Stochastic Mortality Projections: A Comparison of the Lee-Carter and the Cairns-Blake-Dowd models Using Italian Data.

Samuel Nocito *

Abstract

The work proposes a comparison between the Lee-Carter and the Cairns-Blake-Dowd mortality models, employing Italian data. The mortality projections span the period 2016-2060. The mortality data come from the Italian National Statistics Institute (ISTAT) database, and span the period 1980-2012. The results were simply numerically and graphically compared showing differences between models that evidence the existence of a model risk in annuity pension market. Moreover the life-expectancy projections of the model were compared to the ISTAT benchmark, showing divergences in prevision for the dynamic of the mortality gap between genders. The latter result demonstrates the presence of a data-choosing process risk, and reveals the importance of disclosure in using stochastic models. Moreover, the comparison is observed in the pension policy scenario, taking into consideration the role played by mortality models with the transformation coefficients.

Introduction

The present work attempts to show how the choice of the model, of the dataset and of the application methods adopted, influence the outcomes. For this reason, I implemented two different stochastic mortality models, using the same dataset referred to the Italian population, and I compared the results taking into consideration the projections provided by ISTAT as a benchmark. In particular, I chose to compare among others, the Lee-Carter and the Cairns-Blake-Dowd models, since they represent two different parametric families of mortality models; one considering the logarithms of the central rate of mortality as the dependent variable, and the other the *logit* transformation of the mortality odds. Moreover, I considered ISTAT projections obtained with the Lee-Carter model: this allowed me to scrutinize the role of dataset decisions. I estimated and projected the parameters of the models, in order to forecast the future mortality trends for

*Graduated student at the University of Studies of Turin (Italy),
Master degree in Quantitative Finance and Insurance.
Electronic address: samuel.nocito[at]gmail.com

the Italian male and female populations. More specifically, on one hand I used the Excel spreadsheet named 'CBD model M5 estimation.xlsm', provided by A. Cairns, D. Blake and K. Dowd on their model web site[1]. On the other hand, I applied the Lee-Carter model using Matlab; a user-friendly code is presented in the Appendix, indeed. The comparison of results stressed the differences due to the features of the models and to the stream of data chosen: the outcomes of the models when compared to the benchmark showed better improvements in the mortality gap between genders than the one proposed by ISTAT. Furthermore, the analysis is observed considering the pension policy scenario, without having any presumption of completeness, but instead aiming to be one of the first little steps for future researches in the field of mortality model risk in pensions.

1 The role of mortality models in pensions

The past two decades have seen an extraordinary evolution in European pension policy, characterized by new pension reforms [Whitehouse and Queisser, 2012] aiming to fit within the economic and social changes, having the purposes of income redistribution, financial sustainability, and risk sharing among generations. For this reason, forecast of the pensioners' remaining life expectancy [Whitehouse, 2007] became an important information for estimating the pension expenditure. In fact, the latter is influenced by the longevity risk, that derives from improvements in mortality trend, which determine a systematic deviations of the number of deaths from its expected values. In the early 90's, the Italian pension system was organized following a *pay-as-you-go* defined-benefit (DB) pension plan, without any specific treatment for the longevity risk. The retirement ages were 60 for men and 55 for women, and only with the *Riforma Amato* (1992) did the retirement ages skip to 65 and 60, respectively for men and women. Since the pension structure was unfair and financially unsustainable, in 1995 - under the technical government of Dini (Law 335/95)- the Italian pension system partially switched from a DB to a *notional defined-contribution* (NDC) plan, keeping constant the *pay-as-you-go* structure [Brambilla, 2012]. This reform was a kind of revolution for the Italian pension system due also to the fact that, by the introduction of the *notional accounts*, the retirement annuities were for the first time correlated with the evolution of the average life-expectancy. The NDC scheme allows employees to accumulate - at the expected GDP growth rate - their working contributions on a notional account that will be converted into an annuity at retirement. In particular, the annuity is computed by adopting the so-called transformation coefficients [Fornero and Castellino, 2001], which take into consideration the average remaining life expectancy.

Piscopo [2011] presented the following formulation for the computation of NDC pension benefits at retirement age x:

$$P(x) = \left[c_a + \sum_{i=1}^{a-1} c_i \prod_{j=1}^{a-1} (1 + \bar{g}_j) \right] \delta_x \tag{1}$$

[1] http://cbdmodel.com/

where c_i is the contribution paid by the worker at seniority i, a is the seniority at retirement (i.e., the number of years of the working life), $\bar{g}_j$ is the geometric mean of the nominal GDP growth rate calculated according to the past 5 years observations preceding the seniority year j; finally, δ_x is the retirement transformation coefficient at age x, generalized by the formula:

$$\delta_x = \left[\sum_{t=0}^{\Omega - x} E[\mathbb{1}_{T_{x \leq t}}] P_0^{t+1} \right]^{-1} \tag{2}$$

representing the inverse of the expected present value of an unitary annuity revertible to the spouse. In the formula above T_x is the remaining lifetime for an individual aged x, with Ω representing the conventional final age. Furthermore, $\mathbb{1}_{T_{x \leq t}}$ represents the indicator random variable that expresses whether an individual aged x at time 0 will survive at least t more years. In particular:

$$E[\mathbb{1}_{T_{x \leq t}}] = E[{}_t p_x] = \frac{l_{x+t}}{l_x} \tag{3}$$

where ${}_t p_x$ is the survival probability at age $x + t$, conditional on being alive at age x and l_x is the number of survivors at age x. In the Italian case, the formulae adopted for computing pension benefits are slightly different from the ones presented above. However, the Italian presentation will not be discussed here, since it is not strictly necessary for understanding the role of projected death probabilities in pensions[2]. It is noticeable that mortality projections directly alter the annuity values [Tang et al., 2015].

Projected death probabilities are usually provided by stochastic mortality models. For this reason, the choice of the model which could better fit the data is truly important. Cairns et al. [2009] compared eight existing mortality models on the basis of the Bayesian Information Criterion (BIC); showing that on the same dataset, one model can fit better than others. Also Yang et al. [2010] studied the impact of model risk on annuity pricing by comparing existing mortality models, taking in consideration the BIC and also the Mean Absolute Percentage Error (MAPE) when making comparison. Moreover, the magnitude of the stream of data (i.e., the time interval of the dataset) is another important variable affecting results as well as the approach adopted in the application of mortality models [Danesi et al., 2015]. The comparison of the Lee-Carter and the Cairns-Blake-Dowd model that will be presented hereafter, has been conducted without considering the BIC and the MAPE. The results of the models were simply compared in order to analyze the existence and the quality of the differences. Moreover, the outcomes of the models were compared to the life-expectancy projection provided by ISTAT, since the institution have probably adopted a different dataset from the one used for the application of the models. This, it seems, could pose interesting outlooks on the role of data choosing process.

[2]Detailed information on Italian formuale are provided by Piscopo [2011].

2 Data and analysis presentation

The datasets related to the Italian population were provided by ISTAT[3] (Mortality Tables), for both men and women. In particular, I considered data over the time horizon 1980-2012; I chose to discard older data in order to avoid biases due to previous historical events that could have affected results (e.g. discoveries on cardiovascular diseases, occurred in the seventies).

In the original form, the data were provided in the form of the common life tables. However, the LC and the CBD models required to collect the data into matrices having as row index the age x and as column index the year t. In particular, the Lee-Carter model requires the actuarial variables d_x and L_x; whereas the Cairns-Blake-Dowd model needs only the variable q_x, referring to the probability of death. Furthermore, the last age w is considered, which will be different for each model. In particular, $w = 105$ for the Lee-Carter model and $w = 110$ for the Cairns-Blake-Dowd model. The life tables referring to the two models have been closed at different ages, since the CBD model is a good mortality predictor at higher ages. However, this choice does not have implications on the analysis and comparison of results since the life-expectancies are observed at the age $x = 65$. Moreover, the male and female populations were separately considered with exception to the European legislation that instead suggested the use of a genders unified life table.

3 An application of the Lee-Carter model

I took into consideration the original formulation of Lee and Carter [1992], represented by the following model equation:

$$m_x(t) = e^{\alpha_x + \beta_x k_t + \varepsilon_{x,t}} \tag{4}$$

where $m_x(t)$ represents the central rate of mortality at age x and at time t, and it has been computed by the formula:

$$m_x(t) = \frac{d_x(t)}{L_x(t)}$$

with $d_x(t)$ and $L_x(t)$ variables available from the dataset provided by ISTAT. Furthermore, these variables were considered for each age x in the interval $0 \leq x \leq 105$, and for each year t in the interval $1980 \leq t \leq 2012$.

For the sake of transparency, the model was implemented by adopting its logarithms transformation:

$$\ln m_x(t) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

with the following parameter interpretations:

[3] http://demo.istat.it/unitav2012/index.html?lingua=ita (Observed last time in January 2015)

- k_t the time index representing the level of mortality at time t;
- α_x representing the average trend of mortality on the time horizon at age x;
- β_x representing a measure of the sensitiveness in movement from the parameter k_t;
- the homoskedastic error therm $\varepsilon_{x,t}$, that incorporates the historical trends not considered by the model. It is normally distributed with mean 0 and variance σ_ε^2.

Parameter estimations

The parameter estimation was computed with respect to the *Ordinary Least Square* (OLS) estimation method. Moreover, two constraints were considered in order to be able to find a unique solution for the parameters.

Constraints:

$$\sum_{x=x_1}^{x_m} b_x = 1 \qquad and \qquad \sum_{t=t_1}^{t_n} k_t = 0 \tag{5}$$

In order to obtain the estimation for the variable $\hat{\alpha}_x$, it was necessary to compute the partial derivative of the equation $LS(\alpha, \beta, k) = \sum_x \sum_t (\ln[m_x(t)] - \alpha_x - \beta_x k_t)^2$, with respect to α_x. Then, by setting the partial derivative equal to 0 we get:

$$\hat{a}_x = \frac{1}{t_n - t_1 + 1} \sum_t \ln m_x(t) \tag{6}$$

with $t = t_1, ..., t_n = 1980, ..., 2012$ and $x = x_1, ..., x_m = 0, ..., 105$.

As it is expressed by the equation (6), the estimation for the first parameter α_x was given by the average of the logarithms of the central rate of mortality over time t. Furthermore, the estimations of $\hat{\beta}_x$ and $\hat{k}_t$ for the parameters β_x and k_t were obtained by adopting the singular value decomposition of the matrix Z of elements $(\ln[m_{x_i}(t_j)] - \alpha_{x_i})$, with $i = 0, ...105$ and $j = 1980, ...2012$. In particular, given the decomposition $Z = USV^T$, the parameters have been estimated as it follows:

$$\hat{b}_x = \frac{\vec{u}_1}{\sum_{j=1}^{x_m - x_1 + 1} u_{1j}} \tag{7}$$

$$\hat{k}_t = \vec{s}_1 \left(\sum_{j=1}^{x_m - x_1 + 1} u_{1j} \right) \vec{v}_1^T \tag{8}$$

where:

- $\vec{u}_1$ represents the eigenvector corresponding to the biggest eigenvalue of the matrix ZZ^T;
- $\vec{v}_1$ is the eigenvector corresponding to the biggest eigenvalue of the matrix $Z^T Z$;

- finally, $\vec{s}_1$ is a single-element vector corresponding to the biggest eigenvalue of either U or V matrices.

At this point, many scholars [Barugola and Maccheroni, 2014] usually recalibrate the time index in order to reproduce the observed number of deaths in a given year. However, the recalibration was avoided here without implication on results. Finally, the estimated parameters were computed as follows, due to the fact that the estimated parameters $\hat{k}_t$ did not satisfy the constraint at (5), and therefore needed to be adjusted:

$$a_x^* = \hat{a}_x + \hat{b}_x\bar{k} \tag{9}$$

$$b_x^* = \hat{b}_x\left(\sum_{j=1}^{x_m-x_1+1}\hat{b}_{1j}\right) \tag{10}$$

$$k_t^* = (\hat{k}_t - \bar{k})(\sum_{j=1}^{x_m-x_1+1}\hat{b}_{1j}) \tag{11}$$

where $\bar{k} = \frac{1}{t_n-t_1+1}\sum_{t=1}^{t_n}\hat{k}_t$ is the arithmetic average of $\hat{k}_t$ with respect to time t, and $\left(\sum_{j=1}^{x_m-x_1+1}\hat{b}_{1j}\right)$ is simply the sum of all the estimated $\hat{b}$, which sum to 1.

For the sake of simplicity, all the numbers referring to the parameter estimations will not be reported. However, the figures below will show the trend and the path of the estimated parameters α_x^*, β_x^* and k_t^*. In the Figures (1) and (2) below, the decreasing path of the graphs referred to the parameter k_t^*, shows the overall mortality improvement occurred over the considered historical period. Moreover, the graphs referred to the parameter β_x^* show the differences between genders. In particular, the male population had a mortality improvement, starting from the age $x = 30$ to the age $x = 65$ (i.e. increasing path of the male curve at adult ages), whereas the female population had greater improvements in the range of ages $60 \leq x \leq 85$ (i.e. old ages). Moreover, both genders present significative variations for the parameter at birth.

Figure 1: The LC estimated trend of α_x^*, β_x^* and k_t^* parameters (Male).

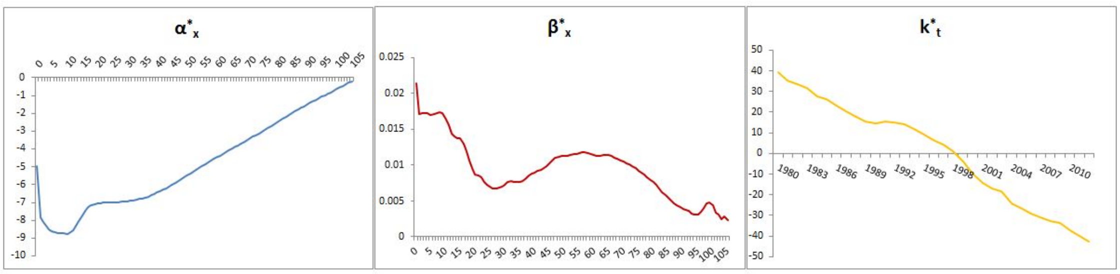
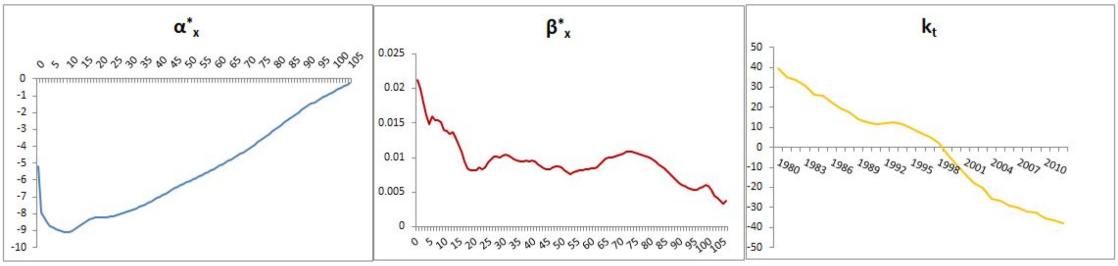


Figure 2: The LC estimated trend of α_x^*, β_x^* and k_t^* parameters (Female).



Forecasting mortality

The following *Random Walk with Drift* equation was adopted in order to project the time index k_t:

$$k_t = k_{t-1} + d + \varepsilon_t \qquad with \qquad \varepsilon_{x,t} \sim \mathcal{N}(0,1); E(\varepsilon_s, \varepsilon_t) = 0 \tag{12}$$

where the drift d was estimated by the formula:

$$\hat{d} = \frac{(k_T^* - k_1^*)}{t_n - t_1} \tag{13}$$

with k_T^* and k_1^* respectively given by the elements k_{1n} and k_{11} of the estimated vector $k_t^* = [k_{t1}^*, ..., k_{tn}^*]$.

After having solved the equation (12) of the RWD model, I projected the parameter k_t at time $T + \Delta t$ as it follows:

$$\hat{k}_{T+\Delta t} = k_T^* + (\Delta t)\hat{d} + \sqrt{\Delta t}\varepsilon_t \tag{14}$$

and taking the expected value we get:

$$E\left[\hat{k}_{T+\Delta t} \middle| k_1^*, ..., k_T^*\right] = k_T^* + (\Delta t)\hat{d}$$

At this point, it was possible to get the equation for the projection of the central rates of mortality as it follows:

$$\ln\left[\hat{m}_{x,T+\Delta t}\right] = a_x^* + b_x^*\left[k_T^* + (\Delta t)\hat{d}\right]$$

and so

$$\hat{m}_{x,T+\Delta t} = e^{a_x^* + b_x^*\left[k_T^* + (\Delta t)\hat{d}\right]} \tag{15}$$

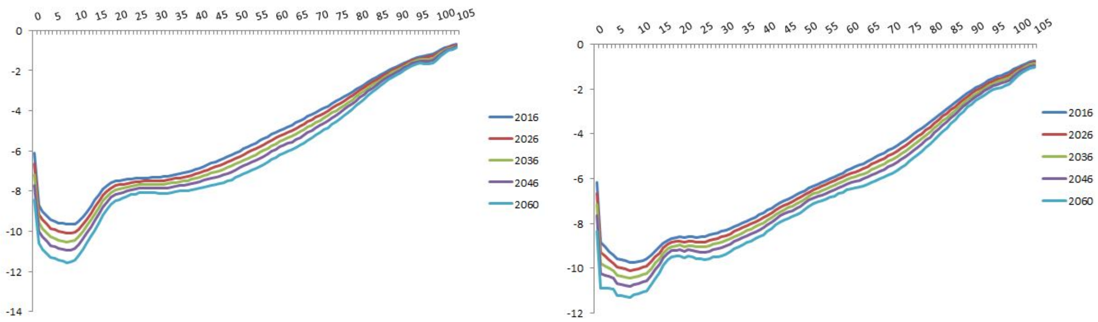
Finally, the central mortality rates were transformed into probabilities, by adopting the Reed and Merrell [1939] method. The relation is expressed by the equation:

$${}_nq_x(t) = 1 - e^{-n\left(m_x(t)\right) - n^3 0.008\left(m_x(t)\right)^2} \tag{16}$$

The results of the Lee-Carter model

The graphs in the Figure 3 below show the forecast mortality trends (logarithmic scale) for different years into the projection interval $2016 \leq t \leq 2060$, respectively for the male and the female populations.

Figure 3: The Lee-Carter Male (left) and Female (right) mortality projection.



The mortality curves, are one under the other following the same order of the legend presented in the figures. This dynamic, shows the forecast mortality improvements over the projection interval. Lower curves are associated to lower death probabilities at each age, indeed.

4 An application of the Cairns-Blake-Dowd model

I considered the original formulation of the model provided by Cairns et al. [2006]. Moreover, since the model is a good predictor of the mortality evolution at higher ages; the age x examined in the interval $0 \leq x \leq 110$ over the time horizon 1980-2012. For this reason, the mortality projections will refer exclusively to every age $x \geq 60$. The model equation adopted is the following:

$$\ln \left[\frac{q_{x,t}}{p_{x,t}} \right] = k_t^{(1)} + k_t^{(2)} (x - \bar{x}) + \varepsilon_{x,t}$$

where

- $k_t^{(1)}$ and $k_t^{(2)}$ are two stochastic processes and represent the two time indexes of the model;
- $q_{x,t}$ and $p_{x,t}$ represent respectively the death and the survival probability, at time t for an individual aged x;
- $\ln \left[\frac{q_{x,t}}{p_{x,t}} \right] = \ln (\phi_x) = \textit{logit}\ q_{x,t}$ is the *logit* transformation of $q_{x,t}$, with ϕ_x representing the mortality odds ;

- $\bar{x}$ is the central age over the range of ages, then $\bar{x} = 55$.
- $\varepsilon_{x,t}$ is the error term that encloses the historical trend that the model does not express. All the error terms are i.i.d following the Normal distribution with mean 0 and variance σ_ε^2.

Moreover, the time index $k_t^{(1)}$ is the intercept of the model, it affects every age in the same way and it represents the level of mortality at time t. More precisely, if it declines over time, it means that the mortality rate have been decreasing over time at all ages. The time index $k_t^{(2)}$ represents the slope of the model: every age is differently affected by this parameter. For instance, if during the fitting period, the mortality improvements have been greater at lower ages than at higher ages, the slope period term $k_t^{(2)}$ would be increasing over time. In such a case, the plot of the logit of death probabilities against age would become more steep as it shifts downwards over time [Pitacco et al., 2009].The model does not require additional constraints. For the estimation and projection of the parameters, refer to the CBD model website[4] provided by Cairns et al.

The Figures (4) and (5) below, will show the trends of the estimated parameters $\hat{k}_t^{(1)}$ and $\hat{k}_t^{(2)}$ over time, for men and women respectively.

Figure 4: The CBD estimated trend of the time indexes $\hat{k}_t^{(1)}$ and $\hat{k}_t^{(2)}$ (Male).

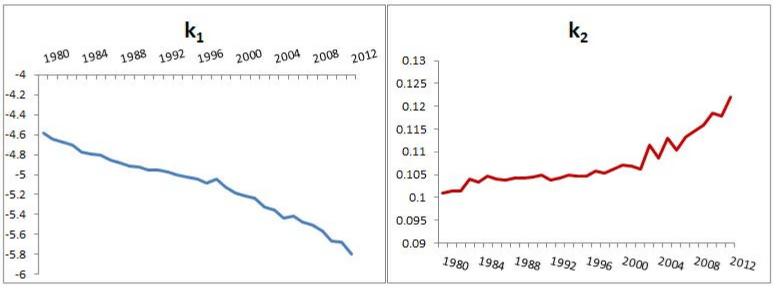
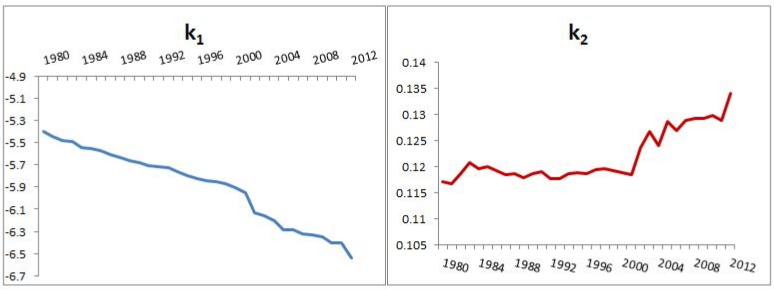


Figure 5: The CBD estimated trend of the time indexes $\hat{k}_t^{(1)}$ and $\hat{k}_t^{(2)}$ (Female).



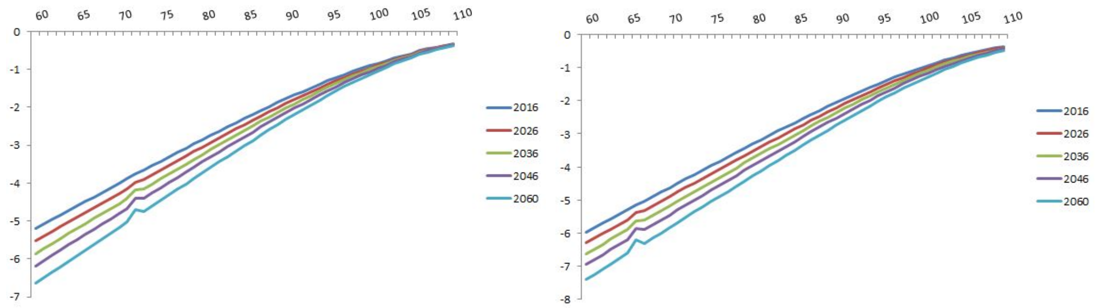
[4] http://cbdmodel.com/estimation.html

As it is shown in the Figure (4) above, the trend of the estimated parameter $\hat{k}_t^{(1)}$ is decreasing over time, which corroborates the idea that the overall level of mortality has been improving over the considered time interval. On the other hand, the trend of the parameter $\hat{k}_t^{(2)}$ had been quite stable until 2002, when it increased. The authors of the model propose th following interpretation of the increasing path of the $\hat{k}_t^{(2)}$ trend: it shows that the mortality improvements - in absolute terms - have been greater at lower ages than at higher ones [Pitacco et al., 2009]. This means that mortality has improved more for ages from 60 to 70 than at higher ages (in absolute terms).

The results of the Cairns-Balke-Dowd model

The graphs in the Figures (6) below, show the forecast mortality trend, respectively for the male and the female populations for every age $60 \leq x \leq 110$.

Figure 6: Cairns-Blake-Dowd mortality projection: Male (left) and Female (right).

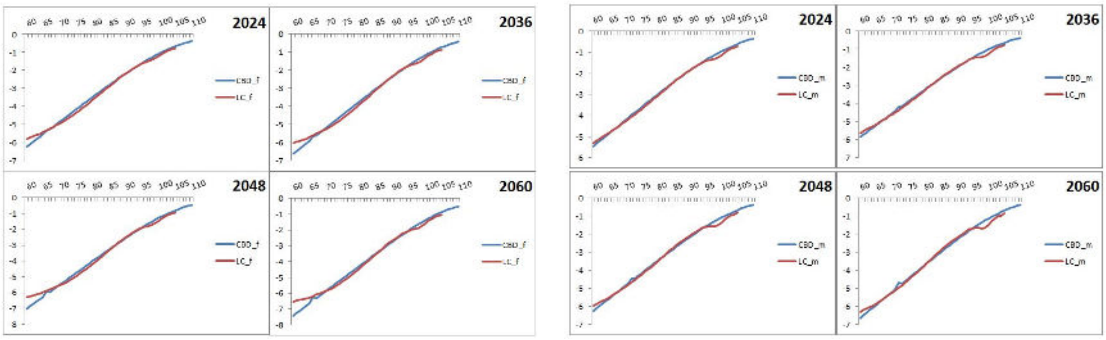


As it can be seen in the figures, the mortality curves for both genders show increasing peaks at the age of 65 for women and at the age of 70 for men. These peaks are anomalies that certainly influence the computation of the forecast life-expectancies, and are probably due to the way by which the model has worked the data. The CBD forecast mortality curves are one under the other, stressing the mortality improvements that occurred at each year over the projection horizon. Moreover, it can be seen that, the curve referring to the year 2016 starts from the level of -6 for women and -5 for men, stressing the differences in mortality between genders.

5 Comparison of results and final considerations

Even though the Lee-Carter mortality trends cover all ages x in the interval $0 \leq x \leq 105$, the comparison of results scrutinizes the interval of ages adopted for the Cairns-Blake-Dowd model application, which is $60 \leq x \leq 110$. For this reason, as demonstrated in the graphs below, the CBD curves will be longer than the LC ones, representing also the ages from 106 to 110 that the Lee-Carter model does not consider. Moreover, I decided to take a sample of four different years in the projection horizon 2016-2060, starting from the year 2024 with intervals of 12 years between them.

Figure 7: CBD vs LC: Female (left) and Male (right) mortality trends



On one hand, the Cairns-Blake-Dowd curves in the female case suggest greater improvements in mortality at the age of 60 than the ones provided by the Lee-Carter curves. Furthermore, the Lee-Carter curves for the set of ages $70 \leq x \leq 80$ are slightly convex compared with the CBD ones. The Lee-Carter convex dynamics in that interval suggest better mortality perspectives with respect to the ones proposed by the Cairns-Blake-Dowd model. On the other hand, the model curves of the male case show errors in the paths, due to the iteration and to the logarithmic transformation processes. The CBD curves present an error peak at the age of 70 which becomes more evident from year to year. The Lee-Carter curves present errors at the age of 93. However, the curves of both models converge at almost all ages, with some divergences at the age of 60 even though smaller than the female case ones.

Concerning the benchmark, it was possible to compare the application results exclusively referring to the life-expectancies, since ISTAT does not disclose the data related to the death probability projections. The life-expectancy projections have been computed by constructing the projected life tables for both men and women, adopting the common actuarial functions.

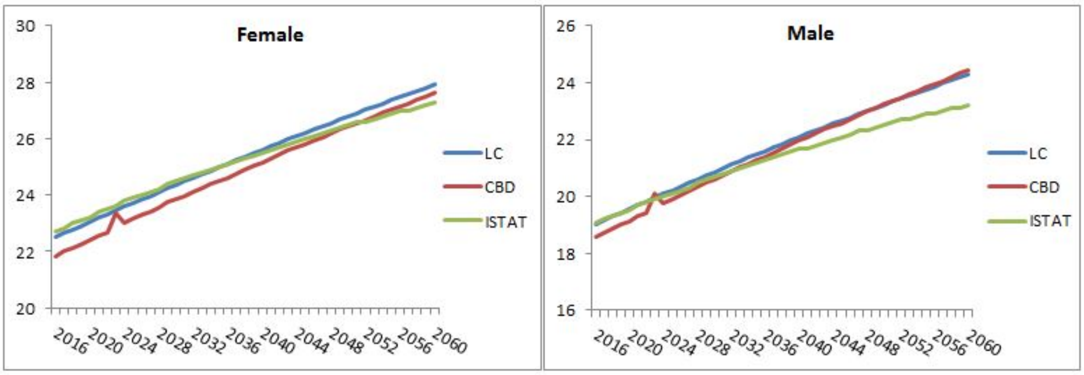
Table (1) below, shows the forecast life-expectancy for women and men respectively. Moreover, the data have been collected considering time steps of eight years, in order to stress the improvements in the life span. In the female case, it is possible to observe that the Lee-Carter curve converges with the ISTAT benchmark particularly in the first half of the projection horizon and starts diverging from the year 2040, with greater improvements suggested by the LC curve. However, it is important to note that both the Lee-Carter and the Cairns-Blake-Dowd models underestimate life-expectancy in the first years of the projection horizon and overestimate in the last ones with respect to ISTAT. Looking at the numbers in Table (1) below, the differences between the models and the benchmark are always less than 11 months which is the highest divergence presented by the CBD model in the year 2016.

Table 1: Life-Expectancy projections at age 65: a comparison of results.

Female (Aged 65)				Male (Aged 65)			
Year	LC	CBD	ISTAT	Year	LC	CBD	ISTAT
2016	22.5	21.8	22.7	2016	19.0	18.6	19.1
2024	23.6	23.0	23.8	2024	20.1	19.8	20.0
2032	24.6	24.1	24.7	2032	21.1	20.9	20.9
2040	25.6	25.2	25.5	2040	22.1	22.0	21.7
2048	26.6	26.2	26.3	2048	23.0	23.0	22.3
2056	27.5	27.2	27.5	2056	23.9	24.0	22.9
2060	27.9	27.6	27.3	2060	24.3	24.4	23.2

Figure (8) below, graphically represents the comparisons of ISTAT benchmark and the life-expectancy projections provided by the models. In the male graph below, it is possible to observe that, even though the CBD curve diverge from the others until the year 2036 with a decreasing difference, it is also noticeable that from that year, it converges almost perfectly with the Lee-Carter curve. Moreover, starting from the year 2030 (2028 for the LC curve) the curves of both the models increasingly diverge -with a similar dynamic- with respect to the ISTAT curve. This result buttresses the evidence that the gap between women and men is expected to decrease over the projection horizon. These divergences between the benchmark and the models emphasize the existence of a data choosing process risk (i.e. a different dataset leads to different results). Moreover it is reasonable to ask: which is the correct dataset to use!?

Figure 8: Life-Expectancy comparisons



Finally, it is important to recall that the divergences between the models are due to the differences between them, and this underlines the existence of the model risk in mortality projections. Keeping constant the chosen dataset, different models lead to different outcomes and so the choice of the result that better fit the reality represent a risk in itself. The differences between the results of models and the ISTAT benchmark could be related to the fact that ISTAT does not provide the model specification, the procedure adopted in order to estimate and forecast the parameters neither does it give information about the dataset used (intervals of ages and years) for the Lee-Carter model implementation. Needless to say that with no certainty, it is possible to assert that all these features perfectly coincide with the ones characterizing the applications of both the Lee-Carter and the Cairns-Blake-Dowd models that were proposed. In this case, the compared results showed how the choice of the dataset directly influence the quality and the features of the model outcomes (different datasets lead to different outcomes *ceteris paribus*), reveling the existence of a data-choosing process risk. The importance of methodology disclosure is usually underestimated and conversely, this could help researchers to improve models and people in general for better comprehension. It is for this reason therefore that I chose to disclose meticulously (sometimes tediously) all the passages of the procedures.

Appendix

The Lee-Carter Matlab code

The data related to the variables $L(x,t)$ and $d(x,t)$ were imported to Matlab, using the function *xlsread*. Then, the users need to first collect the data into an excel spreadsheet. The first two lines of the code represent an example of what I had explained above. The variable t and T represent the extremes of the projection horizon (2012-2060 in our case), so the users can properly change them as they need.

Figure 9: Lee-Carter Matlab code: Part 1

```
L_xt = xlsread('female-dataset.xls', 'L(x,t)', 'B2:AH107');
d_xt = xlsread('female-dataset.xls', 'd(x,t)', 'B2:AH107');
t = 2012;
T = 2060;

deltaT = 1:(T-t);
[m,n]=size(d_xt);
I = ones(m,n);
Ia = ones(1,n);
Ip = ones(1,(T-t));

m_xt=d_xt./L_xt;
lnm_xt=log(m_xt);

a_x1=mean(lnm_xt,2);

Z=lnm_xt-(a_x1*Ia);
[U S V] = svd(Z);
u1=U(:,1);
v1=V(:,1);
sum_u1=sum(u1);
s1=S(1,1);

b_x1=u1/sum_u1;
k_t1=s1*sum_u1*v1';
```

```
sumb_x1=sum(b_x1);
while sumb_x1 ~=1
    disp('It does not satisfies the model constraint sum(b_x1)=1.' )
    disp('It needs to be recalibrated.')
    break
end

sumk_t1=sum(k_t1);
while sumk_t1 ~=0
    disp('It does not satisfies the model constraint sum(k_t1)=0.')
    disp('It needs to be recalibrated.')
    break
end

meank_t1=mean(k_t1);
a_x=a_x1 + b_x1*meank_t1;
b_x=b_x1/sumb_x1;
k_t=(k_t1 - meank_t1*Ia)*sum(b_x1);

sumb_x=sum(b_x);
sumk_t=sum(k_t);

kT = k_t(1,length(k_t));
k1= k_t(1,1);
d= (kT - k1)/length(k_t);
k_projected= kT*Ip + deltaT*d;
lnm_xt_proj = a_x*Ip + b_x*k_projected;
m_xt_proj = exp(lnm_xt_proj);
```

Finally, it will be asked the user to choose the name of the Excel file on which all the outputs will be located. Please, manage the range of the excel data locations in accordance with the stream of data imported.

Figure 10: Lee-Carter Matlab code: Part 2

```
file = input('Please, choose the excel file name for the model results: ');
xlswrite(file,Z,'B2:AH107');
xlswrite(file,a_x,2,'B2:B107');
xlswrite(file,b_x,2,'C2:C107');
xlswrite(file,k_t,3,'B2:AH2');
xlswrite(file,k_projected,4,'B2:AW2');
xlswrite(file,lnm_xt_proj,5,'B2:AW107');
xlswrite(file,m_xt_proj,6,'B2:AW107');
```

References

Tiziana Barugola and Carlo Maccheroni. Le differenze di genere alla morte in italia: evoluzione recente e tendenze in atto. *Quaderni Europei*, 20, 2014.

Alberto Brambilla. *Istituzioni, economia e gestione di previdenza pubblica*. Vita e pensiero, 2012.

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, 2006.

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.

Ivan Luciano Danesi, Steven Haberman, and Pietro Millossovich. Forecasting mortality in subpopulations using lee–carter type models: A comparison. *Insurance: Mathematics and Economics*, 62:151–161, 2015.

Elsa Fornero and Onorato Castellino. La riforma del sistema previdenziale italiano. *Il Mulino, Bologna*, 2001.

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

Gabriella Piscopo. The impact of longevity risk on annuitization factors in the italian state defined contribution pension scheme. *Insurance Markets and Companies: Analyses and Actuarial Computation*, 2(1):31–39, 2011.

Ermanno Pitacco, Michel Denuit, Steven Haberman, and Annamaria Olivieri. *Modelling longevity dynamics for pensions and annuity business*. Oxford University Press, 2009.

Lowell J Reed and Margaret Merrell. A short method for constructing an abridged life table. *American Journal of Epidemiology*, 30(2):33–62, 1939.

Chen Tang, Bridget Browne, and Aaron Bruhn. Valuing annuities based on alternative mortality projections. *Volume 3 Australian Journal of Actuarial Practice*, page 23, 2015.

Edward Whitehouse and M Queisser. Recent pension reforms in oecd countries, 2012.

Edward R Whitehouse. Life-expectancy risk and pensions. 2007.

Sharon S Yang, Jack C Yue, and Hong-Chih Huang. Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. *Insurance: Mathematics and Economics*, 46(1):254–270, 2010.