STUDIUM UNIVERSITATIS BERGOMENSIS
UNIVERSITY OF BERGAMO

Doctor of Philosophy in
ANALYTICS FOR ECONOMICS AND BUSINESS
XXX Cycle

An Essay on Longevity and Pension Systems

Co–ordinator:
Marida Bertocchi †
Adriana Gnudi

Supervisor:
Francesco Menoncin

Candidate:
Michele TOGNAZZO
Candidate ID: 1036280

To my mother who supports me every day.
To my father, he can now remain close to us every day.
To Angela and Alessandro my new family.

Acknowledgment

I wanted to thank Prof. Menoncin for the constant and indispensable help in the realization of the thesis. For two and a half years I have seen him devote his time to his greatest passions: teaching and research. A special thanks to Prof. Regis for having dedicated his time for this work and to Prof. Devolder for the help and hospitality given to me at the UCL.

Contents

1

An Evolving Framework

1.1 The demographic environment

1.1.1 The fertility rate

The fertility rate is one of the most important demographic indicators that we need to take into consideration in order to describe the evolution of a country population's size and distribution. In the past decades the fertility rates of the OECD countries are dramatically declined. Assuming no net migration and unchanged mortality, a total fertility rate of 2.1 ensures a broadly stable population (OECD site). Considering the trends of the fertility rates of some European countries from 1960 to 2015 plotted in figure 1.1 we can see that starting from 1977 all the European countries considered have the fertility rates lower than 2.1.

The decreasing value of fertility rates is a common factor influencing the populations all over the world. The force of this phenomena has not the same intensity as described by the figures 1.2 and 1.3. The common factor is that the fertility rates drop of more than 1, with some exceptions, in all the developed countries all over the world and they drop of 2 points in the less developed countries. The starting points are different: in the most industrialized countries the upper bound of the fertility rates was around 3 in 1960 and it dropped reaching 1.9 in 2015, on the other side, the less developed countries had a lower bound of the fertility rates equal to 6.32 and in 2015 it reached 4.32.

1.1.2 The life expectation

In this section we underline the evolution of the second demographic trend that we need to take into consideration to outline the problem.

In figures 1.4 and 1.5 we present the behavior of the life expectancy at birth from 1872 to 2014 for men, women and their weighted average for the Italian and Belgian populations.

The increasing in the life expectation at birth is quite clear. In their work Oeppen and Vaupel [2002] underline that the female life expectancy has been increasing for 160

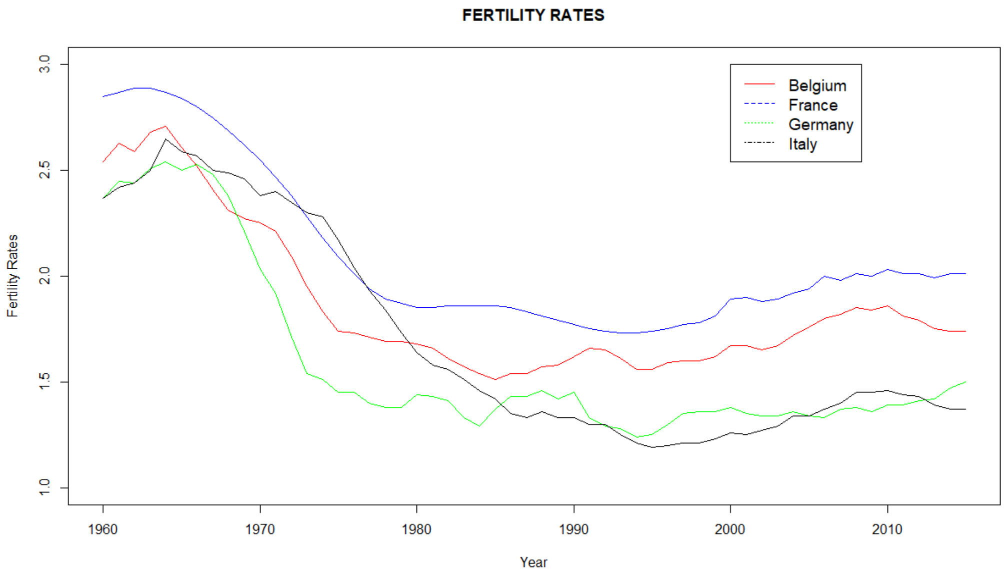

Figure 1.1 Fertility rates from 1960 to 2015. Source: The World Bank

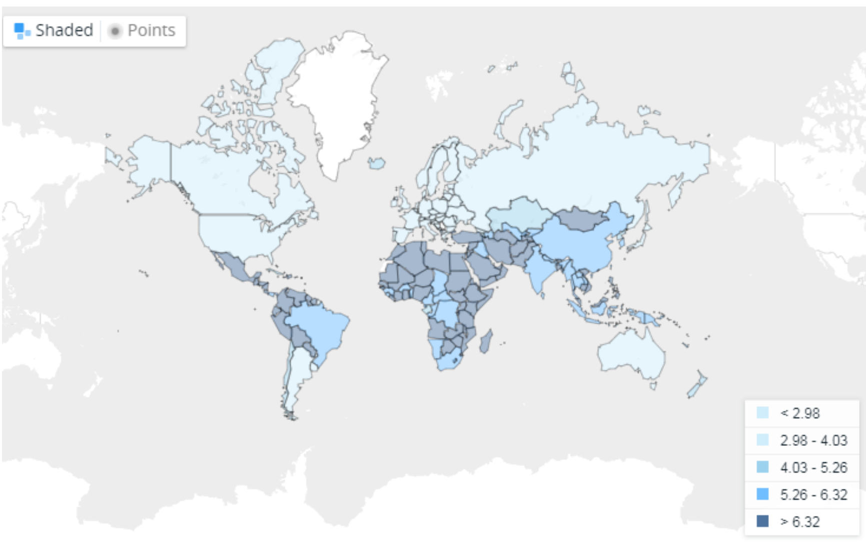

Figure 1.2 Fertility rates in 1960. Source: The World Bank

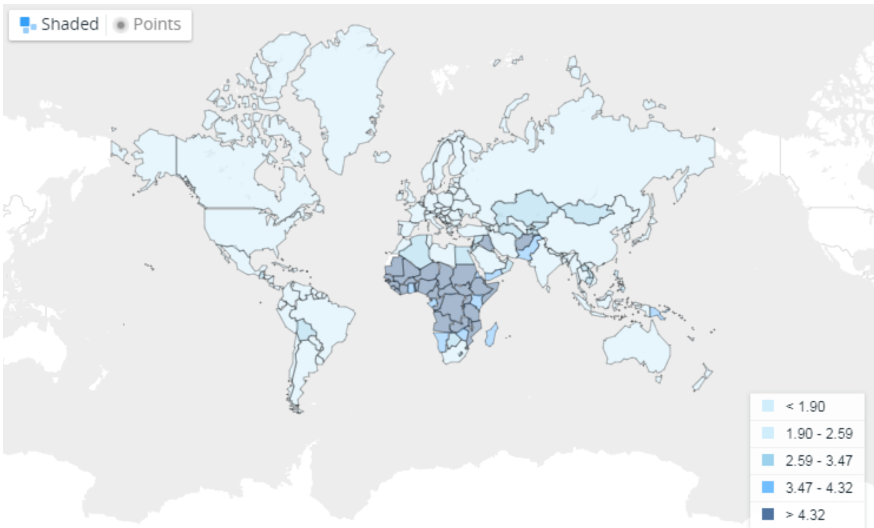

Figure 1.3 Fertility rates in 2015. Source: The World Bank

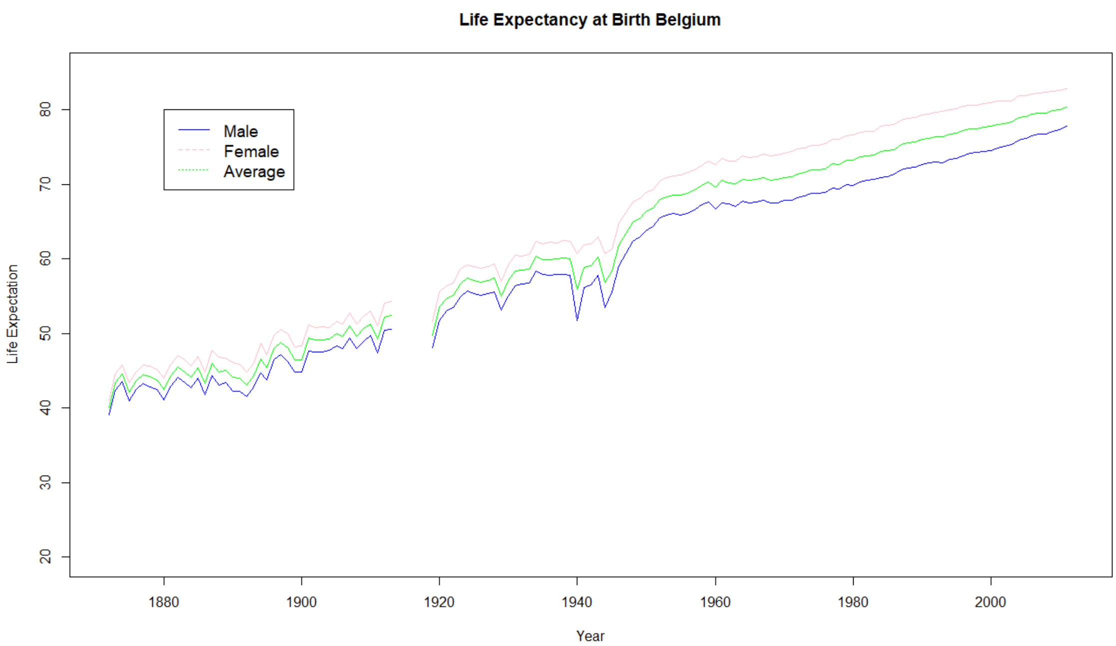

Figure 1.4 Life Expectancy at Birth - Belgian population. Source: HMD

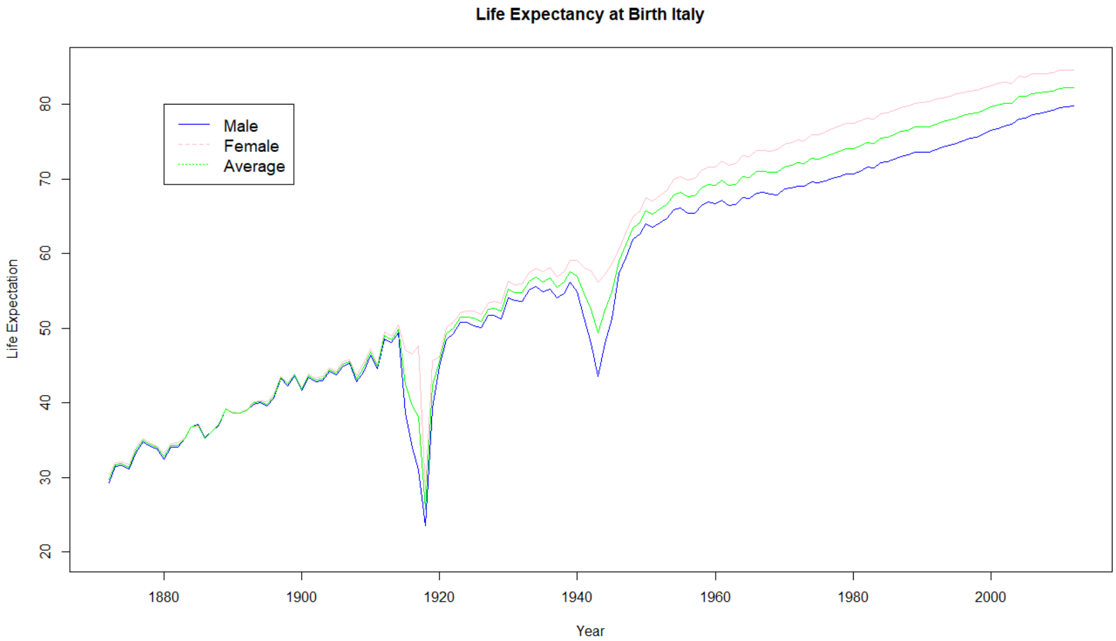

Figure 1.5 Life Expectancy at Birth - Italian population. Source: HMD

years at a steady pace of almost 3 months per year. The increasing in life expectation is not a characteristic of the newborns but it is a common feature during the entire life. The figures 1.6 and 1.7 show the surfaces composed of the log mortality rates for the Italian and Belgian populations for the age 0-100 and from 1872 to 2012. In figures 1.8 and 1.9 we plot the population's log mortality rates along the age for some years (i.e. "slices" of the previous graphs) to underline the common decreasing trend in mortality during the entire life.

In the OECD countries the increasing longevity trends have not slowed down recently [OECD, 2016]. The life expectation of the Italian population from 2015 to 2016 is increased by 5 months both for man and woman [ISTAT, 2016].

1.1.3 The population composition

In this section we consider the population composition in order to describe the result of the demographic trends described in the previous sections. The decreasing of fertility rates and the increasing of the life expectations conduct to the aging populations. A measure of this trend is the dependency ratio. The dependency ratio is the number of people aged more than 65 (pensioners) over the number of people aged 15-64 at the same instant. We can consider it as the number of pensioners over the number of workers: $DR(t) = \frac{\#Pensioners(t)}{\#Workers(t)}$. In Figure 1.10 the increasing of the dependency ratios is clear.

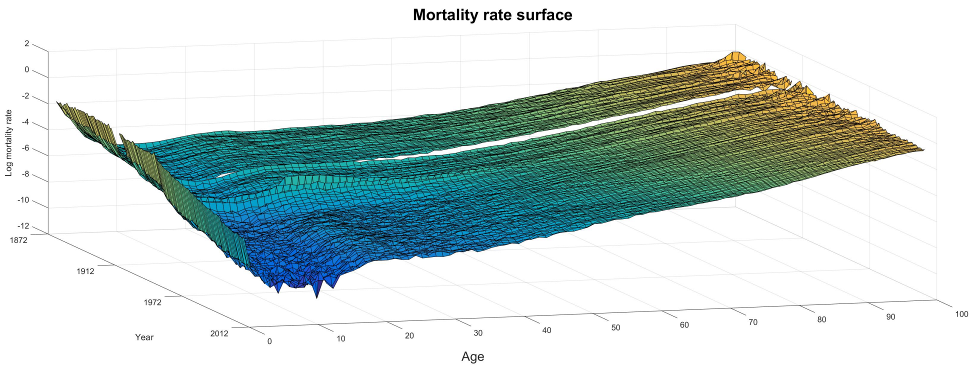

Figure 1.6 Log Mortality Surface: Belgian population. Source: HMD

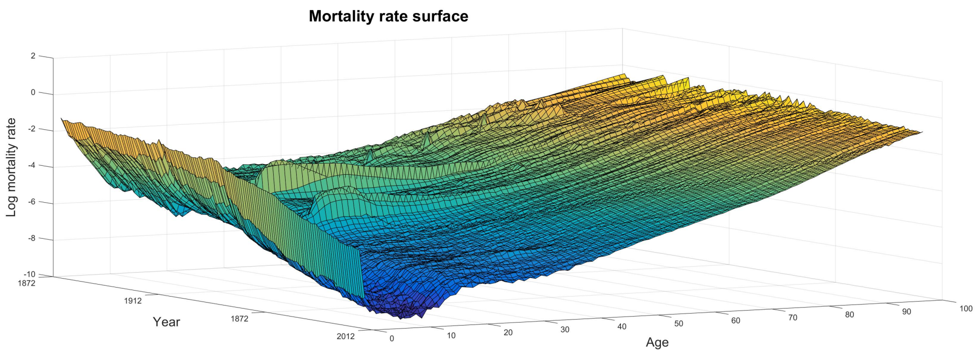

Figure 1.7 Log Mortality Surface: Italian population. Source: HMD

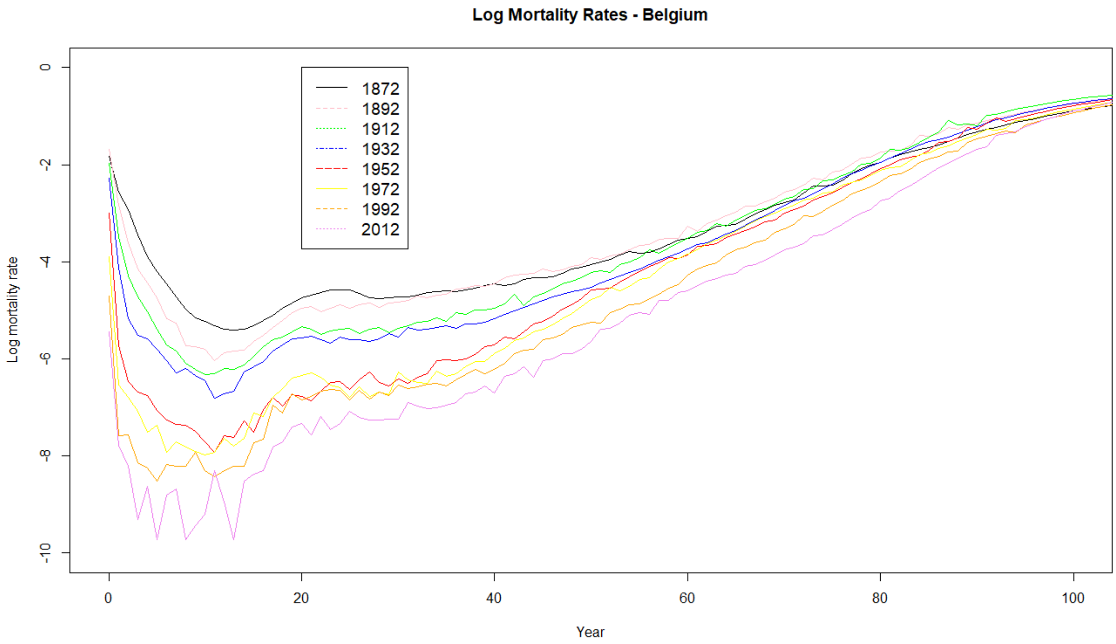

Figure 1.8 Log Mortality Rates: Belgian population

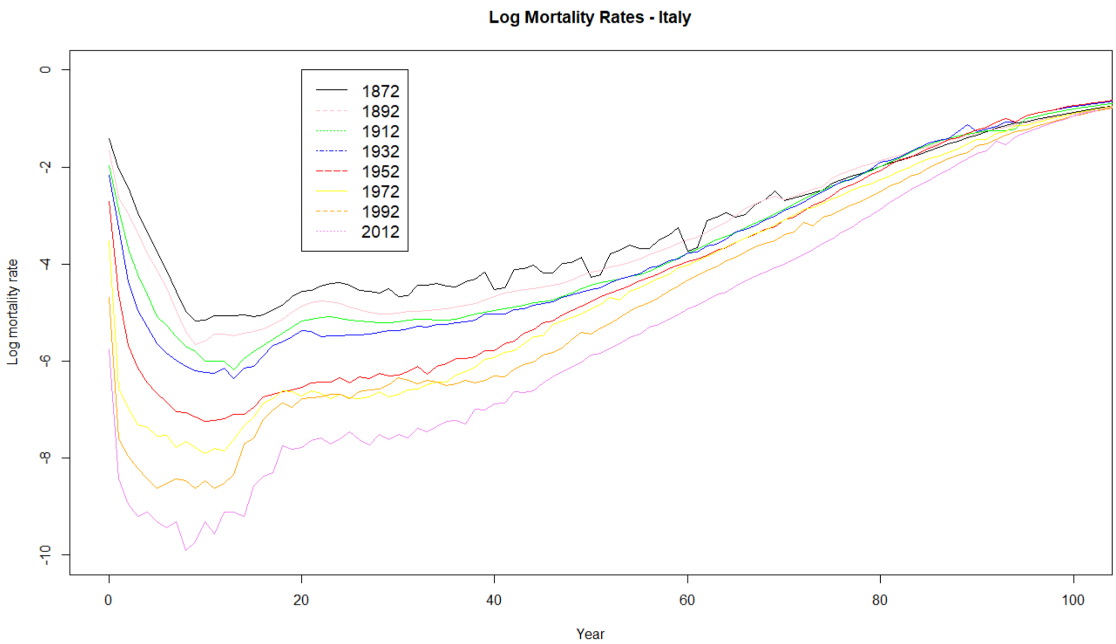

Figure 1.9 Log Mortality Rates: Italian population

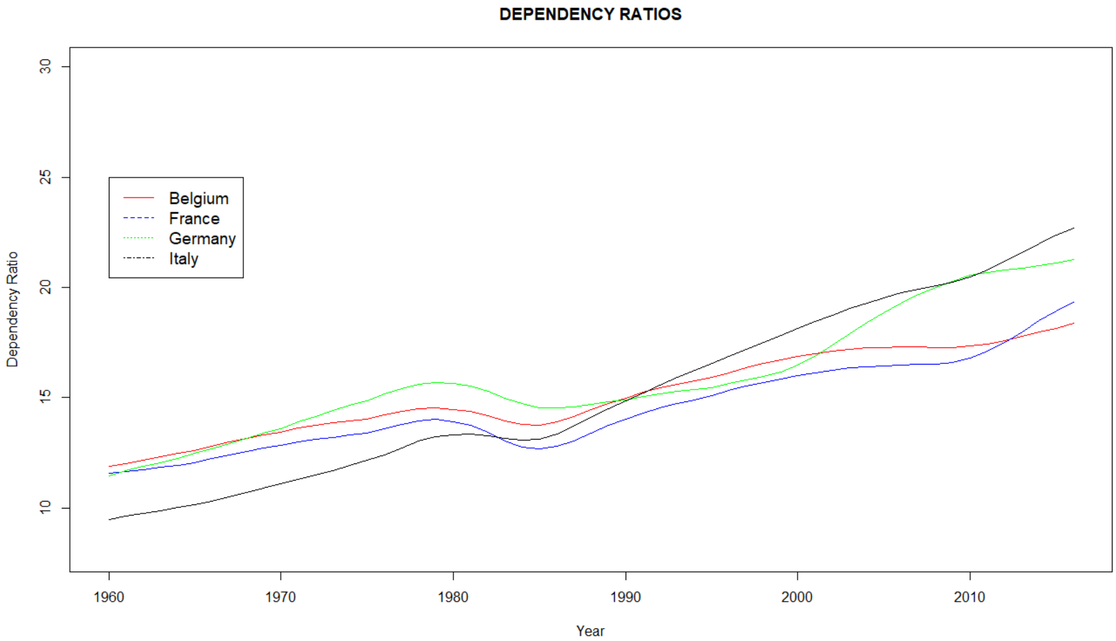

Figure 1.10 Dependency ratios from 1960 to 2016. Source: The World Bank

1.2 The pension systems

The long-term financing of pensions becomes a more and more challenging issue due to both actual and forecast population aging. The most developed countries have pension systems based on the so-called Pay-As-You-Go (PAYG) and they are facing long-term sustainability problems. In the PAYG, indeed, at any time the pensions due to retirees are paid by the active workers. The PAYG is a system that works (well) in a steady state situation where the dependency ratio remains constant. Thus, this system is in equilibrium either with a constant and low dependency ratio or with small fluctuations of $DR(t)$ that are hedged by the wages movements. Furthermore, in the long-term with stable conditions the rate of return of a PAYG scheme is linked with the growth rate of the salaries. The PAYG entails an intergenerational contract. The increasing in $DR(t)$ creates a sustainability problem that can be faced throughout the following levers:

- the increasing of the contributions;
- the increasing of taxes (in the first pillar);
- the reduction of pension benefits;

- the increasing of the retirement age(in the first pillar).

The PAYG, with different characteristics, is the base scheme of the so-called first pillar all over the world and is always managed by the Government.

The fully-funded (or funded) schemes are the base of the so-called second and third pillars. In this kind of schemes the contributions of a cohort are invested in the financial markets in order to finance the same cohort's future pension. Historically, the funded pension systems have higher returns than the unfunded systems. Maddison [2007] and Dimson et al. [2009] find that the average real return on equity in the last century for the main OECD countries was twice or more than the growth rate of the real GDP. If we take into account just the rate of return, the funded pension systems seem to be preferable.

As highlighted in Alonso-GarcÃa and Devolder [2016], there are many reasons to invest a certain percentage of the PAYG even in a low-yield scenario:

- the high variability of the funded mechanism could impose the presence of a PAYG percentage in the pension systems [Persson, 2002];
- the switch from PAYG to fully funded has high transition costs [Fajnzylber and Robalino, 2012];
- if returns on PAYG and funded are not positively correlated, there are benefits for both cohorts entering the system and a multiple cohort coexisting at the same period of time Alonso-GarcÃa and Devolder [2016].

In the literature we can find some other works based on the idea of mixing different pension systems. In Dutta et al. [2000] a mix between PAYG and funded schemes in a mean-variance framework is proposed with the aim to diversify risk. Bilancini and D'Antoni [2012] find that the PAYG is more attractive because it insures pensioners against the risk of being outperformed but it is less effective against the financial risks. In Devolder and Melis [2015], the authors analyze a combination of funded and unfunded pension schemes using portfolio theory; in a deterministic framework they find that the choice between the funded and unfunded depends only on the returns of the demographic and financial variables, i.e. the decision rule of Samuelson-Aaron (Samuelson, 1958, Aaron, 1966). Instead, in a stochastic framework, they find that the diversification is useful and they calculate the optimal level of diversification in several stochastic models with different demographic structures and different assets.

2

APC mortality models an applied literature review

2.1 Introduction

Forecasting mortality trends is important because they affect a wide number of areas with huge economic impacts. The new regulatory system of Solvency 2 implies a risk-based approach to obtain the Solvency Capital Requirement (SCR). Since both the mortality and the longevity risk are unhedgeable, a cost of capital rate should be applied in order to obtain the Market Value Margin. A fair quantification of the longevity and mortality risks leads to a fair quantification of the reserves.

The increased attention towards the demographic projections led to a sharp increase in the number of new mortality models. Unfortunately, many mortality models are over-parametrized or contain terms that cannot be justified in terms of demographic significance [Hunt and Blake, 2015b]. Many models analyze the historical evolution of the mortality trends through the decomposition across three dimensions: age, period, and year of birth (cohort). A deep dissertation of the Age-Period-Cohort models (APC) is presented in Hunt and Blake [2015b], where the authors investigate the structure of APC models, introduce a classification scheme for existing models and list the key factors analysts should consider when constructing a new model in this class. Hopefully, any component of the APC models has a strong demographic, biological, medical, and socio-economic meaning (Cairns et al. 2009, Hunt and Blake 2015b,a). Accordingly, we can easily conclude that different populations would need different APC models.

In this Chapter we describe the most important APC mortality models considered in literature. We describe the structure which encompasses the vast majority of them (Haberman and Renshaw 2011, Cairns et al. 2006b, 2009, Lee and Carter 1992, Hunt and Blake 2015b). We describe the standard terminology introduced by Currie [2016] for generalizing linear and non-linear models. We implement a data analysis of the Belgian population and Italian population, in a specific interval of time and for a specific age range, in order to: (i) describe the differences between the models, (ii) consider their common features, and (iii) try to describe the economic impact that the choice of a specific model will have with the data fitted.

The analysis is conducted using the statistical software R and the StMoMo package (https://CRAN.R-project.org/package=StMoMo) which exploits the unifying framework of the Generalized Age-Period-Cohort family in order to implement many of the stochastic mortality models (Villegas et al. [2015]).

2.2 Notation

We indicate as $D_{x,t}$ the number of deaths in a population aged x at last birthday during the calendar year t .The (central) number of people exposed to the death risk aged x last birthday during the last calendar year t are represented by $E^c_{x,t}$ ($E^0_{x,t}$). With the hypothesis of uniform distribution of the deaths over time we can write:

$$E^0_{x,t} \approx E^c_{x,t} + \frac{1}{2} D_{x,t}.$$

We take into consideration two mortality rates: the central mortality (death) rate,$m_{x,t}$, and the one-year mortality rate,$q_{x,t}$. The central mortality rate is empirically estimated as:

$$\hat{m}_{x,t} = \frac{D_{x,t}}{E^c_{x,t}}.$$

The one-year mortality rate (or mortality rate) is the probability that an individual aged exactly x at the initial time t will die during the year (from t to $t+1$). The link between the two mortality measures is as follows:

$$q_{x,t} \approx 1 - e^{-m_{x,t}}. \tag{2.1}$$

An empirical estimation of $q_{x,t}$ is:

$$\hat{q}_{x,t} = \frac{D_{x,t}}{E^0_{x,t}}.$$

Actually, Equation (2.1) takes into account a third mortality measure called force of mortality, $\mu_{x,t}$, that is the instantaneous death rate exactly considered in t for an individual aged x. If the population is stationary (i.e., the population size of all ages is constant over time), the force of mortality remains constant over time and ages, and we can write:

$$m_{x,t} = \mu_{x,t}.$$

If we assume that the deaths $D_{x,t}$ follow a Poisson distribution (i.e. $D_{xt} \sim \text{Poisson}\left(E^c_{xt}\mu_{xt}\right)$), the central mortality rate is used because $\mathbb{E}\left(D_{xt}/E^c_{xt}\right) = \mu_{xt}$, otherwise if we assume a Binomial distribution (i.e. $D_{xt} \sim \text{Binomial}\left(E^0_{xt}, q_{xt}\right)$) the one-year mortality rate is a better choice because $\mathbb{E}\left(D_{xt}/E^0_{xt}\right) = q_{xt}$.

We recall that this construction is used because when treating mortality the first assumption is that the force of mortality (instantaneous rate) is the same for all the members of population aged x at time t. Considering the real data, stationary population in practice means that we have a number of members and deaths that are high and significant from a statistical point of view. For a small population (in general above the age of 100) this assumption could not be true.

2.3 Data

The data used for the statistical analysis come from the Human Mortality Database (HMD) http://www.mortality.org/ that collects the life tables containing the central mortality rates for the most industrialized countries all over the word. In particular, we take into consideration the Belgian and Italian male populations during 1980-2014 and for the age interval 55:95. This choice is not an "easy" choice cause some models are particularly suitable for the older ages and this range is extended.

2.4 The APC models

2.4.1 Stochastic mortality model: requirements and desired features

After estimation or calibration of the parameters we need to check the fitness of the stochastic mortality model. In this paragraph we recall the main criteria listed in Cairns et al. [2011] and Cairns [2008] (see also Plat, 2009) to check the reliability of a model:

- the mortality rate should remain positive;
- the model should be consistent with the historical data;
- long-term dynamics entailed by the model should be biologically reasonable;
- parameter estimates should be robust relative to the period of data and the range of ages in the sample;
- model forecast should be robust with respect to the sample;
- forecast levels of uncertainty and central trajectories should be plausible and consistent with historical trends and variability in mortality data;
- the model should be straightforward to implement using analytical methods or fast numerical algorithms;
- the model should be relatively parsimonious;

- it should be possible to use the model to generate sample paths and calculate confidence intervals;
- the structure of the model should make it possible to incorporate parameter uncertainty in simulations;
- at least for some countries, the model should incorporate a stochastic cohort effect;
- the model should have a non-trivial correlation structure.

2.4.2 Age-Period-Cohort models: a generalization

In this section we describe the main contributions given to the definition of a general structure in the Age-Period-Cohort (APC) stochastic mortality models. The dynamic structure of the mortality in APC models is obtained by considering the mortality measure of a population linked with a series of factors depending on the age x, the period considered as calendar year t, and the cohort (that is the year of birth) calculated as $y = t - x$. The structure is linear or bilinear and can be generally described (Hunt and Blake 2015b, Villegas et al. 2015) as:

$$\eta_{x,t} = \alpha_x + \sum_{i=1}^{N} \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x},$$

where $\eta_{x,t}$ is the predictor of the mortality (described afterwards), α_x is the age function, $\beta_x^{(i)} \kappa_t^{(i)}$ are the age/period functions, and γ_{t-x} is the cohort effect, all described in the next sections.

The predictor can be computed as:

$$\eta_{x,t} = g\left(\mathbb{E}\left(\frac{D_{x,t}}{E_{x,t}}\right)\right),$$

where g is a linked function used to map the raw data into a better form for modeling purposes. One can choose different link functions but it is more convenient to use the classical or canonical link functions and match them with the distribution of the random component. Indeed, if $D_{x,t}$ is assumed to follow a Poisson distribution, it must be coupled with the log link function, while when if follows a Binomial distribution a logit link function should be used.

The age function α_x captures the behavior of mortality across the age range, and remains constant over time. The age function may not be considered static, for example some models, like the CBD family and the Aro-Pennanen model Aro and Pennanen [2011], assume that α_x is described by a simple function contained in an age/period

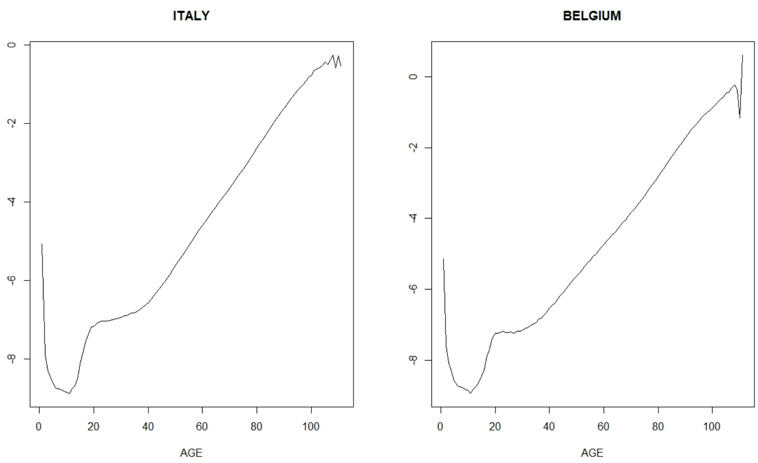

Figure 2.1 α_x in the Lee-Carter model calculated for Italian and Belgian population from 1920 to 2011

term. For example, in the model mentioned before, the age function (static) is a linear combination of the other age function in the model (Cairns et al. 2006b):

$$\alpha_x = \sum_{i=1}^{N} \alpha_x^{(i)} \beta_x^{(i)}.$$

As underlined in Hunt and Blake [2014], this structure improves the parsimony of the model but only for the higher age where the assumption is approximately valid. In figure 2.1 we show an example of the α_X-behavior in the Belgian and Italian population. The behaviors are qualitatively very similar showing a decreasing of mortality during the first years of life and a sharply increasing during the first adult ages, a plateau around the 30 years, and a constant increasing after the age of 40.

Age/period terms The age/period terms,$\sum_{i=1}^{N} \beta_x^{(i)} \kappa_t^{(i)}$, are constructed considering two terms $\beta_x^{(i)}$ and $\kappa_t^{(i)}$ describing respectively, the mortality distributed across ages (age functions) and the mortality evolution along time (period functions). The age/period terms usually can be able to capture and describe the majority of the mortality behavior inside the data. One of the most important difference in the APC models is the "type" of the age period terms considered inside the model structure. The age/period terms can be constructed in two ways: 1) without considering any a priori structure of the age functions $\beta_x^{(i)}$ and fitting the data at the different ages without any assumption on the shape of it, 2) considering a specific functional form for the age functions. Hunt and Blake [2015b] and Hunt and Blake [2014] indicate the first type of models as "non-parametric" (models) or as "factorial" age functions and indicate the second type models as "parametric" or "formulaic". The non-parametric models have common positive and negative characteristics:

- they are bilinear;
- the age/functions maximize the data fitting;
- the functions generated have scarce demographic significance;
- the chosen data interval can change the form of β_x;
- the non-parametric forms do not have to be continuous.

The last three characteristics are crucial to avoid getting projections that are not biologically reasonable, as suggested in the guidelines listed in the previous paragraph..

The parametric models consider a specific form of the age function in order to capture the evolution of the mortality across the age. The parametric models have the following features (Hunt and Blake, 2015b):

- most of them have a linear predictor structure;
- they need less free parameters respect to the non-parametric models to fit the age/period terms;
- knowing the shapes of the age functions permit the users to assign a specific demographic significance;
- parametric age functions are often suitable only over limited age ranges and we need to add additional age/period terms to construct a model able to capture the variability of mortality rates across the ages;
- they give a poorer fit to the data compared to models with the same number of non-parametric age/period terms.

Cohort effects The cohort effects are aimed at describing the difference in two mortality rates belonging to two populations born in different years. Despite the fact that natural experiments do not show clear evidence of such a cohort effect (Murphy [2009, 2010]), the "data from a number of countries appear to exhibit cohort features and so it is prudent to allow for these when modeling mortality" Hunt and Blake [2015b]. In this last paper, the authors give their intuition regarding the general peculiarities that the cohort effect should have:

- a small impact with respect to that of the age and period terms;
- the absence of systematic trends in the expected value or variability;
- a mean equal to zero across cohorts;
- some autocorrelation;
- the absence of indefinite persistence;

Models	Period Functions	Age Functions	Cohort Term
CBD, Aro Pennanen	Stochastic	Parametric	No
APC, M6, M7	Stochastic	Parametric	Yes
LC, RHn	Stochastic	Non Parametric	No
RH	Stochastic	Non Parametric	Yes
P-splines	Deterministic	-	-

Table 2.1 Classification of APC models. Source:Hunt and Blake [2015b]

- ideally be mean reverting;
- be demographic significant.

The autocorrelation is required since cohort born in successive years are more likely to present the same mortality behavior. In the same way, we need the no indefinite persistence to disentangle the life of cohort from the grandparent's cohort. If neither of the two previous hypotheses are verified, then the mean reverting behavior cannot be achieved.

Classification In this section we resume the classification made in Hunt and Blake [2015b]. The authors classify the most commonly used mortality models by important differences in structure between them. They consider the terms within the models and how they achieve the aims of the model user. In this paper, we consider only the stochastic mortality models and we disentangle them considering:

- the presence of the cohort term;
- the type of the age functions (parametric or non-parmetric).

In this chapter we consider six of the models in table 2.1 (LC, APC, CBD, M6, M7, RH) plus the Plat model.

Lee-Carter model The Lee and Carter [1992] model (LC) is the most widely used mortality model. The original model considered a log-bilinear form with a single age-period term:

$$\ln m_{x,t} = \alpha_x + \beta_x \kappa_t. \tag{2.2}$$

The model presents an identifiability problem in parameters estimation due to the overparametrization. We can easily see that for any triplet (α, β, κ) that solves (2.2), it exists a non zero constant c such that also the triplet $(\alpha - \beta c, \beta, \kappa + c)$ is a solution. Instead, if two constraints are imposed, the model becomes identifiable. The constraints are:

$$\sum_t \kappa_t = 0, \qquad \sum_x \beta_x = 1.$$

The intuition behind the first constraint is quite strong since it imposes that, for each age x, α_x is the mean over time of the $\ln m_{x,t}$. Instead, the second constraint is necessary in order to obtain the full model identification. We underline that the second constraint is not unique and in the literature other choices can be found depending on the applications. However, as suggested by Wilmoth [1993], a different constraint like

$$\sum_t \kappa_t = 1,$$

is not a very good choice. In fact, the standardization in terms of κ_t can give similar results for two different groups (for example men and women) and the speed of mortality may change between the two groups. Instead, the standardization of β_x leads to "distinctly different slopes when the speed of mortality changes for the two groups" (Lee and Carter, 1992, p. 661). Lee and Carter [1992] use a two-stage process to fit the data because the model cannot be fitted by ordinary regression methods due to the lack of regressors (i.e. there is no observed variable on the right-hand side). In the first stage the authors use a single value decomposition (SVD) to estimate the parameters. In the second stage the parameters κ_t are re-estimated using an iterative procedure based on 2.2, with α_x and β_x given by the values obtained in the first step. The value of the parameters κ_t found applying the two steps are different because the younger population (with low death rates) receives the same weight and the older population (with low size per years) during the second stage.

Further development in fitting techniques are implemented in Wilmoth [1993], where two fitting methods are proposed. The first method is a weighted SVD with weights equal to the observed number of data in each cell of the data (for a single cohort method is the number of death corresponding to a certain age x in a defined year t : D_{xt}). This method is statistically better than the previous since the variance of $\ln(m_{xt})$ is approximately $\frac{1}{D_{xt}}$. Wilmoth [1989, 1993] provide estimations of the standard errors of α_x and β_x coefficients (Lee, 2000). The second fitting method proposed is the Maximum Likelihood Estimation (MLE) with a Poisson distribution of the deaths (with mean $\lambda_{xt} = m_{xt}E_{xt}$). The likelihood function for a single age-time combination is:

$$L\left(d_{x,t};\lambda_{x,t}\right) = \frac{\lambda_{x,t}^{d_{x,t}} e^{-\lambda_{x,t}}}{d_{x,t}!} \tag{2.3}$$

where d_{xt} is the realization of the random variable D_{xt}. The logarithm of 2.3 is:

$$\log\left(L\left(d_{x,t};\lambda_{x,t}\right)\right) = d_{x,t}\log\left(\lambda_{x,t}\right) - \lambda_{x,t} - \log\left(d_{x,t}!\right),$$

and with the hypothesis of the independence between the observations, we can write:

$$l = \log\left(L\left(d_{xt};\lambda_{xt}\right)\right) = \sum_x\sum_t\left[d_{xt}\log\left(\lambda_{xt}\right) - \lambda_{xt} - \log\left(d_{xt}!\right)\right].$$

Another method based on the MLE is proposed in Brouhns et al. [2002]. The authors consider the Poisson distribution and solve the maximization problem by the

LEM program proposed by Vermunt [1997a,b] . The LEM is a unidimensional Newton method where in every iteration step, all the n parameters in the vector $\theta$are calculated considering the previous estimated values Brouhns et al. [2002] as follows:

$$\hat{\theta}^{(n)} = \hat{\theta}^{(n-1)} - \frac{\partial l^{(n-1)}/\partial\theta}{\partial^2 l^{(n-1)}/\partial\theta^2}$$

with $l^{(n-1)} = l^{(n-1)}\left(\hat{\theta}^{(n-1)}\right)$. From random values it is possible to converge to the values that are the best fit of the data.In Brouhns et al. [2002] the stop criterion is the reduction of the likelihood function increase (the recommended value is 10^{-10}). In figures 2.2 and 2.3 we show the fits of parameters for Belgian and Italian populations.

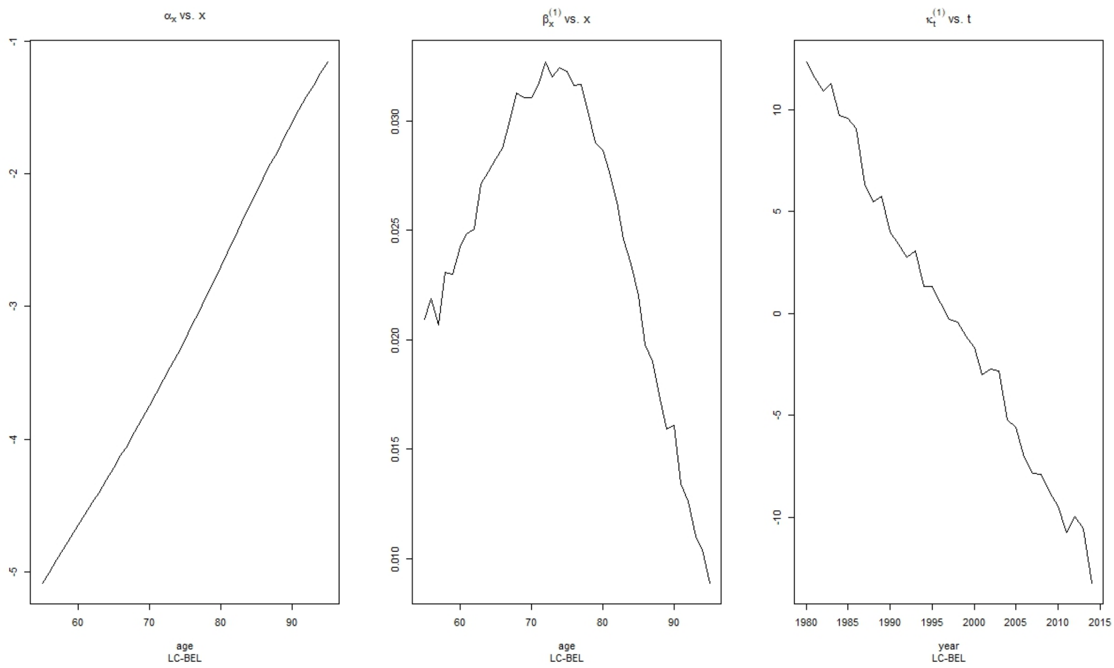

Figure 2.2 LC parameters-Belgian population

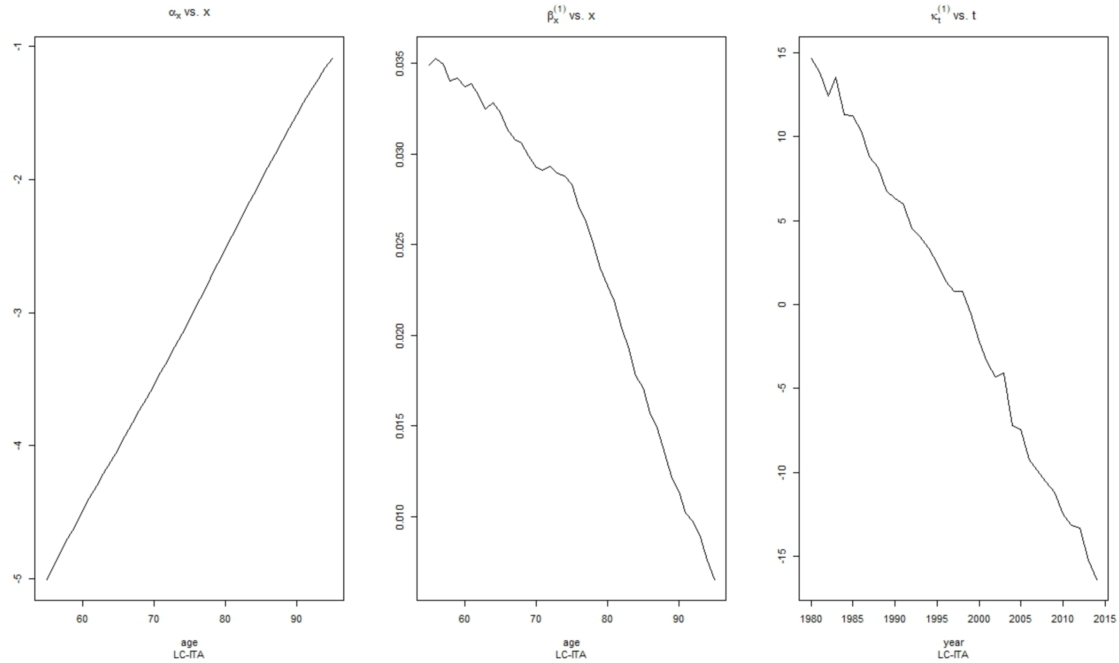

Figure 2.3 LC parameters-Italian population

Cairns-Blake-Dowd M6 and M7 models In Cairns et al. [2006b] the authors propose a two-factors model where the first factor affects the dynamics of the mortality rates at all ages in the same way, whereas the second factor increases the mortality-rates for the higher ages more respect to the lower ages. The model was introduced in order to overcome the perfect correlation in the projected mortality rates obtained in models with term for age and period. Another feature of this kind of models is the parsimony respect to the LC models due to the absence of the static age function α_x. The CBD model can be written in the following form:

$$\text{logit}\,(q_{x,t}) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}$$

where (see Cairns et al., 2006b):

$$\beta_x^{(1)} = 1 \text{ and } \beta_x^{(2)} = (x - \bar{x})$$

with $\bar{x} = \frac{\sum_i x_i}{\#\text{of age take into consideration}}$ (i.e. the mean age in the sample take into consideration). In the data considered in this work $\bar{x} = 75$. The model obtained is:

$$\text{logit}\,(q_{x,t}) = \kappa_t^{(1)} + (x - \bar{x})\,\kappa_t^{(2)}. \tag{2.4}$$

2.4 has no identifiability problems.

The M6 model (Cairns et al., 2009) is a generalization of the CBD model with the cohort effect. The general version of this model can be write as:

$$\text{logit}\,(q_{x,t}) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)}.$$

Following Cairns et al. [2009] we assume the following parametric forms:

$$\beta_x^{(1)} = 1,\ \beta_x^{(2)} = (x - \bar{x}) \text{ and } \beta_x^{(3)} = 1.$$

The model becomes:

$$\text{logit}\,(q_{x,t}) = \kappa_t^{(1)} + (x - \bar{x})\,\kappa_t^{(2)} + \gamma_{t-x}. \tag{2.5}$$

In 2.5 the identifiability problem is present. In order to fit the data we need estimations of $\gamma_{t-x}^{(3)}$ centered around zero without a linear trend (up or down). Considering the identifiability problem we have a model invariant respect the transformation:

$$\left(\kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_{t-x}\right) \to \left(\kappa_t^{(1)} + \phi_1, \kappa_t^{(2)} - \phi_2, \gamma_{t-x} - \phi_1 + \phi_2\,(t - \bar{x})\right)$$

where ϕ_1, ϕ_2 and ϕ_3 are real constants. From Cairns et al. [2011], Appendix A, we can regress γ_{t-x} on $(t - x)$:

$$\gamma_{t-x} = \phi_1 + \phi_2\,(t - x) + \varepsilon_{t-x}$$

with

$$\varepsilon_{t-x} \sim N\left(0, \sigma^2\right) \qquad \text{i.i.d.}$$

the constraints obtained are:

$$\sum_{c=c_e}^{c_l} \gamma_c = 0$$

$$\sum_{c=c_e}^{c_l} c\gamma_c = 0.$$

where $c_e = t_1 - x_k$ and $c_l = t_n - x_1$.

Another generalisation of the CBD, with an additional age effect and a cohort effect, is called M7 Cairns et al. [2009]: :

$$\text{logit}\,(q_{x,t}) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\kappa_t^{(3)} + \beta_x^{(4)}\gamma_{t-x}^{(4)}.$$

The age effect considered in the M7 model is quadratic, as before we can define:

$$\beta_x^{(1)} = 1,\ \beta_x^{(2)} = (x - \bar{x})\,,\ \beta_x^{(3)} = \left((x - \bar{x})^2 - \hat{\sigma}_x^2\right) \text{ and } \beta_x^{(4)} = 1$$

giving:

$$\text{logit}\,(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}\,(x - \bar{x}) + \kappa_t^{(3)}\left((x - \bar{x})^2 - \hat{\sigma}_x^2\right) + \gamma_{t-x},$$

where $\hat{\sigma}^2 = \frac{(x - \bar{x})}{\#\text{number of age take into consideration}}$. This model, like the M6 model, presents an identifiability because the model is not identifiable and the predictor remains unchanged if we apply the transformation:

$$\left(\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}\right) \to \left(\kappa_t^{(1)} + \phi_1 + \phi_2\,(t - \bar{x}) + \phi_3\left((x - \bar{x})^2 - \hat{\sigma}^2\right),\right.$$

$$\kappa_t^{(2)} - \phi_2 - 2\phi_3 (t - \bar{x}), \kappa_t^{(3)} + \phi_3, \gamma_{t-x} - \phi_1 - \phi_2 (t - x) - \phi_3 (t - x)^2 \Big) \tag{2.6}$$

where ϕ_1, ϕ_2 and ϕ_3 are real constants. As in the previous case, the authors decide to leave the cohort effect free to fluctuate around zero without trends. To obtain it three constraints are required. They are obtained using transformation in 2.6 with constants ϕ_1, ϕ_2 and ϕ_3 obtained by regressing γ_{t-x} on $(t-x)$ and $(t-x)^2$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2 (t - x) + \phi_3 (t - x)^2 + \varepsilon_{t-x}$$

with

$$\varepsilon_{t-x} \sim N\left(0, \sigma^2\right) \quad \text{i.i.d.}$$

For further explanations see Haberman and Renshaw [2011].The resulting constraints are:

$$\sum_{c=c_e}^{c_l} \gamma_c = 0$$

$$\sum_{c=c_e}^{c_l} c\gamma_c = 0$$

$$\sum_{c=c_e}^{c_l} c^2\gamma_c = 0.$$

In the Figures 2.4, 2.5, 2.6, 2.7, 2.8 and 2.9 we show the values of the parameters obtained calibrating the data described in the paragraph.

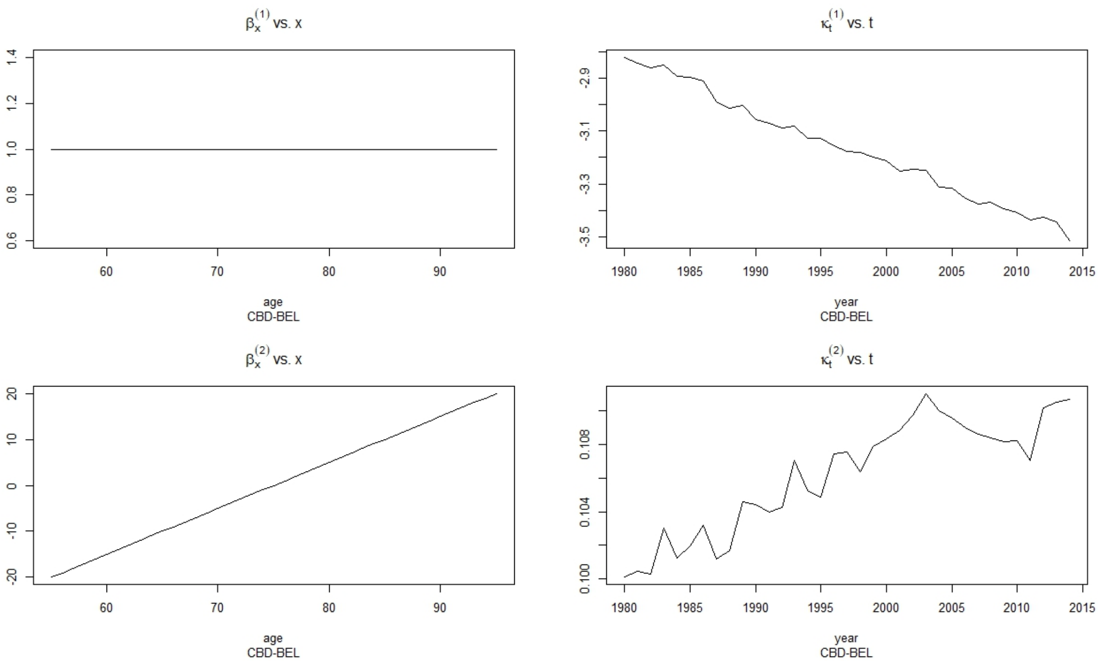

Figure 2.4 CBDparameters: Belgian population

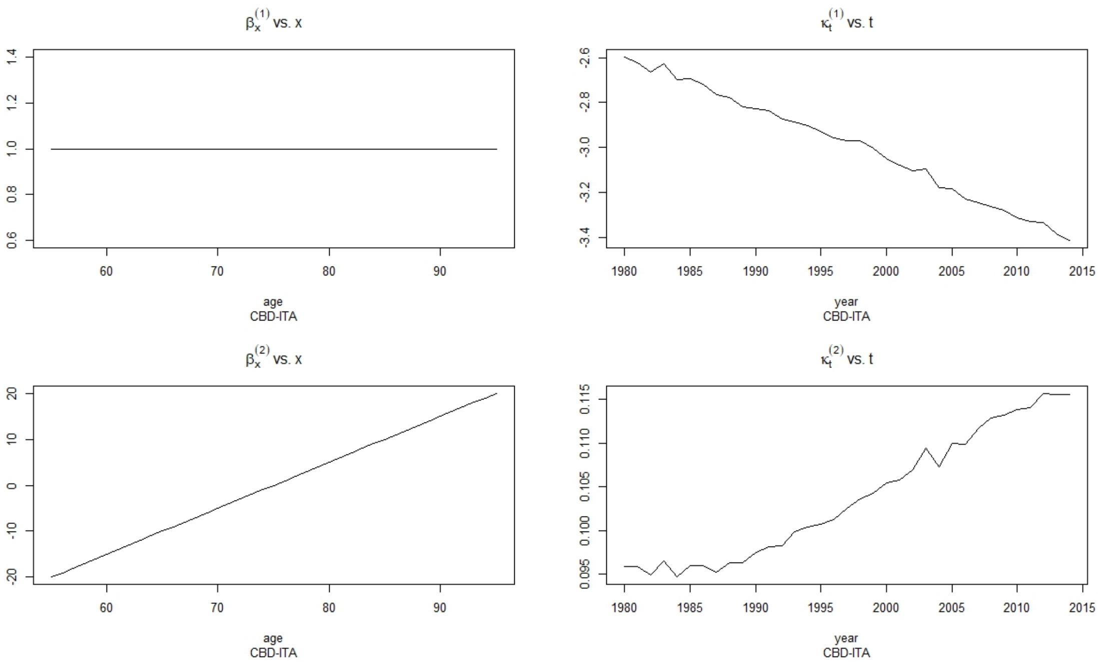

Figure 2.5 CBD parameters: Italian population

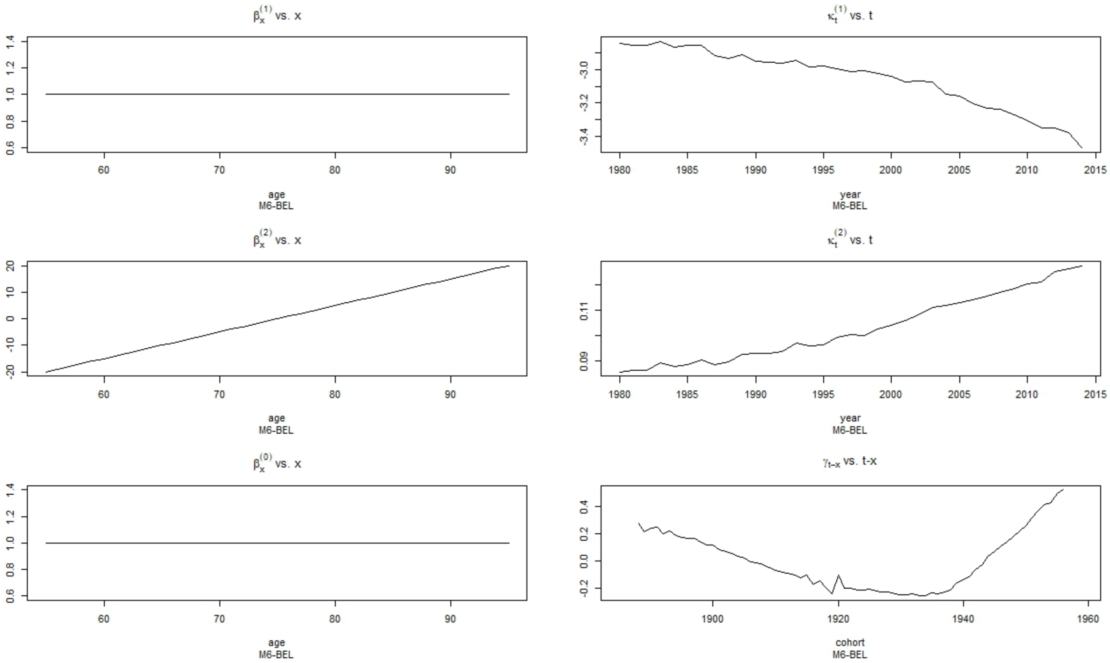

Figure 2.6 M6 parameters: Belgian population

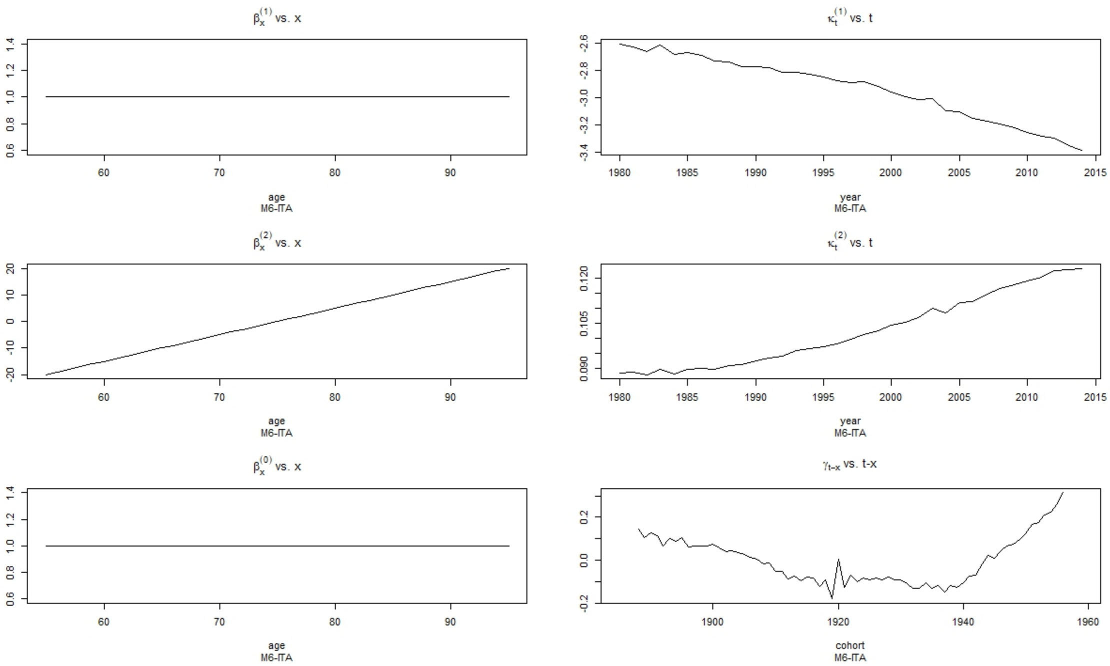

Figure 2.7 M6 parameters: Italian population

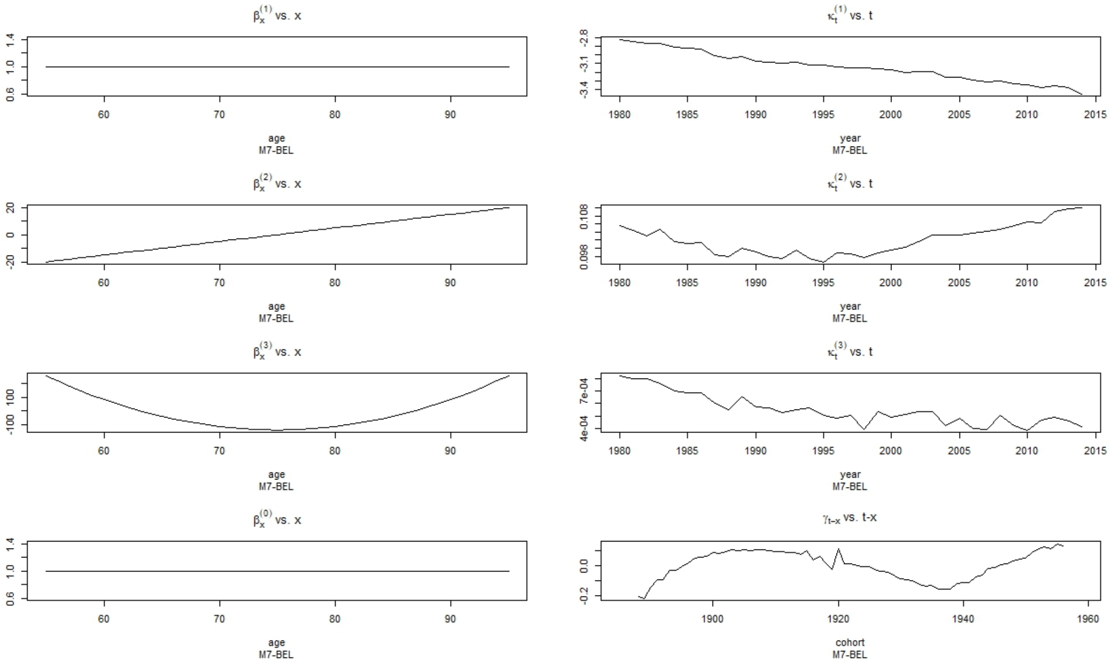

Figure 2.8 M7 parameters: Belgian population

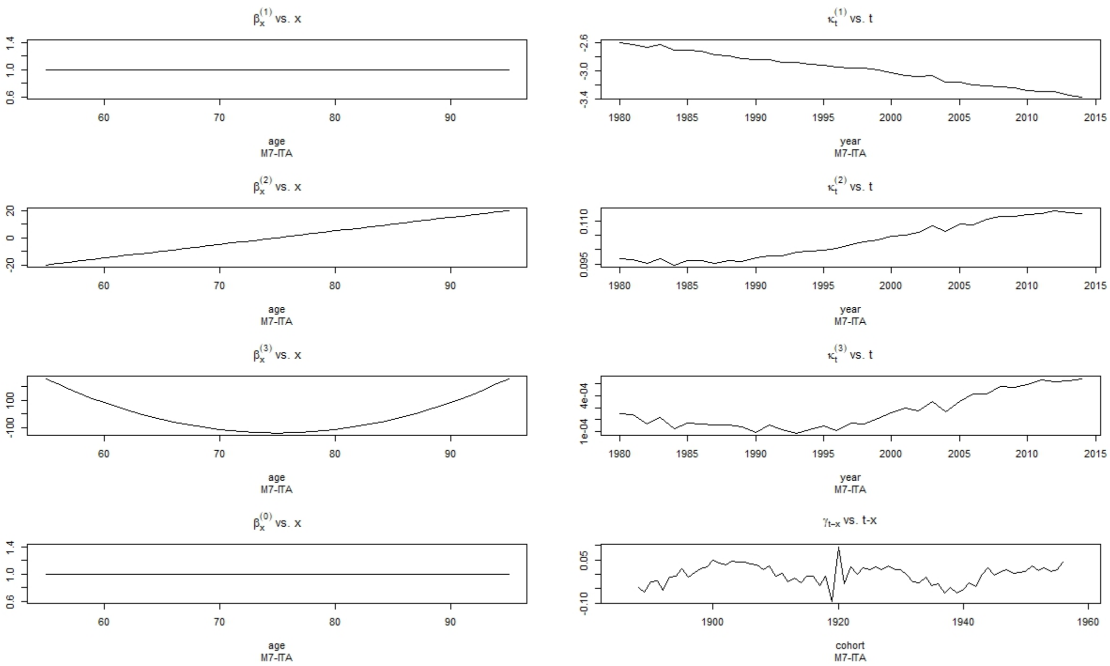

Figure 2.9 M7 parameters: Italian population

Renshaw-Haberman model (RH) The main contribution given by Renshaw and Haberman [2006] is the extension of the modelization and projection methods considering the LC model added with a cohort effect.

$$\ln\left(m_{x,t}\right) = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\gamma_{t-x}.$$

The model has identifiability problem because the following transformation:

$$\left(\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}, \beta^{(2)}, \gamma_{t-x}\right) \rightarrow \left(\alpha_x + \phi_1\beta_x^{(1)} + \phi_2\beta_x^{(2)}, \frac{1}{\phi_3}\beta_x^{(1)}, \phi_3\left(\kappa_t^{(1)} - \phi_1\right), \frac{1}{\phi_4}\beta_x^{(2)}, \phi_4\left(\gamma_{t-x} - \phi_2\right)\right)$$

leaves the predictor $\ln\left(m_{x,t}\right)$ unchanged with ϕ_1,ϕ_2,$\phi_3$andϕ_4 real constants and $\phi_3 \neq 0$ and $\phi_4 \neq 0$. Identifiability can be ensured using the following constraints (see Cairns et al., 2009):

$$\sum_t \kappa_t^{(1)} = 0$$

$$\sum_x \beta_x^{(1)} = 1$$

$$\sum_x \beta_x^{(2)} = 1$$

$$\sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0.$$

The first and the last constraints force α_x to be approximately equal to the mean over t of $\ln\left(m_{x,t}\right)$. The presence of other constraints allows for full identification but the choice is not relevant in terms of the quality of fit. As underline in Currie [2016] the RH model could present some convergence issues when using the r gnm package. Although we consider the starting values suggested in Villegas et al. [2015], the algorithm is not able to converge in the Belgian case. Therefore we consider the fitting result only for Italian case (figure 2.10).

A model derived from the RH model is the Age-Period-Cohort model (APC) introduced in the actuarial literature by Currie [2006]. The model is:

$$\ln\left(m_{x,t}\right) = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x},$$

which is the RH model with $\beta_x^{(1)}$and $\beta_x^{(2)}$ posed equal to one. To further information about how to manage the cohort constraints needed for the identifiability we suggest to read Haberman and Renshaw [2011]. In the figures 2.11 and 2.12 we show the parameters estimated for the datasets.

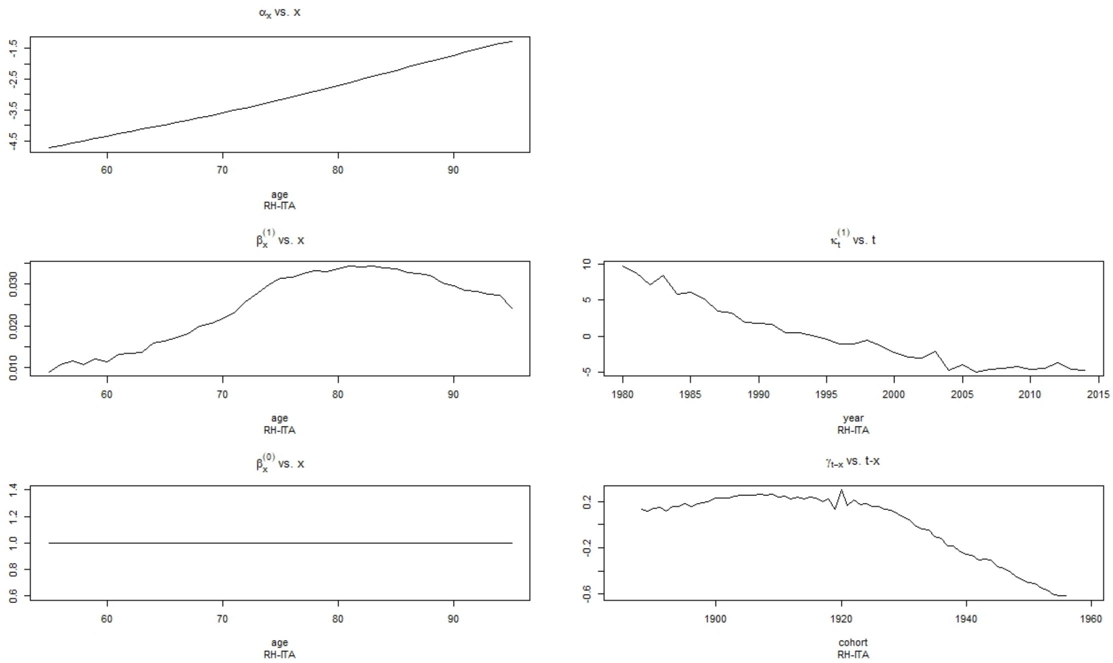

Figure 2.10 RH parameters: Italian population

Plat model The Plat model [Plat, 2009] considers four stochastic factors:

$$\text{logit}\left(q_{x,t}\right) = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}\left(\bar{x} - x\right) + \kappa_t^{(3)}\left(\bar{x} - x\right)^{+} + \gamma_{t-x}.$$

A "simplified" Plat model can be considered for higher ages, Plat [2009] suggests from 60 years old. The simplified version does not consider the factor $\kappa_t^{(3)}$:

$$\text{logit}\left(m_{x,t}\right) = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}\left(\bar{x} - x\right) + \gamma_{t-x}.$$

In this work we consider the simplified version in order to obtain better data for the older ages and to be parsimonious. The parameters estimated for our datasets are shown in figures 2.13 and 2.14.

2.5 Fitting the models

Following Hunt and Blake [2015b] we estimate the parameters of the APC stochastic mortality models via maximization of the model log-likekihood functions. In the Poisson case the log-likelihood is given by:

$$\log L\left[\left(d_{xt}, \hat{d}_{xt}\right)\right] = \sum_{x}\sum_{t}\omega_{xt}\left[d_{xt}\log\hat{d}_{xt} - \hat{d}_{xt} - \log\left(d_{xt}!\right)\right]$$

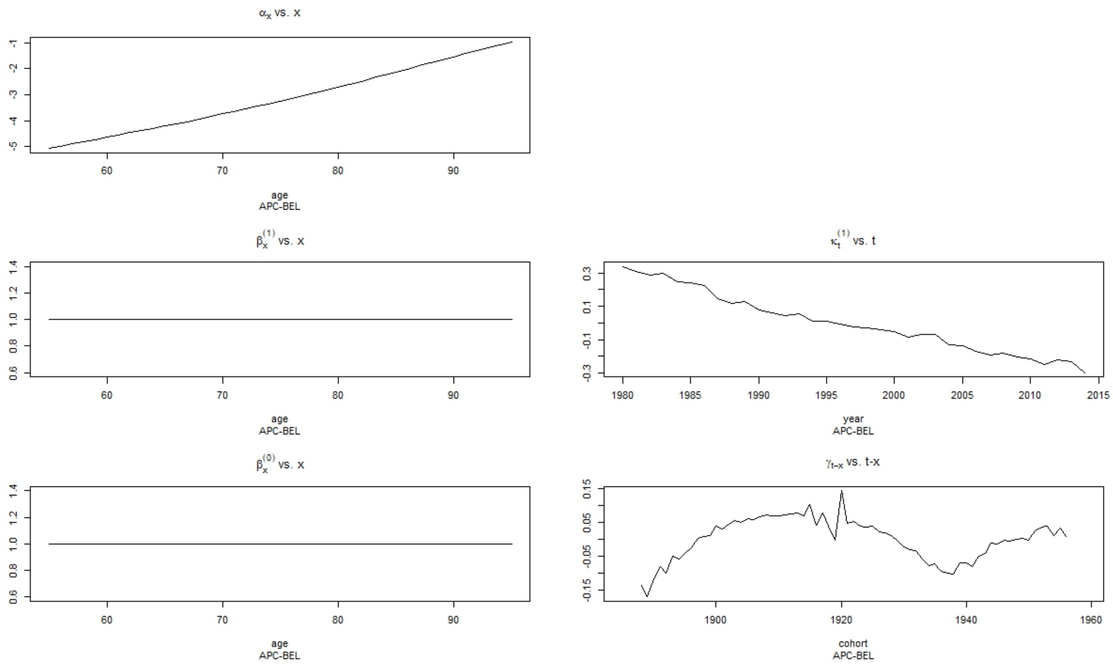

Figure 2.11 APC parameters: Belgian population

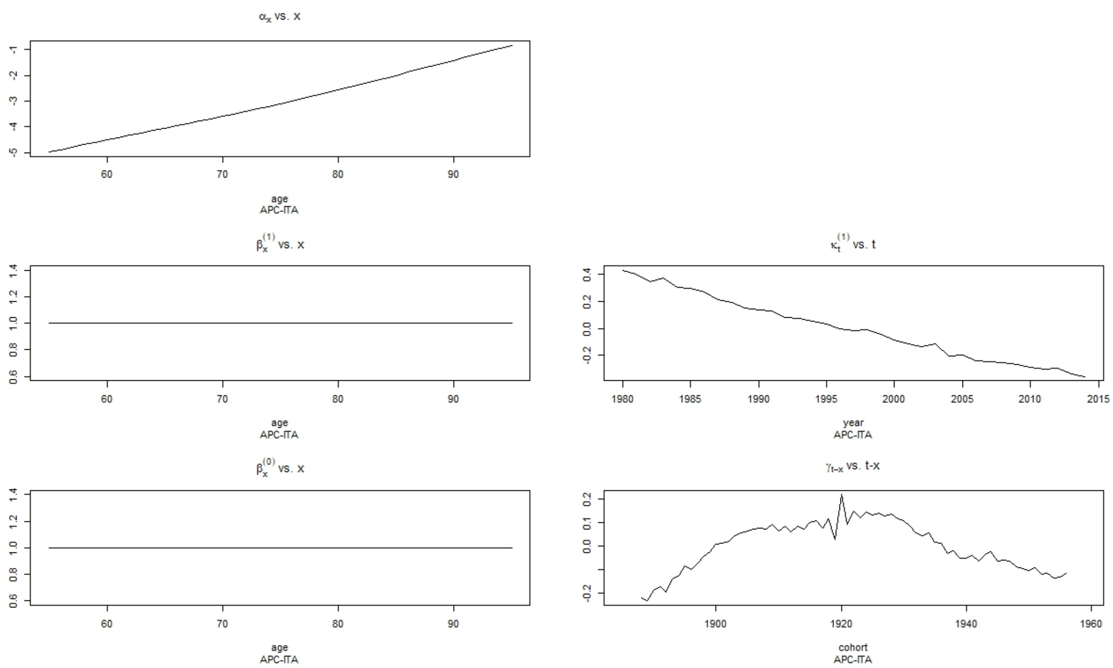

Figure 2.12 APC parameters: Italian population

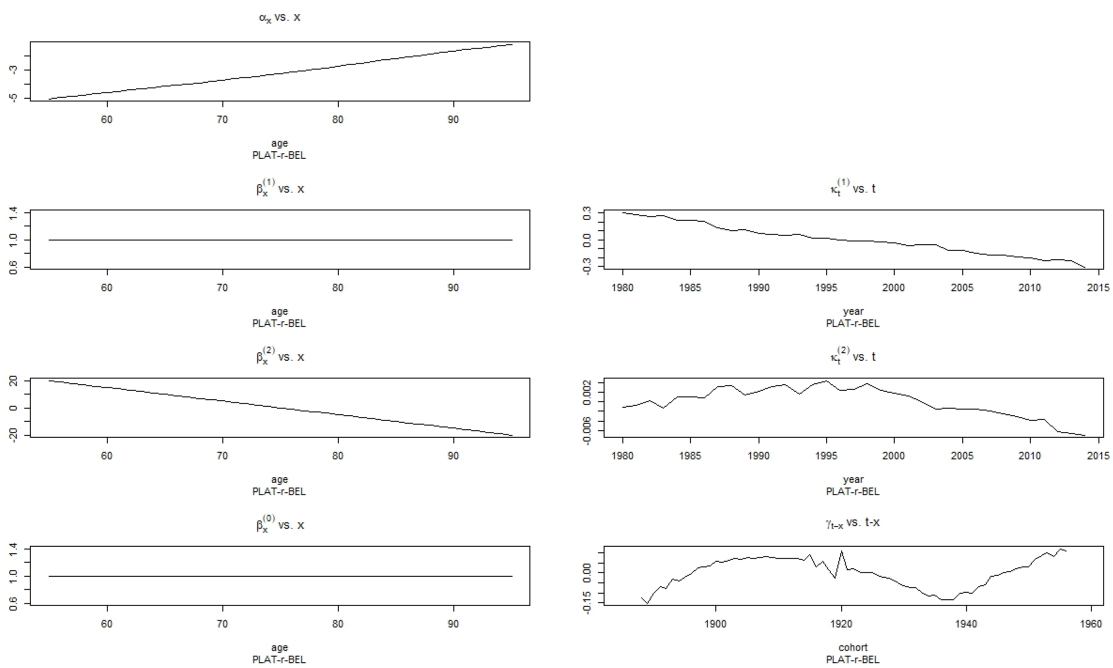

Figure 2.13 Plat parameters: Belgian population

and in the Binomial case is:

$$\log L\left[\left(d_{xt},\hat{d}_{xt}\right)\right]=\sum_{x}\sum_{t}\omega_{xt}\left[d_{xt}\log\left(\frac{\hat{d}_{xt}}{E^0_{xt}}\right)+\left(E^0_{xt}-d_{xt}\right)\log\left(\frac{E^0_{xt}-\hat{d}_{xt}}{E^0_{xt}}\right)+\binom{E^0_{xt}}{d_{xt}}\right]$$

where ω_{xt} is the weight, 0 or 1, if the cell(x,t) is omitted or not respectively Villegas et al. [2015].

2.5.1 Goodness-of-fit

The goodness-of-fit analysis is a building block to consider in order to decide the suitability of a certain model to the data. We compared the models using standard measures described in the literature (Hunt and Blake 2015b, Cairns et al. 2009, Plat 2009, Haberman and Renshaw [2011]). In particular we want to take into consideration the residuals of the fitted model and the information criteria considering the maximum likelihood criterion penalized with more parameters. (Brouhns et al. 2002)

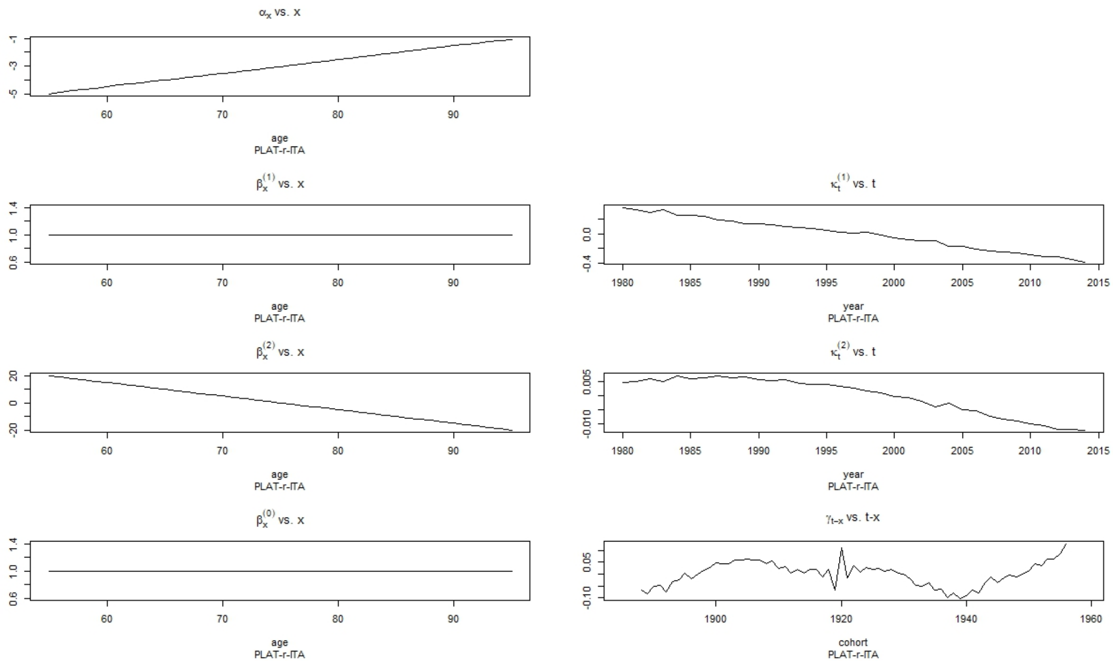

Figure 2.14 Plat parameters: Italian population

2.5.2 Residuals

We can write the scaled deviance residuals that we should consider with a Poisson or Binomial random component as (Renshaw and Haberman 2006,Villegas et al. 2015):

$$r_{xt} = sign\left(d_{xt} - \hat{d}_{xt}\right)\sqrt{\frac{\text{dev}\,(x,t)}{\hat{\phi}}}, \qquad \hat{\phi} = \frac{D\left(d_{xt}, \hat{d}_{xt}\right)}{K - v}$$

where for a Poisson random component:

$$\text{dev}\,(x,t) = 2d_{xt}\text{log}\left(\frac{d_{xt}}{\hat{d}_{xt}}\right) - \left(d_{xt} - \hat{d}_{xt}\right)$$

and for a Binomial random component:

$$\text{dev}\,(x,t) = 2d_{xt}\text{log}\left(\frac{d_{xt}}{\hat{d}_{xt}}\right) + \left(E^0_{xt} - d_{xt}\right)\text{log}\left(\frac{E^0_{xt} - d_{xt}}{E^0_{xt} - \hat{d}_{xt}}\right).$$

Where

$$D\left(d_{xt}, \hat{d}_{xt}\right) = \sum_x \sum_t \omega_{xt}\text{dev}\,(x,t)$$

is the total deviance of the model, $K = \sum_x \sum_t \omega_{xt}$ is the number of observations in the data and v is the effective number of parameters in the model. In order to obtain

Model	Log-likelihood	AIC	BIC	N.Parameters
LC	-12872.04	25974.07	26579.03	115
RH	-8256.26	16878.51	17841.19	183
APC	-9539.94	19363.87	20110.87	142
CBD	-16902.54	33945.08	34313.31	70
M6	-8781.05	17836.10	18556.79	137
M7	-8583.85	17509.69	18409.24	171
Plat	-8335.27	17020.54	17941.13	175

Table 2.2 Italian values

a better visual comparison of our results we consider two different kind of graphic visualizations: i) the scatter plots of residual by age, period and cohort (Renshaw and Haberman, 2006, Cairns et al., 2009), ii) the color maps of residuals. In figures 2.15, 2.16, 2.17, 2.18, 2.19, 2.20, 2.21, 2.22 and 2.23 we show the residuals of the mortality models applied to our datasets.

From both the heat-maps and scatter plots of the two populations we can observe that:

- LC and CBD show highlighted residual patterns. The heat maps of these models show marked diagonal patterns. The scatter plots explain these behaviors showing the inability to capture the cohort-effects, clearly present in the two populations.
- Some light residual patterns are display by the heat maps of the M6, M7 and, in particular, APC models. These results are due to different reasons. The scatter plots of M6 and M7 show a low ability to fit the improvement rates with age, similar behaviors are shown by the CBD models. The scatter plots of APC model instead are more disperse.
- The RH, in the Italian case, and the Plat models give better results.

2.5.3 Information criteria

The information criteria used in this paper are the Bayesian Information Criteria (BIC) and the Akaike Information Criteria (AIC) defined as:

$$\text{BIC} = 2v - 2\mathcal{L}$$

$$\text{AIC} = v \log K - 2\mathcal{L}.$$

We will prefer the lower values of this indicators. The following tables present the values of the log-likelihood functions, AICs, BICs (obtained for the models mentioned before) and the number of free parameters.

Considering just the information criteria we can say that the models with lower BICs and AICs are Plat, M6, M7 and the RH models for both the Italian and Belgian

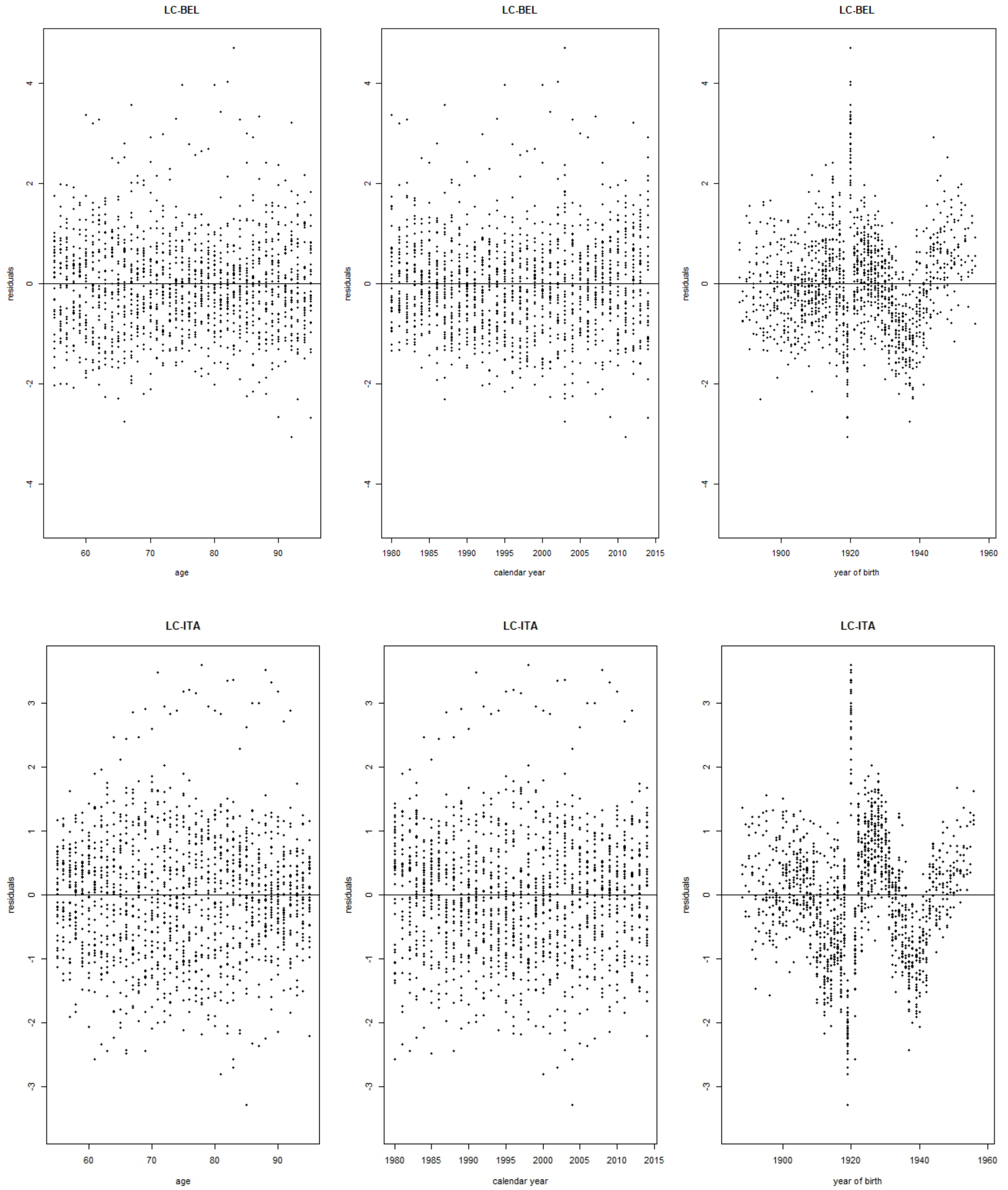

Figure 2.15 LC scatter plots for Belgian and Italian populations

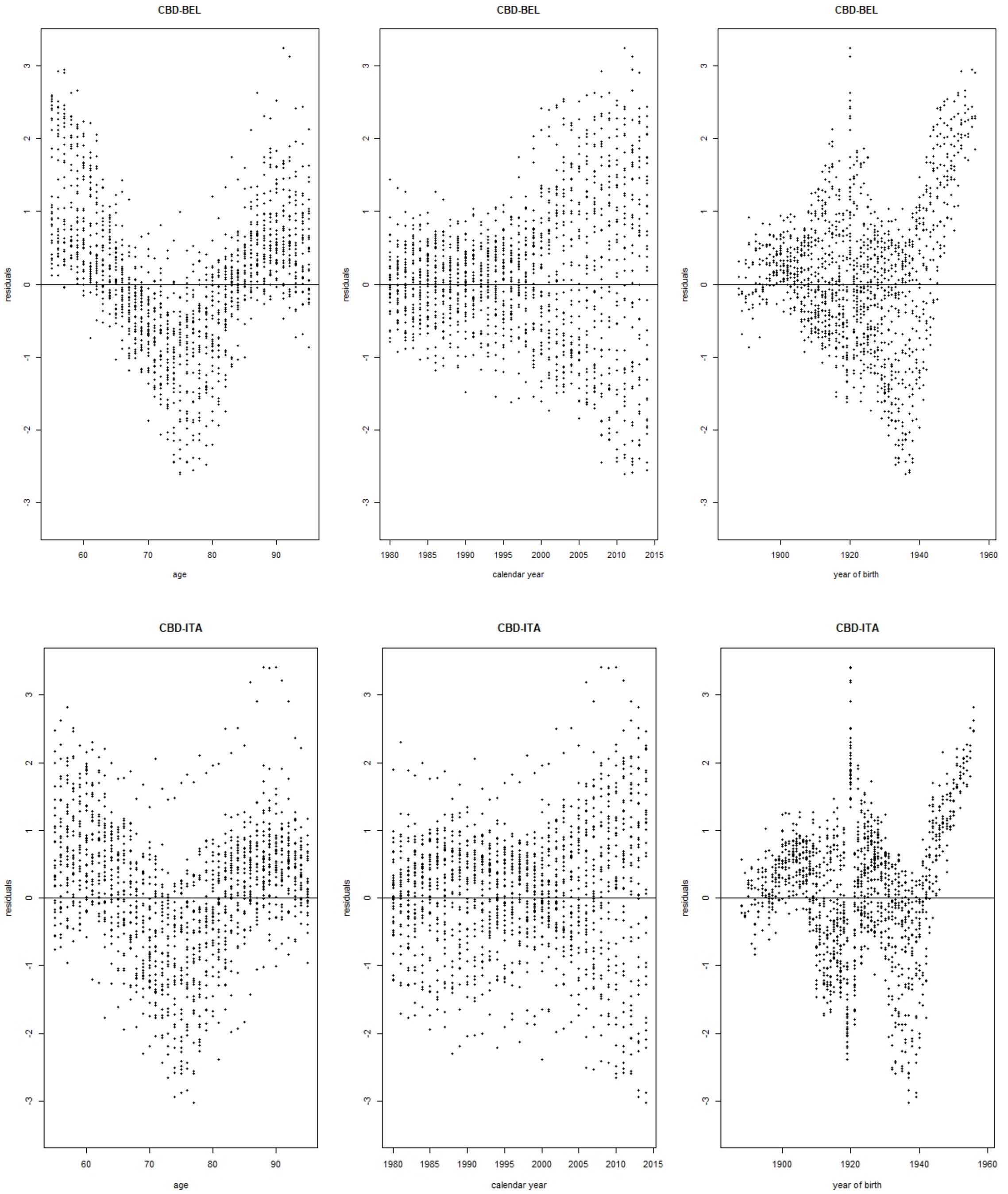

Figure 2.16 CBD scatter plots for Belgian and Italian populations

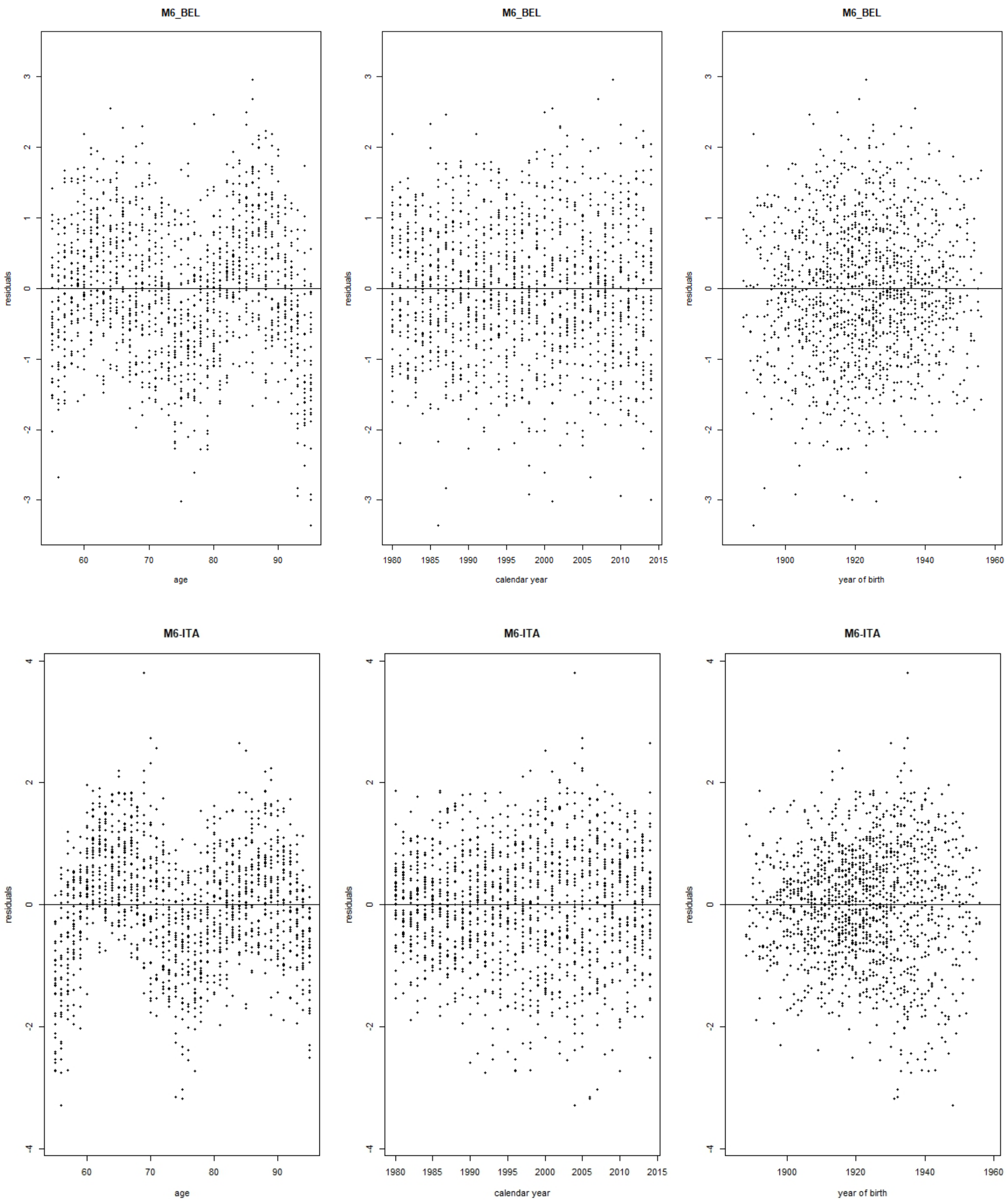

Figure 2.17 M6 scatter plots for Belgian and Italian populations

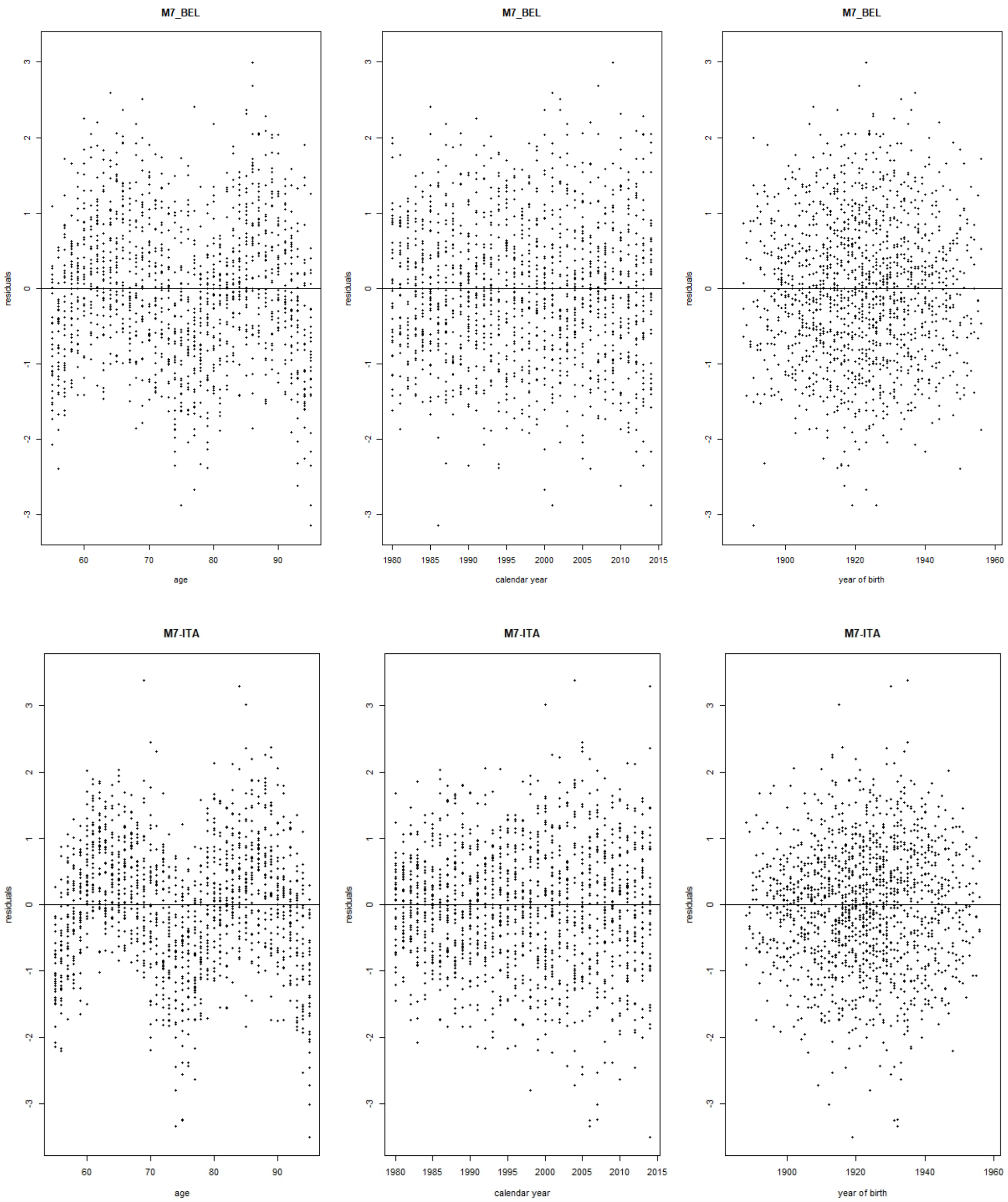

Figure 2.18 M7 scatter plots for Belgian and Italian populations

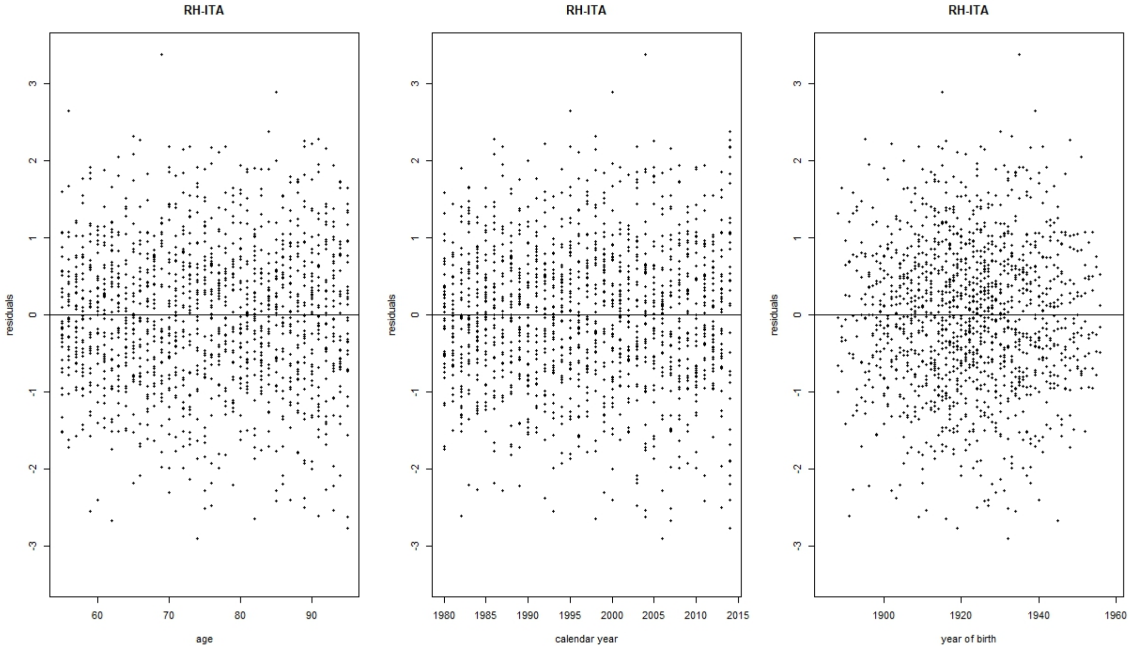

Figure 2.19 RH scatter plot for Italian population

Model	Log-likelihood	AIC	BIC	N.Parameters
LC	-8654.63	17539.25	18144.21	115
APC	-7832.95	15949.89	16696.89	142
CBD	-15213.69	30567.37	30935.61	70
M6	-7521.5	15317.00	16037.69	137
M7	-7462.34	15266.68	16166.23	171
Plat	-7363.83	15077.67	15998.26	175

Table 2.3 Belgian values

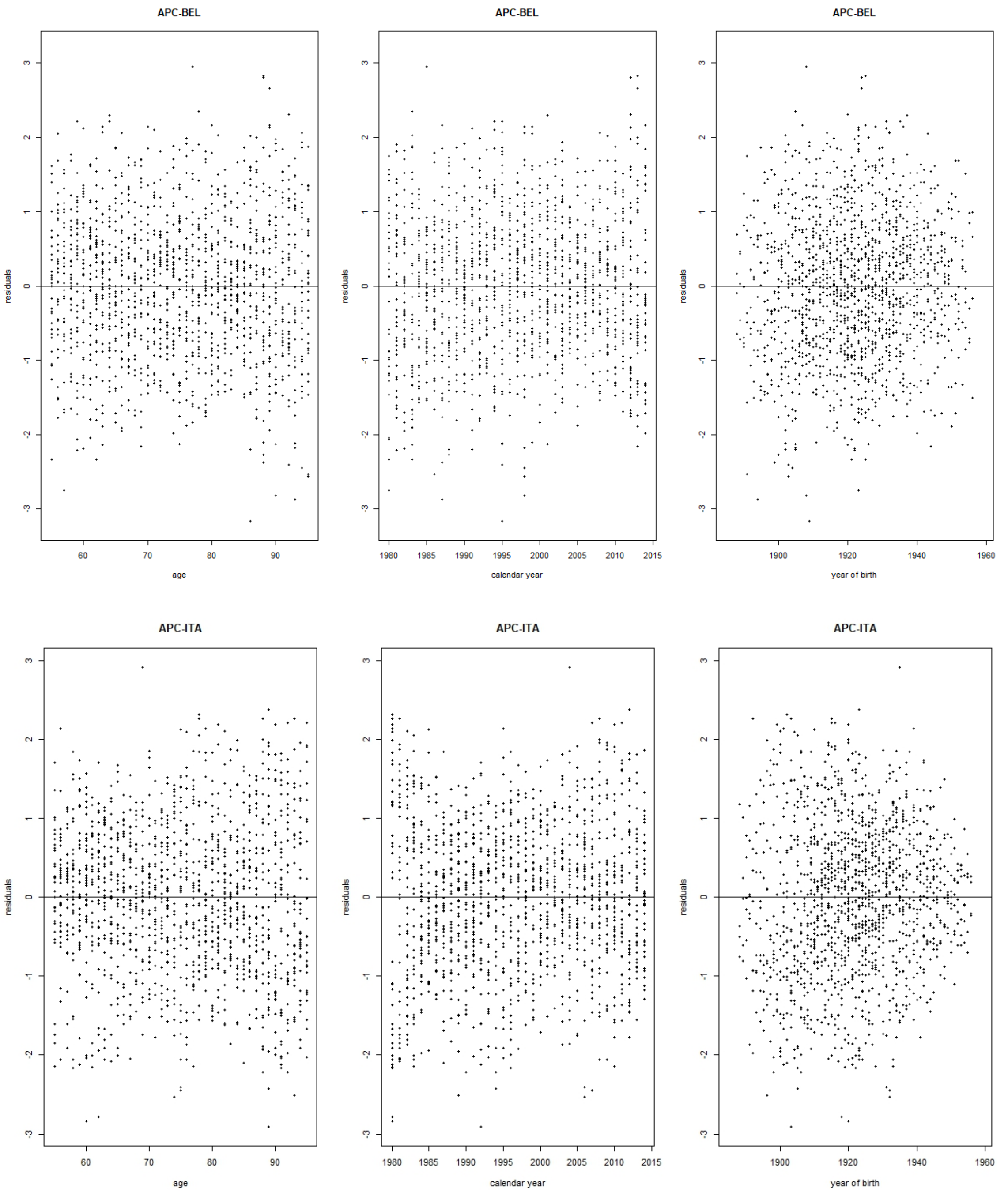

Figure 2.20 APC scatter plots for Belgian and Italian populations

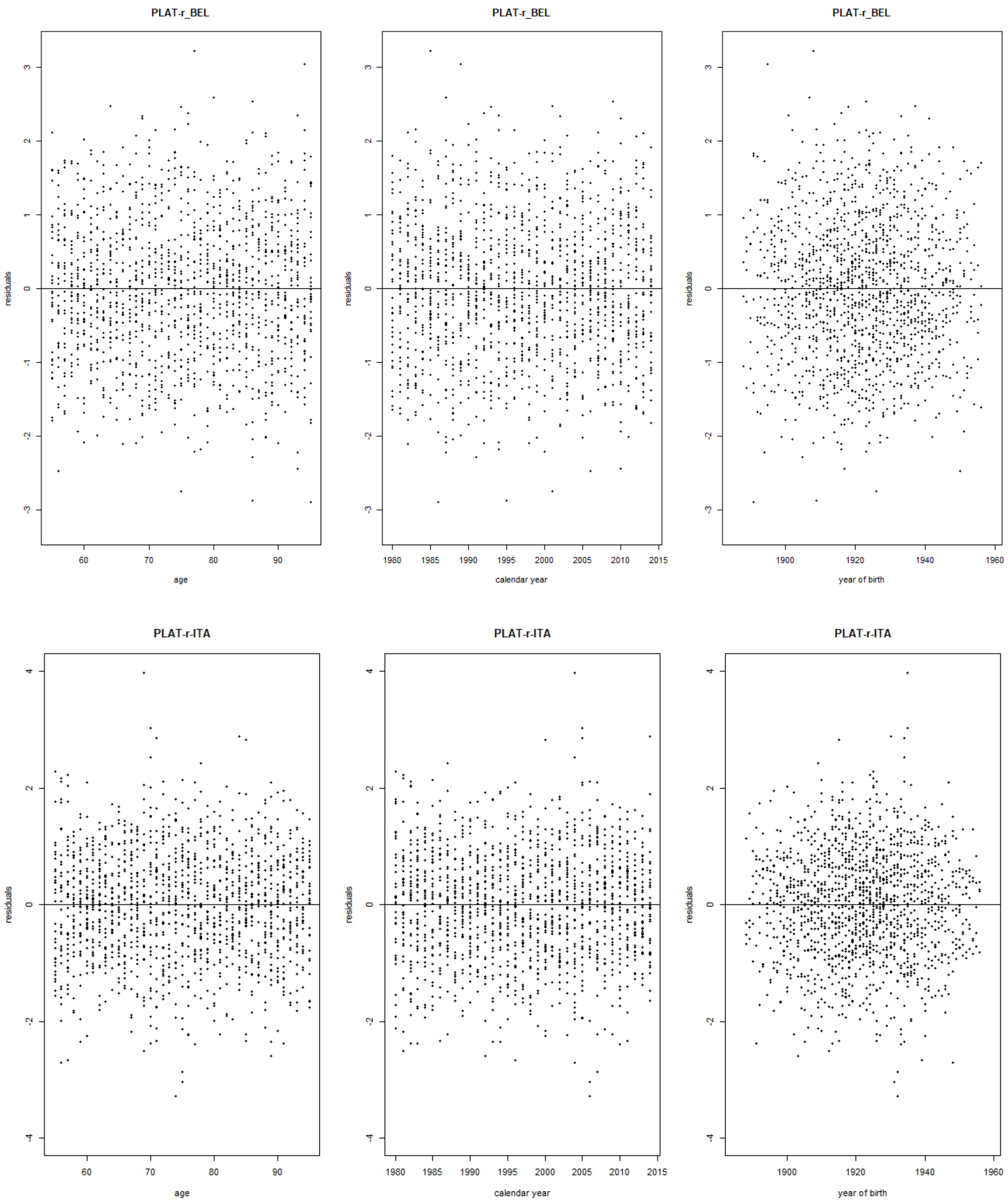

Figure 2.21 M6 scatter plots for Belgian and Italian populations

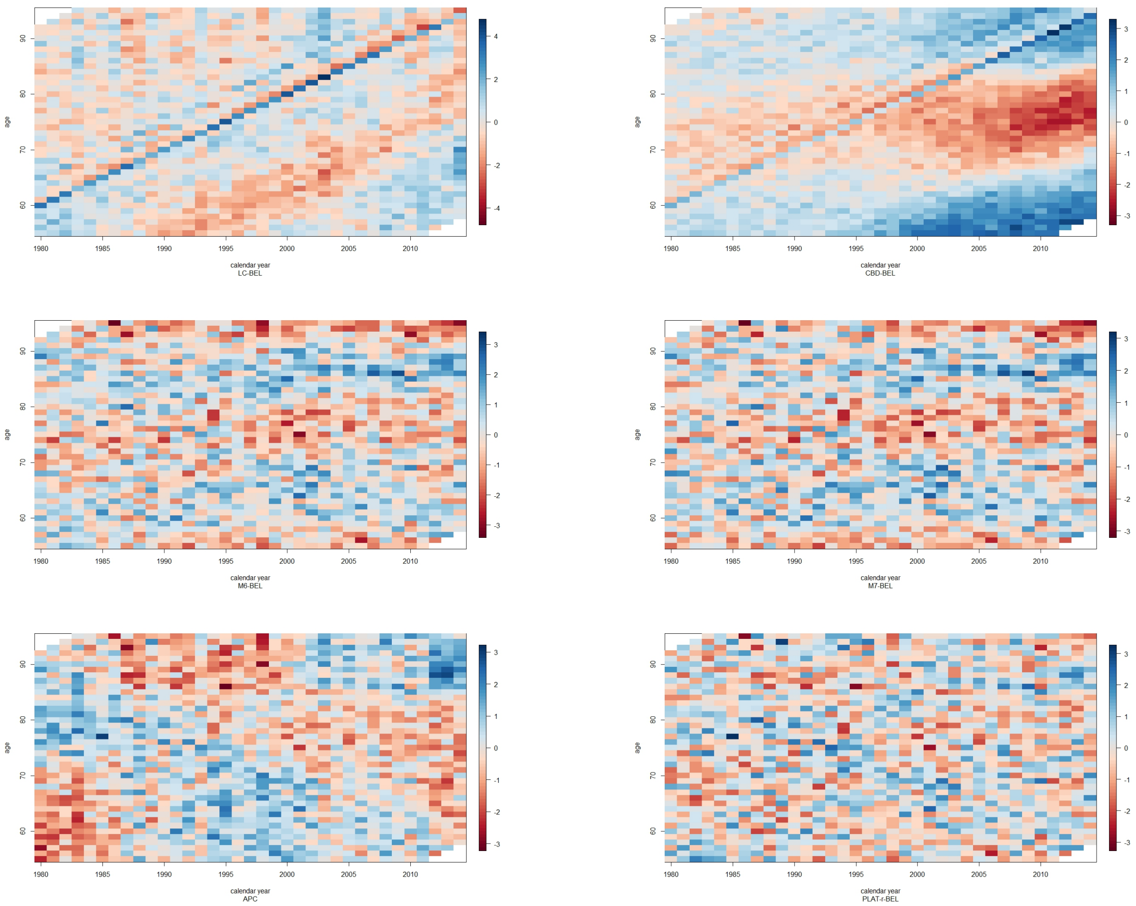

Figure 2.22 Belgian population: heat-maps

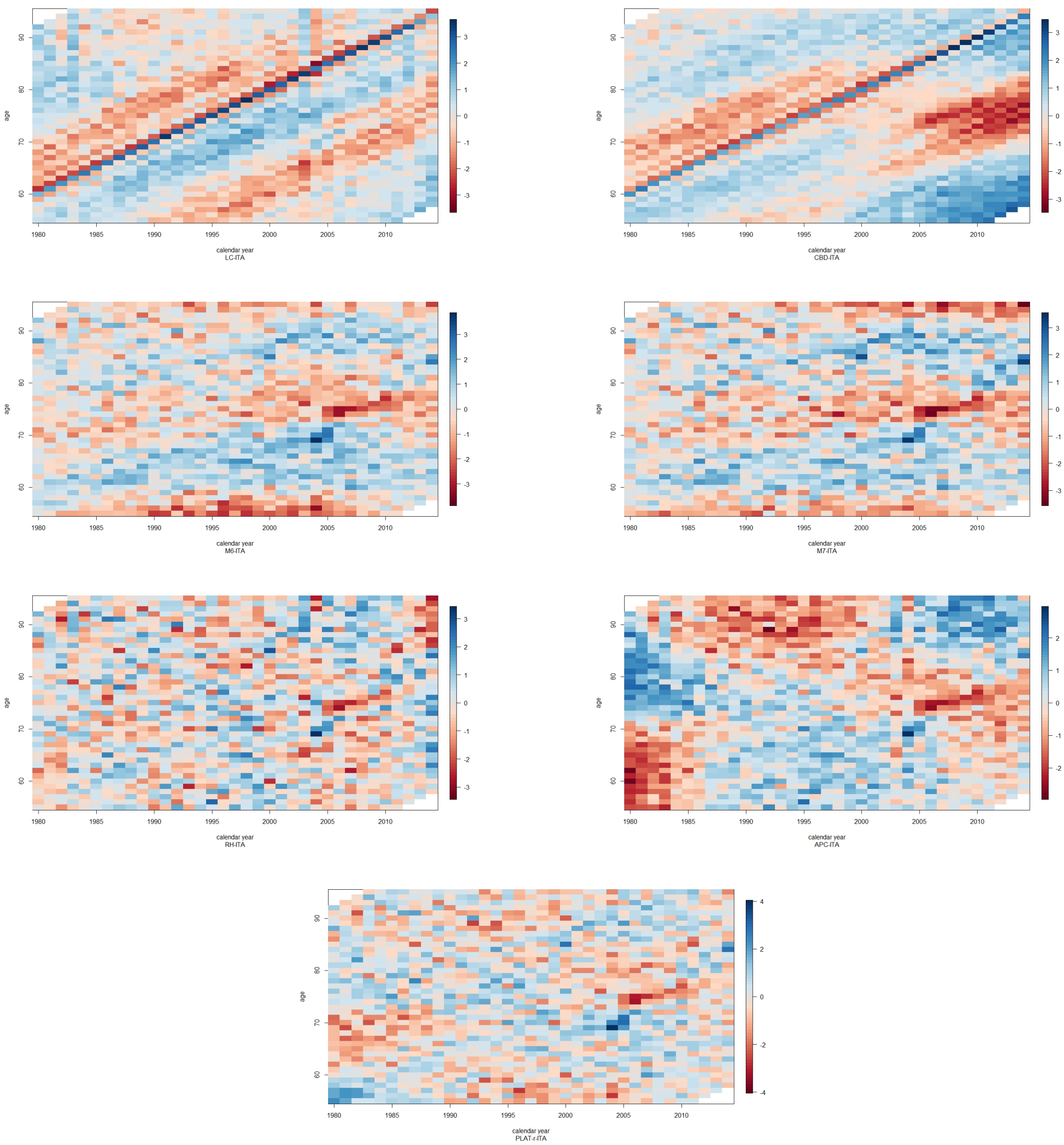

Figure 2.23 Italian population: heat-maps

Model	Model Cohort effect
APC	ARIMA(1,1,0)
RH	ARIMA(1,1,0)
M6	ARIMA(1,1,0)
M7	ARIMA(1,1,0)
Plat	ARIMA(0,1,0)

Table 2.4 Parameters of the ARIMA models chosen for the cohort effect - Belgian population

datasets. The downturn is that the models mentioned are not parsimonious and they used a number of parameters sometimes double respect to the other ones. To obtain a better information we consider the forecasting results in the next section.

2.6 Forecasting

The models described in the previous sections can give projections useful to consider the future impacts. The forecasting can be obtained simulating via time series technique the period terms $\kappa_t^{(j)}$ with $j = 1, ..., N$ and the cohort term γ_{t-x}. The period terms can be projected (see Haberman and Renshaw 2011) assuming a multivariate random walk with drift:

$$\underline{\kappa}_t = \underline{\upsilon} + \underline{\kappa}_{t-1} + \underline{\varepsilon}_t, \qquad \underline{\kappa}_t = \begin{pmatrix} \kappa_t^{(1)} \\ \vdots \\ \kappa_t^{(N)} \end{pmatrix} \qquad \underline{\varepsilon}_t = N\left(0, \Sigma\right),$$

with $\underline{\upsilon}$ a drift parameters vector Nx1 and $\underline{\varepsilon}_t$ a multivariate white noise (i.e. mean zero and a variance-covariance matrix Σ NxN).Alternatively it is possible to implement a multivariate adaption of the Algorithm 2 in Haberman and Renshaw [2009] to simulate $\underline{\kappa}_t$ and simulate the cohort index with an ARIMA process. We assume that the cohort index follows an univariate ARIMA(p,d,q) process with the parameters (p,d,q) indicated in table 2.6 for the Italian population and ARIMA(1,1,0) for the Belgian population.

The choice of parameters is obtained (like in Renshaw and Haberman, 2006,Plat 2009 and Cairns et al., 2011) considering all relevant ARIMA(p,d,q) processes (with p,q=0,1,2 and d= 0,1,2) and selecting the most favorable process in terms of BIC and in terms of forecasting. Figures 2.24 and 2.25 show the results of forecasting both for Italian and Belgian populations. The simulations are performed with 1000 trajectories. The plots show the confidence intervals of 50%, 60%, 90%, 95%.

From 2.24 and 2.25 we can obtain useful informations about the ability of the models to provide forecasts with the data considered. In particular:

- The LC and APC give wider fans at age 90 than to the age 60 both for Italian and Belgian populations. The M7 model has the same behaviors in the ages 60,

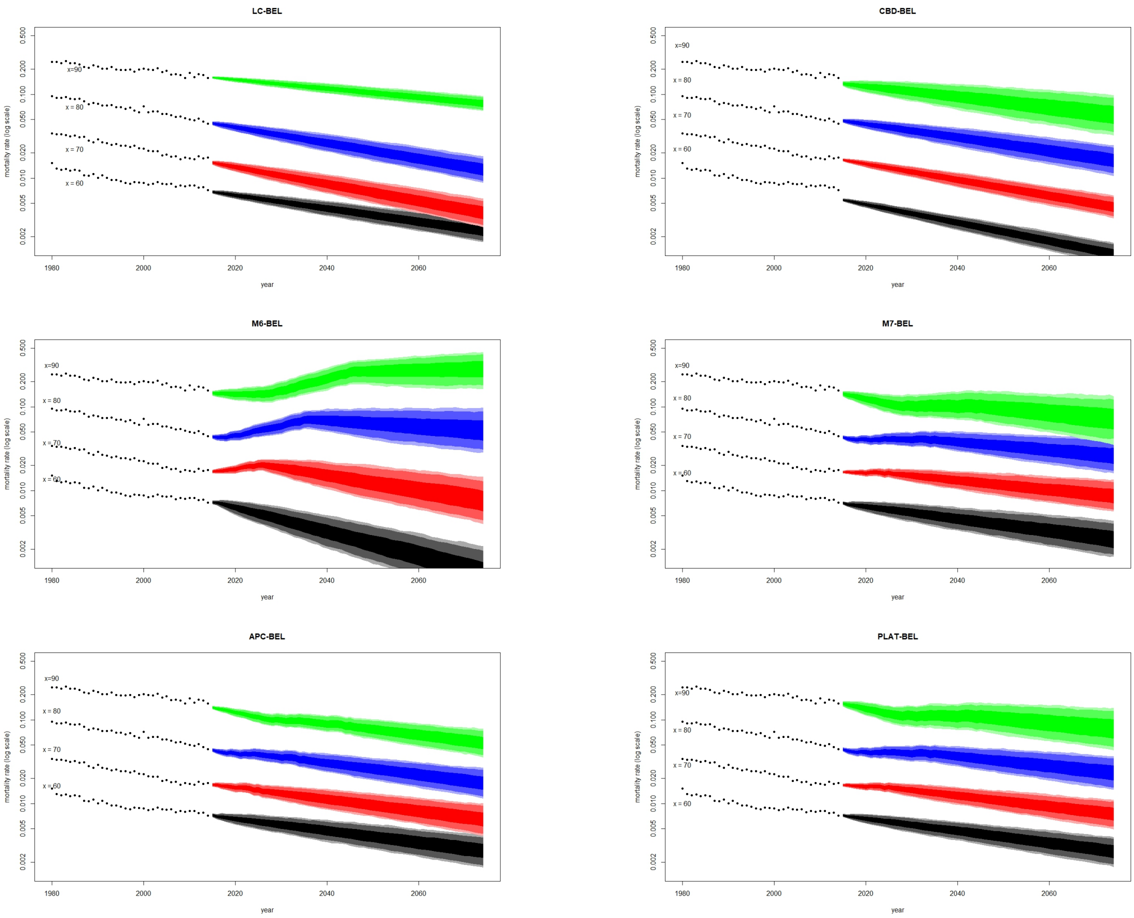

Figure 2.24 Simulations: Belgian population

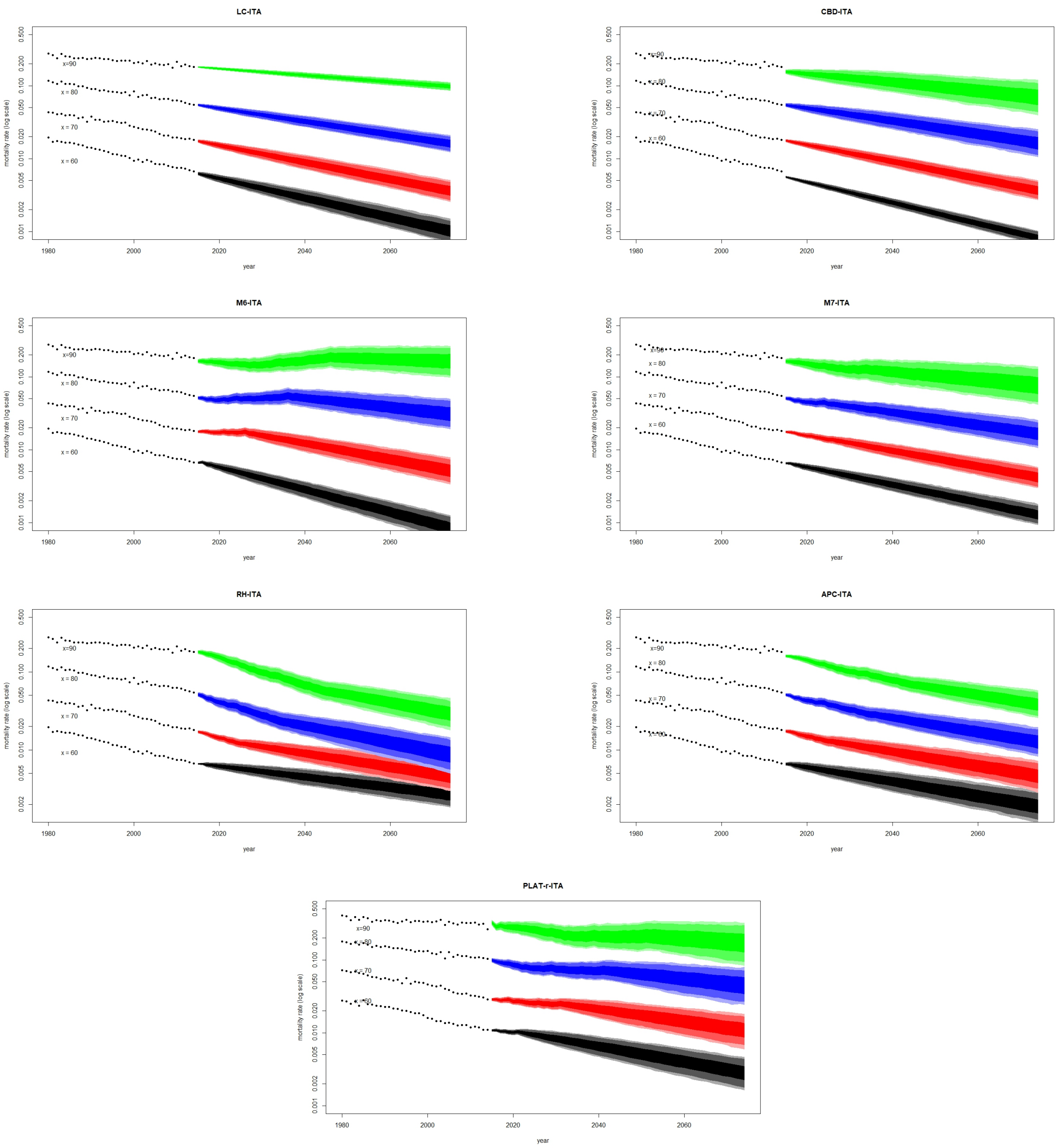

Figure 2.25 Simulations: Italian population

70 and 80 for the Belgian population. These behaviors are not coherent with the data used in this work. LC and APC could be not used to obtain forecasts for the Belgian and Italian population for the age and years range used.

- The M6 model both in Belgian and Italian case conduct to an increase of the mortality rates along the calendar years. The Plat model has the same behavior for the Italian population at the age 60. The result of the Plat model can be explained remembering that in this work we use the simplified version, particularly suitable for the older ages.

- The models produce different fans in terms of central trends and of uncertainty levels.

2.6.1 Financial results

In this section we use the previous forecasts to calculate the the premiums of temporary annuities for a person who wants to receive one euro per year until the age 95. The interest rate is set to zero to take into consideration only the demographic effects. We calculate the premiums for different subscription ages (60, 70, 80, 90) and for different percentile of the survivor probabilities obtained.

2.7 Conclusions

In this work we compare six different mortality models in the Belgian population case and seven mortality models in the case of Italian population. This difference is due to the impossibility for the gnm algorithm to converge in the case of RH model for the Belgian dataset. The first conclusion that we can take is that not all mortality models can be used to estimate the mortality of a population. Some models can fit better, in terms of BICs and AICs, but could not provide reliable simulations. Models with better results in terms of BICs and AICs could be over-parametrized. In the Italian population, for example, the CBD have higher values in terms of BIC and AIC respect to the Plat model but the number of parameters used is quite low and the simulations provided seem to be more reliable. On the other hand the CBD doesn't take into account the cohort effect that is clearly important for the two populations.

All this results have impacts in terms of pricing as we can see in 2.6.1. We know that the M6 model is not reliable for the data used because it provides non feasible forecasts. In particular it overestimates the mortality given systematic underestimated premiums.

In terms of forecasting and pricing the M7, simplified Plat (only for higher ages) and CBD (that not considers the cohort effect, seem to be the better solutions for the Italian population. The CBD and the Plat seem to be better for the Belgian population.

60	97.5	95	50	5	2.5
LC	26.43	26.26	25.39	24.46	24.26
APC	26.21	26.01	24.98	23.86	23.63
CBD	26.75	26.52	25.38	24.16	23.93
M6	22.24	22.01	20.79	19.67	19.46
M7	25.36	25.17	23.96	22.65	22.40
PLAT	25.41	25.19	23.98	22.69	22.44

70	97.5	95	50	5	2.5
LC	16.93	16.81	16.23	15.62	15.50
APC	17.15	17.02	16.31	15.58	15.42
CBD	17.11	16.95	16.10	15.22	15.08
M6	15.45	15.27	14.38	13.53	13.36
M7	16.99	16.85	15.99	15.07	14.90
PLAT	16.91	16.73	15.86	14.97	14.79

80	97.5	95	50	5	2.5
LC	8.70	8.65	8.38	8.12	8.07
APC	9.48	9.40	9.04	8.67	8.60
CBD	9.26	9.17	8.67	8.17	8.09
M6	9.05	8.95	8.44	7.94	7.85
M7	9.63	9.54	9.02	8.48	8.38
PLAT	9.39	9.28	8.78	8.27	8.17

90	97.5	95	50	5	2.5
LC	2.90	2.88	2.84	2.80	2.79
APC	3.20	3.18	3.10	3.03	3.01
CBD	3.32	3.30	3.17	3.05	3.04
M6	3.13	3.11	2.98	2.86	2.84
M7	3.23	3.21	3.06	2.92	2.89
PLAT	3.08	3.06	2.92	2.77	2.74

Table 2.5 Annuity premiums: Belgium

60	97.5	95	50	5	2.5
LC	25.49	25.35	24.61	23.84	23.67
RH	28.08	27.91	26.94	25.80	25.59
APC	27.16	26.99	26.02	24.98	24.78
CBD	26.25	26.03	24.86	23.62	23.39
M6	23.88	23.62	22.42	21.22	21.00
M7	26.19	25.95	24.73	23.46	23.25
PLAT	25.25	25.05	23.84	22.61	22.38

70	97.5	95	50	5	2.5
LC	15.85	15.76	15.29	14.82	14.72
RH	17.90	17.76	16.94	16.03	15.86
APC	17.22	17.09	16.40	15.70	15.56
CBD	16.50	16.32	15.45	14.57	14.39
M6	15.75	15.58	14.72	13.85	13.69
M7	16.65	16.48	15.59	14.69	14.53
PLAT	16.46	16.3	15.38	14.50	14.36

80	97.5	95	50	5	2.5
LC	7.97	7.93	7.73	7.53	7.50
RH	9.06	8.97	8.49	8.02	7.92
APC	8.94	8.88	8.52	8.17	8.11
CBD	8.68	8.59	8.10	7.60	7.52
M6	8.63	8.54	8.05	7.56	7.47
M7	8.88	8.79	8.27	7.75	7.65
PLAT	8.74	8.64	8.10	7.59	7.50

90	97.5	95	50	5	2.5
LC	2.7	2.69	2.66	2.63	2.62
RH	2.82	2.8	2.70	2.59	2.57
APC	3.00	2.98	2.91	2.83	2.82
CBD	3.08	3.06	2.94	2.82	2.79
M6	2.98	2.95	2.83	2.71	2.69
M7	3.00	2.98	2.84	2.69	2.66
PLAT	2.88	2.85	2.70	2.55	2.52

Table 2.6 Annuity premiums: Italy

3

Optimal mix between PAYG and Funded Pension Schemes

3.1 Introduction

We develop a continuous time model to approach the consumer's problem of optimal mix between two pension systems. We consider two stylized pension systems (corresponding to the PAYG and the fully funded) that pay lump sums at the retirement time. We consider a consumer who wants to maximize the inter-temporal utility of his/her consumption during the working life and the utility of the lump sums obtained at the retirement time. The two pension systems accrue the contributions at different rates: a demographic rate in the case of PAYG and a financial rate in the case of the funded pension system.

3.2 Model structure

In this section we describe the building blocks of our model: the consumer's utility function, the financial market, and the two pension systems.

3.2.1 The consumer

The consumer wants to maximize the expected utility of his/her inter-temporal consumption and final wealth during the entire working life $[t_0, \tau]$ where t_0 denotes the time of entrance in the labor market and τ denotes the retirement time. We define the following utility of consumption:

$$U\left(c_1\left(t\right)\right) = \chi_{c_1} \frac{\left(c_1\left(t\right) - \alpha_{c_1}\right)^{1-\delta}}{1-\delta}. \tag{3.1}$$

The utility of wealth, instead, is:

$$U(R(t)) = \chi_R \frac{(R(t) - \alpha_R)^{1-\delta}}{1-\delta}, \tag{3.2}$$

where $c_1(t)$ is the consumption, α_{c_1} is the consumption subsistence level, α_R is the minimum wealth level, and δ is a constant describing the risk aversion. We suppose α_{c_1} and α_R to be constant during the entire life.

3.2.2 The financial market

In a filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\in[0,\tau]}, \mathbb{P}\right)$ we define a frictionless continuously opened arbitrage free and complete financial market over the (fixed) time interval $[t_0, \tau]$. The financial market is composed of two asset classes:

- a riskless asset whose price process $G(t, z(t))$ is given by:

  $$dG(t, z(t)) = G(t, z(t))\, r(t, z(t))\, dt \quad G(t_0) = G_0, \tag{3.3}$$

  where $z(t)$ is the vector of the state variables describing any stochastic variable different from wealth (for instance the riskless interest rate, the market price of risk, the asset volatility, and the demographic rates) defined as:

  $$dz(t) = \mu_z(t, z(t))\, dt + \Omega(t, z(t))'\, dW(t). \tag{3.4}$$

  It is quite common in literature (Battocchio and Menoncin, 2004, Brigo and Mercurio, 2007, Duffie, 2010) to interpret $G(t, z(t))$ as a bank account paying the instantaneous interest rate $r(t, z(t))$ without any default risk.

- n risky assets whose price dynamics are described (in matrix form) by:

  $$dS(t, z(t)) = \mu(t, z(t))\, S(t, z(t))\, dt + \Sigma(t, z(t))\, S(t, z(t))\, dW(t), \tag{3.5}$$

  or

  $$I_S^{-1} dS(t, z(t)) = \mu(t, z(t))\, dt + \Sigma(t, z(t))'\, dW(t),$$

  where

  $$S(t_0) = S_0,$$

  and I_s is the diagonal matrix containing the elements of the vector $S(t, z_t)$:

  $$I_s(t, z(t)) = \begin{bmatrix} S_1(t, z(t)) & 0 & ... & 0 \\ 0 & S_2(t, z(t)) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & S_n(t, z(t)) \end{bmatrix},$$

  and $W(t)$ is a k-dimensional Wiener process.

We underline that the main assumptions are the arbitrage free and the completeness of the market. The assumptions imply the existence of $\Sigma(t, z(t))^{-1}$ and the Girsanov's theorem allows us to switch from the historical probability ($\mathbb{P}$) to the the risk-neutral probability ($\mathbb{Q}$) by using $dW^{\mathbb{Q}}(t) = \xi(t, z(t))\,dt + dW(t)$. Where $\xi(t, z(t))$ is defined as:

$$\Sigma(t, z(t))' \,\xi(t, z(t)) = \mu(t, z(t)) - r(t, z(t)\,1)$$

and 1 is the vector of ones such that $I_s(t, z(t))\,1 = S(t, z(t))$.

3.2.3 The unfunded (PAYG) pension system

The PAYG pension system, as described in the first chapter, is based on the principle that the contributions paid by the workers are immediately used to pay pensions. We can define it as a “notional” (non-financial) system where contributions are accrued in a virtual account (**?**). We define an unfunded system where:

- the individuals can accrue the contributions, $c_2(t, z(t))$, in a notional wealth, $R_2(t, z(t))$;
- $R_2(t, z(t))$ accrues at a demographic rate $d(t, z(t))$ (Devolder and Melis, 2015).

The contributions $c_2(t, z(t))$ are treated as nominal consumptions. Indeed, we need a sort of investment in the PAYG that can not be withdrawn by the consumer. Furthermore, given two different instants of time, $t_1 < t_2$, it must hold: $R_2(t_1) \leq R_2(t_2)$. The differential of the total wealth invested in the PAYG pension scheme at t is:

$$dR_2(t, z(t)) = \left(R_2(t, z(t))\,d(t) + c_2(t, z(t))\right) dt, \tag{3.6}$$

which is the first constrain of our maximization problem.

3.2.4 The funded pension system

The funded pension system considered for contributors, as discussed in the first chapter, can be approximated by an individual account managed by a professional investor. We define a funded system where:

- the individuals invest the remaining part of their salary at time t, $l(t)$, in the financial market described in 3.2.2;
- the wealth accumulated, $R_1(t, z(t))$, accrues with a financial rate given by the portfolio composition.

In every t, the static constraint of the wealth invested in the funded scheme is:

$$R_1(t, z(t)) = \theta_G(t)\,G(t, z(t)) + \theta_S(t)\,S(t, z(t)) \tag{3.7}$$

where $\theta_G(t)$ is the quantity of risk-free assets hold in portfolio and $\theta_S(t)$ is the the vector with the quantities of the risky assets hold in portfolio. The dynamic constraint of the wealth invested in the funded scheme is:

$$\begin{aligned} dR_1(t, z(t)) &= \theta_G(t)\, dG(t, z(t)) + \theta_S(t)\, dS(t, z(t)) \\ &\quad - c_1(t, z(t))\, dt - c_2(t, z(t))\, dt + l(t, z(t))\, dt \end{aligned} \tag{3.8}$$

The differential of the total wealth invested in the PAYG pension scheme at t is (see Appendix A) :

$$\begin{aligned} dR_1(t, z(t)) = &\Big(R_1(t, z(t))\, r(t, z(t)) + \theta_s(t)' I_s(t, z(t)) (\mu(t, z(t)) - r(t, z(t))\, 1)\Big) dt \\ &(-c_1(t) - c_2(t) + l(t, z(t)))\, dt + \theta_s(t)' I_s \Sigma(t, z(t))' dW(t), \end{aligned} \tag{3.9}$$

The 3.9 is another constraint of our maximization problem. For the sake of simplicity we neglect any functional dependences but the time.

3.3 The maximization problem

In this section we define the maximization problem of the consumer. We consider a consumer who wants to maximize his/her utility of consumption during the working period. The 3.4, 3.6 and 3.9 are the constraints of our maximization problem. The objective function is:

$$\mathbb{E}_{t_0}\left[\int_{t_0}^{\tau} I_{s<T} U_c(c_1(s))\, e^{-\int_{t_0}^{s}\rho(u)du} ds + I_{T\geq\tau} U_R(R_1(\tau) + R_2(\tau))\, e^{-\int_{t_0}^{\tau}\rho(u)du}\right] \tag{3.10}$$

with ρ_t the subjective discount rate describing the time's value of an agent, $I_{s<T}$ is an indicator function valued 1 if the death time, T, occurs later than s and 0 otherwise, $I_{T\geq\tau}$ is an indicator function valued 1 if the death time T occurs later than the retirement time and 0 otherwise. Starting from 3.4, 3.6, 3.9 and 3.10 we can write the following maximization problem:

$$\begin{aligned} \max_{\{c_1(t),\theta_s(t)\}_{t\in[t_0,\tau]}} \mathbb{E}_{t_0} &\left[\int_{t_0}^{\tau} U_c(c_1(s))\, e^{-\int_{t_0}^{s}\rho(u)+\lambda(u)du} ds \right. \\ &\left. + I_{T\geq\tau} U_R(R_1(\tau) + R_2(\tau))\, e^{-\int_{t_0}^{\tau}\rho(u)+\lambda(u)du}\right] \\ s.t.\quad & dz(t) = \mu_z(t)\, dt + \Omega(t)' dW_t \\ & dR_1(t) = \Big(R_1(t)(r(t) + \lambda(t)) + \theta_s' I_s(\mu(t) - r(t)\, 1)\Big) dt \\ & + (w(t) - c_1(t) - c_2(t))\, dt \\ & + \theta_s(t)' I_s \Sigma(t)' dW(t) \\ & dR_2(t) = (R_2(t)(d(t) + \lambda(t)) + c_2(t))\, dt \end{aligned} \tag{3.11}$$

Where $\lambda(t)$ indicates the force of mortality.
Proposition 1 The solution of the maximization problem 3.11 is (see appendix B):

$$\begin{aligned} I_s\theta_s(t)^* = & \frac{(\beta R_1 + vR_2 - H)}{\delta\beta}\left(\Sigma'\Sigma\right)^{-1}(\mu(t) - r(t)\,1) \\ & + \frac{(\beta R_1 + vR_2 - H)}{F\beta}\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega F_z \\ & - \frac{1}{\beta}\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega(\beta_z R_1 + v_z R_2 - H_z) \\ & + \frac{(\beta R_1 + vR_2 - H)}{\delta\beta^2}\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\beta_z \end{aligned} \tag{3.12}$$

$$\begin{aligned} c_1(t)^* &= \alpha_{c_1} + \left(\frac{1}{\chi_{c_1}}F^\delta\beta(\beta R_1 + vR_2 - H)^{-\delta}\right)^{-\frac{1}{\delta}} \\ &= \alpha_{c_1} + \left(\frac{\beta}{\chi_{c_1}}\right)^{-\frac{1}{\delta}}\frac{\beta R_1 + vR_2 - H}{F} \end{aligned} \tag{3.13}$$

with:

$$\begin{aligned} F(t,z_t) = & \mathbb{E}_t^{\mathbb{Q}_\delta}\left[\int_t^\omega \chi_{c_1}^{\frac{1}{\delta}}\beta^{1-\frac{1}{\delta}}e^{-\frac{1}{\delta}\int_t^s\left(\rho(u)+\lambda(u)+\frac{\delta-1}{\delta}\xi'(u)\xi(u)+\frac{\delta-1}{\delta\beta}\xi'(u)\Omega\beta_z+(\delta-1)(r(u)+2\lambda(u)+d(u))\right)du}ds\right] \\ & + \mathbb{E}_t^{\mathbb{Q}_\delta}\left[\chi_R^{\frac{1}{\delta}}e^{-\frac{1}{\delta}\int_t^s\left(\rho(u)+\lambda(u)+\frac{\delta-1}{\delta}\xi'(u)\xi(u)+\frac{\delta-1}{\delta\beta}\xi'(u)\Omega\beta_z+(\delta-1)(r(u)+2\lambda(u)+d(u))\right)du}\right] \end{aligned}$$

$$\begin{aligned} \beta(t,z_t) = & \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^\omega(d(s)+\lambda(s))ds}\right] \\ \nu(t,z_t) = & \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^\omega(r(s)+\lambda(s))ds}\right] \\ H(t,z_t) = & \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\omega((\beta(u)-v(u))\,c_2(u) - \beta(u)\,l(u) + \beta(u)\,\alpha_{c_1})\,e^{-\int_t^u(r(s)+2\lambda(s)+d(s))ds}du\right] \\ & + \mathbb{E}_t^{\mathbb{Q}}\left[\alpha_R e^{-\int_t^\omega(r(s)+2\lambda(s)+d(s))ds}\right], \end{aligned}$$

where we used the new probability measure

$$dW(t)^{\mathbb{Q}_\delta} = dW(t) - \frac{\delta-1}{\delta}\xi(t)\,dt,$$

that has two relevant proprieties:

- for $\delta = 1$, i.e. for a log-utility agent, $\mathbb{Q}_\delta$ coincides with the historical probability;
- for $\delta \to \infty$, i.e. an infinitely risk averse agent, the probability $\mathbb{Q}_\delta$ coincides with $\mathbb{Q}$.

As described in Menoncin and Regis [2017] we can define the new probability measure as:

$$dW(t)^{\mathbb{Q}_\delta} = \frac{1}{\delta} dW(t) + \left(\frac{1}{\delta} - 1\right) dW(t)^{\mathbb{Q}},$$

which is a linear combination of the Wiener processes under the risk neutral and the historical probabilities.

The two functions $H(t, z(t))$ and $F(t, z(t))$ have the following meaning:

- $H(t, z(t))$ is the expected discounted value (in both financial and actuarial terms), under $\mathbb{Q}$, of: (i) the minimum wealth, (ii) the minimum consumption level, (iii) the consumption c_2 weighted by $(\beta - \nu)$, net of (iv) the wage;

- $F(t, z(t))$ can be defined as a "global" discount factor. Indeed, it is the expected value, under $\mathbb{Q}_\delta$, of the sum of all the discounted factors for both the consumption stream and the final wealth.

4

Optimal Retirement Time in a Double Stochastic Pension Schemes

4.1 Introduction

In this chapter we consider the continuous time problem of a consumer who can decide when to retire in order to maximize his\her inter-temporal utility of consumption over the entire life. As in the previous chapter we consider two stylized pension systems (PAYG and funded). The contribution to the PAYG is assumed to be constant , whereas the funded is viewed as a third pillar pension system where the consumer can choose the amount of contributions to charge. The funded pension system doesn't pay a lump sum at the retirement time and it continues until the death. The PAYG will pay a lump sum at the retirement time.

4.2 The Model

4.2.1 The consumer

A representative consumer wants to maximize the inter-temporal utility of his/her consumption over the entire life. The utility function of consumption belongs to the Hyperbolic Absolute Risk Aversion (HARA) family:

$$U\left(c_1\left(t\right)\right)=\frac{\left(c_1\left(t\right)-\alpha_{c_1}\right)^{1-\delta}}{1-\delta},$$

where $c_1\left(t\right)$ is the consumption, α_{c_1} is the consumption subsistence level that we suppose constant during the entire life and δ is a constant that measures the risk aversion. The utilities of the funded pension and unfunded pension schemes belong to the HARA family too. We define:

$$U\left(R_1\left(t\right)\right)=\frac{\left(R_1\left(t\right)-\alpha_{R_1}\right)^{1-\delta}}{1-\delta}$$

$$U(R_2(t)) = \frac{(R_2(t) - \alpha_{R_2})^{1-\delta}}{1-\delta}$$

where α_{R_1} and α_{R_2} are respectively the minimum wealth level of the funded and the PAYG pension systems.

4.2.2 The financial market

We consider the same financial market of the previous chapter.

4.2.3 Funded Pension Scheme

The revenue of the consumer comes from both the salary $l(t)$ and the return on portfolio of the funded system. The revenue is used for financing consumption $c_1(t)$, the contributions to the funded pension scheme, and the (constant) contribution c_2 to the PAYG system. We consider a funded pension system where the consumer can invest the difference between the salary an the c_2. The funded system will pay a lump sum at the death time, T. If $T < \tau$ the lump sum is $\kappa_1 R_1(s)$ otherwise, $T \geq \tau$, the lump sum is $\kappa_2 R_1(s)$. The utility function of $R_1(t)$ is:

$$I_{s<\tau} U(\kappa_1 R_1(s)) + I_{s\geq\tau} U(\kappa_2 R_1(s)), \tag{4.1}$$

which is

$$\begin{cases} U(\kappa_1 R_1(s)), & s < \tau \\ U(\kappa_2 R_1(s)), & s \geq \tau \end{cases}$$

This function can alternatively be written as

$$U((I_{s<\tau}\kappa_1 R_1(s) + I_{s\geq\tau}\kappa_2 R_1(s))), \tag{4.2}$$

which is

$$\begin{cases} U(\kappa_1 R_1(s)), & s < \tau \\ U(\kappa_2 R_1(s)), & s \geq \tau \end{cases}$$

and so we have demonstrated that (4.1) and (4.2) are equivalent.

Thus, the whole problem can be simplified by defining

$$\kappa \equiv I_{s<\tau}\kappa_1 + I_{s\geq\tau}\kappa_2$$

which is the coefficient of $R_1(t)$. The expected value in t_0 of R_1 can be written as:

$$R_1(t_0) = \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega} m(t_0,s)\left(\kappa R_1(s)\lambda(s) + c_1(s) + (c_2 - l(s)) I_{s<\tau}\right) e^{-\int_{t_0}^{s}(r(u)+\lambda(u))du} ds\right] \tag{4.3}$$

where $m(t_0,t)$ is the stochastic discount factor defined as (see Menoncin, 2011):

$$m(t_0,t) = e^{-\frac{1}{2}\int_{t_0}^{t}\xi'(s)\xi(s)ds - \int_{t_0}^{t}\xi'(s)dW(s)},$$

with $\xi(t) = \left(\Sigma(t)'\right)^{-1}(\mu(t) - r(t)\mathbf{1})$. 4.3 is the first constraint of our maximization problem.

4.2.4 The PAYG Pension Scheme

The amount $R_2(t)$ is the total amount contributed in the PAYG pension scheme at time t and is accumulated, at the demographic rate $\mathrm{d}(t)$, by the contribution c_2. At the start of working life we can write the actuarial equivalence principle (under the risk neutral probability) as

$$\begin{aligned}\mathbb{E}_{t_0}^{\mathbb{Q}}\left[R_2(\tau)\,e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du}\right] &= \mathbb{E}_{t_0}^{\mathbb{Q}}\left[\int_{t_0}^{\tau} c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right]\\ &= \mathbb{E}_{t_0}^{\mathbb{Q}}\left[\int_{t_0}^{\omega} I_{s<\tau}c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right].\end{aligned} \tag{4.4}$$

We can rewrite 4.4 under the historical probability as:

$$\mathbb{E}_{t_0}\left[R_2(\tau)\,m(t_0,\tau)\,e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du}\right] = \mathbb{E}_{t_0}\left[m(t_0,\tau)\int_{t_0}^{\omega} I_{s<\tau}c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right]. \tag{4.5}$$

From 4.5 we can write c_2 as function of $R_2(\tau)$ in τ:

$$c_2 = \frac{R_2(\tau)\,e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du}}{\mathbb{E}_{\tau}^{\mathbb{Q}}\left[\int_{t_0}^{\omega} I_{s>\tau}e^{-\int_{t_0}^{s}((u)+\lambda(u))du}ds\right]}. \tag{4.6}$$

The 4.5 can be written in the following form:

$$R_2(t_0) = \mathbb{E}_{t_0}\left[R_2(\tau)\,m(t_0,\tau)\,e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du} - \int_{t_0}^{\omega} I_{s<\tau}m(t_0,s)\,c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right] \tag{4.7}$$

which is the second constraint of our maximization problem.

4.3 The maximization problem

The maximization problem is:

$$\begin{aligned}&\max_{\{c_1(s),R_1(s)\}_{s\in[t_0,\omega]},R_2(\tau)}\mathbb{E}_{t_0}\left[\begin{array}{c}\int_{t_0}^{\omega}\left(\frac{\left(c_1(s)-\alpha_{c_1}\right)^{1-\delta}}{1-\delta}+\frac{\left(\kappa R_1(s)-\alpha_{R_1}\right)^{1-\delta}}{1-\delta}\lambda(s)\right)e^{-\int_{t_0}^{s}(\rho(u)+\lambda(u))du}ds\\ +\frac{\left(R_2(\tau)-\alpha_{R_2}\right)^{1-\delta}}{1-\delta}e^{-\int_{t_0}^{\tau}(\rho(u)+\lambda(u))du}\end{array}\right]\\ &\text{s.t. } R_1(t_0) = \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega} m(t_0,s)\left(\kappa R_1(s)\,\lambda(s)+c_1(s)+(c_2-l(s))\,I_{s<\tau}\right)e^{-\int_{t_0}^{s}(r(u)+\lambda(u))du}ds\right]\\ &\quad R_2(t_0) = \mathbb{E}_{t_0}\left[R_2(\tau)\,m(t_0,\tau)\,e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du} - \int_{t_0}^{\omega} I_{s<\tau}m(t_0,s)\,c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right]\end{aligned} \tag{4.8}$$

Proposition 2 The solution of 4.8 is (see Appendix C):

$$c_1^*(t) = \alpha_{c_1} + \frac{R_1^*(t) - H_1(t,z(t))}{F_1(t,z(t))}$$

$$R_1^*(t) = H_1(t, z(t)) + \phi_1^{-\frac{1}{\delta}} m(t_0, t)^{-\frac{1}{\delta}} e^{\frac{1}{\delta}\int_{t_0}^{t}(r(u)-\rho(u))du} F_1(t, z(t))$$

$$R_2^*(\tau) = \alpha_{R_2} + m(t, \tau)^{-\frac{1}{\delta}} e^{\frac{1}{\delta}\int_t^{\tau}(d(u)-\rho(u))du} \frac{R_2^*(t) - H_2(t, z(t))}{F_2(t, z(t))}.$$

with:

$$\begin{gathered} H_1(t, z(t)) \equiv \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{\omega} \alpha_{R_1}\lambda(s) e^{-\int_t^s(r(u)+\lambda(u))du} ds + \int_t^{\omega} \alpha_{c_1} e^{-\int_t^s(r(u)+\lambda(u))du} ds \right. \\ \left. - \int_t^{\omega} e^{-\int_t^s(\rho(u)+\lambda(u))du} I_{\tau>s} l(s)\, ds + \int_t^{\omega} I_{s<\tau} c_2 e^{-\int_t^s(r(u)+\lambda(u))du} ds\right] \end{gathered}$$

$$\begin{gathered} F_1(t, z(t)) \equiv \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{\omega} \lambda(s)\, m(t, s)^{-\frac{1}{\delta}} e^{-\int_t^s\left(\frac{\delta-1}{\delta}r(u)+\frac{1}{\delta}\rho(u)+\lambda(u)\right)du} ds \right. \\ \left. + \int_t^{\omega} m(t, s)^{-\frac{1}{\delta}} e^{-\int_t^s\left(\frac{\delta-1}{\delta}r(u)+\frac{1}{\delta}\rho(u)+\lambda(u)\right)du} ds\right] \end{gathered}$$

$$\begin{gathered} H_2(t, z(t)) \equiv \mathbb{E}_t^{\mathbb{Q}}\left[\alpha_{R_2} e^{-\int_t^{\tau}(d(u)+\lambda(u))du} - \int_t^{\tau} c_2 e^{-\int_t^s(d(u)+\lambda(u))du} ds\right] \\ F_2(t, z(t)) \equiv \mathbb{E}_t^{\mathbb{Q}}\left[m(t, \tau)^{-\frac{1}{\delta}} e^{-\int_t^{\tau}\left(\frac{\delta-1}{\delta}d(u)+\frac{1}{\delta}\lambda(u)\right)du}\right]. \end{gathered}$$

The optimal portfolio is:

$$I_S\theta_S^*(t) = \frac{R_1(t) - H_1(t)}{\delta}\Sigma^{-1}\xi(t) + \Sigma^{-1}\Omega\frac{\partial H_1(t)}{\partial z(t)} + \frac{R_1(t) - H_1(t)}{F_1(t)}\Sigma^{-1}\Omega\frac{\partial F_1(t)}{\partial z(t)}. \quad (4.9)$$

With the optimal decision rules presented above we can obtain the indirect utility/value function:

$$\begin{aligned} V(t_0; \tau) = {} & \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega} e^{-\int_{t_0}^{s}(\rho(u)+\lambda(u))du} \frac{1}{1-\delta}\left(\phi_1^{-\frac{1}{\delta}} m(t_0, s)^{-\frac{1}{\delta}} e^{\frac{1}{\delta}\int_{t_0}^{s}(r(u)-\rho(u))du}\right)^{1-\delta} ds\right] \\ & + \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega}\left[\frac{1}{1-\delta}\left(\phi_1^{-\frac{1}{\delta}} m(t_0, s)^{-\frac{1}{\delta}} e^{\frac{1}{\delta}\int_{t_0}^{s}(r(u)-\rho(u))du}\right)^{1-\delta}\right]\lambda(s) e^{-\int_{t_0}^{s}(\lambda(u)+\rho(u))du} ds\right] \\ & + \mathbb{E}_{t_0}\left[e^{-\int_{t_0}^{\tau}(\rho(u)+\lambda(u))du}\frac{1}{1-\delta}\left(\phi_2^{-\frac{1}{\delta}} m(t_0, \tau)^{-\frac{1}{\delta}} e^{\frac{1}{\delta}\int_{t_0}^{\tau}(d(u)-\rho(u))du}\right)^{1-\delta}\right] = \\ & \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega}\frac{1}{1-\delta}\phi_1^{\frac{\delta-1}{\delta}} e^{-\int_{t_0}^{s}\left(\frac{\delta-1}{\delta}r(u)+\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)\right)du-\frac{\delta-1}{\delta}\int_{t_0}^{s}\xi(u)dW(u)} ds\right] \\ & + \mathbb{E}_{t_0}\left[\int_{t_0}^{\omega}\frac{1}{1-\delta}\phi_1^{\frac{\delta-1}{\delta}}\lambda(s) e^{-\int_{t_0}^{s}\left(\frac{\delta-1}{\delta}r(u)+\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)\right)du-\frac{\delta-1}{\delta}\int_{t_0}^{s}\xi(u)dW(u)} ds\right] \quad (4.10) \\ & + \mathbb{E}_{t_0}\left[\frac{1}{1-\delta}\phi_2^{\frac{\delta-1}{\delta}} e^{-\int_{t_0}^{\tau}\left(\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)+\frac{\delta-1}{\delta}d(u)\right)du-\int_{t_0}^{\tau}\xi(u)dW(u)}\right]. \end{aligned}$$

To obtain the optimal retirement time,$\tau*$, we consider the first-order derivative of the value function at $t = 0$ with respect to τ:

$$\begin{aligned}
\frac{\partial V(t_0;\tau)}{\partial \tau} &= \mathbb{E}_{t_0}\left[\frac{1}{1-\delta}\phi_2^{\frac{\delta-1}{\delta}} e^{-\int_{t_0}^{\tau}\left(\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)+\frac{\delta-1}{\delta}d(u)\right)du-\int_{t_0}^{\tau}\xi(u)dW(u)}\right] \\
&= \mathbb{E}_{t_0}\left[\frac{1}{\delta-1}\phi_2^{\frac{\delta-1}{\delta}} e^{-\int_{t_0}^{\tau}\left(\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)+\frac{\delta-1}{\delta}d(u)\right)du-\int_{t_0}^{\tau}\xi(u)dW(u)}\right. \\
&\quad \left.\times\left(\frac{1}{\delta}\rho(\tau)+\lambda(\tau)+\frac{\delta-1}{2\delta}\xi^2(\tau)+\frac{\delta-1}{\delta}d(\tau)+\xi(\tau)\,dW(\tau)\right)\right] \\
&= \mathbb{E}_{t_0}\left[\frac{1}{\delta-1}\phi_2^{\frac{\delta-1}{\delta}} e^{-\int_{t_0}^{\tau}\left(\frac{1}{\delta}\rho(u)+\lambda(u)+\frac{\delta-1}{2\delta}\xi^2(u)+\frac{\delta-1}{\delta}d(u)\right)du-\int_{t_0}^{\tau}\xi(u)dW(u)}\right. \\
&\quad \left.\times\left(\frac{1}{\delta}\rho(\tau)+\lambda(\tau)+\frac{\delta-1}{2\delta}\xi^2(\tau)+\frac{\delta-1}{\delta}d(\tau)\right)\right]
\end{aligned} \tag{4.11}$$

We can see that 4.11 is positive and the consumer utility is increasing over the working period. The consumer will delay the retirement time until the legal retirement age.

This result is obtained starting from a utility that doesn't take into consideration leisure. Chen et al. [2017] demonstrate that (in an optimal consumption and asset allocation problem) early retirement is optimal if either the leisure gain through earlier retirement is highly appreciated or the standard of living (habit level) is low. In our model the consumer has a trade-off between the bequest motivation and the consumption. Choosing the retirement means to obtain $R_2(\tau)$, that can be invested in $R_1(\tau)$ and left in the future as $kR_1(T)$. From the other point of view, retirement implies giving up $l(t)$ that generates cash flows for both the pension systems and finances a part of the consumption $c_1(t)$.

Finally, we underline that the two optimal portfolios 3.12 and 4.9 coincide when the two functions:

$$\begin{aligned}
\beta(t,z_t) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{\omega}(d(s)+\lambda(s))ds}\right], \\
\nu(t,z_t) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{\omega}(r(s)+\lambda(s))ds}\right],
\end{aligned}$$

are constant.

A

Differential of the total wealth invested in the PAYG

Starting from the static constraint we can explicit $\theta_G(t)$ in 3.7

$$\theta_G(t) = \frac{R_1(t) - \theta_S S(t)}{G(t)} \tag{A.1}$$

and we can substitute it in 3.8 obtaining:

$$dR_1(t) = \left(\frac{R_1(t) - \theta_S S(t)}{G(t)}\right) dG(t) + \theta_S dS(t) + (l(t) - c_1(t) - c_2(t))\, dt, \tag{A.2}$$

substituting $dG(t)$ and $dS(t)$ respectively from 3.3 and 3.5 in A.2 we can obtain :

$$\begin{aligned} dR_1(t) = &\left(\frac{R_1(t) - \theta_S' I_s 1}{G(t)}\right) G(t)\, r(t)\, dt + \theta_S' \left(\mu(t)\, I_s dt + \Sigma' I_s dW\right) \\ &+ (l(t) - c_1(t) - c_2(t))\, dt, \end{aligned}$$

that can be written as:

$$\begin{aligned} dR_1(t) = &\left(R_1(t)\, r(t) - \theta_s'(t)\, I_s(\mu - r(t)\, 1) - c_1(t) - c_2(t) + l(t)\right) dt \\ &+ \theta_s' I \Sigma_s' dW \end{aligned}$$

B

Maximization Problem's Solution

We define $R(t) = R_1(t) + R_2(t)$. The value function $J(t, R(t), z(t))$, i.e the function which gives the value of the objective function after maximization, can be split into two sub-problems describing respectively two sub-periods $[t, t+dt]$ and $[t+dt, T]$:

$$\begin{cases} \max_{\{c_1(t),\theta_s(t)\}_{t\in[t_0,\tau]}} \mathbb{E}_{t_0}\left[\int_{t_0}^{\tau} U_c(c_1(s))\, e^{-\int_{t_0}^{s}\rho(u)+\lambda(u)du} ds \right. \\ \left. +I_{T\geq\tau} U_R(R_1(\tau)+R_2(\tau))\, e^{-\int_{t_0}^{\tau}\rho(u)+\lambda(u)du}\right] \\ dz(t) = \mu_z(t)\,dt + \Omega(t)'\,dW_t \\ dR_1(t) = \big(R_1(t)(r(t)+\lambda(t)) + \theta_s' I_s(\mu(t) - r(t)1) - c_1(t) - c_2(t) + l(t)\big)\,dt \\ +\theta_s(t)'\, I_s \Sigma(t)'\, dW(t) \\ dR_2(t) = (R_2(t)(d(t)+\lambda(t)) + c_2(t))\,dt \end{cases}$$

$$\begin{aligned} J(t, R(t), z(t)) = \max_{\{c_1(s),\theta_s(s)\}_{s\in[t,\tau]}} \mathbb{E}_t & \left[\int_t^{t+dt} U_c(c_1(s))\, e^{-\int_{t_0}^{s}\rho(u)+\lambda(u)du} ds \right. \\ & + \int_{t+dt}^{\tau} U_c(c_1(s))\, e^{-\int_t^{t+dt}\rho(u)+\lambda(u)du} e^{-\int_{t+dt}^{t+s}\rho(u)+\lambda(u)du} ds \\ & \left. + U_R(R(T))\, e^{-\int_t^{t+dt}\rho(u)+\lambda(u)du} e^{-\int_{t+dt}^{\tau}\rho(u)+\lambda(u)du}\right] \end{aligned} \tag{B.1}$$

In the Bellman's framework we consider a backward induction optimization. The period $[t+dt, T]$ is already optimized. We need to solve respect to consumption and portfolio the period $[t, t+dt]$ assuming the value of the subproblem in $t+dt$ is equal to $J(t+dt, R_{t+dt}, z_{t+dt})$, and starting from B.1 we can write:

$$\begin{aligned} J(t, R(t), z(t)) = \max \mathbb{E}_t & \left[\int_t^{t+dt} U_c(c_1(s))\, e^{-\int_{t_0}^{s}\rho(u)+\lambda(u)du} ds \right. \\ & \left. + e^{-\int_t^{t+dt}\rho(u)+\lambda(u)du} J(t+dt, R(t+dt), z(t+dt))\right]. \end{aligned}$$

Subtracting $J(t, R(t), z(t))$ from both sides, dividing by dt and taking the limit for $dt \to 0$, we obtain:

$$0 = \max_{[c(t),\theta_s(t)]} \left[U_c(c_1(t)) - (\rho(t) + \lambda(t)) J(t, R(t), z(t)) + \frac{1}{dt}\mathbb{E}_t [dJ(t, R(t), z(t))] \right].$$

We obtain dJ using the Ito's Lemma:

$$\begin{aligned} dJ = \Bigg(\frac{\partial J}{\partial t} + \frac{\partial J}{\partial R_1(t)} \Big(& R_1(t)(r(t) + \lambda(t)) + \theta_s(t)' I_s (\mu(t) - r(t) 1) \\ & -c_1(t) - c_2(t) + w(t) \Big) \Bigg) dt \\ & + \left(\frac{\partial J}{\partial R_2(t)} (R_2(t)(d(t) + \lambda(t)) + c_2(t)) + \mu_z(t)' \frac{\partial J}{\partial z(t)} \right) dt \\ + \Bigg(\frac{1}{2}\frac{\partial^2 J}{\partial R_1(t)^2}\theta_s(t)' I_s \Sigma' \Sigma I_s \theta_s(t) & + \frac{1}{2}\mathrm{tr}\left(\Omega'\Omega \frac{\partial^2 J}{\partial z(t)' \partial z(t)} \right) + \theta_s(t)' I_s \Sigma' \Omega \frac{\partial^2 J}{\partial z(t) \partial R_1(t)} \Bigg) dt \\ & + \frac{\partial J}{\partial R_1(t)} \theta_s(s)' I_s \Sigma' dW(t) + \left(\frac{\partial J}{\partial z(t)} \right)' \Omega' dW(t). \end{aligned} \tag{B.2}$$

The expected value of B.2 coincides with the drift term, so we can write the maximization problem in this form:

$$0 = \max_{\{c_1(t),\theta_s(t)\}} \left\{ \begin{array}{c} U_{c_1}(c_1(t)) - (\rho(t) + \lambda(t)) J + \frac{\partial J}{\partial t} \\ + \frac{\partial J}{\partial R_1(t)} \Big(R_1(t)(r(t) + \lambda(t)) + \theta_s(t)' I_s (\mu(t) - r(t) 1) \\ -c_1(t) - c_2(t) + l(t) \Big) \\ + \frac{\partial J}{\partial R_2(t)} (R_2(t)(d(t) + \lambda(t)) + c_2(t)) + \mu_z(t)' \frac{\partial J}{\partial z(t)} \\ + \frac{1}{2}\frac{\partial^2 J}{\partial R_1(t)^2}\theta_s(t)' I_s \Sigma' \Sigma I_s \theta_s(t) \\ + \frac{1}{2}\mathrm{tr}\left(\Omega'\Omega \frac{\partial^2 J}{\partial z(t)' \partial z(t)} \right) + \theta_s(t)' I_s \Sigma' \Omega \frac{\partial^2 J}{\partial z(t) \partial R_1(t)} \end{array} \right\} \tag{B.3}$$

At time horizon τ we impose the value function equal with the final value of the utility function:

$$J(\tau, R(\tau), z(\tau)) = U_R(R(\tau)). \tag{B.4}$$

The first order conditions are:

$$\frac{\partial U_{c_1}(c_1(t)^*)}{\partial c_1(t)} - \frac{\partial J}{\partial R_1(t)} = 0, \tag{B.5}$$

$$\frac{\partial J}{\partial R_1(t)} I_s (\mu(t) - r(t) 1) + \frac{\partial^2 J}{\partial R_1(t)^2} I_s \Sigma' \Sigma I_s \theta_s(t) + I_s \Sigma' \Omega \frac{\partial^2 J}{\partial z_t \partial R_1(t)} = 0. \tag{B.6}$$

Taking into account the3.1 the optimal consumption is:

$$c_1(t)^* = \alpha_{c_1} + \left(\frac{1}{\chi_{c_1}}\frac{\partial J}{\partial R_1(t)}\right)^{-\frac{1}{\delta}}, \tag{B.7}$$

The optimal portfolio is:

$$I_s\theta_s(t)^* = -\frac{\frac{\partial J}{\partial R_1(t)}}{\frac{\partial^2 J}{\partial R_1(t)^2}}\left(\Sigma'\Sigma\right)^{-1}(\mu_t - r_t 1) - \frac{1}{\frac{\partial^2 J}{\partial R_1(t)^2}}\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\frac{\partial^2 J}{\partial z(t)\,\partial R_1(t)}. \tag{B.8}$$

Substituting B.8 and B.7 into B.3 we obtain:

$$0 = \left\{\begin{array}{c}
\frac{\delta}{1-\delta}\chi_{c_1}^{\frac{1}{\delta}}\left(\frac{\partial J}{\partial R_1(t)}\right)^{1-\frac{1}{\delta}} - \rho(t)J - \lambda(t)J + \frac{\partial J}{\partial t} + \frac{\partial J}{\partial R_1(t)}R_1(t)(r(t)+\lambda(t)) \\
-\frac{\left(\frac{\partial J}{\partial R_1(t)}\right)^2}{\frac{\partial^2 J}{\partial R_1(t)^2}}(\mu(t) - r(t)1)'\left(\Sigma'\Sigma\right)^{-1}(\mu_t - r_t 1) \\
-\frac{\frac{\partial J}{\partial R_1(t)}}{\frac{\partial^2 J}{\partial R_1(t)^2}}(\mu(t) - r(t)1)'\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\frac{\partial^2 J}{\partial z(t)\partial R_1(t)} \\
-\frac{\partial J}{\partial R_1(t)}\alpha_{c_1} - \frac{\partial J}{\partial R_1(t)}c_2(t) + \frac{\partial J}{\partial R_1(t)}l(t) + \frac{\partial J}{\partial R_2(t)}R_2(t)(d(t)+\lambda(t)) + \frac{\partial J}{\partial R_2(t)}c_2(t) \\
-\frac{1}{2}\frac{1}{\frac{\partial^2 J}{\partial R_1(t)^2}}\left(\frac{\partial^2 J}{\partial z(t)\partial R_1(t)}\right)'\Omega'\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\frac{\partial^2 J}{\partial z(t)\partial R_1(t)} \\
+\mu_z(t)'\frac{\partial J}{\partial z(t)} + \frac{1}{2}\mathrm{tr}\left(\Omega'\Omega\frac{\partial^2 J}{\partial z(t)'\partial z(t)}\right)
\end{array}\right\} \tag{B.9}$$

To solve B.9 we consider a guess function with the same functional form of the utility function **??**:

$$J = F(t, z_t)^{\delta}\frac{(\beta(t, z(t))R_1 + \nu(t, z(t))R_2 - H(t, z(t)))^{1-\delta}}{1-\delta}.$$

The boundary condition in T is:

$$J(T, R_T) = \chi_R\frac{(R_T - \alpha_R)^{1-\delta}}{1-\delta}$$

implies that F and H must satisfy:

$$F(T, z_T)^{\delta} = \chi_R \Longrightarrow F(T, z_T) = \chi_R^{\frac{1}{\delta}}$$

$$H(T, z_T) = \alpha_R.$$

In order to simplify the notation we define:

$$F(t, z_t) \equiv F,\ \frac{\partial F(t,z_t)}{\partial t} \equiv F_t,\quad \frac{\partial F(t,z_t)}{\partial z_t} \equiv F_z,\quad \frac{\partial^2 F(t,z_t)}{\partial z_t'\partial z_t} \equiv F_{zz},$$

$$H\left(t,z_t\right)\equiv H,\ \tfrac{\partial H(t,z_t)}{\partial t}\equiv H_t,\ \ \tfrac{\partial H(t,z_t)}{\partial z_t}\equiv H_z,\ \ \tfrac{\partial^2 H(t,z_t)}{\partial z_t'\partial z_t}\equiv H_{zz},$$

$$\alpha\left(t,z_t\right)\equiv\alpha,\frac{\partial\alpha\left(t,z_t\right)}{\partial t}\equiv\alpha_t,\frac{\partial\alpha\left(t,z_t\right)}{\partial z_t}\equiv\alpha_z,\frac{\partial^2\alpha\left(t,z_t\right)}{\partial z_t'\partial z_t}\equiv\alpha_{zz},$$

$$\beta\left(t,z_t\right)\equiv\beta,\frac{\partial\beta\left(t,z_t\right)}{\partial t}\equiv\beta_t,\frac{\partial\beta\left(t,z_t\right)}{\partial z_t}\equiv\beta_z,\frac{\partial^2\beta\left(t,z_t\right)}{\partial z_t'\partial z_t}\equiv\beta_{zz}.$$

The derivatives we are interested in are:

$$\frac{\partial J}{\partial t}=\delta F^{\delta-1}F_t\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}+F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\beta_t R_1+v_t R_2-H_t\right)$$

$$\frac{\partial J}{\partial R_1}=F^{\delta}\beta\left(\beta R_1+vR_2-H\right)^{-\delta}$$

$$\frac{\partial^2 J}{\partial R_1^2}=-\delta F^{\delta}\beta^2\left(\beta R_1+vR_2-H\right)^{-\delta-1}$$

$$\frac{\partial J}{\partial R_2}=F^{\delta}v\left(\beta R_1+vR_2-H\right)^{-\delta}$$

$$\frac{\partial J}{\partial z}=\delta F^{\delta-1}F_z\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}+F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\beta_z R_1+v_z R_2-H_z\right)$$

$$\begin{aligned}\frac{\partial^2 J}{\partial z\partial R_1}=&\ \delta F^{\delta-1}F_z\beta\left(\beta R_1+vR_2-H\right)^{-\delta}-\delta F^{\delta}\beta\left(\beta R_1+vR_2-H\right)^{-\delta-1}\left(\beta_z R_1+v_z R_2-H_z\right)\\&+F^{\delta}\beta_z\left(\beta R_1+vR_2-H\right)^{-\delta}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 J}{\partial z'\partial z}=&\ \delta\left(\delta-1\right)F^{\delta-2}F_z^2\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}+\delta F^{\delta-1}F_{zz}\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}\\&+2\delta F^{\delta-1}F_z\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\beta_z R_1+v_z R_2-H_z\right)\\&-\delta F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta-1}\left(\beta_z R_1+v_z R_2-H_z\right)^2\\&+F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\beta_{zz}R_1+v_{zz}R_2-H_{zz}\right)\end{aligned}$$

Substituting the extended derivatives in (B.9) we obtain:

$$\begin{aligned}0=&\ \frac{\delta}{1-\delta}\chi_{c_1}^{\frac{1}{\delta}}F^{\delta-1}\beta^{1-\frac{1}{\delta}}\left(\beta R_1+vR_2-H\right)^{1-\delta}\\&-\left(\rho\left(t\right)+\lambda\left(t\right)\right)F^{\delta}\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}\\&+\delta F^{\delta-1}F_t\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}\end{aligned}$$

$$+\frac{1}{\delta}F^{\delta}\left(\beta R_1+vR_2-H\right)^{1-\delta}\left(\mu\left(t\right)-r\left(t\right)1\right)'\left(\Sigma'\Sigma\right)^{-1}\left(\mu_t-r_t1\right)$$
$$+F^{\delta-1}\left(\beta R_1+vR_2-H\right)^{1-\delta}\left(\mu\left(t\right)-r\left(t\right)1\right)'\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega F_z$$
$$+\frac{1}{\delta\beta}F^{\delta}\left(\beta R_1+vR_2-H\right)^{1-\delta}\left(\mu\left(t\right)-r\left(t\right)1\right)'\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\beta_z$$
$$+\frac{1}{2}\delta F^{\delta-2}F_z\left(\beta R_1+vR_2-H\right)^{1-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega F_z$$
$$+\frac{1}{2}\frac{1}{\beta}F^{\delta-1}F_z\left(\beta R_1+vR_2-H\right)^{1-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega\beta_z$$
$$+\frac{1}{2}\frac{1}{\beta}F^{\delta-1}\beta_z\left(\beta R_1+vR_2-H\right)^{1-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega F_z$$
$$+\frac{1}{2}\frac{1}{\delta\beta^2}F^{\delta}\left(\beta R_1+vR_2-H\right)^{1-\delta}\beta_z\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega\beta_z$$
$$+\mu_z\left(t\right)'\delta F^{\delta-1}F_z\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}$$
$$+\frac{1}{2}\delta F^{\delta-1}\frac{\left(\beta R_1+vR_2-H\right)^{1-\delta}}{1-\delta}\mathrm{tr}\left(\Omega'\Omega F_{zz}\right)$$
$$+F^{\delta}\left(\beta R_1+vR_2-H\right)^{1-\delta}\left(r\left(t\right)+\lambda\left(t\right)\right)$$
$$+F^{\delta}\left(\beta R_1+vR_2-H\right)^{1-\delta}\left(d\left(t\right)+\lambda\left(t\right)\right)$$
$$-\frac{1}{2}F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\mathrm{tr}\left(\Omega'\Omega\left(\beta_{zz}R_1+v_{zz}R_2-H_{zz}\right)\right)$$
$$+F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\beta_tR_1+v_tR_2-H_t\right)$$
$$-F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\left(\mu\left(t\right)-r\left(t\right)1\right)'\left(\Sigma'\Sigma\right)^{-1}\Sigma'\Omega\left(\beta_zR_1+v_zR_2-H_z\right)$$
$$+F^{\delta}\beta\left(\beta R_1+vR_2-H\right)^{-\delta}\left(l\left(t\right)-\alpha_{c_1}-c_2\left(t\right)\right)$$
$$+F^{\delta}v\left(\beta R_1+vR_2-H\right)^{-\delta}c_2\left(t\right)$$
$$-\frac{1}{2}\delta F^{\delta-1}\left(\beta R_1+vR_2-H\right)^{-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega F_z\left(\beta_zR_1+v_zR_2-H_z\right)$$
$$-\frac{1}{2}\delta F^{\delta-1}\left(\beta R_1+vR_2-H\right)^{-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega F_z\left(\beta_zR_1+v_zR_2-H_z\right)$$
$$-\frac{1}{2}\frac{1}{\beta}F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega\beta_z\left(\beta_zR_1+v_zR_2-H_z\right)$$
$$-\frac{1}{2}\frac{1}{\beta}F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}\Omega'\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma'-I\right)\Omega\beta_z\left(\beta_zR_1+v_zR_2-H_z\right)$$

$$+\mu_z(t)' F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (\beta_z R_1 + v_z R_2 - H_z)$$
$$-F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (vR_2 - H)(r(t) + \lambda(t))$$
$$-F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (\beta R_1 - H)(d(t) + \lambda(t))$$
$$+\frac{1}{2}\delta F^\delta (\beta R_1 + vR_2 - H)^{-\delta-1} \Omega' \left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma' - I\right)(\beta_z R_1 + v_z R_2 - H_z)^2 \quad \text{(B.10)}$$

where the term $F^\delta \beta (\beta R_1 + vR_2 - H)^{-\delta} R_1(t)(r(t) + \lambda(t))$ is written as:
$F^\delta \beta (\beta R_1 + vR_2 - H)^{-\delta} R_1(t)(r(t) + \lambda(t))$
$= F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (\beta R_1 + vR_2 - H - vR_2 + H)(r(t) + \lambda(t))$
$= F^\delta (\beta R_1 + vR_2 - H)^{1-\delta} (r(t) + \lambda(t)) - F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (vR_2 - H)(r(t) + \lambda(t))$
and, in the same way, the term $F^\delta v (\beta R_1 + vR_2 - H)^{-\delta} R_2(t)(d(t) + \lambda(t))$ is written as:
$F^\delta v (\beta R_1 + vR_2 - H)^{-\delta} R_2(t)(d(t) + \lambda(t))$
$= F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (\beta R_1 + vR_2 - H - \beta R_1 + H)(t)(d(t) + \lambda(t))$
$= F^\delta (\beta R_1 + vR_2 - H)^{1-\delta} (d(t) + \lambda(t)) - F^\delta (\beta R_1 + vR_2 - H)^{-\delta} (\beta R_1 - H)(d(t) + \lambda(t))$
The complete market assumption implies:

$$\left(\Sigma\left(\Sigma'\Sigma\right)^{-1}\Sigma' - I\right) = 0.$$

In B.10 we obtain three sets of terms containing $(\beta R_1 + vR_2 - H)^{1-\delta}$,$(\beta R_1 + vR_2 - H)^{-\delta}$and $(\beta R_1 + vR_2 - H)^{-\delta-1}$but we have two undetermined functions (i.e F and H) that can be used for solving the equations. In the case of complete market the vector of the market price of risk is

$$\xi(t) = \left(\Sigma'\right)^{-1} (\mu(t) - r(t)1).$$

Accordingly we can rewrite, some matrix products:

$$(\mu(t) - r(t)1)' \left(\Sigma'\Sigma\right)^{-1} (\mu_t - r_t 1) = \xi'(t)\xi(t)$$

$$(\mu(t) - r(t)1)' \left(\Sigma'\Sigma\right)^{-1} \Sigma' = \xi'(t).$$

The differential Equation corresponding to the terms containing $(\beta R_1 + vR_2 - H)^{1-\delta}$is

$$0 = \frac{\delta}{1-\delta}\chi_{c_1}^{\frac{1}{\delta}} F^{\delta-1}\beta^{1-\frac{1}{\delta}} - \frac{(\rho_t + \lambda_t) F^\delta}{1-\delta} + \delta F^{\delta-1} F_t \frac{1}{1-\delta}$$
$$+\frac{1}{\delta} F^\delta \xi'(t)\xi(t) + F^{\delta-1}\xi'(t)\Omega F_z + \frac{1}{\delta\beta} F^\delta \xi'(t)\Omega\beta_z$$

$$+\mu_z(t)' \delta F^{\delta-1} F_z \frac{1}{1-\delta} + \frac{1}{2}\delta F^{\delta-1}\frac{1}{1-\delta}\mathrm{tr}\left(\Omega'\Omega F_{zz}\right) + F^\delta (r(t) + \lambda(t)) + F^\delta (d(t) + \lambda(t))$$

that can be further simplified by multiplying it by $F^{1-\delta\frac{1-\delta}{\delta}}$:

$$0=\chi_{c_1}^{\frac{1}{\delta}}\beta^{1-\frac{1}{\delta}}-\frac{(\rho_t+\lambda_t)\,F}{\delta}+F_t+\frac{1-\delta}{\delta^2}F\xi^{'}(t)\,\xi(t)+\frac{1-\delta}{\delta}\xi^{'}(t)\,\Omega F_z+\frac{1-\delta}{\delta^2\beta}F\xi^{'}(t)\,\Omega\beta_z$$
$$+\mu_z(t)^{'}F_z+\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega F_{zz}\right)+\frac{1-\delta}{\delta}F\left(r(t)+\lambda(t)\right)+\frac{1-\delta}{\delta}F\left(d(t)+\lambda(t)\right). \quad \text{(B.11)}$$

The Equation B.11 can be written as:

$$0=F_t+\left(\mu_z(t)^{'}-\frac{\delta-1}{\delta}\xi^{'}(t)\,\Omega\right)F_z+\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega F_{zz}\right)$$
$$-\frac{F}{\delta}\left(\rho(t)+\lambda(t)+\frac{\delta-1}{\delta}\xi^{'}(t)\,\xi(t)+\frac{\delta-1}{\delta\beta}\xi^{'}(t)\,\Omega\beta_z+\left((\delta-1)\left(r(t)+2\lambda(t)+d(t)\right)\right)\right)$$
$$+\chi_{c_1}^{\frac{1}{\delta}}\beta^{1-\frac{1}{\delta}}. \quad \text{(B.12)}$$

The differential equation corresponding to the terms containing $F^{\delta}\left(\beta R_1+vR_2-H\right)^{-\delta}$is

$$0=\left(\beta_t R_1+v_t R_2-H_t\right)-\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega\left(\beta_{zz}R_1+v_{zz}R_2-H_{zz}\right)\right)$$
$$+\left(\mu_z(t)^{'}-\xi^{'}(t)\,\Omega\right)\left(\beta_z R_1+v_z R_2-H_z\right)+(v-\beta)\,c_2(t)+\beta\left(l(t)-\alpha_{c_1}\right)$$
$$-\left(vR_2-H\right)\left(r(t)+\lambda(t)\right)-\left(\beta R_1-H\right)\left(d(t)+\lambda(t)\right). \quad \text{(B.13)}$$

from B.13 we can write:

$$\begin{cases}\left(\mu_z(t)^{'}-\xi^{'}(t)\,\Omega\right)\beta_z R_1-\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega\beta_{zz}R_1\right)-\beta R_1\left(d(t)+\lambda(t)\right)+\beta_t R_1=0\\ \left(\mu_z(t)^{'}-\xi^{'}(t)\,\Omega\right)v_z R_2-\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega\left(v_{zz}R_2\right)\right)-vR_2\left(r(t)+\lambda(t)\right)+v_t R_2=0\\ -\left(\mu_z(t)^{'}-\xi^{'}(t)\,\Omega\right)H_z+(v-\beta)\,c_2(t)+\beta\left(l(t)-\alpha_{c_1}\right)-H_t+\frac{1}{2}\mathrm{tr}\left(\Omega^{'}\Omega\left(H_{zz}\right)\right)\\ \qquad +H\left(r(t)+2\lambda(t)+d(t)\right)=0.\end{cases} \quad \text{(B.14)}$$

The system B.14 can be solved setting the following boundary conditions:

$$\begin{cases}\beta(T,z_T)=1\\ \nu(T,z_T)=1\\ H(T,z_T)=\alpha_R.\end{cases} \quad \text{(B.15)}$$

The results of B.14 and B.15 are:

$$\beta(t,z_t)=\mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{\omega}(d(s)+\lambda(s))ds}\right]$$
$$\nu(t,z_t)=\mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{\omega}(r(s)+\lambda(s))ds}\right]$$
$$H(t,z_t)=\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{\omega}\left((\beta(u)-v(u))\,c_2(u)-\beta(u)\,l(u)+\beta(u)\,\alpha_{c_1}\right)e^{-\int_t^{u}(r(s)+2\lambda(s)+d(s))ds}du\right]$$
$$+\mathbb{E}_t^{\mathbb{Q}}\left[\alpha_R e^{-\int_t^{\omega}(r(s)+2\lambda(s)+d(s))ds}\right],$$

We can obtain the Feynman-Kac solution of Equation B.12 considering the modified stochastic differential equation for the state variable z_t as:

$$\begin{aligned} dz_t &= \left(\mu_z - \frac{\delta-1}{\delta}\xi'(t)\,\Omega\right)dt + \Omega' dW_t \\ &= \mu_z dt + \Omega'\left(dW_t - \frac{\delta-1}{\delta}\xi(t)\,dt\right) \\ &= \mu_z dt + \Omega' dW^{\mathbb{Q}_\delta}(t) \end{aligned}$$

where we consider the probability measure from the following version of the Girsanov's theorem:

$$dW^{\mathbb{Q}_\delta}(t) = dW(t) - \frac{\delta-1}{\delta}\xi(t)\,dt$$

$$\begin{aligned} F(t, z_t) = {} & \mathbb{E}_t^{\mathbb{Q}_\delta}\left[\int_t^{\omega} \chi_{c_1}^{\frac{1}{\delta}} \beta^{1-\frac{1}{\delta}} e^{-\frac{1}{\delta}\int_t^s \left(\rho(u)+\lambda(u)+\frac{\delta-1}{\delta}\xi'(u)\xi(u)+\frac{\delta-1}{\delta\beta}\xi'(u)\Omega\beta_z+(\delta-1)(r(u)+2\lambda(u)+d(u))\right)du} ds\right] \\ & + \mathbb{E}_t^{\mathbb{Q}_\delta}\left[\chi_R^{\frac{1}{\delta}} e^{-\frac{1}{\delta}\int_t^s \left(\rho(u)+\lambda(u)+\frac{\delta-1}{\delta}\xi'(u)\xi(u)+\frac{\delta-1}{\delta\beta}\xi'(u)\Omega\beta_z+(\delta-1)(r(u)+2\lambda(u)+d(u))\right)du}\right] \end{aligned}$$

C

Appendix C

The Lagrangian of the problem 4.8 is:

$$\mathcal{L} = \mathbb{E}_{t_0}\left[\begin{array}{c} \int_{t_0}^{\omega}\left(\frac{\left(c_1(s)-\alpha_{c_1}\right)^{1-\delta}}{1-\delta}+\frac{\left(\kappa R_1(s)-\alpha_{R_1}\right)^{1-\delta}}{1-\delta}\lambda\left(s\right)\right)e^{-\int_{t_0}^{s}(\rho(u)+\lambda(u))du}ds \\ +\frac{\left(R_2(\tau)-\alpha_{R_2}\right)^{1-\delta}}{1-\delta}e^{-\int_{t_0}^{\tau}(\rho(u)+\lambda(u))du} \end{array}\right]$$
$$+\phi_1 R_1\left(t_0\right)-\phi_1\mathbb{E}_{t_0}\left[\int_{t_0}^{\omega} m\left(t_0,s\right)\left(\kappa R_1\left(s\right)\lambda\left(s\right)+c_1\left(s\right)+\left(c_2-l\left(s\right)\right)I_{s<\tau}\right)e^{-\int_{t_0}^{s}(r(u)+\lambda(u))du}ds\right]$$
$$+\phi_2 R_2\left(t_0\right)-\phi_2\mathbb{E}_{t_0}\left[R_2\left(\tau\right)m\left(t_0,\tau\right)e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du}-\int_{t_0}^{\omega} I_{s<\tau}m\left(t_0,s\right)c_2 e^{-\int_{t_0}^{s}(d(u)+\lambda(u))du}ds\right]$$

and the first order conditions (FOCs) are:

FOC on $c_1\left(s\right)$

$$c_1^*\left(s\right)=\alpha_{c_1}+\left(\phi_1 m\left(t_0,s\right)\frac{e^{-\int_{t_0}^{s}(r(u)+\lambda(u))du}}{e^{-\int_{t_0}^{s}(\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}}$$

FOC on $R_1\left(s\right)$

$$R_1^*\left(s\right)=\frac{\alpha_{R_1}}{\kappa}+\frac{1}{\kappa}\left(\phi_1 m\left(t_0,s\right)\frac{e^{-\int_{t_0}^{s}(r(u)+\lambda(u))du}}{e^{-\int_{t_0}^{s}(\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}}$$

FOc on $R_2\left(\tau\right)$

$$R_2^*\left(\tau\right)=\alpha_{R_2}+\left(\phi_2 m\left(t_0,\tau\right)\frac{e^{-\int_{t_0}^{\tau}(d(u)+\lambda(u))du}}{e^{-\int_{t_0}^{\tau}(\rho(u)+\lambda(u))du}}\right).^{-\frac{1}{\delta}}$$

The Equation 4.3 can be written as:

$$R_1(t) = \mathbb{E}_t\left[\int_t^\omega m(t,s)\left(\kappa R_1^*(s)\lambda(s) + c_1^*(s) + (c_2 - l(s)) I_{s<\tau}\right) e^{-\int_t^s (r(u)+\lambda(u))du} ds\right]$$

after substituting the optimal values we obtain:

$$\begin{aligned} R_1(t) &= \mathbb{E}_t\left[\int_t^\omega m(t,s)\left(\lambda(s)\alpha_{R_1} + \alpha_{c_1} + (c_2 - l(s)) I_{s<\tau}\right) e^{-\int_t^s (r(u)+\lambda(u))du} ds\right] \\ &+ \left(\phi_1 m(t_0,t)\frac{e^{-\int_{t_0}^t (r(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} \\ &\times \mathbb{E}_t\left[\int_t^\omega m(t,s)^{1-\frac{1}{\delta}}(\lambda(s)+1)\left(\frac{e^{-\int_t^s (r(u)+\lambda(u))du}}{e^{-\int_t^s (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} e^{-\int_t^s (r(u)+\lambda(u))du} ds\right] \end{aligned}$$

where we can define

$$\begin{aligned} H_1(t) &\equiv \mathbb{E}_t\left[\int_t^\omega m(t,s)\left(\lambda(s)\alpha_{R_1} + \alpha_{c_1} + (c_2 - l(s)) I_{s<\tau}\right) e^{-\int_t^s (r(u)+\lambda(u))du} ds\right] \\ &= \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\omega \left(\lambda(s)\alpha_{R_1} + \alpha_{c_1} + (c_2 - l(s)) I_{s<\tau}\right) e^{-\int_t^s r(u)+\lambda(u))du} ds\right] \end{aligned}$$

$$F_1(t) \equiv \mathbb{E}_t\left[\int_t^\omega m(t,s)^{1-\frac{1}{\delta}}(\lambda(s)+1)\left(\frac{e^{-\int_t^s (r(u)+\lambda(u))du}}{e^{-\int_t^s (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} e^{-\int_t^s (r(u)+\lambda(u))du} ds\right]$$

and so

$$R_1(t) = H_1(t) + \left(\phi_1 m(t_0,t)\frac{e^{-\int_{t_0}^t (r(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} F_1(t). \tag{C.1}$$

The Equation 4.7 can be written as:

$$R_2(t) = \mathbb{E}_t\left[R_2^*(\tau) m(t,\tau) e^{-\int_t^\tau (d(u)+\lambda(u))du} - \int_t^\omega I_{s<\tau} m(t,s) c_2 e^{-\int_t^s (d(u)+\lambda(u))du} ds\right]$$

after substituting the optimal value we obtain:

$$\begin{aligned} R_2(t) = &\left(\phi_2 m(t_0,t)\frac{e^{-\int_{t_0}^t (d(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} \mathbb{E}_t\left[m(t,\tau)^{1-\frac{1}{\delta}}\left(\frac{e^{-\int_t^\tau (d(u)+\lambda(u))du}}{e^{-\int_t^\tau (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} e^{-\int_t^\tau (d(u)+\lambda(u))du}\right] \\ &+ \mathbb{E}_t\left[\alpha_{R_2} m(t,\tau) e^{-\int_t^\tau (d(u)+\lambda(u))du} - \int_t^\omega I_{s<\tau} m(t,s) c_2 e^{-\int_t^s (d(u)+\lambda(u))du} ds\right] \end{aligned}$$

where we can define

$$H_2(t) \equiv \mathbb{E}_t\left[\alpha_{R_2} m(t,\tau)\, e^{-\int_t^\tau (d(u)+\lambda(u))du} - \int_t^\omega I_{s<\tau} m(t,s)\, c_2 e^{-\int_t^s (d(u)+\lambda(u))du} ds\right]$$
$$= \mathbb{E}_t^{\mathbb{Q}}\left[\alpha_{R_2} e^{-\int_t^\tau (d(u)+\lambda(u))du} - \int_t^\omega I_{s<\tau} c_2 e^{-\int_t^s (d(u)+\lambda(u))du} ds\right]$$
$$F_2(t) \equiv \mathbb{E}_t\left[m(t,\tau)^{1-\frac{1}{\delta}} \left(\frac{e^{-\int_t^\tau (d(u)+\lambda(u))du}}{e^{-\int_t^\tau (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} e^{-\int_t^\tau (d(u)+\lambda(u))du}\right]$$

and so

$$R_2(t) = H_2(t) + \left(\phi_2 m(t_0,t)\, \frac{e^{-\int_{t_0}^t (d(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} F_2(t).$$

To obtain the optimal portfolio we calculate the differential of $R_1^*(t)$ from (C.1) is

$$dR_1^*(t) = (...)\, dt + \frac{1}{\delta}\left(\phi_1 m(t_0,t)\, \frac{e^{-\int_{t_0}^t (r(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} F_1(t)\, \xi(t)'\, dW(t)$$
$$+ \left(\frac{\partial H_1(t)}{\partial z(t)} + \left(\phi_1 m(t_0,t)\, \frac{e^{-\int_{t_0}^t (r(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}} \frac{\partial F_1(t)}{\partial z(t)}\right)' \Omega' dW(t)$$

and substituting

$$\frac{R_1(t) - H_1(t)}{F_1(t)} = \left(\phi_1 m(t_0,t)\, \frac{e^{-\int_{t_0}^t (r(u)+\lambda(u))du}}{e^{-\int_{t_0}^t (\rho(u)+\lambda(u))du}}\right)^{-\frac{1}{\delta}}$$

we have

$$dR_1^*(t) = (...)\, dt + \frac{R_1(t) - H_1(t)}{\delta} \xi(t)'\, dW(t)$$
$$+ \left(\frac{\partial H_1(t)}{\partial z(t)} + \frac{R_1(t) - H_1(t)}{F_1(t)} \frac{\partial F_1(t)}{\partial z(t)}\right)' \Omega' dW(t).$$

The dynamics of wealth that must be replicated is

$$dR_1(t) = (...)\, dt + \theta_S' I_S \Sigma' dW(t)$$

and so the optimal portfolio is

$$\theta_S' I_S \Sigma' = \frac{R_1(t) - H_1(t)}{\delta} \xi(t)' + \left(\frac{\partial H_1(t)}{\partial z(t)} + \frac{R_1(t) - H_1(t)}{F_1(t)} \frac{\partial F_1(t)}{\partial z(t)}\right)' \Omega'$$
$$I_S \theta_S = \frac{R_1(t) - H_1(t)}{\delta} \Sigma^{-1} \xi(t) + \Sigma^{-1} \Omega \frac{\partial H_1(t)}{\partial z(t)} + \frac{R_1(t) - H_1(t)}{F_1(t)} \Sigma^{-1} \Omega \frac{\partial F_1(t)}{\partial z(t)}.$$

Bibliography

Henry Aaron. The social insurance paradox. *The Canadian Journal of Economics and Political Science/Revue canadienne d'Economique et de Science politique*, 32(3): 371–374, 1966.

J. Alonso-GarcÃa and P. Devolder. Optimal mix between pay-as-you-go and funding for dc pension schemes in an overlapping generations model. *Insurance: Mathematics and Economics*, 70(C):224–236, 2016. URL `https://EconPapers.repec.org/RePEc:eee:insuma:v:70:y:2016:i:c:p:224-236`.

Christoph Anders and Max Groneck. The optimal portfolio of payg benefits and funded pensions in germany. *FinanzArchiv: Public Finance Analysis*, 73(3):255–291, 2017.

Helena Aro and Teemu Pennanen. A user-friendly approach to stochastic mortality modelling. *European Actuarial Journal*, 1(2):151–167, 2011.

Paolo Battocchio and Francesco Menoncin. Optimal pension management in a stochastic framework. *Insurance: Mathematics and Economics*, 34(1):79–95, 2004.

Paolo Battocchio, Francesco Menoncin, et al. *Optimal pension management under stochastic interest rates, wages and inflation*. Universite catholique de Louvain, 2002.

Ennio Bilancini and Massimo D'Antoni. The desirability of pay-as-you-go pensions when relative consumption matters and returns are stochastic. *Economics Letters*, 117(2):418–422, 2012.

Heather Booth, John Maindonald, and Len Smith. Applying lee-carter under conditions of variable mortality decline. *Population studies*, 56(3):325–336, 2002.

Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

Natacha Brouhns, Michel Denuit, and Jeroen K Vermunt. A poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and economics*, 31(3):373–393, 2002.

Andrew JG Cairns, David Blake, and Kevin Dowd. Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin: The Journal of the IAA*, 36(1):79–120, 2006a.

Andrew JG Cairns, David Blake, and Kevin Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718, 2006b.

Andrew JG Cairns, David Blake, Kevin Dowd, Guy D Coughlan, David Epstein, Alen Ong, and Igor Balevich. A quantitative comparison of stochastic mortality models using data from england and wales and the united states. *North American Actuarial Journal*, 13(1):1–35, 2009.

Andrew JG Cairns, David Blake, Kevin Dowd, Guy D Coughlan, David Epstein, and Marwa Khalaf-Allah. Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3):355–367, 2011.

K. Cairns, A.J.G. Blake D. Dowd. Modelling and management of mortality risk: A review,. *Working paper, Heriot-Watt University and Pnsions Institute Discussion Paper PI-0814*, 2008.

E. Castellino, O. Fornero. From payg to funding in italy: a feasible transaction? *Geneva Papers on Risk & Insurance - Issues & Practice. Oct99, Vol. 24 Issue 4, p473. 15p.*, 1999.

An Chen, Felix Hentschel, and Xian Xu. Optimal retirement time under habit persistence: what makes individuals retire early? *Scandinavian Actuarial Journal*, pages 1–25, 2017.

Iain D Currie. On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal*, 2016(4):356–383, 2016.

ID Currie. Smoothing and forecasting mortality rates with p-splines. *Talk given at the Institute of Actuaries*, 2006.

Pierre Devolder and Roberta Melis. Optimal mix between pay as you go and funding for pension liabilities in a stochastic framework. *ASTIN Bulletin: The Journal of the IAA*, 45(3):551–575, 2015.

Elroy Dimson, Paul Marsh, and Mike Staunton. *Triumph of the optimists: 101 years of global investment returns*. Princeton University Press, 2009.

Darrell Duffie. *Dynamic asset pricing theory*. Princeton University Press, 2010.

Jayasri Dutta, Sandeep Kapur, and J Michael Orszag. A portfolio approach to the optimal funding of pensions. *Economics Letters*, 69(2):201–206, 2000.

Eduardo Fajnzylber and David A Robalino. Assessing fiscal costs and pension distribution in transitions to defined contribution systems: A retrospective analysis for chile. *Nonfinancial Defined Contribution Pension Schemes in a Changing Pension World*, 1:281–315, 2012.

Jean-Daniel Guigou, Bruno Lovat, and Jang Schiltz. Optimal mix of funded and unfunded pension systems: the case of luxembourg. *Pensions: An International Journal*, 17(4):208–222, 2012.

Steven Haberman and Arthur Renshaw. On age-period-cohort parametric mortality rate projections. *Insurance: Mathematics and Economics*, 45(2):255–270, 2009.

Steven Haberman and Arthur Renshaw. A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, 48(1):35–55, 2011.

A. Hunt and D Blake. Identifiability in age/period/cohort mortality models. *Work in Progress*, 2015a.

A Hunt and D Blake. On the structure and classification of mortality models. *Work in Progress*, 2015b.

Andrew Hunt and David Blake. A general procedure for constructing mortality models. *North American Actuarial Journal*, 18(1):116–138, 2014.

ISTAT. Bilancio demografico nazionale. Technical report, ISTAT, 2016.

Markus Knell. The optimal mix between funded and unfunded pension systems when people care about relative consumption. *Economica*, 77(308):710–733, 2010.

Ronald Lee. The lee-carter method for forecasting mortality, with various extensions and applications. *North American actuarial journal*, 4(1):80–91, 2000.

Ronald D Lee and Lawrence R Carter. Modeling and forecasting us mortality. *Journal of the American statistical association*, 87(419):659–671, 1992.

Angus Maddison. *The world economy volume 1: A millennial perspective volume 2: Historical statistics*. Academic Foundation, 2007.

Francesco Menoncin. Death bonds with stochastic force of mortality. In *Actuarial and Financial Mathematics Conference–Interplay between finance and insurance*, 2009.

Francesco Menoncin. *Analisi e gestione dei rischi di mercato, di credito e operativo*. Accademia Universa Press, 2011.

Francesco Menoncin and Luca Regis. Longevity-linked assets and pre-retirement consumption/portfolio decisions. *Insurance: Mathematics and Economics*, 76:75–86, 2017.

Franco Modigliani. Life cycle, individual thrift, and the wealth of nations. *The American Economic Review*, 76(3):297–313, 1986.

Michael Murphy. The âgolden generationsâ in historical context. *British Actuarial Journal*, 15(S1):151–184, 2009.

Michael Murphy. Reexamining the dominance of birth cohort effects on mortality. *Population and development review*, 36(2):365–390, 2010.

OECD. Organisation for economic co-operation and development: Mortality assumptions and longevity risk, implications for pension funds and annuity providers. Technical report, OECD, 2016.

Jim Oeppen and James W Vaupel. Broken limits to life expectancy. *Science*, 296(5570): 1029–1031, 2002.

M. Persson. Five fallacies in the social security debate. *Social Security Reform in Advanced Countries: Evaluating Pension Finance*, pages 39–51, 2002. URL `www.scopus.com`. Cited By :1.

Richard Plat. On stochastic mortality modeling. *Insurance: Mathematics and Economics*, 45(3):393–404, 2009.

Arthur Renshaw and Steven Haberman. Lee–carter mortality forecasting: A parallel generalized linear modelling approach for england and wales mortality projections. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(1):119–137, 2003a.

Arthur E Renshaw and Steven Haberman. On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics*, 32(3):379–401, 2003b.

Arthur E Renshaw and Steven Haberman. A cohort-based extension to the lee–carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38 (3):556–570, 2006.

Paul Samuelson. An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66, 1958. URL `https://EconPapers.repec.org/RePEc:ucp:jpolec:v:66:y:1958:p:467`.

Bernard MS Van Praag and Pedro Cardoso. The mix between pay-as-you-go and funded pensions and what demography has to do with it. *CESifo Working Paper, No 865, 2003.*, 2003.

Jeroen K Vermunt. Log-linear models for event histories, advanced quantitative techniques in the social sciences series, 1997a.

Jeroen K Vermunt. Lem: A general program for the analysis of categorical data. *Department of Methodology and Statistics, Tilburg University*, 1997b.

Andres M Villegas, Vladimir K Kaishev, and Pietro Millossovich. Stmomo: an r package for stochastic mortality modelling. *RCRAN*, 2015.

John R Wilmoth. Fitting three-way models to two-way arrays of demographic rates. *University of Michigan, Population Studies Center*, 1989.

John R Wilmoth. Computational methods for fitting and extrapolating the lee-carter model of mortality change. Technical report, Technical report, Department of Demography, University of California, Berkeley, 1993.