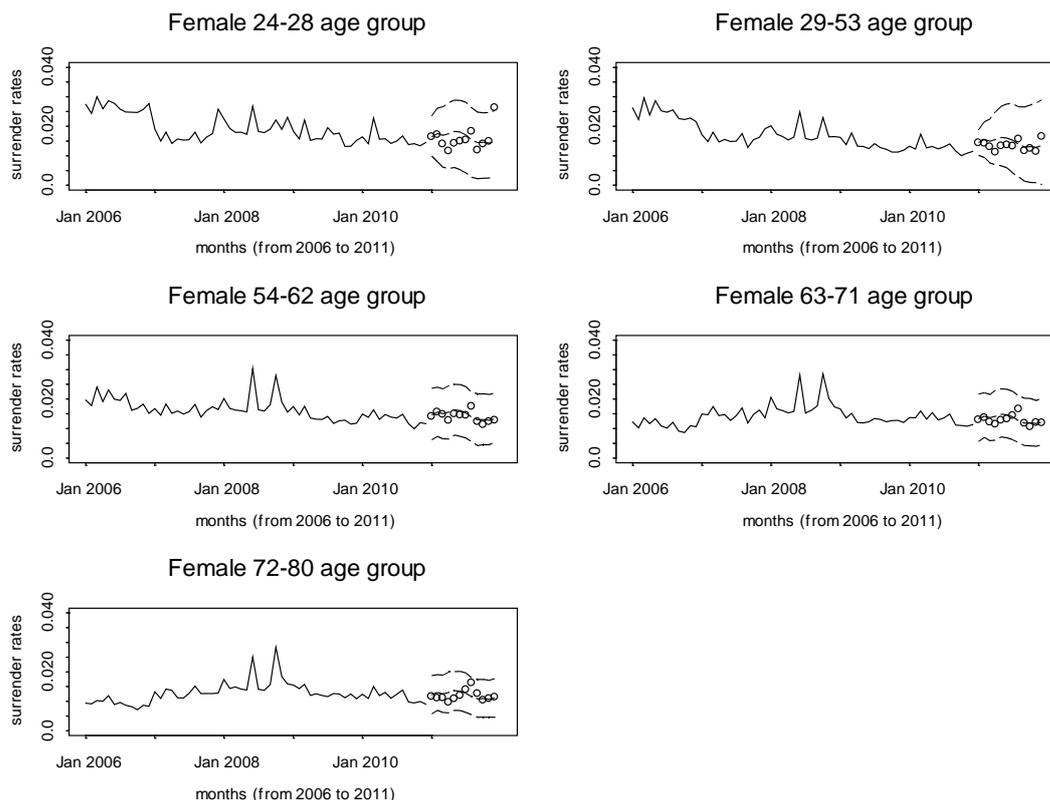
Figure 4.6.Observed and Forecasted Surrender Rates for Females.

Note: Observed monthly surrender rates up to 2010, for each age group, in solid lines; predicted value of surrender rates for 2011 and its confidence interval of 95% in dashed lines; and observed surrender rates in 2011 as circles.

Finally, we used our conditional copula approach to simulate surrender rates conditional on the given values of the monthly Ibovespa return observed in 2011, by applying the algorithms described at the end of section 4.2.1. These results show a slight MAPE increase in comparison to the unconditional copulas (16.7% by MGC and 16.8% by MTC), but the number of observations outside the 95% confidence interval decreased. In fact, using the Kupiec coverage test, we could not reject the null hypothesis of the right coverage for the chosen confidence level, with the same p-value for both copulas. From these results and given that the MTC converge to the MGC in our application, we chose to use conditional MGC to simulate the surrender rates assuming a crisis period.

Figures 4.7 and 4.8 show the observed surrender rates and the predicted values conditioned on the Ibovespa returns for the months of 2011, including a confidence interval of 95%.
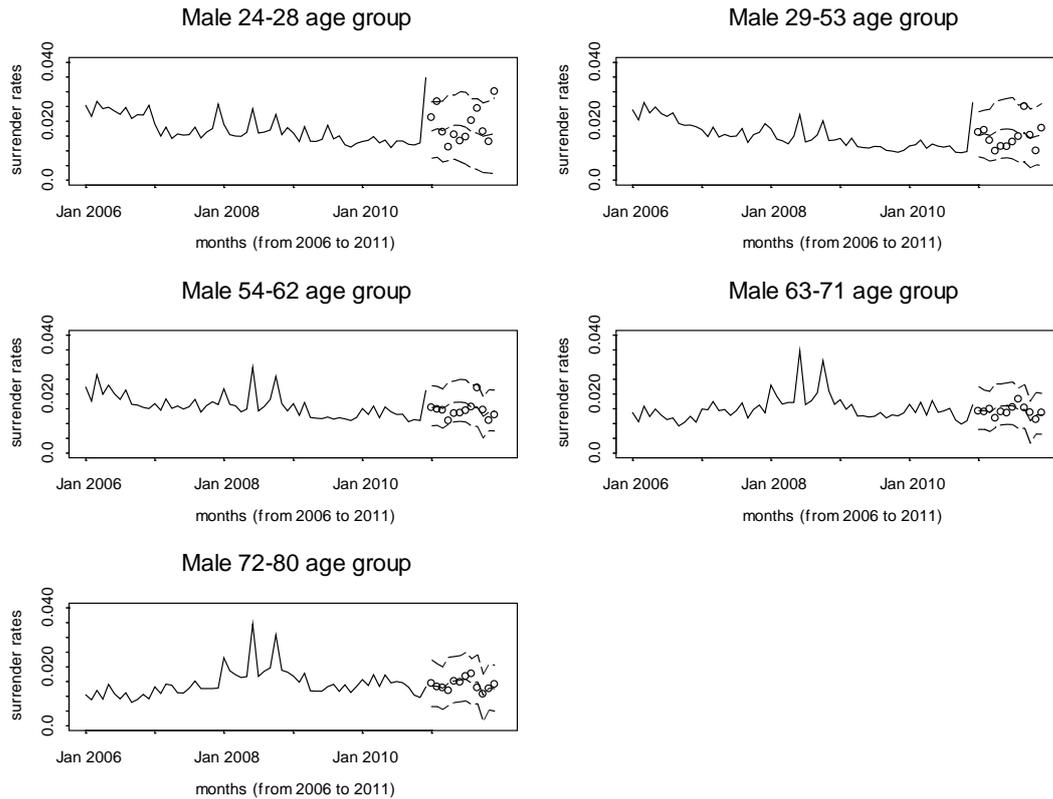
Figure 4.7.Observed and Forecasted Surrender Rates by Conditional Copula for Males.

Note: Observed monthly surrender rates up to 2010, for each age group, in solid lines; predicted value of surrender rates conditioned on the Ibovespa returns for 2011 and its confidence interval of 95% in dashed lines; and observed surrender rates in 2011 as circles.
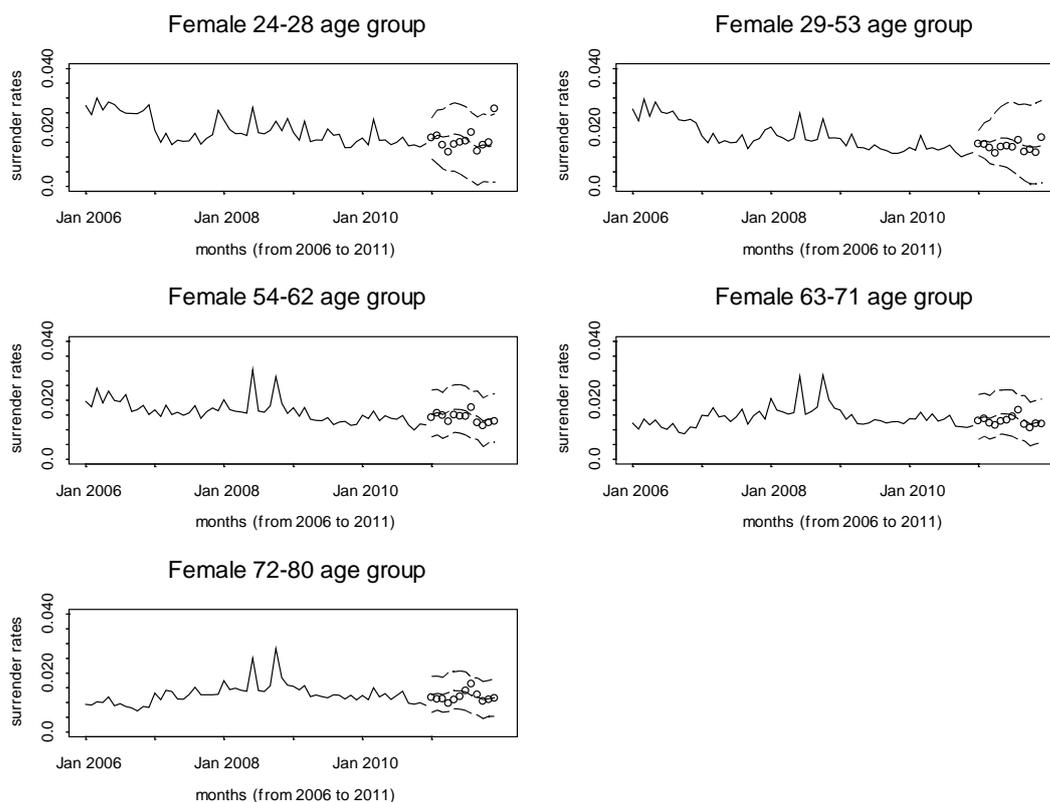
Figure 4.8. Observed and Forecasted Surrender Rates by Conditional Copula for Females.

Note: Observed monthly surrender rates up to 2010, for each age group, in solid lines; predicted value of surrender rates conditioned on the Ibovespa returns for 2011 and its confidence interval of 95% in dashed lines; and observed surrender rates in 2011 as circles.

The usefulness of our conditioned simulation through copulas can be corroborated by observing the surrender rates in September 2011, when the smallest Ibovespa value during the year was observed. The monthly return in this period was only - 0.077, which represents the quantile of 0.065 of the cumulative distribution function of the residuals of the Ibovespa returns. When fitting the MGC, one observes that the corresponding surrender rate lies outside the 95% simulated confidence interval for the male 29-53 and 54-62 age groups, as seen in Figure 4.5. But when the conditional MGC is applied to this same period and age/gender groups, the 95% confidence interval now contains the true observed surrender value, as can be seen in Figure 4.7.

One then can conclude that our approach can be applied to forecast surrender rates several steps ahead without operational difficulty, which is essential to evaluate liabilities and actuarial risk in the insurance industry.

4.4
Simulation with a financial stress scenario

Through our approach it is feasible to simulate the distribution of surrender rates assuming the occurrence of a crisis period. This type of simulation is useful in the insurance industry when one needs to evaluate the insurer's commitments presuming a financial market situation (stress test), in order to improve risk management.

To simulate surrender rates given a financial stress scenario, we assumed values for Ibovespa residuals - or quantiles of their cumulative distribution function - during the crisis period. After that, we applied the simulation algorithm described at the end of section 4.2.1. For instance, using conditioned MGC, we simulated values of surrender rates for one year assuming a crisis period in the last four months of that year. We assumed that in this stress scenario the quantiles of the standard Gaussian marginal distributions are fixed at 0.02, 0.01, 0.005 and 0.03, respectively for the last four months, which represent big negative Ibovespa returns. The predicted surrender rates are compared with the values obtained using the unconditional MGC, i.e., without fixing a stress scenario, whose results were presented in Figures 4.5 and 4.6.

Figure 4.9 shows surrender rate point forecasts (means) and corresponding 95% confidence intervals assuming both a no-stress scenario and a crisis period (dashed lines) for those age/gender groups containing policyholders of retirement ages. During the first eight months, before the crisis, the results of both approaches coincide. But, in the crisis period one can easily notice a significant increase in the forecasted surrender rates. In order to better grasp the effect of the crisis period on the forecasted surrender rates, Table 4.7 presents the percentage increase in the predicted surrender rates during the crisis scenario in relation to those that would be obtained by considering a MGC model in which no-stress scenario is assumed.

Male 54-62 age group

Male 63-71 age group

Male 72-80 age group

Female 54-62 age group

Female 63-71 age group

Female 72-80 age group
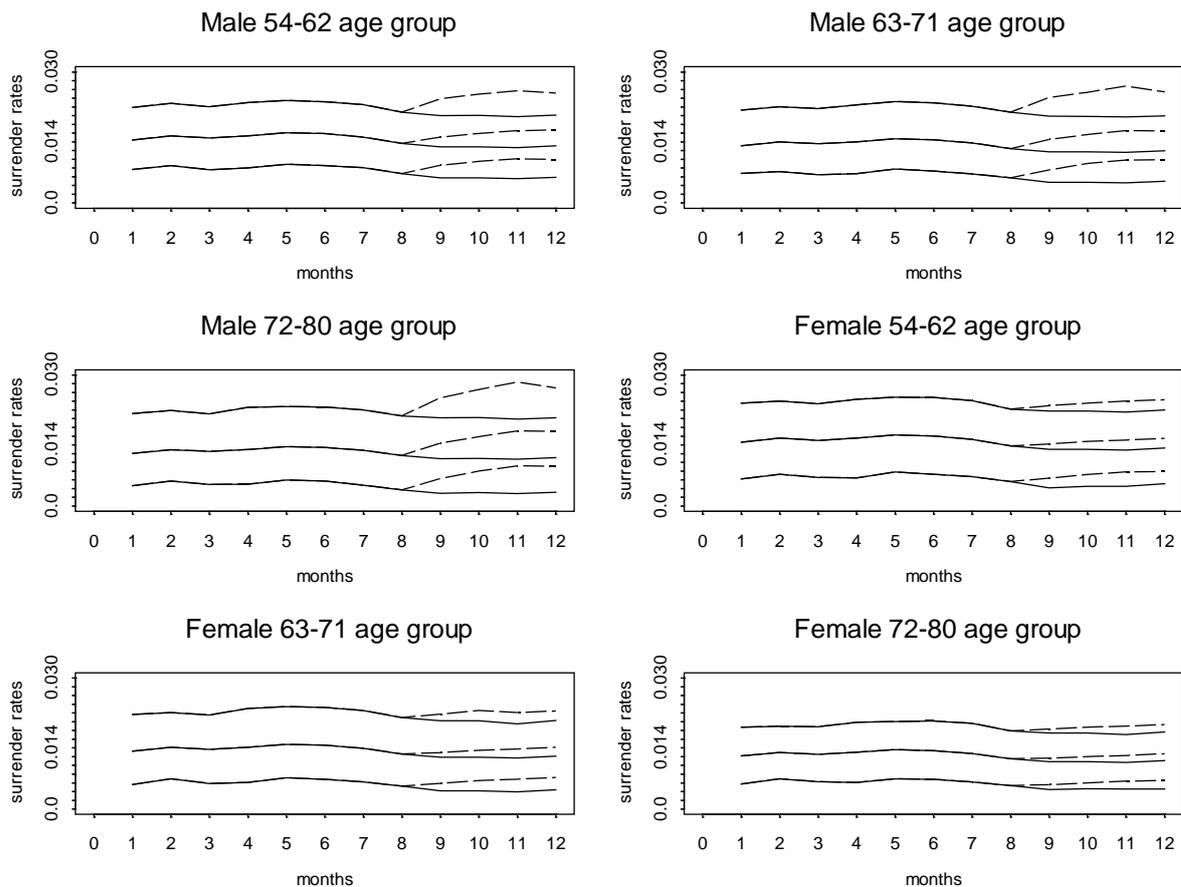
Figure 4.9. Stress test: Forecasted Surrender Rate for 2011.

Note: Means and confidence intervals of 95% of the forecasted surrender rates without fixing a stress scenario in the solid lines; and in the dashed lines those obtained by the financial stress scenario.

Table 4.7. Percentages of Increase.

Age/gender groups	Months			
	9th	10th	11th	12th
54-62 / male	17.24%	23.78%	30.34%	27.12%
63-71/ male	24.00%	30.30%	42.78%	37.51%
72-80 / male	32.06%	45.74%	59.97%	54.06%
54-62 / female	9.36%	13.53%	17.81%	16.65%
63-71/ female	8.86%	13.24%	17.63%	16.77%
72-80 / female	6.92%	10.56%	14.22%	13.93%

By analyzing Figure 4.9 and Table 4.7 in conjunction with Table 4.5, one can see that the surrender rate increments are higher in the crisis period for those age/gender groups that have higher measures of dependence with the Ibovespa residuals.

4.5
Conclusion

In this paper a multi-stage stochastic model is proposed to forecast time series of surrender rates using both policyholder variables (gender and age) and financial variables, such as short-term interest rates and a proxy for the stock market. Such a model can be useful to insurers and pension funds to realistically evaluate their liabilities and actuarial risks under scenarios of financial stress. Our modeling procedure is illustrated using time series of surrender rates obtained from a Brazilian insurance company.

In the first stage of the modeling process, time series for age specific surrender rates were modeled by a GLM framework, in which gender and interest rates are used as predictors via a binomial model with a logit link function. ARMA-GARCH processes were then fitted to these GLM residuals, generating a second set of residuals. Since our aim is to obtain forecasts for surrender rates conditional on market variables (the residuals of an ARMA-GARCH model fitted to the returns of a stock index), we joined the surrender rate residual series together with stock index residuals via multivariate Gaussian and Student's t copulas. From this multivariate model one can forecast surrender rates conditional on a specific financial stress scenario. More specifically, it is possible to simulate future surrender rates from the multivariate elliptical copulas (Gaussian or Student's t) conditioned on a specific percentile of the cumulative distribution function of the stock index (the residuals of the Ibovespa model). Using this framework it was possible to infer that most of the correlations between the stock index (the residuals from the GARCH model fitted to the Ibovespa returns) and each of the surrender rates (the residuals from the GLM fitted to the age/gender group rates) were negative. Therefore one would expect that during crisis periods, when stock indexes are at their lowest, the forecasted surrender rates should attain their highest values. More specifically, our results show that for the Brazilian insurance market this correlation is both stronger and statistically more significant for male groups with age closer to retirement. This in turn can suggest that Brazilian policyholders tend to be more risk averse as they grow older.

One of the advantages of using this method in practice is that if insurers want to forecast future surrender rates several steps ahead, it will not be necessary

to predict interest rates or stock returns. Interest rates future values are obtained using the current term structure and stock returns only enter in the model through its residuals whose distribution is used to compose a multivariate copula from which future surrender rates are simulated. Furthermore, an insurer or pension fund can apply stress tests using our model to simulate conditional copulas to analyze its capacity to manage its portfolios during a financial crisis.

5

Paper 4: Embedded options in Brazilian Unit-linked Plans: evaluation using a real-world probability measure

We propose a model for evaluating the value of embedded options in the Brazilian unit-linked plans, such as the guaranteed annuity option, which includes deferment option and switching the type of annuity, growth option, surrender option and shutdown option. We describe the main characteristics of these options and shows that the Brazilian annuity market is incomplete and not free from arbitrage, illustrated by three practical examples, in the context of which we use the real-world probability measure in this evaluation. The proposed model allows one to consider both the rational and irrational decision of the policyholder to surrender before the date set for retirement. Additionally, the model considers the possibility of self-annuitization by means of partial surrenders, in addition to changing the type of income and deferring the date of conversion into income. The surrender options, as well as other embedded options, such as growth and cancelation, are modelled using jump processes in the stochastic differential equation that describes the evolution of the unit-linked fund.

5.1
Introduction

Brazil has two types of unit-linked plans: VGBL ("Vida Gerador de Benefícios Livres", corresponding to Redeemable Life Insurance) and the PGBL ("Plano Gerador de Benefícios Livres", or literally Benefit Generator Plan). The only difference between them is the tax benefit. Consumers can opt for purchasing a PGBL or VGBL according their gross annual income in order to maximize their tax benefit. Both are annuity plans and allow the conversion of the unit-linked fund into income annuities using the technical bases (mortality tables and interest rate) fixed at the time of purchase. These products represent some 90% of the total technical provisions, about 300 billion reais, 92% of the total annual premiums,

89% of policyholders, and 14% of beneficiaries of the Brazilian annuity market[19]. The last percentage is low due to the fact that these products are relatively new, the PGBL was created in 1997 and VGBL in 2001.

Policyholders of these plans are afforded several contractual options, such as guaranteed annuity option, the option to defer the conversion into income (annuitization), the surrender option, switching the type of annuity, interruption of premium payment, the option to transfer funds to/from another insurer, or the option to increase the income by the payment of additional premiums. Insurers should seek to price correctly the set of these embedded options in order to determine their liabilities, as well as their need for risk capital.

There are several published articles dealing with the assessment of these options. Grasselli and Silla (2009) classified these studies using two approaches: financial and actuarial. Most follow the financial approach, based on the principle of no-arbitrage. Ballotta and Haberman (2003) introduced a theoretical model for pricing the guaranteed annuity conversion option included in deferred unit-linked contracts in the United Kingdom, purchased by a single premium. These authors assumed a rational behavior of the policyholders and a risk-neutral measure, pricing this embedded option as a European call option. Later, Ballotta and Haberman (2006) extended the model incorporating mortality risk by means of a Monte Carlo simulation.

Following the same line of thought, Boyle and Hardy (2003) explored the pricing of that option, as well as its risk management. In turn, Biffis and Millossovich (2006) considered a complete market and used an affine process for dynamic mortality. Ziveyi et al. (2013) also priced the guaranteed annuity option through the no-arbitrage approach and modelled the evolution of the mortality table, but they derived the pricing partial differential equation and the corresponding transition density. Several articles have been published with the objective of pricing U.S. variable annuities, known as Guaranteed Minimum Withdrawal Benefit (GMWB) plans, among which we can mention Milevsky and Salisbury (2006) who presented two approaches in relation to policyholder behavior in order to price these variable annuities. First, they assumed that policyholders behave passively with respect to utilizing the embedded guaranteed.

---

[19] Information provided by the Brazilian Insurance Supervisor (SUSEP), database: December 2012.

In the second hypothesis, they assumed policyholders are completely rational, seeking to maximize the embedded option through the surrender of the product at the optimal time. A martingale measure was applied in both approaches.

In general, the pricing of embedded options in private pension contracts and life insurance uses an approach that assumes a risk neutral evaluation. However, the fundamental theorem of pricing ensures that there is a unique martingale measure if, and only if, the market is complete, which happens when all assets can be perfectly hedged. When the market is incomplete, but is arbitrage free, there will be more than one risk-neutral measure. Some authors present strategies to replicate the portfolio of an insurer assuming the insurance market is incomplete, such as Møller (2001) and Consiglio and De Giovanni (2010). The first specified a hedging strategy that minimizes the risk for equity-linked life insurance contracts. In turn, the second adopted a super-replication model to determine the portfolio replication of a surrender option. However, as stated by Gatzert and Kling (2007), in many countries the insurer's asset allocation is subject to various regulatory limitations. So, in practice, they cannot adopt the hedging strategies proposed in the literature.

In turn, Solvency II specifies that insurers must evaluate their assets and liabilities in a manner consistent with the market. Thus, the liabilities should be valued on the amount for which they can be transferred or settled. Technical provisions are calculated as a best estimate of this value plus a risk margin. Moreover, when the future cash flows associated with the obligations of the insurer can be reliably replicated using financial instruments with an observable market value, the value of the technical provisions associated with these future cash flows is determined based on the market value of those instruments.

Analyzing the behavior of the policyholder is extremely important in the evaluation of embedded options. Solvency II, rightly, recommends that this behavior should be appropriately based on statistical and empirical evidence with respect to variations in the financial market, explicitly mentioning the fact of options being in or out of the money. Some authors, such as Ballotta and Haberman (2003 and 2006), Boyle and Hardy (2003) and Biffis and Millossovich (2006), have assumed that policyholders take decisions optimally. On the other hand, De Giovanni (2010) built a rational expectation model to describe the surrender rate of insurance contracts and found that the behavior of the

policyholder is far from optimal. In addition, Kling et al. (2014) analyzed the impact of the behavior of the policyholder on the pricing and hedging of variable annuity contracts and presented different assumptions about this behavior.

Our main objective is to discuss the embedded options in Brazilian unit-linked plans and propose a model to obtain the best estimate of their value, given by the average value of future cash flows, considering the time value of money. Our model for evaluation of the options is based on a real-world measure, because the Brazilian annuities market, similar to other insurance markets, is incomplete. Furthermore, this market is not arbitrage free and does not have a risk-neutral measure (martingale), according to the central theorem of pricing (Harrison and Pliska, 1983). Our study presents realistic examples of arbitrage in the Brazilian annuities market.

Our approach does not assume optimal policyholder behavior. In lieu of this, we model the surrender rates in the period up to the predetermined retirement date in order to allow rational and irrational surrenders, applying the model proposed by Neves et al. (2014) – chapter 4 of this thesis, as part of the same concept presented by Giovanni (2010). The model used is stochastic and considers the dependency between the surrender rates and real short-term interest rates. Furthermore, using multivariate elliptical copulas, the model assumes that the surrender rates are affected by a financial crisis. This is achieved using the returns of the Brazilian stock market index (Ibovespa) as one of the marginal distributions used in dependence modelling through multivariate copulas.

In addition, to obtain the best estimate of embedded options, we model the decision of the fund's conversion into annuity considering the policyholder has the right to change the type of annuity at the time of retirement, the option of postponing the date of retirement, as well as the possibility of choosing a self-annuitization strategy[20], modelled by a jump process. Moreover, policyholders can increase the value of embedded options if they continue to pay regular premiums, pay additional premiums, or transfer their funds from other plans or from other insurers to their unit-linked plans. These movements are also modelled by means of jump processes.

---

[20] Opt for partial surrender after retirement date, postponing the default date of retirement.

The rest of this study is organized as follows. In section 5.2, we describe the embedded options in the Brazilian unit-linked plans. In section 5.3, we present the characteristics of the Brazilian annuities market and demonstrate through examples that there are arbitrage opportunities. In section 5.4, the proposed model to evaluate the best estimate of the embedded options is presented. In section 5.5, we apply the model by means of a Monte Carlo simulation and perform a sensitivity analysis. In the last section, we draw our conclusions.

5.2
Embedded options in Brazilian unit-linked plans

The PGBL and VGBL products contain several embedded options that should be evaluated for a correct determination of the insurer's solvency. To study these, we turn to both the life insurance literature and literature of real options (Boyle and Irwin, 2004, and Li et al., 2007). In fact, some of the embedded options can be classified as real options, as indicated by Milevsky and Young (2001), which identified the option of deferring annuitization as a real option, due to its irreversibility and its non-negotiable and personal nature.

The embedded options in the Brazilian unit-linked contracts are:

5.2.1
Guaranteed Annuity Option

The mortality table and interest rate used for calculating the value of income are predefined in the annuity contract at the time of purchase. In addition, the policyholder predetermines the exact date of retirement in the contract, but can change it at any time. With that, this option bears a strong resemblance to an American option, whose main feature is able to be exercised at any time.

However, in Brazil, people tend to retire at the age stipulated by the State pension system (INSS), which generally coincides with the date fixed in the annuity plans. Nevertheless, policyholders commonly decide to postpone, at no cost, the conversion date fixed in the annuity contracts. So, we have not assumed the hypothesis of early retirement and, consequently, have modelled the guaranteed annuity option as a European option added to the option to defer the

date of retirement, being evaluated in conjunction with the other embedded options for calculating the value of the best estimate of the options.

5.2.2
Switch Option

Some time prior to the conversion of the Fund into annuity, the policyholder has the right to switch, without cost, the type of annuity type determined in the contract. The PGBL and VGBL plans offer several types of annuity, such as:

- simple, where the lifetime annuity shall be paid to the policyholder;

- temporary, where the policyholder can define the time of annuity payment;

- reversion to an indicated beneficiary, where the policyholder designates a beneficiary, independent of age or sex, and the percentage of reversion;

- reversion to spouse and minor children, with the possibility of choosing the percentage of reversion;

- certain income, when the policyholder converts the fund into financial annuities, defining the payment period; and

- certain income plus a deferred annuity (guaranteed minimum period), whose period of payment of certain income is set by the policyholder.

Hu and Scott (2007) and Milevsky and Young (2001) identified reasons related to bequest often compel policyholders not to convert the fund into annuities, since, after conversion into income, the fund belong to the insurer. As seen above, the unit-linked plans offer various types of annuity at the moment of conversion, including the possibility of third parties beneficiaries, which can reduce policyholders' resistance to conversion into annuity.

5.2.3
Growth option

With unit-linked plans, policyholders can pay regular premiums according to the frequency defined in contract, and have the right to change such values during the period of the contract. In addition, they may choose to pay additional (supplementary) premiums, in the amount and the date of their choosing. Given

this contractual feature, if the guaranteed annuity option is in the money, each premium paid increases the value of that option. Hence, one can see that the value of the growth option depends on the behavior of the policyholder.

According to this option, if the behavior of the policyholder were truly optimal, when the guaranteed annuity option was in the money, the policyholder could conduct a series of additional contributions that would substantially increase the value of the guaranteed annuity option, increases the value of future benefits, and may lead to the insurer having solvency problems.

Moreover, the difficulty of pricing embedded options is due to the fact that the policyholder could also increase the value of the contract through transfer (portability), at no cost, of their funds/provision from other insurers (or pension funds), or other annuity plans of the same insurer, including defined contribution and defined benefit plans, to the unit-linked plan. In short, the value of contract may be increased by regular premiums and additional payments and portability of resources to the unit-linked fund.

5.2.4
Surrender option (cancelation option)

Policyholders can surrender, in whole or in part, at any time after the purchase of the plan, following a short grace period. The total surrender, according to the terminology of real options, characterises the abandonment option. The exercise of this option entails the cancelation of the plan and causes the values of all other embedded options to become zero, since the policyholder only surrenders the value of the unit-linked fund, which corresponds to the value of the mathematical provision of benefits to be granted.

On the other hand, a partial surrender reduces the value of other embedded options because the value of the unit-linked fund is reduced. The surrenders may also be related to portability to another plan or insurer. The partial surrender can also occur after the predetermined retirement age, in which case the policyholder decides to make scheduled withdrawals, rather than convert the fund into annuities. For this, the policyholder postpones his retirement and then determines the rate of withdrawals, i.e. he opts for a self-annuitization strategy. In the

approach proposed in this text, this alternative is considered only for dates after the predetermined date of retirement.

5.2.5
Shutdown option (payment interruption option)

Policyholders have the right to temporarily or permanently discontinue the payment of premiums. Exercising this option does not increase the value of the other embedded options, but can affect revenue and asset and liability management of the insurer.

As the evaluation of guaranteed annuity option depends on the options of switching, growth and surrender, the model to obtain the best estimate of the embedded options should cover all those options. It is important to note that most of these options are also included in defined contribution plans.

5.3
Brazilian Annuity Market

By definition, a market is complete if, and only if, all contingent rights are attainable. Additionally, one can describe the no-arbitrage theory via martingales, as summarized by Brigo and Mercurio (2006):

- "The market is free of arbitrage if (and only if) there exists a martingale measure;
- The market is complete if and only if the martingale measure is unique;
- In an arbitrage-free market, not necessarily complete, the price of any attainable claim is uniquely given, either by the value of the associated replicating strategy, or by the risk-neutral expectation of the discounted claim payoff under any of the equivalent (risk-neutral) martingale measures."

In Brazil, there is no relevant longevity reinsurance market. In addition, there are no assets connected to the Brazilian mortality rate available in the financial market. Finally, there is a strict regulation of insurers' asset allocation policies by limiting the range of assets that insurance companies can buy to cover their commitments to policyholders and beneficiaries. So, we can conclude that

the annuity market is incomplete and insurers cannot apply optimal hedging strategies.

As stated by Brigo and Mercurio (2006), a market presents the absence of arbitrage when there exists the impossibility of investing zero today and receiving tomorrow an amount that is greater than zero with positive probability, i.e. two portfolios of assets with the same returns on a future date must have the same price today. In the next subsections, we present three examples of arbitrage opportunity in the Brazilian annuities market to prove the non-existence of a martingale measure in this market. For this, it is worth mentioning the evidence of Sutcliffe (2015), which examined the occurrence of arbitrage in annuity markets and demonstrated that it is possible for an annuity to be under-priced, but not overpriced.

5.3.1
Loans and life insurance

We assume a frictionless market, with continuous trading, no taxes or transaction costs and no borrowing restrictions or short selling and perfectly divisible assets or securities.

As shown in section 5.2, policyholders can exercise the option of growth and pay additional premiums, including immediately before converting the fund into annuities. This procedure is equivalent to buying an additional annuity. Thus, an insured of a unit-linked plan can simultaneously apply for a bank loan and with this money, buy an immediate annuity, by means of an additional premium, and also a life insurance policy. The latter will have to have decreasing values, always corresponding to the remaining amount of the loan. With this strategy, assuming the guaranteed annuity is under-priced, it is possible to obtain immediate arbitration. This strategy is similar to that shown by Sutcliffe (2015), although, we demonstrate it using a temporary annuity.

Suppose that an insurer guarantees an annuity based on the 1983 Individual Annuity Mortality Table (AT 1983) and a real interest of 4% per annum - conditions commonly offered by Brazilian unit-linked plans. To calculate the fair value of this annuity, i.e. with realistic technical bases, we will assume that the mortality rates of retirement plans in Brazil match those found in the BR-

EMS 2010 mortality table for survival coverage (Oliveira et al., 2012) and that future longevity gains are obtained according to the model developed in Neves et al. (2016) – chapter 3 of this thesis. Additionally, we assume that the real interest rate in the market is 4% p.a. and that life insurance is sold for its fair price.

Given its technical bases, the insurer charges a man of 65 years R\$ 18,488.97 for a temporary annuity of R\$ 10,000.00 paid for two years, whose fair price is R\$ 18,571.90. The policyholder can then apply for a loan of this last value, amortizing it in two payments on the same date of receipt of the annuities. The annual cost of the loan is R\$ 9,846.75, the fair price of the annual premium of a temporary life insurance is R\$153.25 and the sum of these values is exactly the value of the temporary annuity. With that, the policyholder gets an instant profit of R\$ 82.92 (= R\$ 18,571.90 - R\$ 18,488.97), which shows the possibility to invest zero today and get a positive return without risk in the Brazilian annuity market. In other words, there is an opportunity for arbitrage.

Tables 5.1, 5.2 and 5.3 illustrate the cash flows of this strategy. In Table 5.1, we assume that the policyholder does not die during the payment of annuities over the two years. In Table 5.2, the policyholder dies in the first year and in Table 5.3, dies in the second year. In all cases there is profit in the operation due to arbitrage.

Table 5.1: Example of arbitrage in the Brazilian annuity market, assuming the policyholder does not die during the period of payment of annuities.

Years	Transactions	Inflows (R\$)	Outflows (R\$)
0	Premium Payment		18,488.97
0	Loan	18,571.90	
0	**Profit**		**82.92**
1	Annuity	10,000.00	
1	Cost of Loan		9,846.75
1	Life Insurance Premium		153.25
2	Annuity	10,000.00	
2	Cost of Loan		9,846.75
2	Life Insurance Premium		153.25

Table 5.2. Example of arbitrage in the Brazilian annuities market, assuming the policyholder dies in the first year of the contract.

Years	Transactions	Inflows (R$)	Outflows (R$)
0	Premium Payment		18,488.97
0	Loan	18,571.90	
0	**Profit**		**82.92**
1	Life Insurance Claim	18,571.90	
1	Loan payment		18,571.90

Table 5.3. Example of arbitrage in the Brazilian annuity market, assuming the policyholder dies in the second year of the contract.

Years	Transactions	Inflows (R$)	Outflows (R$)
0	Premium Payment		18,488.97
0	Loan	18,571.90	
0	**Profit**		**82.92**
1	Annuity	10,000.00	
1	Cost of Loan		9,846.75
1	Life Insurance Premium		153.25
2	Life Insurance Claim	9,846.75	
2	Loan Payment		9,846.75

5.3.2

Transfer of funds from another insurance company or plan

Brazilian legislation allows policyholders to perform transfers (portability) of their funds or provisions from other insurers or between plans within the same insurer, without paying taxes. A policyholder may therefore transfer a fund to an insurer that has a higher guaranteed annuity rate[21], i.e., pay a higher income with the same resource or pay the same amount of income by making use of a smaller fund. With the transfer from an insurance company with a lower annuity rate, the policyholder can then obtain the same annuity with a smaller amount, generating a profit without risk. For example, suppose that the policyholder has two unit-linked funds, one with a guaranteed annuity rate based on AT 1983 and on real interest rate of 4% per annum and the other using the Annuity 2000 Mortality Table (AT 2000) with real interest rate of 2% per annum. The policyholder is a man of 65 years and has R$ 500,000.00 in each plan at the moment of annuity conversion.

---

[21] The rate to be multiplied by the fund on the date of retirement to calculate the value of the income, i.e., the inverse of the guaranteed annuity value.

The annuity for the first plan will be R$ 41,875.12 and R$ 33,026.53 for the second, totalling R$ 74,901.65. Calculating them on realistic bases, both annuities generate the same fair value.

This policyholder can, therefore, keep the same value of annuity (R$74,901.65) by transferring R$394,345.45 from the plan that uses AT 2000 to the plan that uses AT 1983 and exercising the option of guaranteed annuity in the latter. The policyholder then exercises the option of cancelation and surrenders the amount of R$105,654.55 that is remaining in the AT 2000, yielding an instant profit, which characterizes the arbitrage.

5.3.3
Switching the Type of Annuity

Suppose that a policyholder contracts a simple lifetime annuity and, at the time of conversion of the fund into income, exercises the option of switching the type of annuity. Assuming that annuities are under-priced, there is the possibility of a partial surrender while maintaining the same fair value of annuities, i.e. there are arbitrage opportunities.

To examine this possibility, consider once again that the policyholder is a 65-year-old man under the same conditions of the example presented in section 5.3.1. This policyholder can change a simple lifetime annuity for an annuity with 50% reversion to his wife, who is 10 years younger than him. Assuming this policyholder had R$ 500,000.00 accumulated in the plan, he would receive a lifetime annuity of R$ 41,875.12, the fair value of this annuity being R$ 561,659.00 assuming realistic bases defined in section 5.3.1.

However, if the policyholder exercises the option to change the type of income contracted for a reversible annuity to his wife, with the same fair value he could purchase an annuity of R$ 41,814.51. Moreover, according to the contract, the price of that reversible annuity is R$ 498,496.40. This strategy allows, therefore, a partial surrender of R$ 1,503.60 maintaining the same fair annuity value, i.e. R$ 561,659.50. Again, what we observe are two types of annuity with the same future flows, but with different prices, raising the possibility of arbitrage.

5.4

Model

In this section, we propose a model to obtain the best estimate of the embedded options in unit-linked plans. To this end, we assume that the best estimate shall be equal to the average of future cash flows from these options, taking into account the time value of money and the uncertainties present in the operation, including mortality rates, surrender and the behavior of the policyholder. The evaluation is made using a Monte Carlo simulation.

Such an evaluation depends on the evolution over time of unit-linked funds. So, in the next section we describe this evolution.

5.4.1

Evolution of the unit linked fund

We assume that the investment equity fund always keeps the same portfolio of assets and assume a frictionless market, with continuous trading, without tax and without transaction costs.

As in Ballotta and Haberman (2003), Ballotta and Haberman (2006), Graselli and Silla (2009), De Giovanni (2010) and others, the returns on the fund follow a geometric Brownian motion with constant drift and volatility, but under the real-world probability measure and considering that policyholders pay regular and additional premiums prior to the date of retirement. In addition, policyholders can choose to stop paying premiums and resume payments later, as well as surrender part of their funds and transfer resources of others plans to the unit-linked plan.

Since Brazilian funds usually invest the majority of their assets in Federal Government bonds, we have adopted a correlation between the financial return of the unit-linked fund and the variation in the risk-free short-term interest rate. The development of the fund depends, therefore, on the average and variance of the return of the fund, as well as the correlation. The process is defined in a filtered probability space $\left(\Omega, F, \{F_t\}_{t \geq 0}, P\right)$ and the dynamics under the objective measure is described by the following stochastic differential equation:

$$dS_t = \mu_S S_t dt + \sigma_S S_t dZ_t + I_{(t<T)} a_t d\pi_t + I_{(t<T)} b_t d\xi_t + I_{(t<T)} c_t d\gamma_t - I_{(t<T)} e_t d\delta_t - I_{(t\geq T)} f_t d\zeta_t$$

(5-1)

where

S_t is the investment fund at time t;

μ_S is the instantaneous expected return of the fund (drift), this is the real return;

σ_S is the instantaneous variance of the return of the fund;

dZ_t is the P-Brownian motion;

$I_{(A)}$ is an indicator function, assuming the unit value if A occurs and zero otherwise;

a_t is the regular premium, being zero after time T;

b_t is the additional premium, being zero after time T;

c_t is the amount (provision) transferred from another insurer or another plan of the same insurer to the fund;

e_t is the amount surrendered or transferred from the funds to another insurer or plan;

f_t is the partial surrender related to the self-annuitization process;

T is the time remaining to the predefined date of retirement in the contract;

$\pi_t, \xi_t, \gamma_t, \delta_t$ and ζ_t are independent Poisson processes;

$d\pi_t, d\xi_t, d\gamma_t, d\delta_t, d\zeta_t$ and dZ_t are independent; and

$d\pi_t, d\xi_t, d\gamma_t, d\delta_t$ and $d\zeta_t$ assume values equal to zero if the related Poisson event does not occur, and 1 (one) in case it does occur.

As in De Giovanni (2010), we assume that the short-term risk free interest rate evolves according to the classical model CIR (Cox et al., 1985), which is a known affine model with a factor and follows the stochastic differential equation below:

$$dr_t = \kappa(\mu_r - r_t)dt + \sigma_r \sqrt{r_t} dW_T$$

(5-2)

where

r_t is the short-term risk-free interest rate in time t;

μ_r is the central location or long term interest rate;

σ_r is the volatility;

κ is the speed of adjustment;

μ_r, σ_r and κ are constants, with $\kappa\mu_r \geq 0$ and $\sigma_r^2 > 0$;

Is imposed $2\kappa\mu_r \geq \sigma_r^2$ for non-negative rates;

dW_t is the P-Brownian motion correlated with dZ_t, such that $dW_t dZ_t = \rho dt$, and for $\rho \neq 0$ we have:

$$Z_t = \rho W_t + \sqrt{(1-\rho^2)} W'_t \text{; and} \tag{5-3}$$

dW'_t is a P-Brownian motion independent of dW_t.

The regular premiums (a_t) are fixed in the contract, but the policyholder can change the value at any time, by simply giving notice to the insurer. Yet, due to the interruption of payment option, policyholders can stop paying premiums and resume at any time. So, there is a probability that the regular premium is not paid within a specified period of time, or that there is no change in its value over the course of the contract. Given these characteristics, as in Coleman et al. (2006) and Coleman et al. (2007), we use a Poisson jump process to model the evolution of unit-linked funds. Thus, this payment process is modelled by adopting a model similar to the Merton jumps model (Merton, 1976) used by Wu and Yen (2007) to model real growth options. But, we assume that a_t, represents the amplitude of the jump, as an amount rather than as a percentage of the fund. In the model, we assume that the random variable has an independent lognormal distribution with parameters μ_a and σ_a. In turn, the intensity of the jump (π_t) is an independent Poisson process with parameters λ_a. Furthermore, we assume that the regular premiums can only be paid up to the time T.

We use the same strategy to model the additional premiums (b_t), given that the policyholder contributes these premiums as and when, and at a value they see fit. In the model, we consider that these premiums can only be invested until time T. So, we take the amplitude of the jump (b_t) as an independent lognormal distribution with parameters μ_b and σ_b, the intensity of the jump being (ξ_t) an independent Poisson process with parameter λ_b.

The process of fund transfer (portability) to the unit-linked fund can also be framed as a growth option and is equally modelled by means of a jump process. So, c_t has an independent lognormal distribution with parameters μ_c and σ_c, and γ_t is an independent Poisson process with parameter λ_c.

In turn, the surrender option allows policyholders to recover part of the funds or transfer it to another insurer or another plan of the same insurer before the guaranteed annuity option is exercised. This surrender or transfer, prior to the date of retirement, is modelled by a Poisson jump process, where e_t also has an independent lognormal distribution with parameters μ_e and σ_e, and δ_t is an independent Poisson process with parameter λ_e.

In the model, we can also assume that the mentioned amplitudes and intensities of jumps vary over time. For example, it is reasonable to assume that these parameters change when the amount of the fund grows, or when the retirement date nears.

However, after time T, when, in our approach, policyholders may decide their annuitization strategies, the surrenders are related to self-annuitization. The decision to exercise the option of guaranteed annuity is irreversible, making it complex. Hu and Scott (2007) and Milevsky and Young (2001) describe reasons for the policyholder avoiding annuitization, citing some other studies. The main reasons are bequest and the fear of illiquidity of assets. As an example, the last authors stated that only 2% of the amount invested in variable annuities in the US were converted into income during the period studied, according to National Association of Variable Annuities (June 30, 2001).

Hence, equation (5-1) permits policyholders to opt for self-annuitization. Each surrender lowers the value of the embedded options, considering the reduction in the unit-linked fund. These surrenders are modelled in the same way as the previous jumps, f_t being an independent lognormal distribution with parameters μ_f and σ_f, and ζ_t is an independent Poisson process with parameter λ_f.

The jump amplitudes can also be defined as a percentage of the fund. For example, one may assume the value of the partial surrender after T as a perpetual income until the optimal moment for exercising the guaranteed annuity option.

We emphasize that the parameters of the jumps depend on whether the annuity is under-priced or not, i.e. the technical bases of the contract and, mainly, the behavior of the policyholder. Thus, to assess the best estimate of the embedded options, those parameters must be established on the basis of the insurer's statistical and empirical evidence.

In the following sections, before the presentation of the model to evaluate embedded options, we describe how we address the uncertainties related to this assessment.

5.4.2
Mortality Risk

We work with mortality rates in the real world probability measure. First, we assume that the probability of the policyholder covered by the Brazilian unit-linked plans dying within a year is found in BR-EMS 2010 mortality table for survival coverage (Oliveira et al., 2012), which was constructed from the experience of the Brazilian insurance market.

To estimate the future longevity gains, we apply the multivariate structural model that uses the seemingly unrelated time series equation (SUTSE) proposed by Neves et al. (2016). This model estimates the longevity gains from a population with a short time series of observed mortality rates, which is the case of the Brazilian population. The model also admits that there is a population whose central rates of mortality show some similarity with those of the population studied. So, the authors used the concept of common trends, working with the idea that the mortality rates of the two populations are affected by common factors, as adopted by Li and Lee (2005), Jarner and Kryger (2011), and Li and Hardy (2011), Cairns et al. (2011) and Dowd et al. (2011). Furthermore, the structure of dependency between mortality rates of different ages is also captured by the covariance matrix of the errors in observation equation of the state space model.

In Neves et al. (2016), the longevity gain factors are established by the following expression:

$$G_{x,t+s} \sim \log normal\left(E\left(\log\left(m_{b,x,t+s}\right)\mid F_t\right) - \log\left(m_{b,x,t}\right), V\left(\log\left(m_{b,x,t+s}\right)\mid F_t\right)\right) \qquad (5\text{-}4)$$

where

t is the moment of forecast;

x is the age of the policyholder;

s is number of steps forward in the forecast;

$G_{x,t+s}$ is the longevity gain factor for age x of the Brazilian population in time $t+s$;

$m_{b,x,t}$ is the central rate of mortality for age x for the Brazilian population in time t; and

$E\left(\log\left(m_{b,x,t+s}\right)\mid F_t\right)$ and $V\left(\log\left(m_{b,x,t+s}\right)\mid F_t\right)$ are estimated in Neves et al. (2016).

As in Neves et al. (2016), here we assume that the longevity gain distribution of the policyholders of unit-linked plans is equal to that of the Brazilian population. Thus, we take an approximation to achieve the forecast of mortality rates of those policyholders, which also have lognormal distribution, as follows:

$$m_{x,t+s} = m_{x,t}.G_{x,t+s} \qquad (5\text{-}5)$$

where

$m_{x,t+s}$ is the distribution of the forecasted central rate of mortality for age x in time $t+s$; and

$m_{x,t}$ is the central rate of mortality for age x in time t, obtained by the formula below (Bowers et al.,1997):

$$m_{x,t} = \frac{q_{x,t}}{1 - \left(\frac{q_{x,t}}{2}\right)} \qquad (5\text{-}6)$$

where $q_{x,t}$ is the probability of a policyholder of age x, in time t, dying during a year, from the BR-EMS 2010 mortality table for survival coverage.

Therefore, to discount the cash flows and obtain the values of future annuities, given the information up to the evaluation date, we work with stochastic mortality rates. We assume that mortality rates are independent of financial risk and behavior of the policyholder. As a result, in the following formula, we define the probability that the policyholder of age $x+s$ will remain alive until age $x+s+z$:

$$_z p_{x+s} = P(\tau_{x+s} > z \mid F_t) = E\left(e^{-\int_0^z \mu_{x+s+i}\, d_i} \mid F_t \right) \cong E\left(\prod_{i=0}^{z-1} (1 - q_{x+s+i,s+i}) \mid F_t \right) \qquad (5\text{-}7)$$

where

x is the age of the policyholder on the date of evaluation t;

τ_{x+s} is a random variable that represents the remaining life time of a policyholder of age $x+s$;

μ_{x+s+i} is the force of mortality at age $x+s+i$; and

$q_{x+s+i,s+i}$ is the probability of the death of a policyholder of age $x+s+i$, scheduled for $s+i$ periods after t, obtained by applying the equations (5-5) e (5-6).

5.4.3
Policyholder Behavior

Policyholders of unit-linked plans may surrender their funds at some time after the purchase of the plan. The grace period is set in the contract. The PGBL plan allows tax deferral, so many uses the product to reduce their tax liabilities, i.e. they contribute to the plan in a calendar year and surrender some or all of their funds in another year. Another important characteristic is that, for the most part, these unit-linked products are marketed through the banking network, being sold as another financial investment fund rather than as a retirement plan. Due to these characteristics, policyholders can choose to exercise the option of abandoning regardless of the value of the guaranteed annuity option. To understand the dynamics of the surrender rates of the Brazilian annuity plans, Neves et al. (2014) proposed a multistage stochastic model to predict the surrender rates by means of a Monte Carlo simulation after executing the following processes: generalized linear models (GLM), ARMA-GARCH and multivariate copulas. In the GLM, assuming a logit link function, the explanatory variables are: risk free real short-term interest rate, gender and age. The GLM residuals for each age/gender group are then modelled by applying the ARMA-GARCH to generate i.i.d. residuals. After that, the dependency structure between these residuals is modelled by the multivariate Gaussian and T-student copulas. To explain the unusual and high

surrender rates observed during economic crises, the residuals from the ARMA-GARCH model adjusted to the returns of the Brazilian stock exchange index (Ibovespa) is used as one of the marginal distributions of multivariate copulas.

The article showed that the surrender rate odds for women is 1.35% greater than for men. Additionally, it can be seen in Table 5.4, from Neves et al. (2014), the odds for the four first age groups are greater than for the older group (72-80 years), and that the younger the group the greater the odds. Thus, the results showed that surrender rates are higher among women and younger policyholders. Therefore, analysing such a result, it is concluded that the policyholder behavior is far from optimal, which is the same conclusion taken by De Giovanni (2010), who affirmed that Kuo et al. (2003) and Kim (2005) presented strong evidence to support such a conclusion.

Table 5.4: Percentage change in the odds of age groups compared to older group (72-80 years), where w_x is the surrender rate for age x. Source: Neves et al. (2014).

Age Groups (years)	$\Delta\left(\dfrac{w_x}{1-w_x}\right)\%$
24-28	36.40%
29-53	19.87%
54-62	20.31%
63-71	9.19%

Table 5.5, which is also extracted from Neves et al. (2014), illustrates the sensitivity of the monthly variation of the surrender rate in relation to variation of the real short-term interest rate. The quoted article showed that when there is an increase in the interest rate, the surrender rate also rises. This behavior is related to the value of the guaranteed annuity option, given that when the interest rate increases there is a greater chance that the option is out of the money. Thus, there is a rational behavior in relation to the variation of the short-term interest rate.

Table 5.5: Percentage variation in odds given a variation in the real short-term interest rate. Source: Neves et al. (2014).

Δ Annual Interest Rate	$\Delta\left(\dfrac{w_x}{1-w_x}\right)\%$
2%	16.04%
1%	7.72%
0.75%	5.74%
0.50%	3.79%
0.25%	1.88%
-0.25%	-1.84%
-0.50%	-3.65%
-0.75%	-5.43%
-1%	-7.17%
-2%	-13.82%

In order to relate the decision to surrender with the financial situation, Neves et al. (2014) obtained the estimates for the values of the measures of dependence between the residuals of the Brazilian stock market index returns and the residuals of each age/gender group. Table 5.6 illustrates the results obtained by those authors.

Table 5.6: Measures of dependence among the Ibovespa's return residuals and residuals from each age/sex group. Source: Neves et al. (2014).

Age/Gender Group	Pearson's Correlation	Kendall's Tau	Spearman's Rho
24-28 / male	-0.078	-0.045	-0.058
29-53 / male	-0.159	-0.141	-0.187
54-62 / male	-0.405	-0.202	-0.274
63-71/ male	-0.506	-0.202	-0.299
72-80 / male	-0.514	-0.216	-0.307
24-28 / female	0.072	0.048	0.067
29-53 / female	0.042	0.054	0.080
54-62 / female	-0.104	-0.096	-0.155
63-71/ female	-0.141	-0.096	-0.131
72-80 / female	-0.135	-0.076	-0.124

These authors concluded that the negative correlations of Table 5.6 can be explained by noting that when there are negative returns in the Ibovespa, policyholders tend to switch their investments for products with less risk, increasing the surrender fees. The fact that the absolute values of the measures of dependence are higher for older ages is because groups that are close to retirement have less time to recover from financial losses incurred during the crisis and opt for surrender. Furthermore, the study shows that correlations are stronger for men. The relationship between financial crises and surrender rates was also studied by Loisel and Milhaud (2011). The results of these two studies showed that policyholders can make irrational decisions when faced with financial crises.

Based on the characteristics of the model used for prediction of the surrender rates, as well as De Giovanni (2010), we assumed a mix of a rational and irrational behavior of the policyholder in respect of the decision to surrender a plan. So, we apply the model of Neves et al. (2014) for a prediction of surrender rates. As required by Solvency II, the adopted model assumes that these rates are not independent of the financial market.

In our study, the surrender rates, obtained in the objective measure, are used to discount the values of embedded options at time T until the date of the evaluation of the best estimate. We determine the probability of the policyholder of age $x+s$ remaining in the plan until age $x+s+z$ according to the formula below:

$$_z h_{x+s} = E\left(e^{-\int_0^z w'_{x+s+i}, d_i} \mid F_t \right) \cong E\left(\prod_{i=0}^{z-1} \left(1 - w_{x+s+i,s+i}\right) \mid F_t \right) \tag{5-8}$$

where

w'_{x+s+i} is the force of surrender at the age $x+s+i$; and

w_{x+s+i} is the probability of policyholder of age $x+s+i$, $s+i$ times after the date of evaluation, surrenders within a year, which is obtained by applying Neves et al. (2014).

5.4.4
Annuitization

We assume that Brazilian policyholders, in general, take the decision to exercise the guaranteed annuity option on or after the date of retirement specified in contract. So, to evaluate the best estimate of the embedded options, our assumption is that policyholders will convert funds into income exactly on or after the time T.

Unlike Milevsky and Young (2001), Grasselli and Silla (2009) and Hu and Scott (2007), we do not apply the utility theory in the decision making process. Instead, our model assumes that policyholders will exercise the option of converting to income if the option is in the money at any time from T, considering the switch option, the option of deferral and the jump process related to self-annuitization. In our approach, the value of the embedded options at time T is the maximum value that the guaranteed annuity option reaches at T, taking into account the time value of money and assuming a similar approach to the Bellman equation or fundamental equation of optimality (Dixit and Pindyck, 1994).

We believe the bequest motive to avoid annuitizing the unit-linked fund is reduced because of the possibility to choose the type of annuity. Thus, in the

model, to be conservative, we make an assumption that policyholders choose the type of income that maximizes the value of the guaranteed annuity option. In addition, the model assumes that the policyholder can opt for self-annuitization for a predetermined period or perpetually. Thus, the fear of illiquidity of assets is also considered when we permit these surrenders after T.

As assumed before T, we admit that after that date the behavior of the policyholder is a combination of rational and irrational behavior. Rational because the policyholder chooses to convert the fund into income at the moment when the value of the option reaches its maximum, assuming the type of income that generates the highest value, and irrational because it admits the possibility of partial surrenders through jump processes. (see equation (5-1)) after T, regardless of whether the option is in the money or not.

5.4.5
Evaluation of Embedded Options

Initially, we obtain the best estimate in T. So, options of deferral and switching are evaluated at the same time as the guaranteed annuity option. Then, we find the best estimate at the time of the evaluation, taking into account the instantaneous expected return of the fund between that date and the date of annuitization, the probability of the policyholder remaining alive until the transformation into income and the probability of staying in the fund until the time T. Since the unit-linked plans allow the insurer to change the fund where the resource of the policyholder is applied after conversion into income, we use a forecast of a relevant risk-free short term interest rate term structure for calculation of annuities, obtained using the CIR model. We do not consider the hypothesis of financial surplus after conversion into income.

To obtain the value of the best estimate at the time T, we analyse the switching option, assuming that policyholders will opt for the type of income that maximizes the guaranteed annuity option. Due to the possibility of deferring the date of conversion, we need to measure that option of T until $(t + \omega - x)$, where ω is the last age for policyholders living in the plan. The value of the embedded options when $t + s \geq T$ is given by:

$$O'_{x,t+s} = \max_i \left[E\left(\frac{S_{t+s}}{g_{x+s,i}} a_{x+s,i} - S_{t+s} \mid F_t \right) \right]^+ \qquad (5\text{-}9)$$

where

$O'_{x,t+s}$ is the best estimate of the embedded options in time $t+s$ for a policyholder of age x on the date of evaluation t;

t is the date of evaluation;

S_{t+s} is obtained by equation (5-1);

$t+s \geq T$;

i is the type of annuity offered in the plan, $i = 1, 2, 3, \ldots$;

$g_{x+s,i}$ is the value of annuity of type i for age $x+s$ calculated considering the technical bases laid down in the plan; and

$a_{x+s,i}$ is the value of the annuity of type i for the age $x+s$, that is, the expected present value of the income of type i in time $t+s$ calculated considering realistic technical bases (premises).

We represent the main annuities offered in unit-linked plans, assuming that mortality rates and interest rates are independent, as follows:

a) Simple Life Annuity:

$$a_{x+s,i} = \sum_{j=1}^{\omega-(x+s)} \left(P(t+s, t+s+j)\, {}_j p_{x+s} \right) \qquad (5\text{-}10)$$

where

$P(t+s, t+s+j)$ is the expected price, in time $t+s$, of a unit of a risk free zero-coupon, with maturity date in $t+s+j$, under the real-world probability measure[22]. Since the CIR model is an affine process, this value is obtained by using the mathematical expressions given in Cox et al. (1985); and

${}_j p_{x+s}$ is obtained by the equation (5-7).

b) Temporary:

$$a_{x+s,i} = \sum_{j=1}^{n} \left(P(t+s, t+s+j)\, {}_j p_{x+s} \right) \qquad (5\text{-}11)$$

onde n is the period of income payment.

---

[22] We assume that the parameter related to the market risk of the CIR model is zero.

c) Certain Income:

$$a_{x+s,i} = \sum_{j=1}^{n} P(t+s,t+s+j) \qquad (5\text{-}12)$$

where n is the payment period of financial income.

d) With guaranteed minimum period:

$$a_{x+s,i} = \sum_{j=1}^{n} P(t+s,t+s+j) + \sum_{j=n+1}^{\omega-(x+s)} \left(P(t+s,t+s+j)_j p_{x+s} \right) \qquad (5\text{-}13)$$

where n is the period in which the financial income (certain income).

e) Reversion to an indicated beneficiary:

$$a_{x+s,i} = \sum_{j=1}^{\infty} \left(P(t+s,t+s+j) \left({}_j p_{x+s} + \beta \left({}_j p_{y+s} - {}_j p_{x+s} \, {}_j p_{y+s} \right) \right) \right) \qquad (5\text{-}14)$$

where

β is the reversion percentage; and

y is the age in t, of the indicated beneficiary.

f) Reversion to a spouse and minor children:

$$a_{x+s,i} = \beta \sum_{j=1}^{m} \left(P(t+s,t+s+j) \right) + (1-\beta) \sum_{j=1}^{m} \left(P(t+s,t+s+j)_j p_{x+s} \right) +$$
$$+ \sum_{j=m+1}^{\infty} \left[P(t+s,t+s+j) \left({}_j p_{x+s} + \beta \left({}_j p_{y+s} - {}_j p_{x+s} \, {}_j p_{y+s} \right) \right) \right] \qquad (5\text{-}15)$$

where

β is the reversion percentage;

y is the age of the spouse in t; and

m is the time remaining, in $t+s$, for the youngest child to reach adulthood.

To estimate the expected value, at time T, of expected future cash flows of embedded options, we have adopted an approach similar to the Bellman equation (Dixit and Pindyck, 1994), as follows:

$$O_{x,T} = \max_{(v \geq 0)} \left[{}_v p_{x+T} \, e^{-\mu_s v} O'_{x,T+v} \right] \qquad (5\text{-}16)$$

where

$O_{x,T}$ is the best estimate of the embedded options at time T for policyholder of age x on the date of evaluation t, considering the estimates at all times $T+v$; $v \in [0, \omega-(x+T-t)]$;

μ_s is the instantaneous expected return of the fund (see equation (5-1));

$_{v}p_{x+T}$ is obtained by equation (5-7);

$_{0}p_{x+T} = 1$; and

we assume that μ_S and the mortality rates are independent.

Finally, the best estimate for the value of embedded options on date of evaluation is obtained using the following formula:

$$V_{x,t} = O_{x,T}\; _{T}p_{x}\; _{T}h_{x}e^{-\mu_S T} \tag{5-17}$$

where

$V_{x,t}$ is the value of the best estimate of the embedded options on date of evaluation;

$_{T}p_{x}$ is obtained by equation (5-7);

$_{T}h_{x}$ is obtained by equation (5-8); and

we assume that the instantaneous expected return of the fund, mortality rates and surrender rates are independent.

5.5
Application of the model and sensitivity analysis

5.5.1
Definitions

In this section, we apply sensitivity analysis to the proposed model by varying its financial parameters, the technical bases of the plan, the individual characteristics of the policyholder and the parameters of the jumps processes, in order to address the problem presented. The base date of the calculations is January 2014.

In the Monte Carlo simulation, for the development of the unit-linked fund, we use the traditional method of Euler-Maruyama (Maruyama, 1955) for the approximate numerical solution of the stochastic differential equation of the fund (eq. (5-1)). We opted for working with annual discretization. To simulate the jump processes of eq. (5-1), we used the Bernoulli approach, first introduced by Ball and Tourous (1983). This premise permits each time interval to occur in one jump at the most.

For the simulation it is also necessary to define values for the instantaneous expected return - μ_S and for instantaneous variance of the fund - σ_S. At first, we chose to work with the average of the values obtained by the largest unit-linked funds operated by the three largest insurance companies in Brazil in the period between 2009 and 2013. The values found are $\mu_S = 0,04$ and $\sigma_S = 0.018$. We also assume a correlation - ρ of 0.90 (see eq.(5-3)), that is the average value of the correlations found between the three funds considered and the risk-free short-term interest rate. These values are sensitized in this section. To obtain the real returns of each fund, we use the National Consumer Price Index (IPCA), which measures the official Brazilian rate of inflation and is the most widely used for adjustments to the values of the unit-linked plans.

To estimate the parameters of the CIR model (eq. (5-2)), we apply the generalized moments methods in the form presented by Chan et al. (1992). We found an annual long-term interest rate (μ_r) of 5.5% p.a. Then we use a Monte Carlo simulation to predict risk free short-term interest rates for future periods. The estimation was based on a short-term rate of six months for the IPCA coupon curve, September 2003 to December 2013, obtained in the form presented by Franklin et al. (2012).

To predict the future longevity gains, we have adopted the multivariate structural model of Neves et al. (2016). In this article, the authors estimate the longevity gain from 2010, given that this uses the Brazilian population mortality data until 2009. The Appendix 7.1 presents the longevity factors referred to in that article for age groups of male and female. These factors are applied to central mortality rates from BR-EMS 2010 mortality table for survival coverage to find the probability of death in the year 2014 onwards (see section 5.4.2), assuming that the table reflects the survival of policyholders of unit-linked plans on the first day of 2010. From these results, we get the probability of survival expressed by equation (5-7), for any gender, age and time.

We can see the effect of the longevity gain for both genders by analyzing Figure 5.1, which shows the probability of a person of 60 years surviving 10 more years, considering various base dates. We can see the evolution over time of the probability of survival and that it is always greater for women.
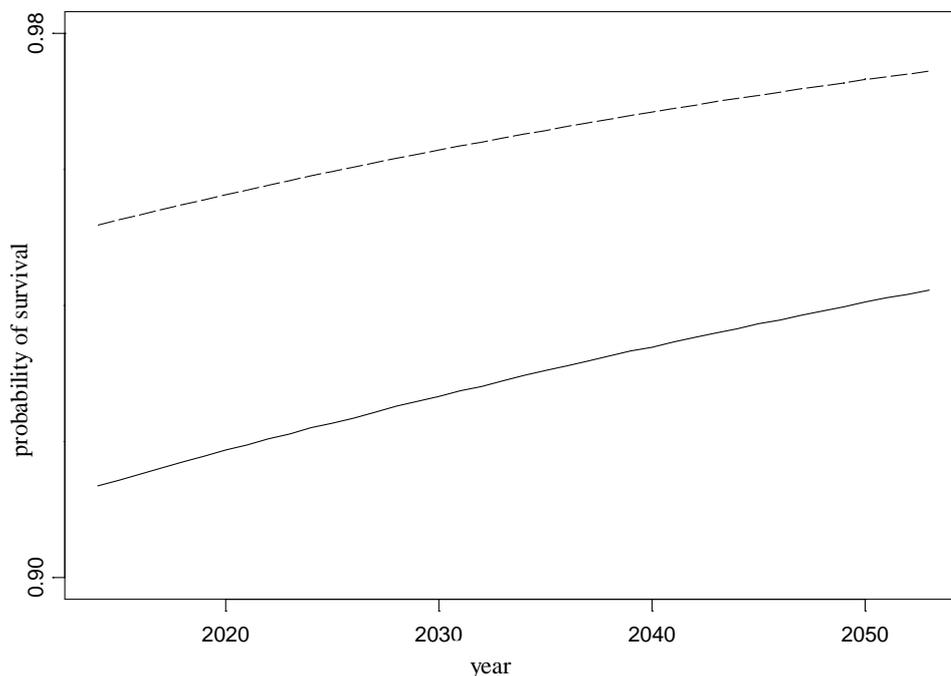
Figure 5.1. Evolution over time of the probability of a person of 60 years surviving 10 more years. Solid line for male and dashed line for female.

To obtain the surrender rates used to calculate the probability of the policyholder on the date of evaluation remaining in the plan until the date of retirement set in the contract (see eq. (5-8)), we applied the model of Neves et al. (2014). The database used consists of the monthly surrender rates from annuity plans of a relevant Brazilian insurer, from January 2006 until December 2011. For the purposes of this application, we take the premise that this data reflects cancelations of Brazilian unit-linked plans. Figure 5.2 presents the probability of people of different ages remaining in the plan until the default date of retirement, which for this illustration we define as 60 years, the base date of the evaluation. These probabilities are fairly low for younger people in the light of the high rates of cancelation.
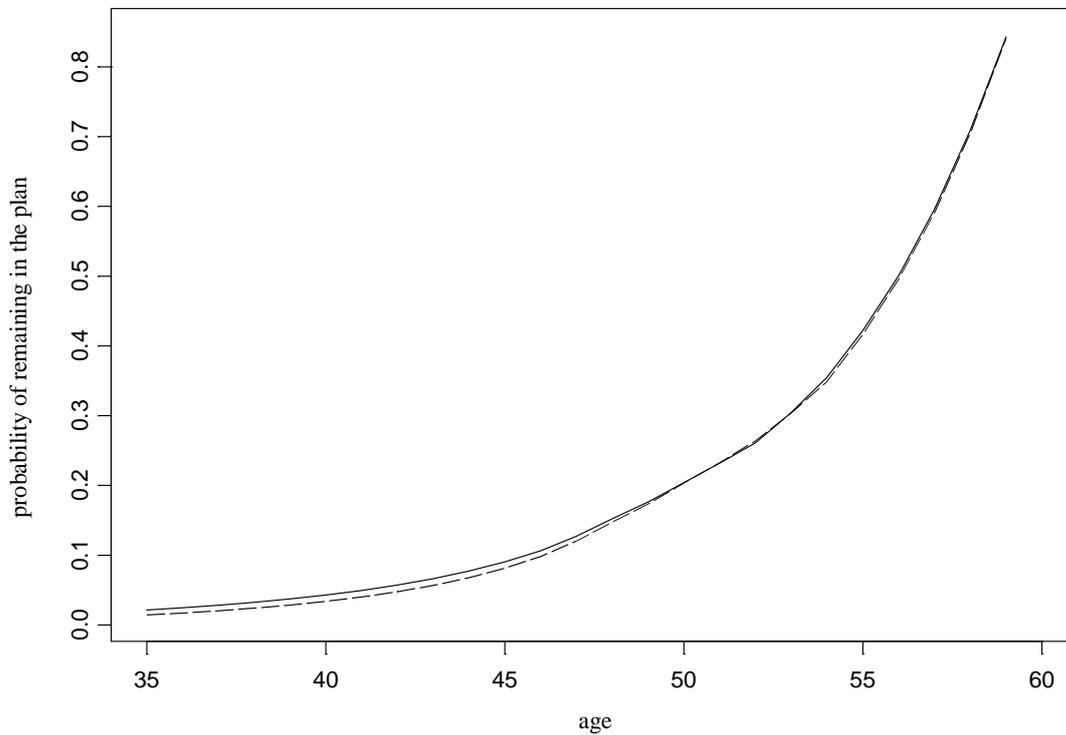
Figure 5.2. Probability of people of different ages not surrendering the plan until the predetermined date of retirement, in the base date of the evaluation. Solid line for male and dashed line for female.

For generalization, we assume that the calculation of the best estimate of the embedded options ($O'_{x,t+s}$) – see equation (5-9) - will be at the beginning of each year from the predetermined date of retirement.

5.5.2
Sensitivity Analysis

In this section, we present sensitivity analysis for parameters used in the calculation of the best estimate of the embedded options. At first, we vary the technical bases of the plan, i.e., mortality table and the rate of interest set out in the policy. To this end, we keep constant the other variables involved in the calculation.

a) Analysis 1:

Our analysis is made based on a reference and the following initial assumptions:

- age of the policyholder on the evaluation date = 40 years;
- gender = male;
- predetermined date of retirement = 60 years;
- unit-linked fund: $\mu_S = 0,04$, $\sigma_S = 0.018$ and $\rho = 0,90$;
- initial amount of the fund = R$ 60,000.00;
- for temporary, certain and guaranteed minimum income term: 15 years of temporality;
- for income reversion to the spouse: female beneficiary 3 years younger than the policyholder;
- for income reversion to the spouse and minors: female beneficiary 3 years younger than the policyholder, and youngest son needing four years to reach adulthood in the predetermined date of retirement; and
- jump process referring to regular contribution: probability of occurrence = 1; mean of the lognormal distribution for the value of this contribution = R$ 5,000.00 per year; and standard deviation of this lognormal distribution = R$ 500.00.

This first analysis does not consider the other jumps, including those relating to self-annuitization. We perform sensitivity tests for annual interest rates between 0% and 6%, the range permitted by Brazilian regulators, and for three mortality tables: AT 2000, AT 1983 and Annuity Mortality Table for 1949 (AT 1949), being that for the first two there is a distinction for gender.

Table 5.7 presents the ratio of the value of the best estimate of the options at the date of evaluation relative to the value of the initial amount of the unit-linked fund, considering various technical bases of plans. The higher the guaranteed interest rate, the higher the value of the best estimate of the embedded options and, therefore, the greater the commitment of the insurer to the policyholder. When we compared guaranteed mortality tables, the higher the mortality rates from the fixed mortality tables, the greater the value of the best estimate of the options. We emphasize that with the more conservative tables – with lower mortality rates (AT 2000), even for low interest rates guaranteed, there

is a need to provision a value pertaining to the embedded options. Therefore, in simulated situations, at some point the option of conversion into income will be in the money because of the longevity gain.

Table 5.7. Ratio between the value of the best estimate of the embedded options, on the date of evaluation, and the initial amount of the unit-linked fund, for different technical bases of plan.

Table	Interest Rate per year						
	0%	1%	2%	3%	4%	5%	6%
AT 2000	0.015%	0.107%	0.275%	0.486%	0.751%	1.079%	1.517%
AT 1983	0.175%	0.344%	0.559%	0.825%	1.137%	1.487%	1.937%
AT 1949	2.512%	2.730%	2.974%	3.214%	3.533%	3.846%	4.205%

It is important to note that the high values for the risk-free interest rate result in relatively low values for the best estimates of the options. These values would increase considerably in a scenario of decrease of the risk-free interest rate.

In Table 5.8, we compared the optimal age for conversion into income. It is clear that the higher the interest rate guaranteed by the unit-linked plan, the faster is this conversion. The decrease in age of conversion in relation to the guaranteed interest rate growth is not as pronounced in a mortality table with mortality rates further away from reality. However, in AT 1983 and AT 2000, when the guaranteed interest rate is 6% per year, there is a greater decrease in the optimal age of conversion, due to this rate exceeding the estimate of the risk-free long term interest rate.

At fairly low interest rates, we found extremely high ages of conversion into income. This is because we assume that the value of the best estimate of the embedded options is the maximum value that the guaranteed annuity option reaches from the predetermined date of retirement, taking into account the time value of money, the deferral option and the option to switch the type of income, and without taking into account any behavior factor.

Table 5.8. Optimal age for conversion into income.

Table	Interest Rate per year						
	0%	1%	2%	3%	4%	5%	6%
AT 2000	105	92	89	88	85	83	78
AT 1983	92	89	88	86	84	83	78
AT 1949	89	89	88	86	86	84	83

We emphasize that for all combinations of interest rates and mortality tables, the type of income that maximizes the value of embedded options is the lifetime annuity (simple annuity). We highlight, as an example, the plan that ensures AT 2000 and 3% per annum interest rate. In this plan, assuming the above premises, the age of annuity conversion will be 88 years old, where the value of the option is at its maximum value. We note that for some ages the option with the highest value is the income with reversion to an indicated beneficiary. But, when applied to equation (5-16) we obtain the maximum value in the conversion to annuity at 88 years. This dynamic happens in other combinations of tables and interest rates.

b) Analysis 2:

Now, we assume the premise of self-annuitization, i.e. include in the Monte Carlo simulation the jumps process that represents the partial surrender after the default date of retirement. It is quite reasonable to assume that after this date the policyholder defers his decision to exercise the option of conversion into income and applies the option of partial surrender. In this analysis, we assume that, while the optimal moment for conversion into income is not attained, each year the policyholder surrenders an income. The income is defined, for the purpose of this sensitivity test, as a perpetual financial income.

So, we add the following assumptions to analysis 1:

- Jump process for self-annuitization: probability of occurrence = 1; mean of the lognormal distribution for surrender each time $t = S_t \mu_s$; and a standard deviation of this lognormal distribution of 10% of the average each time t.

In Table 5.9, we can see the effect of this self-annuitization on the value of the ratio of the best estimate of the options and the initial value of the fund. This is repeated for various combinations of annual interest rates and mortality tables. Also in this case, the lifetime annuity was the option that presented the highest value. When the results are compared with those of Table 5.7, we see that, assuming the premise of partial surrender, the value of embedded options decreases.

Table 5.9. Ratio of the value of the best estimate of the embedded options, the date of evaluation, and the initial amount of the unit-linked funds to different technical bases of plan, assuming partial surrenders after the predetermined date of retirement.

Table	Interest Rate per year						
	0%	1%	2%	3%	4%	5%	6%
AT 2000	0.003%	0.031%	0.090%	0.176%	0.296%	0.525%	1.106%
AT 1983	0.051%	0.113%	0.201%	0.318%	0.476%	0.736%	1.355%
AT 1949	0.860%	0.969%	1.106%	1.265%	1.517%	1.891%	2.656%

In Figure 5.3, the ratio between the results presented in Table 5.9 and Table 5.7 is analyzed. As the average real return on the unit-linked fund used in our premise is of 4% per annum, from this guaranteed interest rate there is a sharp fall in the difference between the results with and without partial surrender. This is because from this value the guaranteed interest rate is greater than the profitability of the fund, and so not worth postponing the date of conversion into income and opt for partial surrender. This becomes clear when we compare Table 5.10 and Table 5.8, from the guaranteed interest rate of 4%, the difference between the optimal ages increase, because the partial surrender is of no interest.
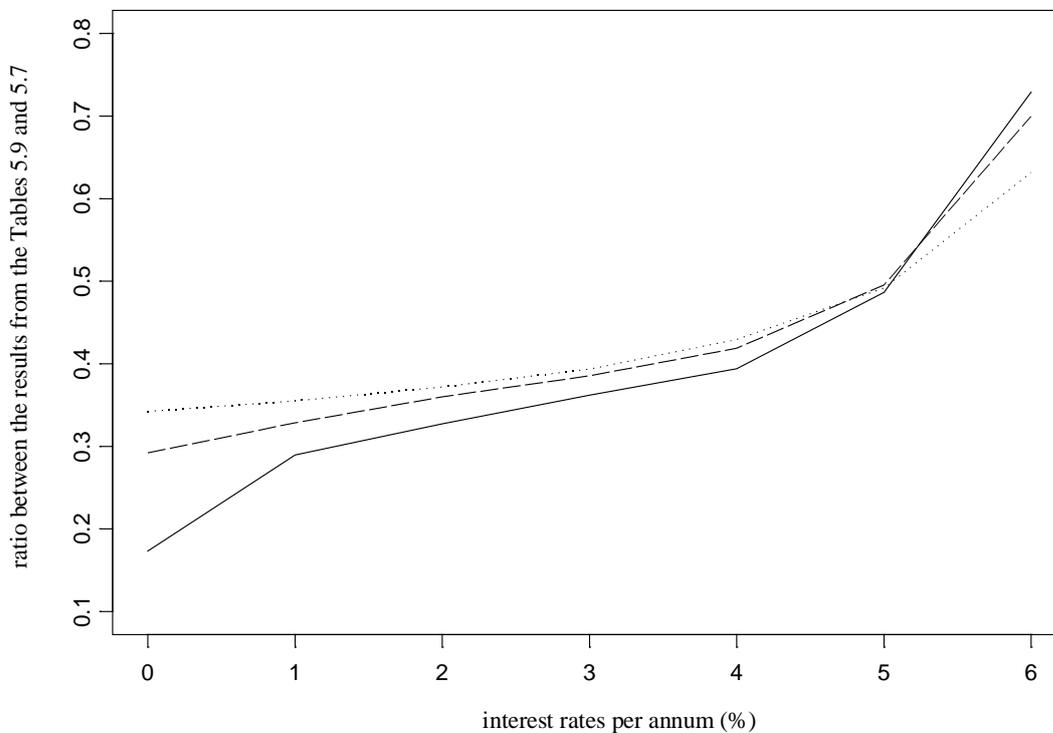
Figure 5.3: Ratio between the results of Table 5.9 and Table 5.7. Solid line for table AT 2000, dashed line for AT 1983 and dotted line for AT 1949.

Table 5.10. Optimal age for conversion into income, considering the premise of partial surrender.

Table	Interest Rate per year						
	0%	1%	2%	3%	4%	5%	6%
AT 2000	105	92	88	86	83	75	60
AT 1983	99	88	86	84	79	75	60
AT 1949	86	84	83	79	77	68	60

c) Analysis 3:

In this analysis, we varied the financial parameters of the unit-linked fund. To do this, we adopt a standard contract, which is a contract that guarantees AT 2000 and interest rate of 3%. We assume all other assumptions set out in the Analysis 2. In Table 5.11, we compare the ratio of the best estimate of the options with the value of the initial fund, for three different scenarios of expected returns for the fund. Logically, when we increased the fund's return there is a reduction in the value of the best estimate of the options, in view of it becoming less attractive to conversion into income. In all cases tested, the lifetime annuity is the highest value option.

Table 5.11. Ratio between the value of the best estimate of the embedded options on date of evaluation and the initial amount of the unit-linked funds, for different returns of unit-linked fund.

Expected Return of the Fund		
2%	4%	6%
0,325%	0,176%	0,096%

How we work with the expected value of the options, changes in values of the instantaneous return variance of the funds and the correlation between P-Brownian motion, does not affect the value of the best estimate of the embedded options. However, if the risk-based capital required for the company to ensure these options were measured, awareness of those parameters would affect the value of the capital.

d) Analysis 4:

In this analysis we vary some individual characteristics of the policyholder. With respect to the standard contract defined in Analysis 3, first we change the gender of the policyholder to female. In this case, the ratio between the

value of the best estimate of the embedded options and the initial amount of the unit-linked fund more than doubles, rising from 0.176% percent to 0.356%, given that the current rates of mortality of women are smaller, and longevity gains far greater. Thus, it is clear that for women the obligation of the insurer concerning embedded options of unit-linked plans is greater.

Now, let's consider a policyholder aged between 30 and 50 years, maintaining the retirement age of the referenced contract. In the first age, the ratio between the best estimate and the initial value goes from 0.176 percent to 0.042%, despite the value of the option on the default date of retirement being 64% greater than in the standard contract, given the greater amount of regular premiums paid. However, this smaller ratio is due to the financial discount and the probability of death and surrender applied to a longer period. For the policyholder of 50 years, the ratio is 0.570%, considering a much shorter period of application of the aforementioned probabilities and the financial discount.

We also altered the age of the beneficiary and spouse for income with reversion to the indicated beneficiary and income with reversion to the spouse and children, respectively. We tested the beneficiary/spouse with two different ages: 20 and 50 years. In these cases, the type of income which maximizes the values of the best estimates of the options is also the lifetime annuity, so no change in these values when compared with the reference standard contract. We increased the reversion to 100% and also there was no change. We found the same conclusion when we sensitized the temporality of income: temporary, guaranteed minimum period and certain, these tests were conducted for 5 and 25 years of temporality.

e) Analysis 5:

In this analysis, we change the initial value of the fund and the regular premium, assuming the reference contract and premises. First, we assume an initial fund equal to zero. In this case, the value of the best estimate of the options at the date of evaluation is R$ 55.32. With that same hypothesis, but assuming that the value of the mean of the regular premium is double the previous, the best estimate also doubles. This is because we are working with the expectations. The linear relationship between the mean of the premium and the best estimate of the options is evidenced.

We assume now that there will be no premium payment and the initial fund is the previously set (R$ 60,000.00). In this case, we will find the best estimate only in accordance with the obligations arising from the initial fund. The value found is R$ 49.27, i.e. 0.082% of the initial value of the fund. Then, after other simulations, assuming different values of regular premium and fixing the initial value of the fund, we have come to the conclusion that, given all the hypotheses assumed, for each R$ 1,000 of regular annual premium, the ratio between the best estimate of the options and initial fund increases by 0.019%. It is shown that the higher the premium paid and the initial fund, the greater the obligation of the insurer with embedded options.

Lastly, we have reduced the probability of payment of premium to 0.75. With that, the best estimate found is exactly equal to 75% of that coming from the regular premium, when we assume probability 1, more from the initial fund. Therefore, in conclusion, considering all the premises assumed, the ratio arising from regular premiums is equal to the probability of 0.019% times the probability of payment of the premium for each thousand reais of annual premium.

f) Analysis 6:

We analyze the jump processes pertaining to the additional premiums and transfers from another insurer or another plan of the same insurer to the fund. The methods of these processes are identical to those submitted for regular premiums. All these are related to the growth option defined in section 5.2.3. The jump process concerning the transfer of resources out of the fund before the default date of retirement also has the same effect, but with negative sign.

For all the jumps, the standard deviations of the amplitudes of the jumps are not relevant, because we work with the average of the simulated values, since we are interested in the best estimate.

For all combinations and assumptions adopted in our simulation, assuming the financial situation at the time of the evaluation, the type of income that maximizes the value of embedded options is the lifetime annuity. Therefore, we do not find value in the option of switching income type into a plan that uses the standard annuity. However, that does not mean it cannot occur in other scenarios. In other cases, the switching option contributes to increase the best estimate of the options.

In turn, the option of postponing the date of conversion into income increases the value of embedded options, given that the policyholder can opt for an optimal conversion date (see Tables 8 and 10). The growth option, defined in section 5.2.3, also increases the value of options, as seen in the analyses in this section. The option of partial surrender, presented in section 5.2.4, both before and after the default date of retirement, reduces the obligation of the insurer related to options. In turn, the option of payment interruption option, described in section 5.2.5, also contributes to reducing the best estimate of the options, because, according to this option, you must reduce the intensity of the jump regarding the payment of the regular premium.

5.6
Conclusion

We present a model for evaluating the value of embedded options in Brazilian unit-linked plans. The main features of these options have been properly described in the course of this work. We show that the Brazilian annuities market is incomplete and is not free of arbitrage, which was duly evidenced by the submission of three examples of arbitrage opportunity. Proven that there is no martingale measure in this market, we use the real-world probability measure in our modelling of the value of the best estimate of the embedded options. In this study, we detail the way we predict the variables involved in the simulation, as well as the premises assumed by the model.

We emphasize that the best estimate of the embedded options in unit-linked contracts must be properly provisioned by the insurer in order to ensure solvency. Thus, the proposed model can be widely used by companies in order to measure the liability resulting from the offer of these options in PGBL and VGBL contracts. In addition, some options described in section 5.2 are also offered in traditional contribution defined plans. Therefore, our model can be extended to these plans.

In the sensitivity tests performed in section 5.5, we analyze the effects of embedded options, and we study the influence of the technical bases of the contract on the value of the best estimate of the options. It is worth highlighting the importance of the adequate definition of the parameters used in the calculation

of the best estimate of the embedded options, which must be established by the actuary on the basis of statistics and empirical evidence of the insurer, with a clear objective to maintain the solvency of the company. The assumptions adopted are on the basis of a Monte Carlo simulation, therefore, the difficulty of having company data or empirical knowledge sufficient for determination of the premises limits the use of the model and, consequently, the assessment of the options. We emphasize also that the parameters of the jumps depend on whether the annuity is under-priced or not, i.e. the technical bases of the contract, and, mainly, the behavior of the policyholder.

It is also noted that a universe of falling interest rates would result in a reduction in the return of the unit-linked fund and the risk-free rate, making the embedded options more costly for companies. Regarding the management of the unit-linked funds, companies should intensify efforts to obtain high returns. This is because high returns, as well as attracting and retaining policyholders, give rise to a reduction in the value of the best estimate of the options, in view of the conversion into income being less advantageous to the policyholder when the return of the fund is attractive, assuming the self-annuitization hypothesis.

The results of these analyses, quantifying the value of arbitrage opportunities, especially concerning the portability between plans, show that the annuities market should study ways of protection against these risks in the contracts they sell. One must avoid that annuities becoming overpriced in the future. For this purpose, well-defined criteria of longevity gain must be stipulated in the contracts, as well the need to perform a correct estimation of the mortality table guaranteed in the plan. For the contracts in force, the actuaries must measure carefully the obligations coming from the possibility of portability in order to correctly calculate the corresponding obligations. One can, as protection against the risk of longevity, try to stimulate the creation of a transfer market for this risk, whether through insurance, reinsurance or issuing bonds or swaps.

As regards the other options offered in the contracts, so as to not leverage their risks, companies should impose limits in the policy on the values of regular and additional premiums. In addition, the insurer should be aware that offering the option to defer the date of retirement and the option of switching the type of annuity may result in costs.

The model proposed in this paper can serve as a base, after the appropriate adjustments, for a future work aimed at obtaining the distribution of the company's obligations with embedded options and, consequently, the value of the capital based on underwriting risk from these options.

Ultimately, the model can be extended assuming a utility function in the decision making process of the policyholder and by applying a methodology of optimization under uncertainty.

6
Conclusion

This thesis presents important contributions for the dynamic modelling of the risk factors present in life insurance and private pensions and for evaluating the embedded options of annuity products.

In the first essay, a multivariate SUTSE framework was proposed to forecast longevity gains for different age groups of a population. In sample and out-of-sample, our SUTSE model outperformed other relevant models found in the mortality rate literature, even when observing the cohort effects in the studied population. Furthermore, as it is a practical model, we believe that it can be used by insurers and pension plans to forecast mortality rates and, consequently, to evaluate their solvency.

In the second essay presented in chapter 3, a SUTSE model was applied to forecast the longevity gains for populations with a short time series of observed mortality rates, using a related population for which there exists long time series of mortality rates, through a common trends framework. Given the better statistical properties of the estimates obtained from our model when compared to those obtained from the well-known Lee-Carter method for a short time series of observed mortality rates, the former should be preferred to forecast mortality rates for a population with a short time series of observed mortality rates. Therefore, insurance companies and pension plans could also use the proposed model to forecast their own mortality rates using a relevant related population.

The third essay, in chapter 4, proposes a multi-stage stochastic model to forecast time series of surrender rates. It is important to note that in addition to presenting a model to predict the surrender rates, we have proposed a specific algorithm for simulation of elliptical copulas conditioned on a marginal distribution. Hence, from this multivariate model, it is possible to forecast surrender rates conditional on a specific financial stress scenario. The model is also useful to companies because they can apply stress tests using our model to simulate conditional copulas to analyze their capacity to manage portfolios during a financial crisis.

In the fourth essay presented in chapter 5, we show a model for evaluating the value of embedded options in Brazilian unit-linked plans. Our model can be widely used by companies in order to measure the liability resulting from the offer of these options in PGBL and VGBL contracts, as well as in traditional contribution defined plans. The model proposed in chapter 5 can serve as a base for future work aimed at obtaining the distribution of the company's obligations with embedded options and, consequently, the value of the calculation of risk-based capital subscription from these options. Moreover, the model can be extended assuming a utility function in the decision making process of the policyholder and by applying a methodology of optimization under uncertainty.

These essays, besides presenting relevant academic innovations, can be used by insurers and pension plans to model the longevity and surrender risks and to evaluate the embedded options. Because of the contributions presented in the essays, we believe that our models will be used as the basis for future studies in modelling insurance risks and actuarial science.

References

BALL, Clifford A.; TOROUS, Walter N. A simplified jump process for common stock returns. *Journal of Financial and Quantitative analysis*, v. 18, n. 01, p. 53-65, 1983.

BALLOTTA, Laura; HABERMAN, Steven. Valuation of guaranteed annuity conversion options. *Insurance: Mathematics and Economics*, v. 33, n. 1, p. 87-108, 2003.

BALLOTTA, Laura; HABERMAN, Steven. The fair valuation problem of guaranteed annuity options: The stochastic mortality environment case. *Insurance: Mathematics and Economics*, v. 38, n. 1, p. 195-214, 2006.

BARRIEU, Pauline et al. Understanding, modelling and managing longevity risk: key issues and main challenges. *Scandinavian actuarial journal*, v. 2012, n. 3, p. 203-231, 2012.

BIFFIS, Enrico; MILLOSSOVICH, Pietro. The fair value of guaranteed annuity options. *Scandinavian Actuarial Journal*, v. 2006, n. 1, p. 23-41, 2006.

BOWERS, Newton L. et al. Actuarial mathematics. Itasca, Ill.: Society of Actuaries, 1997.

BOYLE, Phelim; HARDY, Mary. Guaranteed annuity options. *Astin Bulletin*, v. 33, n. 02, p. 125-152, 2003.

BOYLE, G.; IRWIN, T. A primer on real options. *INFINZ Journal*, v. 1, p. 44-59, 2004.

BRIGO, Damiano; MERCURIO, Fabio. Interest Rate Models—Theory and Practice: With Smile, Inflation and Credit. Springer Finance, 2006.

BROOCK, W.A. et al. A Test for Independence Based on the Correlation Dimension. Department of Economics, University of Wisconsin at Madison, University of Houston, and University of Chicago, 1987.

CAIRNS, Andrew JG et al. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, v. 13, n. 1, p. 1-35, 2009.

CAIRNS, Andrew JG et al. Bayesian stochastic mortality modelling for two populations. *Astin Bulletin*, v. 41, n. 01, p. 29-59, 2011.

CAIRNS, Andrew JG. Modelling and management of longevity risk: Approximations to survivor functions and dynamic hedging. *Insurance: Mathematics and Economics*, v. 49, n. 3, p. 438-453, 2011.

CHAN, Kalok C. et al. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, v. 47, n. 3, p. 1209-1227, 1992.

CHERUBINI, Umberto; LUCIANO, Elisa; VECCHIATO, Walter. Copula methods in finance. John Wiley & Sons, 2004.

COLEMAN, Thomas F.; LI, Yuying; PATRON, Maria-Cristina. Hedging guarantees in variable annuities under both equity and interest rate risks. *Insurance: Mathematics and Economics*, v. 38, n. 2, p. 215-228, 2006.

COLEMAN, T. F. et al. Robustly hedging variable annuities with guarantees under jump and volatility risks. *Journal of Risk and Insurance*, v. 74, n. 2, p. 347-376, 2007.

CONSIGLIO, Andrea; DE GIOVANNI, Domenico. Pricing the option to surrender in incomplete markets. *Journal of Risk and Insurance*, v. 77, n. 4, p. 935-957, 2010.

COX, John C.; INGERSOLL, Jonathan E.; ROSS, Stephen A. A theory of the term structure of interest rates. *Econometrica*, v. 53, n. 2, p. 385-407, 1985.

COX, Samuel H.; LIN, Yijia. Annuity lapse rate modeling: Tobit or not tobit. Society of Actuaries, 2006.

CURRIE, Iain. Smoothing and forecasting mortality rates with P-splines. Talk given at the Institute of Actuaries, June 2006. Available at http://www.macs.hw.ac.uk/~iain/research/talks/London.2007.pdf

D'AMATO, Valeria et al. Modelling dependent data for longevity projections. *Insurance: Mathematics and Economics*, v. 51, n. 3, p. 694-701, 2012.

DE GIOVANNI, Domenico. Lapse rate modeling: a rational expectation approach. *Scandinavian Actuarial Journal*, v. 2010, n. 1, p. 56-67, 2010.

DE JONG, Piet; TICKLE, Leonie. Extending Lee–Carter mortality forecasting. *Mathematical Population Studies*, v. 13, n. 1, p. 1-18, 2006.

DE JONG, Piet; HELLER, Gillian. Generalized linear models for insurance data. Cambridge: Cambridge University Press, 2008.

DEHEUVELS, Paul. Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. *Publications de l'Institut de Statistique de l'Université de Paris*, v. 23, n. 3, p. 1-36, 1978.

DELWARDE, Antoine; DENUIT, Michel; PARTRAT, Christian. Negative binomial version of the Lee–Carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry*, v. 23, n. 5, p. 385-401, 2007.

DIEBOLD, Francis X.; MARIANO, Robert S. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, v. 20, n. 1, p. 134-144, 2002.

DIXIT, Avinash K.; PINDYCK, Robert S. Investment under uncertainty. *Princeton University Press*, 1994.

DOWD, Kevin et al. A gravity model of mortality rates for two related populations. *North American Actuarial Journal*, v. 15, n. 2, p. 334-356, 2011.

DURBIN, James; KOOPMAN, Siem Jan. Time series analysis by state space methods. *Oxford University Press*, 2012.

FANTAZZINI, Dean. Three-stage semi-parametric estimation of T-copulas: Asymptotics, finite-sample properties and computational aspects. *Computational Statistics & Data Analysis*, v. 54, n. 11, p. 2562-2579, 2010.

FERNÁNDEZ, F. Javier; HARVEY, Andrew C. Seemingly unrelated time series equations and a test for homogeneity. *Journal of Business & Economic Statistics*, v. 8, n. 1, p. 71-81, 1990.

FORNI, M. et al. The generalized dynamic factor model: one sided estimation and forecasting. *Journal of the American Statistical Association*, 100(471): 830-840, 2005.

FRANKLIN JR, Sergio L. et al. A estrutura a termo de taxas de juros no Brasil: modelos, estimação e testes. *Economia Aplicada*, v. 16, n. 2, p. 255-290, 2012.

FRENCH, Declan, O'HARE, Colin. A dynamic factor approach to mortality modeling. *Journal of Forecasting*, v. 32, n. 7, p. 587-599, 2013.

GAO, Quansheng, HU, Chengjun. Dynamic mortality factor model with conditional heteroskedasticity. *Insurance: Mathematics and Economics*, v. 45, n. 3, p. 410-423, 2009.

GATZERT, Nadine, KLING, Alexander. Analysis of participating life insurance contracts: a unification approach. *Journal of Risk and Insurance*, v. 74, n. 3, p. 547-570, 2007.

GIROSI, Federico; KING, Gary. A reassessment of the Lee-Carter mortality forecasting method. In: Harvard University Working Paper. 2005.

GRASSELLI, Matheus R.; SILLA, Sebastiano. A policyholder's utility indifference valuation model for the guaranteed annuity option. arXiv preprint arXiv:0908.3196, 2009.

HABERMAN, Steven; RENSHAW, Arthur. On age-period-cohort parametric mortality rate projections. *Insurance: Mathematics and Economics*, v. 45, n. 2, p. 255-270, 2009.

HABERMAN, Steven; RENSHAW, Arthur. A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, v. 48, n. 1, p. 35-55, 2011.

HABERMAN, Steven; RENSHAW, Arthur. Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, v. 50, n. 3, p. 309-333, 2012.

HÁRI, Norbert et al. Estimating the term structure of mortality. *Insurance: Mathematics and Economics*, v. 42, n. 2, p. 492-504, 2008.

HARRISON, J. Michael; PLISKA, Stanley R. A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes and their Applications*, v. 15, n. 3, p. 313-316, 1983.

HARVEY, Andrew C. Forecasting, structural time series models and the Kalman filter. *Cambridge University Press*, 1990.

HARVEY, Andrew C.; KOOPMAN, S.J.M. Multivariate structural time series models. *Systematic Dynamics in Economic and Financial Models*, p. 269-298, 1997.

HU, Wei-Yin; SCOTT, Jason S. Behavioral obstacles in the annuity market. *Financial Analysts Journal*, v. 63, n. 6, p. 71-82, 2007.

IAIS. Insurance Core Principals, Standards, Guidance, and Assessment Methodology. *International Association of Insurance Supervisor*, 2012. Available at www.iaisweb.org.

JAIN, Anil K.; MURTY, M. Narasimha; FLYNN, Patrick J. Data clustering: a review. *ACM computing surveys (CSUR)*, v. 31, n. 3, p. 264-323, 1999.

JARNER, Søren Fiig; KRYGER, Esben Masotti. Modelling adult mortality in small populations: The SAINT model. *Astin Bulletin*, v. 41, n. 02, p. 377-418, 2011.

KIM, Changki. Modeling surrender and lapse rates with economic variables. *North American Actuarial Journal*, v. 9, n. 4, p. 56-70, 2005.

KLING, Alexander; RUEZ, Frederik; RUß, Jochen. The impact of policyholder behavior on pricing, hedging, and hedge efficiency of withdrawal benefit guarantees in variable annuities. *European Actuarial Journal*, v. 4, n. 2, p. 281-314, 2014.

KNOLLER, Christian; KRAUT, Gunther; SCHOENMAEKERS, Pascal. On the propensity to surrender a variable annuity contract: an empirical analysis of dynamic policyholder behavior. *Journal of Risk and Insurance*, 2015.

KOLEV, Nikolai; ANJOS, Ulisses dos; MENDES, Beatriz Vaz de M. Copulas: a review and recent developments. *Stochastic Models*, v. 22, n. 4, p. 617-660, 2006.

KOOPMAN, Siem Jan. Exact initial Kalman filtering and smoothing for nonstationary time series models. *Journal of the American Statistical Association*, v. 92, n. 440, p. 1630-1638, 1997.

KOOPMAN, Siem Jan et al. STAMP 8: Structural Time Series Modeller and Predictor. London: Timberlake Consultants, 2007.

KOOPMAN, Siem Jan; SHEPHARD, Neil; DOORNIK, Jurgen A. Statistical Algorithms for Models in State Space Form: SsfPack 3.0. Timberlake Consultants Ltd, 2008.

KUO, Weiyu; TSAI, Chenghsien; CHEN, Wei Kuang. An empirical study on the lapse rate: the cointegration approach. *Journal of Risk and Insurance*, v. 70, n. 3, p. 489-508, 2003.

KUPIEC, P. Techniques for Verifiying the Accuracy of Risk Management Models. *Journal of Derviatives*, 1995.

LAZAR, Dorina; DENUIT, Michel M. A multivariate time series approach to projected life tables. *Applied Stochastic Models in Business and Industry*, v. 25, n. 6, p. 806-823, 2009.

LEE, Ronald D.; CARTER, Lawrence R. Modeling and forecasting US mortality. *Journal of the American statistical association*, v. 87, n. 419, p. 659-671, 1992.

LI, Johnny Siu-Hang; HARDY, Mary R. Measuring basis risk in longevity hedges. *North American Actuarial Journal*, v. 15, n. 2, p. 177-200, 2011.

LI, Nan; LEE, Ronald. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, v. 42, n. 3, p. 575-594, 2005.

LI, Nan; LEE, Ronald; TULJAPURKAR, Shripad. Using the Lee–Carter Method to Forecast Mortality for Populations with Limited Data. *International Statistical Review*, v. 72, n. 1, p. 19-36, 2004.

LI, Yong et al. Real options: Taking stock and looking ahead. *Advances in Strategic Management*, v. 24, n. 1, p. 31-66, 2007.

LOISEL, Stéphane; MILHAUD, Xavier. From deterministic to stochastic surrender risk models: impact of correlation crises on economic capital. *European Journal of Operational Research*, v. 214, n. 2, p. 348-357, 2011.

MARUYAMA, Gisiro. Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, v. 4, n. 1, p. 48-90, 1955.

MERTON, Robert C. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, v. 3, n. 1, p. 125-144, 1976.

MILEVSKY, Moshe A.; YOUNG, Virginia R. The real option to delay annuitization: it's not now-or-never. *York-Schulich-Finance Working Paper No. MM11-1*, 2001.

MILEVSKY, Moshe A.; SALISBURY, Thomas S. Financial valuation of guaranteed minimum withdrawal benefits. *Insurance: Mathematics and Economics*, v. 38, n. 1, p. 21-38, 2006.

MØLLER, Thomas. Hedging equity-linked life insurance contracts. *North American Actuarial Journal*, v. 5, n. 2, p. 79-95, 2001.

NEVES, César; FERNANDES, Cristiano; MELO, Eduardo. Forecasting Surrender Rates Using Elliptical Copulas and Financial Variables. *North American Actuarial Journal*, v. 18, n. 2, p. 343-362, 2014.

NEVES, César; FERNANDES, Cristiano; VEIGA, Álvaro. Forecasting longevity gains for a population with short time series using a structural SUTSE model: an application to Brazilian annuity plans. *North American Actuarial Journal*, 2016, in process.

OLIVEIRA, Mário et al. Tábuas biométricas de mortalidade e sobrevivência-experiência do mercado segurador brasileiro–2010. Rio Janeiro: Funenseg, 2012.

PATTON, Andrew J. Modelling asymmetric exchange rate dependence. *International Economic Review*, v. 47, n. 2, p. 527-556, 2006.

PLAT, Richard. On stochastic mortality modeling. *Insurance: Mathematics and Economics*, v. 45, n. 3, p. 393-404, 2009.

RENSHAW, Arthur E.; HABERMAN, Steven. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, v. 38, n. 3, p. 556-570, 2006.

RUIZ, Ricardo Machado. Estruturas urbanas comparadas: Estados Unidos e Brasil. *Estudos Econômicos* (São Paulo), v. 35, n. 4, p. 715-737, 2005.

SILVA, Fabiana Lopes da. Impacto do risco de longevidade em planos de previdência complementar. Tese de Doutorado. Universidade de São Paulo, 2010. Disponível: http://www.teses.usp.br/teses/disponiveis/12/12136/tde-29112010-182036/pt-br.php)

SUTCLIFFE, Charles. Trading death: The implications of annuity replication for the annuity puzzle, arbitrage, speculation and portfolios. *International Review of Financial Analysis*, v. 38, p. 163-174, 2015.

TSAI, Chenghsien; KUO, Weiyu; CHEN, Wei-Kuang. Early surrender and the distribution of policy reserves. *Insurance: Mathematics and Economics*, v. 31, n. 3, p. 429-445, 2002.

TSAI, Chenghsien. The Impacts of Surrender Options on Reserve Durations. *Risk Management and Insurance Review*, v. 15, n. 2, p. 165-184, 2012.

TSAY, Ruey S. Analysis of financial time series. John Wiley & Sons, 2005.

WU, Ming Cheng; YEN, Simon H. Pricing real growth options when the underlying assets have jump diffusion processes: the case of R&D investments. *R&D Management*, v. 37, n. 3, p. 269-276, 2007.

ZIVEYI, Jonathan; BLACKBURN, Craig; SHERRIS, Michael. Pricing European options on deferred annuities. *Insurance: Mathematics and Economics*, v. 52, n. 2, p. 300-311, 2013.

7
Appendix

7.1

Chapter 3

Table 7.1. Forecast of the expected longevity gain of the Brazilian male population.

years	<1	1-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	≥ 80
2010	0.9711	0.9631	0.9792	0.9762	0.9759	0.9844	0.9868	0.9916	0.9921	0.9920
2011	0.9411	0.9280	0.9555	0.9506	0.9498	0.9651	0.9722	0.9809	0.9851	0.9876
2012	0.9121	0.8943	0.9323	0.9257	0.9243	0.9461	0.9578	0.9703	0.9782	0.9833
2013	0.8839	0.8618	0.9098	0.9015	0.8996	0.9276	0.9437	0.9598	0.9713	0.9791
2014	0.8567	0.8306	0.8877	0.8779	0.8755	0.9094	0.9297	0.9494	0.9645	0.9750
2015	0.8303	0.8006	0.8662	0.8550	0.8520	0.8916	0.9160	0.9391	0.9577	0.9709
2016	0.8047	0.7718	0.8452	0.8327	0.8292	0.8742	0.9024	0.9290	0.9510	0.9669
2017	0.7799	0.7441	0.8247	0.8110	0.8071	0.8572	0.8891	0.9190	0.9443	0.9630
2018	0.7559	0.7175	0.8048	0.7899	0.7855	0.8405	0.8759	0.9091	0.9377	0.9592
2019	0.7326	0.6920	0.7853	0.7694	0.7645	0.8241	0.8630	0.8993	0.9311	0.9554
2020	0.7101	0.6675	0.7663	0.7495	0.7441	0.8082	0.8502	0.8896	0.9246	0.9517
2021	0.6882	0.6440	0.7477	0.7302	0.7243	0.7925	0.8377	0.8800	0.9181	0.9481
2022	0.6671	0.6215	0.7296	0.7114	0.7050	0.7772	0.8253	0.8705	0.9117	0.9446
2023	0.6465	0.5999	0.7120	0.6932	0.6863	0.7622	0.8131	0.8611	0.9054	0.9411
2024	0.6267	0.5793	0.6947	0.6754	0.6681	0.7476	0.8011	0.8518	0.8990	0.9377
2025	0.6074	0.5594	0.6779	0.6582	0.6503	0.7333	0.7893	0.8427	0.8928	0.9344
2026	0.5888	0.5405	0.6615	0.6415	0.6331	0.7193	0.7776	0.8336	0.8866	0.9312
2027	0.5707	0.5223	0.6455	0.6253	0.6164	0.7056	0.7661	0.8246	0.8804	0.9280
2028	0.5532	0.5049	0.6299	0.6096	0.6001	0.6922	0.7548	0.8158	0.8743	0.9250
2029	0.5363	0.4882	0.6147	0.5943	0.5843	0.6791	0.7437	0.8070	0.8682	0.9220
2030	0.5199	0.4723	0.5998	0.5795	0.5689	0.6663	0.7327	0.7983	0.8621	0.9191
2031	0.5040	0.4570	0.5853	0.5652	0.5540	0.6538	0.7219	0.7898	0.8562	0.9163
2032	0.4886	0.4424	0.5711	0.5512	0.5395	0.6416	0.7113	0.7813	0.8502	0.9135
2033	0.4737	0.4285	0.5573	0.5377	0.5254	0.6297	0.7008	0.7729	0.8443	0.9108
2034	0.4592	0.4152	0.5439	0.5246	0.5116	0.6181	0.6905	0.7646	0.8385	0.9083
2035	0.4452	0.4025	0.5307	0.5119	0.4983	0.6067	0.6803	0.7564	0.8327	0.9058
2036	0.4317	0.3903	0.5179	0.4995	0.4854	0.5956	0.6703	0.7483	0.8269	0.9033
2037	0.4185	0.3787	0.5054	0.4876	0.4728	0.5847	0.6604	0.7403	0.8212	0.9010
2038	0.4058	0.3676	0.4932	0.4760	0.4606	0.5741	0.6507	0.7324	0.8156	0.8988
2039	0.3935	0.3570	0.4813	0.4648	0.4487	0.5638	0.6411	0.7246	0.8099	0.8966
2040	0.3815	0.3469	0.4697	0.4539	0.4371	0.5537	0.6317	0.7168	0.8044	0.8945
2041	0.3700	0.3373	0.4584	0.4433	0.4259	0.5438	0.6224	0.7091	0.7988	0.8925
2042	0.3588	0.3281	0.4473	0.4331	0.4150	0.5342	0.6133	0.7016	0.7933	0.8906
2043	0.3479	0.3194	0.4365	0.4232	0.4044	0.5248	0.6042	0.6941	0.7879	0.8888
2044	0.3374	0.3111	0.4260	0.4136	0.3941	0.5156	0.5954	0.6867	0.7825	0.8870
2045	0.3272	0.3031	0.4157	0.4043	0.3841	0.5067	0.5866	0.6793	0.7771	0.8854
2046	0.3174	0.2956	0.4057	0.3954	0.3744	0.4980	0.5780	0.6721	0.7718	0.8838
2047	0.3078	0.2884	0.3959	0.3866	0.3650	0.4895	0.5695	0.6649	0.7665	0.8823
2048	0.2986	0.2816	0.3864	0.3782	0.3558	0.4812	0.5612	0.6578	0.7613	0.8809
2049	0.2896	0.2752	0.3771	0.3701	0.3469	0.4731	0.5529	0.6508	0.7561	0.8796
2050	0.2809	0.2691	0.3680	0.3622	0.3383	0.4652	0.5448	0.6439	0.7509	0.8784
2051	0.2725	0.2633	0.3592	0.3545	0.3299	0.4575	0.5368	0.6370	0.7458	0.8773
2052	0.2644	0.2578	0.3505	0.3471	0.3217	0.4500	0.5290	0.6302	0.7407	0.8763
2053	0.2565	0.2526	0.3421	0.3400	0.3138	0.4427	0.5212	0.6235	0.7357	0.8753
2054	0.2488	0.2477	0.3339	0.3331	0.3061	0.4356	0.5136	0.6169	0.7307	0.8745
2055	0.2414	0.2431	0.3259	0.3264	0.2986	0.4286	0.5061	0.6104	0.7257	0.8738
2056	0.2343	0.2388	0.3180	0.3199	0.2913	0.4219	0.4987	0.6039	0.7208	0.8731
2057	0.2273	0.2348	0.3104	0.3137	0.2842	0.4153	0.4914	0.5975	0.7159	0.8725
2058	0.2206	0.2310	0.3030	0.3076	0.2774	0.4089	0.4842	0.5911	0.7111	0.8721
2059	0.2141	0.2274	0.2957	0.3018	0.2707	0.4026	0.4772	0.5848	0.7063	0.8717

Table 7.2. Forecast of the expected longevity gain of the Brazilian female population.

years	<1	1-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	≥ 80
2010	0.9576	1.0030	1.0117	0.9725	0.9778	0.9777	0.9833	0.9867	0.9857	0.9911
2011	0.9202	0.9569	0.9683	0.9410	0.9483	0.9534	0.9641	0.9705	0.9724	0.9844
2012	0.8843	0.9130	0.9269	0.9105	0.9198	0.9297	0.9452	0.9545	0.9593	0.9778
2013	0.8499	0.8711	0.8872	0.8810	0.8921	0.9065	0.9267	0.9387	0.9463	0.9713
2014	0.8167	0.8311	0.8494	0.8524	0.8653	0.8840	0.9085	0.9233	0.9336	0.9649
2015	0.7849	0.7930	0.8132	0.8248	0.8393	0.8620	0.8907	0.9080	0.9210	0.9585
2016	0.7543	0.7566	0.7786	0.7981	0.8141	0.8405	0.8733	0.8931	0.9086	0.9523
2017	0.7249	0.7219	0.7456	0.7722	0.7896	0.8196	0.8561	0.8784	0.8964	0.9460
2018	0.6967	0.6889	0.7140	0.7472	0.7659	0.7993	0.8394	0.8639	0.8843	0.9399
2019	0.6695	0.6573	0.6839	0.7230	0.7430	0.7794	0.8229	0.8497	0.8724	0.9338
2020	0.6435	0.6273	0.6552	0.6996	0.7207	0.7601	0.8068	0.8357	0.8607	0.9279
2021	0.6184	0.5986	0.6278	0.6769	0.6992	0.7412	0.7910	0.8219	0.8492	0.9220
2022	0.5944	0.5713	0.6016	0.6550	0.6783	0.7228	0.7756	0.8084	0.8378	0.9161
2023	0.5712	0.5452	0.5766	0.6338	0.6580	0.7049	0.7604	0.7950	0.8266	0.9104
2024	0.5490	0.5204	0.5528	0.6133	0.6384	0.6874	0.7455	0.7820	0.8155	0.9047
2025	0.5277	0.4967	0.5301	0.5934	0.6194	0.6704	0.7309	0.7691	0.8046	0.8992
2026	0.5072	0.4741	0.5084	0.5742	0.6010	0.6538	0.7166	0.7564	0.7939	0.8937
2027	0.4875	0.4526	0.4877	0.5556	0.5832	0.6376	0.7026	0.7440	0.7833	0.8883
2028	0.4685	0.4320	0.4680	0.5376	0.5659	0.6218	0.6889	0.7317	0.7728	0.8829
2029	0.4503	0.4124	0.4492	0.5202	0.5491	0.6064	0.6754	0.7197	0.7625	0.8777
2030	0.4329	0.3938	0.4313	0.5033	0.5329	0.5915	0.6622	0.7079	0.7523	0.8725
2031	0.4161	0.3760	0.4142	0.4870	0.5172	0.5769	0.6493	0.6962	0.7423	0.8675
2032	0.3999	0.3590	0.3979	0.4713	0.5020	0.5626	0.6366	0.6848	0.7324	0.8625
2033	0.3844	0.3428	0.3824	0.4560	0.4872	0.5488	0.6242	0.6735	0.7227	0.8576
2034	0.3695	0.3274	0.3676	0.4412	0.4730	0.5353	0.6120	0.6624	0.7131	0.8527
2035	0.3552	0.3127	0.3535	0.4270	0.4591	0.5221	0.6001	0.6516	0.7036	0.8480
2036	0.3415	0.2986	0.3400	0.4131	0.4457	0.5093	0.5884	0.6408	0.6943	0.8434
2037	0.3283	0.2853	0.3272	0.3998	0.4328	0.4967	0.5769	0.6303	0.6851	0.8388
2038	0.3156	0.2725	0.3150	0.3868	0.4202	0.4846	0.5657	0.6200	0.6760	0.8344
2039	0.3034	0.2603	0.3034	0.3743	0.4080	0.4727	0.5547	0.6098	0.6671	0.8300
2040	0.2917	0.2487	0.2923	0.3622	0.3962	0.4611	0.5439	0.5998	0.6582	0.8257
2041	0.2804	0.2377	0.2817	0.3505	0.3848	0.4498	0.5333	0.5899	0.6495	0.8215
2042	0.2696	0.2271	0.2717	0.3391	0.3738	0.4388	0.5229	0.5802	0.6410	0.8174
2043	0.2592	0.2171	0.2621	0.3281	0.3631	0.4281	0.5127	0.5707	0.6325	0.8133
2044	0.2493	0.2075	0.2530	0.3175	0.3527	0.4177	0.5028	0.5614	0.6242	0.8094
2045	0.2397	0.1983	0.2443	0.3072	0.3426	0.4075	0.4930	0.5522	0.6160	0.8056
2046	0.2305	0.1896	0.2360	0.2973	0.3329	0.3976	0.4834	0.5431	0.6078	0.8018
2047	0.2216	0.1813	0.2281	0.2877	0.3235	0.3879	0.4740	0.5342	0.5998	0.7982
2048	0.2131	0.1733	0.2206	0.2784	0.3144	0.3785	0.4648	0.5254	0.5920	0.7946
2049	0.2049	0.1657	0.2134	0.2694	0.3055	0.3693	0.4558	0.5168	0.5842	0.7911
2050	0.1971	0.1585	0.2066	0.2606	0.2970	0.3604	0.4470	0.5084	0.5765	0.7878
2051	0.1895	0.1516	0.2002	0.2522	0.2887	0.3517	0.4383	0.5000	0.5690	0.7845
2052	0.1823	0.1450	0.1940	0.2440	0.2806	0.3432	0.4298	0.4918	0.5615	0.7813
2053	0.1753	0.1388	0.1881	0.2362	0.2729	0.3349	0.4215	0.4838	0.5542	0.7782
2054	0.1686	0.1328	0.1825	0.2285	0.2653	0.3269	0.4133	0.4759	0.5469	0.7752
2055	0.1622	0.1271	0.1772	0.2211	0.2581	0.3190	0.4053	0.4681	0.5398	0.7723
2056	0.1560	0.1216	0.1722	0.2140	0.2510	0.3114	0.3975	0.4604	0.5327	0.7695
2057	0.1501	0.1164	0.1674	0.2070	0.2441	0.3039	0.3898	0.4529	0.5258	0.7668
2058	0.1444	0.1114	0.1628	0.2003	0.2375	0.2966	0.3823	0.4455	0.5189	0.7642
2059	0.1389	0.1067	0.1585	0.1939	0.2311	0.2896	0.3749	0.4382	0.5122	0.7616

7.2

Chapter 4

7.2.1

Copula

The dependence structure among the random variables $X_1, \ldots, X_n$, which in our approach are i.i.d. residuals from fitted models, is described by their joint cumulative distribution function $F(X_1, \ldots, X_n)$ with marginal distribution function $F_i(X_i)$, $(i = 1, \ldots, n)$, where the generalized inverse of a distribution function is given by:

$$F_i^{-1}(t) = \inf\left\{ u : F_i(u) \geq t, \, 0 < t < 1 \right\} \tag{7-1}$$

By formal definition, an n-dimensional copula is a function $C : [0,1]^n \to [0,1]$ with the following properties:

-   C is grounded, meaning that for every $u = (u_1, \ldots u_n) \in [0,1]^n$, $C(u) = 0$ if at least one coordinate $u_i = 0, i = 1, 2, \ldots, n$;

-   C is n-increasing, i.e., for every $u \in [0,1]^n$ and $v \in [0,1]^n$ such that $u \leq v$, the C-volume $V_C([u,v])$ of the box $[u,v]$ is non-negative; and

-   $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $u_i \in [0,1], i = 1, 2, \ldots, n..$

The n-dimensional extension of Sklar's theorem guarantees that every joint distribution can be represented as a unique copula if the marginals are continuous. Then, by that theorem, the cumulative joint distribution can be written as functions of marginal distributions:

$$F(X_1, X_2, \ldots, X_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) = C(u_1, u_2, \ldots, u_n) \tag{7-2}$$

And, consequently:

$$C(u_1, u_2, \ldots, u_n) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)\right) \tag{7-3}$$

The density of the copula can be related to the density of the distribution *F,* for continuous random variables, by the following expression:

$$f(x_1, x_2, \ldots, x_n) = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \cdot \prod_{j=1}^{n} f_j(x_j) \tag{7-4}$$

where $c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) = \dfrac{\partial^n C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))}{\partial F_1(x_1) \partial F_2(x_2) \cdots \partial F_n(x_n)}$.

In this paper, we tested two multivariate elliptical copulas: Gaussian and Student's t. The multivariate Gaussian copula is defined by the following formula:

$$C_R^G(u_1, u_2, \ldots, u_n) = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right) \tag{7-5}$$

where Φ_R is the standardized multivariate normal distribution with a symmetric and positive definite matrix R, which contains the full set of parameters of the MGC. The density function of this copula is given by:

$$c_R^G(u_1, u_2, \ldots, u_n) = \frac{1}{|R|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \varsigma^T (R^{-1} - I) \varsigma\right) \tag{7-6}$$

with $\varsigma = \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right)^T$.

By these terms, the multivariate Student's t copula, which has two sets of parameters (matrix R and the common degree of freedom υ) and presents dependence in the tails, is defined by:

$$T_{R,\upsilon}(u_1, u_2, \ldots, u_n) = t_{R,\upsilon}\left(t_\upsilon^{-1}(u_1), t_\upsilon^{-1}(u_2), \ldots, t_\upsilon^{-1}(u_n)\right) \tag{7-7}$$

where $t_{R,\upsilon}$ is the standardized multivariate Student's t distribution with correlation matrix R and υ parameters. The related density function is given by:

$$c_{R,\upsilon}(u_1, u_2, \ldots, u_n) = \frac{1}{|R|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{\upsilon+n}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} \left(\frac{\Gamma\left(\frac{\upsilon}{2}\right)}{\Gamma\left(\frac{\upsilon+1}{2}\right)}\right)^n \frac{\left(1 + \frac{1}{\upsilon} \varsigma^T R^{-1} \varsigma\right)^{-\frac{\upsilon+n}{2}}}{\prod_{j=1}^n \left(1 + \frac{\varsigma_j^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}}} \tag{7-8}$$

where $\varsigma_j = t_\upsilon^{-1}(u_j)$.

Following Cherubini et al. (2004), as a consequence of Skalar's theorem, in the bivariate case, the conditional copula is given by:

$$P(X \le x \mid Y = y) = F_{X|Y}(x \mid y) = C_{X|Y}(F_X(x), F_Y(y)) = \frac{\partial(C(F_X(x), F_Y(y)))}{\partial F_Y(y)} \tag{7-9}$$

Then, for a multivariate copula conditioned on one marginal distribution, we have the following joint cumulative distribution function:

$$F(X_2,\ldots,X_n \mid X_1) = C_{X_2,\ldots,X_n\mid X_1}(F_1(x_1), F_2(x_2),\ldots, F_n(x_n)) = \frac{\partial(C(F_1(x_1), F_2(x_2),\ldots, F_n(x_n)))}{\partial F_1(x_1)}$$

(7-10)

More detailed information on conditional copulas can be found in Patton (2006) and Kolev et al. (2006).

7.2.2

Empirical estimation of different measures of dependence

Table 7.1. Person's correlation matrix.

Residuals	Ibovespa returns	24-28 male	29-53 male	54-62 male	63-71 male	72-80 male	24-28 female	29-53 female	54-62 female	63-71 female	72-80 female
Ibovespa returns	1	-	-	-	-	-	-	-	-	-	-
24-28 male	-0.078	1	-	-	-	-	-	-	-	-	-
29-53 male	-0.159	0.902	1	-	-	-	-	-	-	-	-
54-62 male	-0.405	0.641	0.822	1	-	-	-	-	-	-	-
63-71 male	-0.506	0.331	0.514	0.789	1	-	-	-	-	-	-
72-80 male	-0.514	0.279	0.417	0.714	0.952	1	-	-	-	-	-
24-28 female	0.072	0.066	0.0004	-0.124	-0.178	0.140	1	-	-	-	-
29-53 female	0.042	-0.100	-0.070	-0.109	0.108	-0.099	0.761	1	-	-	-
54-62 female	-0.104	-0.216	-0.101	-0.042	0.035	0.026	0.521	0.852	1	-	-
63-71 female	-0.141	-0.137	-0.096	-0.050	0.007	0.027	0.357	0.554	0.772	1	-
72-80 female	-0.135	-0.114	-0.097	-0.065	-0.028	0.017	0.294	0.401	0.622	0.906	1

Table 7.2. Kendall's tau matrix.

Residuals	Ibovespa returns	24-28 male	29-53 male	54-62 male	63-71 male	72-80 male	24-28 female	29-53 female	54-62 female	63-71 female	72-80 female
Ibovespa returns	1	-	-	-	-	-	-	-	-	-	-
24-28 male	-0.045	1	-	-	-	-	-	-	-	-	-
29-53 male	-0.141	0.625	1	-	-	-	-	-	-	-	-
54-62 male	-0.202	0.402	0.633	1	-	-	-	-	-	-	-
63-71 male	-0.202	0.131	0.394	0.471	1	-	-	-	-	-	-
72-80 male	-0.216	0.143	0.325	0.397	0.664	1	-	-	-	-	-
24-28 female	0.048	0.027	0.014	-0.001	-0.089	-0.0497	1	-	-	-	-
29-53 female	0.054	-0.032	-0.002	0.011	0.004	-0.001	0.648	1	-	-	-
54-62 female	-0.096	-0.194	-0.084	0.023	0.111	0.065	0.374	0.619	1	-	-
63-71 female	-0.096	-0.097	-0.048	0.012	0.066	0.047	0.195	0.357	0.484	1	-
72-80 female	-0.076	-0.046	-0.063	-0.020	-0.011	-0.001	0.085	0.190	0.284	0.546	1

Table 7.3. Spearman's rho matrix.

Residuals	Ibovespa returns	24-28 male	29-53 male	54-62 male	63-71 male	72-80 male	24-28 female	29-53 female	54-62 female	63-71 female	72-80 female
Ibovespa returns	1	-	-	-	-	-	-	-	-	-	-
24-28 male	-0.058	1	-	-	-	-	-	-	-	-	-
29-53 male	-0.187	0.809	1	-	-	-	-	-	-	-	-
54-62 male	-0.274	0.567	0.812	1	-	-	-	-	-	-	-
63-71 male	-0.299	0.198	0.535	0.629	1	-	-	-	-	-	-
72-80 male	-0.307	0.210	0.466	0.558	0.842	1	-	-	-	-	-
24-28 female	0.067	0.045	0.029	-0.001	-0.136	-0.082	1	-	-	-	-
29-53 female	0.080	-0.052	0.001	0.022	0.016	0.000	0.839	1	-	-	-
54-62 female	-0.155	-0.303	-0.1258	0.045	0.169	0.108	0.497	0.779	1	-	-
63-71 female	-0.131	-0.159	-0.071	0.014	0.107	0.063	0.279	0.491	0.634	1	-
72-80 female	-0.124	-0.069	-0.096	-0.043	-0.013	-0.006	0.128	0.256	0.373	0.737	1