 doi:10.20944/preprints201701.0065.v1

Backtesting the Lee-Carter and the Cairns-Blake-Dowd Stochastic Mortality Models on Italian Death Rates

Carlo Maccheroni [1,2,*] and Samuel Nocito [2,*]

[1]Dondena Centre for Research on Social Dynamics, Bocconi University, 20100 Milano, Italy
[2]Collegio Carlo Alberto, University of Turin, 10124 Torino, Italy
[*]Correspondence: carlo.maccheroni@unibocconi.it (C.M.); samuel.nocito@unito.it (S.N.)

ABSTRACT

The work proposes a backtesting analysis in comparison between the Lee-Carter and the Cairns-Blake-Dowd mortality models, employing Italian data. The mortality data come from the Italian National Statistics Institute (ISTAT) database and span the period 1975-2014, over which we computed back-projections evaluating the performances of the models in comparisons with real data. We propose three different backtest approaches, evaluating the goodness of short-run forecast versus long-run ones. We find that both models were not able to capture the improving shock on the mortality observed for the male population on the analyzed period. Moreover, the results suggest that CBD forecast are reliable prevalently for ages above 75, and that LC forecast are basically more accurate for this data.

Keywords: lee-carter; cairns-blake-dowd; mortality models; backtesting

1 Introduction

Dowd et al. [2010a] recently performed a backtesting analysis on seven different stochastic mortality models with results providing that the models performed adequately by most backtests. The analysis was applied to English & Welsh male mortality data. We decided to perform a backtesting investigation using Italian mortality data. The decision was motivated by the study of the historical mortality trend, observed on the forty past years horizon for both the male and female populations. The gap between genders deeply decreased over the considered horizon with steep improvements in male mortality. So the first attempt of this paper is to scrutinize the forecast proposed by the models for both sexes, which have experienced different mortality evolutions. Moreover, in the last three decades mortality projections have been widely used by Italian

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

policy makers for taking decisions on public pension reforms. The study of the *mortality risk* - intended as a the uncertainty in future mortality rates - as well as *longevity risk* for the long-term trend in mortality rates [Cairns et al., 2006] played a central role for both public and private annuity providers. Among all the principal stochastic mortality models[1], we chose to compare the Lee-Carter (LC) and the Cairns-Blake-Dowd (CBD) ones.

On the one hand, the Lee-Carter model has signed a deep methodological revolution in the field of demographic forecast, particularly in mortality. The mortality model has been used together with a similar fertility model and deterministic migration assumptions to generate stochastic forecast of the population and its components. These stochastic population forecast, in turn, have been used as the key component of stochastic projections of the finances of the U.S. Social Security system. The stochastic forecast avoid some of the problems inherent to the use of the classic scenario method for the representation of forecast uncertainty [Lee, 2000]. Moreover, the LC model has stimulated a long series of applications involving different countries and institutions: statistical, research and government offices. The great credit of the LC model was to give access to mortality projection not only to demographers but also to many other researchers of different fields. Then in occurrence with the main demographic applications, the LC model suggested:

- an important research front on the problematics related to the parameter estimations [Booth et al., 2006], with many applications also in the actuarial and economics literature [Loisel and Serant, 2007];
- extension of the forecasting analysis with disaggregated projection on demographic subsets trying to maintain consistency at the aggregate level [Lee and Miller, 2001], [Li and Lee, 2005] and [Li, 2010].

On the other hand, the Cairns-Blake-Dowd model even if more recent in its formulation with respect to the LC model, has played an important role in forecasting mortality at higher ages (i.e. from ages starting at 60 and over). The mortality model gave great contributions for pension funds, life-insurance companies and private annuity providers in general. It is mainly used for pricing longevity bonds as suggested also by the authors in the first formulation of the model Cairns et al. [2006]. For these reasons, we chose LC and CBD models among others, since they also represent two of the principal parametric family of mortality models; one considering the logarithms of the central rate of mortality as dependent variable, and the other the logit transformation of the mortality odds.

The second attempt of this work is to analyze the long-term forecast with respect to the short-term, observing potential differences in the parameter estimations [Mavros

[1]Refer to Cairns et al. [2009] for detailed list and quantitative comparison of the principal stochastic mortality models.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

et al., 2014][2] accordingly with changes in the starting point of the database. Chan et al. [2014] have also studied the new-data-invariant property on the quality of the CBD mortality index. For this purpose, we introduced a new backtesting approach named hereafter jumping fixed-length horizon that makes short-run projections of five years, "jumping forward" the historical database by five-year-steps.

Considerations on the backtesting results suggested by the models do not imply a conclusive evaluation of the models, since we perform the analysis considering exclusively the range of ages $57 \leq x \leq 90$. The choice for the interval of ages was motivated by the fact that in Italy, Ragioneria dello Stato compute the so called transformation coefficients for pension annuities, starting from age 57. Moreover, since the CBD model is recommended as a good predictor of mortality at higher ages, we chose such interval of ages in order to make a more prudent and accurate comparison between the models. Furthermore, we decided to take in consideration only death probabilities $q_{x,t}$ among all the others possible biometric functions.

We used death probabilities $q_{x,t}$ provided by the Italian National Statistics Institute (ISTAT) spanning the period 1975-2014. Then, over the designed horizon of historical data, we select the "lookback" and the "lookforward" windows[3] respectively for the parameter estimation and forecast. In particular, the length of the estimation and forecast window will be different for each of the three backtesting approaches proposed by the work:

- *Fixed horizon backtests*: lookback and lookforward windows of 20 years;
- *Jumping fixed-length horizon backtests*: lookback window of 20 years and lookforward window of 5 years (short-term projections);
- *Rolling fixed-length horizon backtests*: lookback window of fixed-length (20 years) and a contracting lookforward window from 20 to 2 years of projections.

The paper is organized as follows. Section 2 briefly presents the models and the adopted terminology, Section 3 shows the historical Italian mortality data, Section 4 and subsections explain methodology and the backtesting results obtained by the different approaches. Section 5 concludes.

2 Model Specifications

2.1 The Lee-Carter Model

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

$$m_x(t) = e^{\alpha_x + \beta_x k_t + \varepsilon_{x,t}} \tag{1}$$

[2] In particular for the case of Cairns-Blake-Dowd model.

[3] For the sake of simplicity, we decided to adopt the same terminology used by Dowd et al. [2010a]

where $m_x(t)$ is the central rate of mortality at age x and at time t, and it is given by the formula:

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

with $d_x(t)$ representing the number of deaths occurred between x and $x+1$, and $L_x(t)$ called the age units lived in x that is simply the mean number of individuals alive between x and $x+1$.

For simplicity, the model was implemented by adopting its logarithm transformation:

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

with the following parameter interpretations:

- k_t is 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 sensitivity in movement from the parameter k_t. In particular, β_x describes the relative speed of mortality changes, at each age, when k_t changes.
- $\varepsilon_{x,t}$ is the homoskedastic error term, that incorporates the historical trends not considered by the model. It is assumed to be $\varepsilon_{x,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

The Appendix A illustrates the method adopted for the estimation and projection of the parameters.

2.2 The Cairns-Blake-Dowd Model

We considered the original formulation of the model provided by Cairns et al. [2006] with the following model equation:

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

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;
- $\ln \left[\frac{q_{x,t}}{p_{x,t}} \right] = \ln(\phi_x) = logit\ q_{x,t}$ is the *logit* transformation of $q_{x,t}$, with ϕ_x representing the mortality odds;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

- $\bar{x}$ is the mean age of the considered interval of ages;
- $\varepsilon_{x,t}$ 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.

The model is full identified, so it does not require additional constraints.

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 Appendix B illustrates the estimation and projection methods involved.

3 Case Study: Italian Mortality Data from 1975 to 2014.

The application of the presented models requires the use of the death probabilities time series for extrapolating mortality forecast. As already mentioned, we use data provided by ISTAT, also because these data are commonly used by private insurance companies and public pension providers. The range of ages is $57 \leq x \leq 90$. In particular, we chose the upper limit for taking into consideration the ISTAT graduation method of ending the life table [Istat, 2001]. The calculation of the probabilities of dying for ages over 95 is performed by extrapolating the $q_{x,t}$ graduated values following the Thatcher et al. [1998] model[4]:

$$q_{x,t} = \frac{\vartheta e^{\gamma x}}{1 + \vartheta e^{\gamma x}} \tag{3}$$

This kind of graduation could have affect the backtesting results, comparing realized data with forecast obtained applying the LC (1) and the CBD (2) models, since they offer a different mortality pattern at old ages. Moreover, we selected the time period from 1975 to 2014, because from the mid-seventies in Italy successfully began the fight against cardiovascular diseases, and more recently against tumors that are still the main cause of death. These successes have contributed to an extraordinary acceleration of growth in life expectancy, especially at higher ages: e.g. from 1975 to 2014, life expectancy at 60 years has seen an average increase of about four hours each day, both for men and women. In the male case, this phenomena extraordinarily occurred. Previously, life expectancy at birth had registered a first significant increase due to the

[4] In the equation (3) ϑ and γ are parameters that need to be estimated. In general, those parameters are estimated applying OLS on the logit transformation of equation (3).

control of infant and child mortality, while from the years under review it has also benefited from the control of adult age mortality.

Table 1: Proportion of persons aged 30 and expected to be alive at selected ages

	Italian Period Life Tables								
Ages	1975	1980	1985	1990	1995	2000	2005	2010	2014
					Male				
50	0.9438	0.9483	0.9554	0.9583	0.9591	0.9662	0.9722	0.9755	0.9777
60	0.8406	0.8487	0.8646	0.8839	0.8951	0.90962	0.9242	0.9324	0.9376
70	0.6292	0.6409	0.6691	0.7081	0.7351	0.7732	0.8060	0.8257	0.8385
80	0.3014	0.3161	0.3539	0.4029	0.4406	0.4936	0.5434	0.5912	0.6188
90	0.0464	0.0527	0.0682	0.0954	0.1170	0.1396	0.1648	0.1996	0.2250
95	0.0080	0.0096	0.0140	0.0235	0.0318	0.0401	0.0491	0.0595	0.0743
					Female				
50	0.9703	0.9739	0.9769	0.9785	0.9796	0.9822	0.9850	0.9865	0.9871
60	0.9194	0.9290	0.9364	0.9427	0.9473	0.9525	0.9585	0.9620	0.9639
70	0.8009	0.8168	0.8337	0.8546	0.8681	0.8828	0.8972	0.9053	0.9087
80	0.5070	0.5403	0.5814	0.6249	0.6576	0.69561	0.7322	0.7540	0.7674
90	0.1154	0.1433	0.1629	0.2141	0.2547	0.2860	0.3297	0.3653	0.3878
95	0.0226	0.0326	0.0380	0.0626	0.0830	0.1030	0.1259	0.1420	0.1654

Currently, the probability of reaching a high age for a young adult is really high: for a 30 years old, the probability of reaching the age of 60 is almost 94% for male and 96.4% for women. However, it remains difficult to reach the threshold of 90 years, especially for men. Table (1) accurately shows[5] how this probability changed starting from age 50. Moreover, it shows how the difference in probability between genders became greater as the age increased.
This process is known as the rectangularization and shift forward of the survival curves, its measure can be derived from the entropy of a life table (4). It was introduced by Keyfitz and Caswell [2005] and it is referred to in this paper as ${}^{t}H_{K,\alpha}$ with α the age by which the survival curve is built, and t the year of the period life table at which the entropy is computed (in our case $t = 1975, 1976, ... 2014$). Then,

$$ {}^{t}H_{K,\alpha} = -\frac{\sum_j (\ln l_j) l_j}{\sum_j l_j} \tag{4} $$

where l_j is the probability of surviving form age[6] α ($\alpha = 0, 1, ..., w$; l_α=1) to age j ($j = \alpha + 1, \alpha + 2, \ldots w$). The entropy index becomes smaller whenever the survivorship curve l_j moves towards a rectangular form; in this limit case ${}^{t}H_{K,\alpha} = 0$.

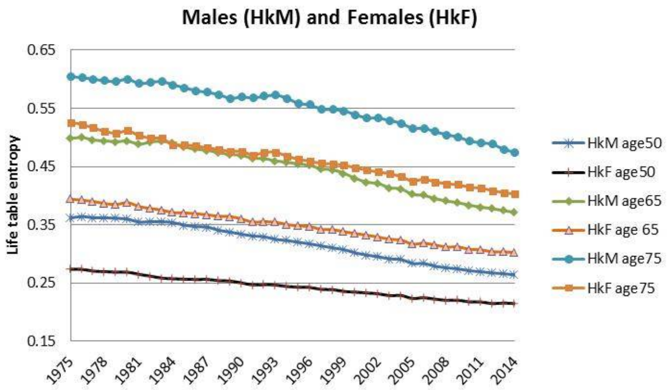
[5] Even though the backtesting analysis will be focused on the interval of ages 57-90, however here we decided to provide information also on ages lower than $x = 57$. In this way, we are able to present the more accurate Italian demographic scenario for the period observed.

[6] The starting point for the final age interval is denoted by w.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

Figure (1) shows how the trend of the rectangularization process has change accordingly to ages (i.e. from $\alpha = 50$ to $\alpha = 65, 75$). As regards women, this process was already in place before 1975. In particular, starting from ages 50 and 65 it is continued with a substantially linear continuity. In the case of men instead, the rectangularization process begins to escalate smoothly after 1984. However, the following trend stresses a deep reduction of mortality, from which is derived an attenuation of the inequality between sexes even though it is not disappeared. From Figure (1), ${}^tH_{k,\alpha}$ shows that the mortality improvement in the elderly population has taken place at different rates over time, particularly with a faster steep decline for both sexes after 1993.

Figure 1: Italian life tables 1975-2014: males and female entropy (${}^tH_{K,\alpha}$)



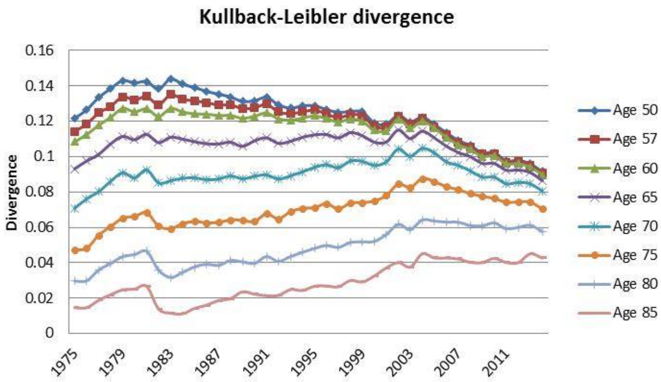
The differentiation of the pace in reducing mortality of both sexes - starting from adult age up to those old - is confirmed by the results of so-called Kullback and Leibler [1951] divergence:

$$ {}^tD_{KL,\alpha}(h_z, g_z) = \sum_{z=0}^{w-\alpha} h_z \ln\left(\frac{h_z}{g_z}\right) \tag{5} $$

where h_z and g_z are the probability distributions of the “time until death" random variable Z_α for a person aged α respectively for males and females. Equation (5) measures the "difference" between these two probability distributions, which in our case is taken as the reference model g_z. The choice is motived not only by the fact that mortality is significantly lower for women than men, but also because the continuous decline of female mortality in the reporting period occurred much more regularly [Maccheroni, 2014]. The divergence in mortality between genders mortality has different characteristics depending on the considered age group indeed.

Figure (2) shows that the divergence in mortality between sexes presents different characteristics, depending on the observed age. In particular, until 1981 the divergence gradually increased on the full range of ages. At a later time, differentials in mortality between sexes decrease whenever x is lower that 60, while it progressively increase at higher ages. These diverging trends make interesting the application of the models, especially for the comparison of results. Needless to say that, the mortality forecast will be more accurate for women than men. This is due to the fact that, women experienced a death risk reduction process with greater regularity than men.

Figure 2: Kullback-Leibler divergence with respect to Z_α at selected ages



4 Backtesting Analysis

In this section we introduce the three different backtesting frameworks studied, and we present the related forecast results.

- The *fixed horizon backtest* uses a fixed twenty-years historical "lookback" interval $1975 \leq t \leq 1994$, and a fixed "lookforward" horizon $1995 \leq t \leq 2014$ (20 years);

- The *jumping fixed-length horizon backtests* make short run projections of five years, and keep fixed the length of the "lookback" horizon (20 years), but making jumps of 5-years-ahead so to cover the "lookforward" interval $1995 \leq t \leq 2014$.
  This analysis is divided in four groups of estimations and forecast, described by the Table (2) below:

  Table 2: Jumping Fixed-Length Horizon Backtests Data Horizon

Lookback Horizon	Lookforward Horizon
1975-1994	1995-1999
1980-1999	2000-2004
1985-2004	2005-2009
1990-2009	2010-2014

- Finally, the *rolling fixed-length horizon backtests* keep fixed the length of the "lookback" horizon (i.e. 20 years) and lets roll it ahead year-by-year. The projection are made over the remaining horizon keeping fixed the last year of projection at $t = 2014$.This analysis is divided in nineteen groups of estimations and forecast, described by the Table (3) below:

Table 3: Rolling Fixed-Length Horizon Backtests Data Horizon

Lookback	Lookforward	Lookback	Lookforward
1975-1994	1995-2014 (20)	1985-2004	2005-2014 (10)
1976-1995	1996-2014 (19)	1986-2005	2006-2014 (9)
1977-1996	1997-2014 (18)	1987-2006	2007-2014 (8)
1978-1997	1998-2014 (17)	1988-2007	2008-2014 (7)
1979-1998	1999-2014 (16)	1989-2008	2009-2014 (6)
1980-1999	2000-2014 (15)	1990-2009	2010-2014 (5)
1981-2000	2001-2014 (14)	1991-2010	2011-2014 (4)
1982-2001	2002-2014 (13)	1992-2011	2012-2014 (3)
1983-2002	2003-2014 (12)	1993-2012	2013-2014 (2)
1984-2003	2004-2014 (11)		

The numbers in parenthesis show the length of the "lookforward" horizon. Moreover, they indicate also the position of the year 2014 over the related projection interval. This will be particularly useful for the analysis of results that will be presented on the related section (4.3).

4.1 Fixed Horizon Backtest (1995-2014)

We begin the backtesting analysis taking into account a forecast horizon that is demographically considered as to be a medium term projection horizon. The comparisons proposed hereafter are among the most likely values of $q^P_{x,t}$ prediction - that is, the central value of the 95% confidence interval derived from the model - and those observed $q^O_{x,t}$; comparisons between the central value and extremes of the confidence interval occurs only for the ages 65 and 85. These are the ages that in the demographic literature mark the entrance in the range of so-called "young-old" and "oldest-old". Unfortunately, due to space limitation it was not possible to present the comparison to the age 75, that divides the old from the "young-old" [Vaupel, 2010].
The $q^O_{x,t}$ can present a strong temporal variability due to the so-called "time effect", that is the time condition that affects mortality from a variety of factors. Among these, the best known is the climatic effect that for instance can cause a rise in mortality at old ages during a very hot summer (e.g. episode occurred in Italy on 2003); or also epidemiological effects that currently arise from flu on winter in low-mortality countries. Needless to say that, the impact of those factors is stronger on the most vulnerable people. For this reason, a rise in mortality due to those factors is generally followed by a decrease in mortality, since those remained alive have a lower frailty level. These mortality shocks can affect short-term forecasts rather than long-term ones, since the latter are usually more capable to capture changes in environmental conditions, socio-economic and people's living styles.

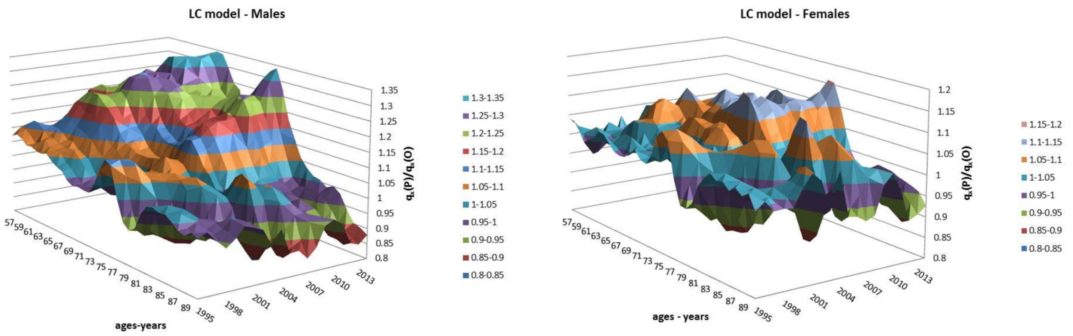
From an applicative point of view, particularly focused on the insurance and social security sector, we were interested in analyzing the performance of the models on assessing the risk of death at various ages. It is from this point of view that we are going to develop our analysis. For this purpose, we make a briefly assessment of forecast errors, which was performed using as an index the sum of the squared errors (SSE), defined as it follows:

$$SSE = \sum_x \sum_t (q^O_{x,t} - q^P_{x,t})^2$$

where $q^O_{x,t}$ and $q^P_{x,t}$ are respectively the death probabilities observed and forecast (projected). Table (4) shows SSE for the first and the second becktesting approach that will be presented in the following section.

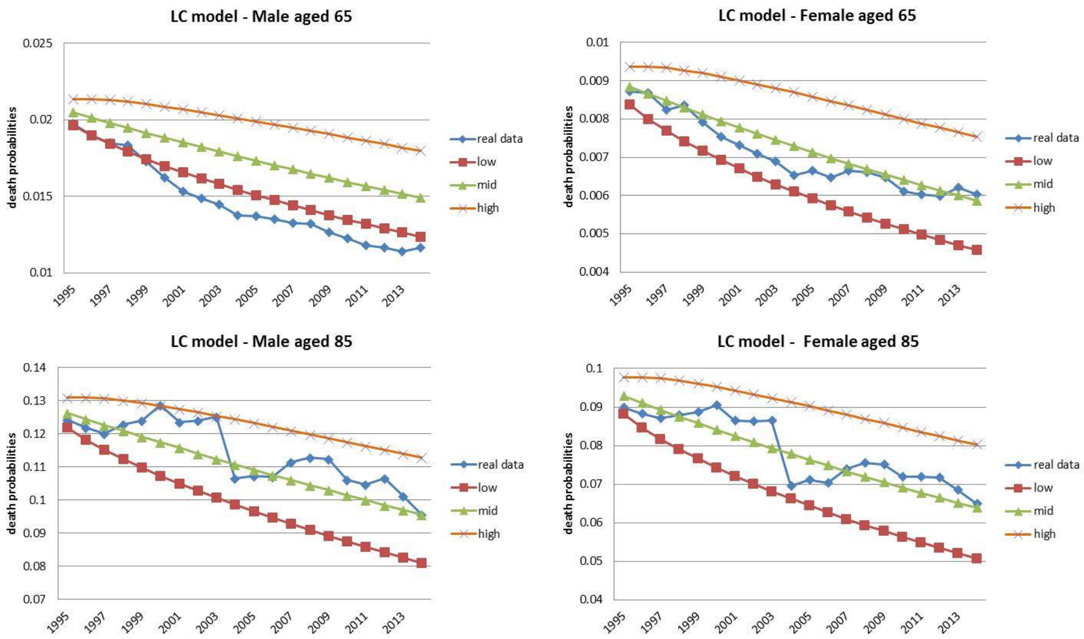
Table (4) shows how the LC model proposes more accurate forecast with respect to the CBD model for the period 1995-20014 in the female case; it is more difficult to judge the models' performances for the male case. As far as men are concerned, LC model initially over-estimates $q^O_{x,t}$ from ages 57 to 70 (approximately), with persistence across years. In particular, the over-estimation errors become greater as the projection is referred to the last year of the forecast horizon. Figure (3) multi-dimensionally shows the ratio between the projected and the real death probabilities. The described LC performance trend is also graphically reported by the figure (4), comparing projections at ages 65 and 85 to the observed data. The over-estimation starts decreasing from age 80, point at which the divergence between $q^O_{x,t}$ and $q^P_{x,t}$ is really close to zero. However, for high ages at the extreme of the interval, LC forecast systematically under-estimate $q^O_{x,t}$

Figure 3: Lee-Carter Fixed Horizon Backtest: $\frac{q^P_x}{q^O_x}$ ratio.



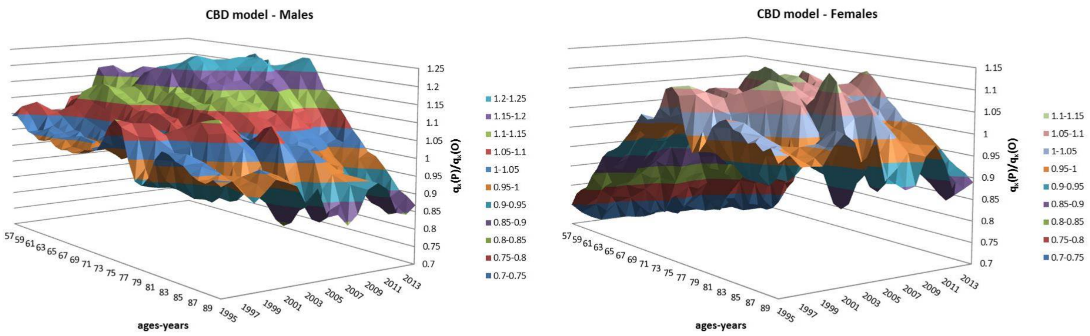
As to women, the divergence between $q^O_{x,t}$ and $q^P_{x,t}$ is sharply smaller than for men. In particular, this is evident on figure (3) where we can see that forecast initially underestimates real data converging at the age 65, then starts overestimating for a wide span of ages. Furthermore, the last part of the age range is again characterized by an underestimation path. However, also the over-estimation experienced at higher ages is smaller than the one observed in the male case. The results obtained with the CBD model are similar to the LC results particularly for the male case. For this reason, it is difficult to privilege the performances of one model with respect to the other over the forecast horizon 1995-2014, also because we have few useful elements to balance the differences in the "goodness of fit" provided by each model.

Figure 4: LC Fixed Horizon Backtest forecast: comparison between observed death rates and the corresponding 95% confidence interval of the forecast based on the time series 1975-1994



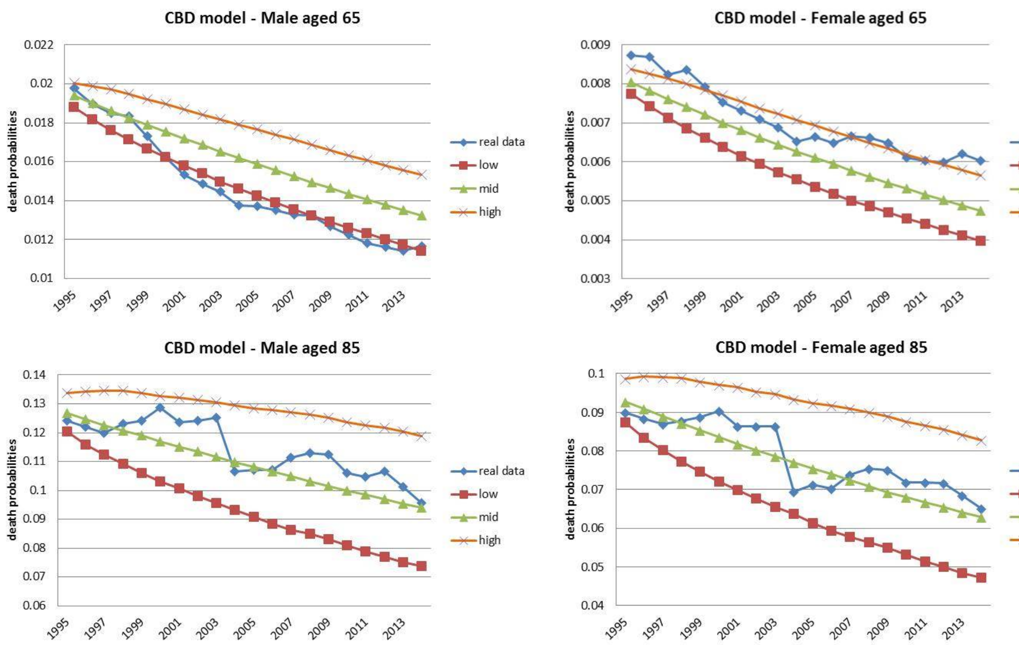
The CBD forecast greatly over-estimates the male mortality historical evolution, particularly for the central and the last years of projection. The error is evident on the full range of ages, however it comes to be smaller at the age 80 after which forecast start under-estimating $q^O_{x,t}$ with an increasing magnitude until the last age and the last projection year (i.e. $x = 90$ and $t = 2014$). This is to be seen in figure (5) below.
When we look at the female case, the accuracy of the CBD forecast is worst. In this case, in fact, we can notice a wide and systematic under-estimation on approximately all the first half of the age range for almost the totality of the forecast horizon. In particular, the forecast error reduces around the age 68, then it starts over-estimating until $x = 85$ after which it under-estimates again. However, at $x = 85$ the forecast is relatively accurate with values of $q^O_{x,t}$ all inside the confidence interval [Figure (6)].

Figure 5: Cairns-Blake-Dowd Fixed Horizon Backtest: $\frac{q_x^P}{q_x^O}$ ratio.



In conclusion, both models make similar forecast errors. On the one hand, regarding males the error is represented by an initial over-estimation that smoothly converge to the real data and then starts under-estimating, though the divergences experienced with the CBD model are characterized by a smaller variability with respect to the LC model. On the other hand, the female case shows an initial under-estimation converging to the real data, and then a fluctuation of over-estimation and final under-estimation. In general, the LC model provides a better fit over a wide range of ages, showing lower variability in both over and under-estimation.

Figure 6: CBD Fixed Horizon Backtest forecast: comparison between observed death rates and the corresponding 95% confidence interval of the forecast based on the time series 1975-1994



Anyway, the choice between the models comes to be difficult at particular ages. Figure (7) shows the high and the low confidence intervals for both the models considered. Even though LC curves are nested into the CBD lines with greater differences shown by the male case, however both models' confidence intervals include the observed data, providing theoretical robustness to the projections.

Figure 7: Fixed Horizon Backtest forecast: comparison between CBD and LC confidence intervals at age 85.

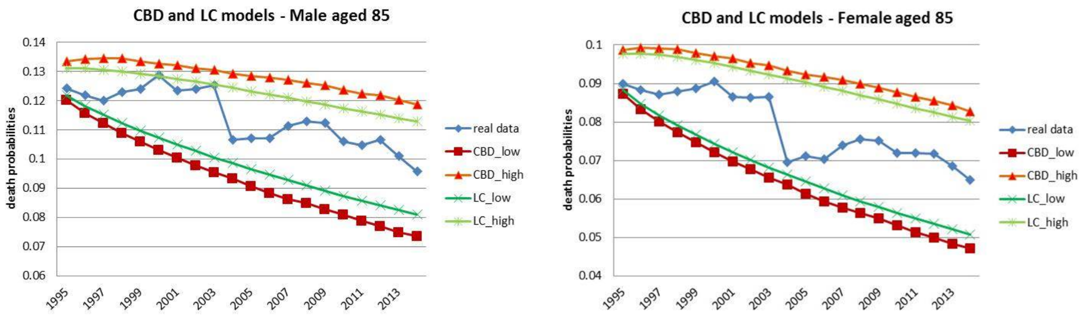


Table 4: Sum of Squared Errors (SSE) between observed q_x^O and forecast q_x^P.

Fixed Horizon Backtest				
	CBD model		LC model	
Prediction Years	Male	Female	Male	Female
1995-2014	0.02653	0.01088	0.02403	0.00507

Jumping Fixed-Length Horizon Backtests				
	CBD model		LC model	
Prediction Years	Male	Female	Male	Female
1995-1999	0.00174	0.00068	0.00249	0.00073
2000-2004	0.00370	0.00283	0.00473	0.00228
2005-2009	0.00234	0.00225	0.00346	0.00151
2010-2014	0.00105	0.00088	0.00149	0.00080

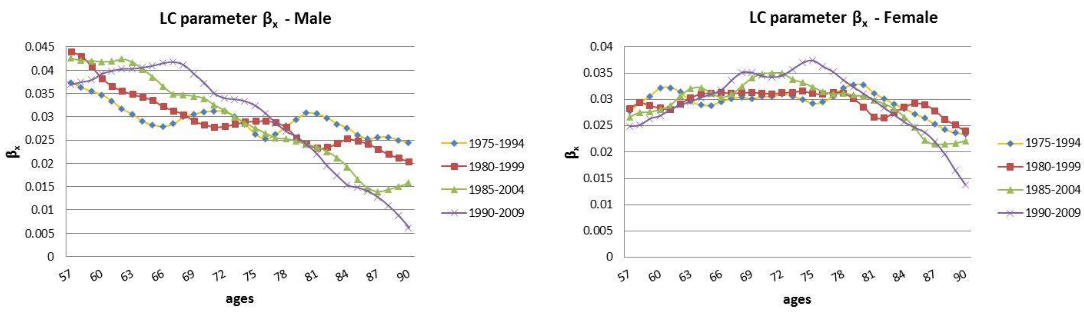
4.2 Jumping Fixed-Length Horizon Backtests

From the results shown in table (4), it is clear that the two models best capture the trend of female mortality. More in detail, the accuracy of the prediction about the next five years, using the periods 1975-1994 and 1990-2009 as database, is far higher than the other two sub-groups of forecast. On the contrary, both models do not show the under-estimating and over-estimating path of the $q_{x,t}^O$ at various ages, which was peculiar in characterizing the result in the previous backtesting case. Only the results of the CBD

model show a similar pattern as already seen, although in this case with overestimates and underestimates staggered by age differently from period to period. This point will be discussed in detail hereafter.
Even though the analyzed models should be assessed on a long-term prediction, however in this case it is particularly noticeable how a change in the starting point of the time series, makes the models to differently incorporate the changes in mortality occurred in the past twenty years. This is generally accomplished through the parameter estimates, which are also reflected throughout the extrapolation process associated with the model. However, an estimation procedure cannot guarantee *a priori* a constant performance of the forecast. This is also due to the fact that, the dynamic of mortality varies in accordance with a multiplicity of social factors that affect the life of every person. Unfortunately, mathematical models are not able to capture such factors.

Now we analyze the immediate effects of these estimates, starting with the LC model (1). The parameters α_x and β_x are time independent age-specific constants, so their estimations will depend on the historical period used as database, and do not need to be predicted. The k_t index captures the time-series common risk factor in that same period, showing the main mortality trend for all ages at time t. Forecast are produced extrapolating the time index k_t, and the mortality projection at each age are all linked together by the product $\beta_x k_t$ (1).
In this backtesting framework, the shift forward of the database shows a continuous decline in mortality provided by the estimates of the parameter α_x and k_t. Moreover, the estimations for the parameter β_x - referred to the male case - show greater values at the beginning of the age range ($57 \leq x \leq 90$) than at the end. This result describes a greater decrease in mortality for those ages with respect to the others, at which β_x presents smaller estimated values [Figure (8), male].

Figure 8: LC Jumping Fixed-Length-Horizon: b_x parameter estimates

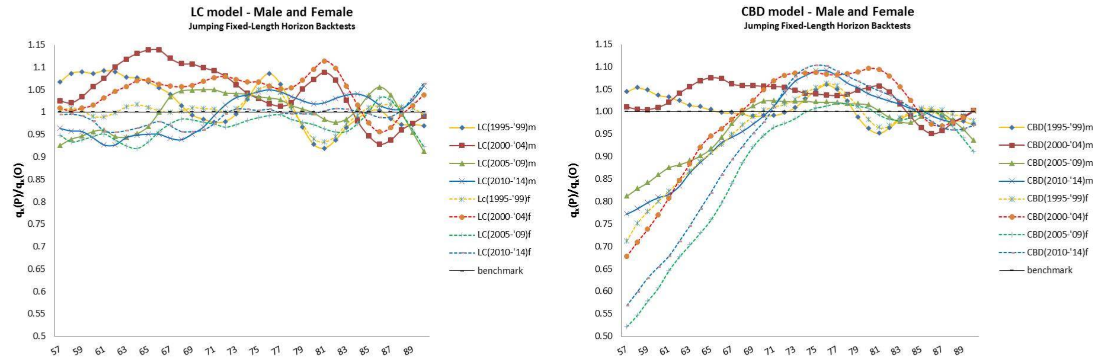


This scenario is in line *ex ante* with the historical experience. However, the forecast referred to the period 2000-2004 show a systematic over-estimation of the $q^O_{x,t}$ for both men and women until the age $x = 80$ [Figure (9), LC model]. Taking into consideration the female case, the estimates of the parameter β_x are more susceptible to changes in the starting point of the time series. This is evident on Figure (8) for the female case.

Needless to say that, the female β_x trend improved the accuracy of the forecast for the periods 1995-1999 and 2010-2014 [Figure (9), LC model].

In the case of CBD model (2), the presence of two time-varying parameters $k_t^{(1)}$ and $k_t^{(2)}$ should increase - at least *a priori* - the forecasting performance with respect to the LC model. This result is evident for the male forecast on the short-run [Table (4)]. As already mentioned, in the CBD model the $k_t^{(1)}$ mortality index represents the level of the mortality curve, after the logit transformation. A reduction in $k_t^{(1)}$ entails a parallel downward shift of the logit-transformed mortality curve, which represents an overall mortality improvement. In particular, this is what occurred in the practice, with greater effects for the female case that are enhanced by the smoothly divergences of $k_t^{(1)}$ trends between sexes. This is clear on the left hand side of Figure (10) below.

Figure 9: LC and CBD $\frac{q_x^P}{q_x^O}$ ratio: comparison between models.



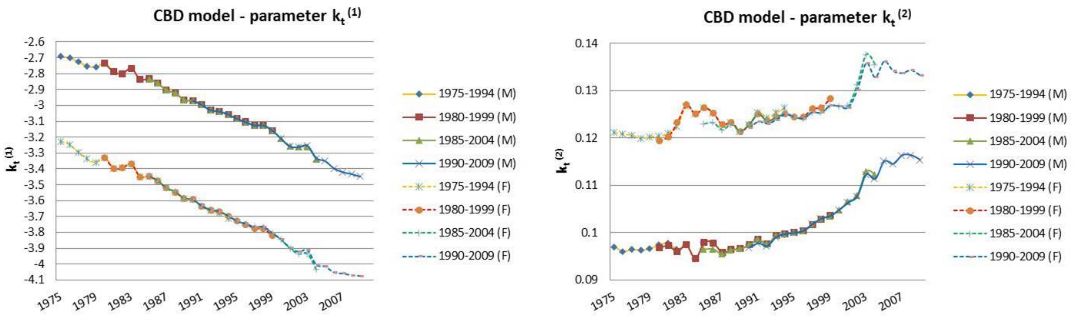
Note: The curves represent the average of the $\frac{q_x^P}{q_x^O}$ ratio over the five-years forecast horizon.

In this case, the jumps of 5-year-ahead do not seem to affect the $k_t^{(1)}$ trend. This is also evidenced by the substantial continuity of the overall reduction in mortality. This is not the case as far as the the $k_t^{(2)}$ mortality index is concerned. Its path drafts the slope of the logit-transformed mortality curve. An increase in $k_t^{(2)}$ entails an increase in the steepness of the logit-transformed mortality curve, which means that mortality at younger ages - i.e. those below the mean age $\bar{x}$ (here $\bar{x} = 73.5$) - improves more rapidly than at older ages. This is clear on the right hand side of the Figure (10) below. Referring to the male case, we find that the speeding of increase in $k_t^{(2)}$ is greater for the periods 1985-2004 and 1990-2009 than for the other two. For this reason, the projected $q_{x,t}^P$ show stronger improvements in mortality for the periods 2005-2009 and 2010-2014 than for the others, particularly for the ages lower that $x = 69$. More in depth, results show an underestimation of the $q_{x,t}^O$ for the ages lower that $x = 69$, and a smooth overestimation path for those higher. Despite the fact that, the growth of $k_t^{(2)}$ between 1980 and 1999 is higher than that of 1975-1994 and that the reduction of $k_t^{(1)}$ is greater, we find that $q_{x,t}^P$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

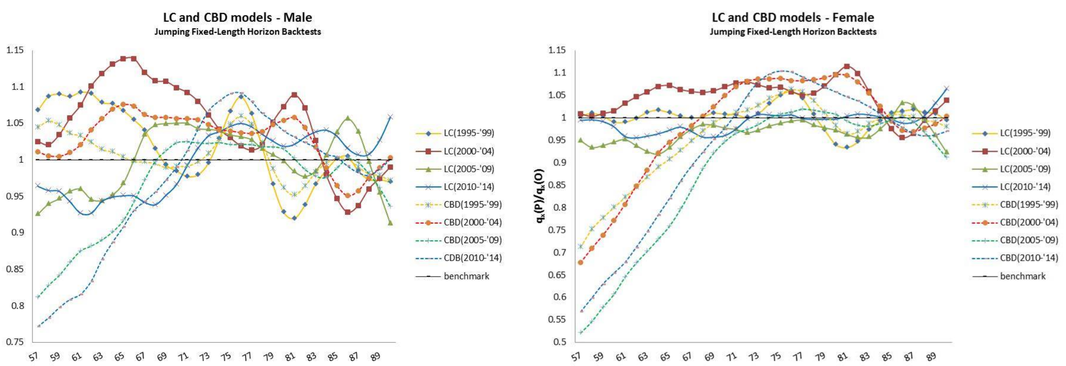
sharply overestimates $q^O_{x,t}$ on the period 1995-1999 and particularly on 2000-2004 for the full range of ages.

Figure 10: CBD Jumping Fixed-Length-Horizon: parameter estimates



As regards to women, $k_t^{(2)}$ presents similar records to men, whereas for the period 1990-2009 the growth rate of $k_t^{(1)}$ is slightly attenuated. In contrast with the male scenario, in this case$q^P_{x,t}$ systematically and significantly underestimates $q^O_{x,t}$ from the age $x = 57$, converging gradually to the observed data as x moves towards $\bar{x}$. Moreover, underestimations become larger as the "lookback" horizon slides forward for 5 years in 5 years. For the ages higher that $\bar{x}$, the overestimation is usually small as large was the previous underestimation in terms of ages [Figure (9), CBD model].

Figure 11: LC and CBD $\frac{q^P_x}{q^O_x}$ ratio: comparison between models on the same gender.



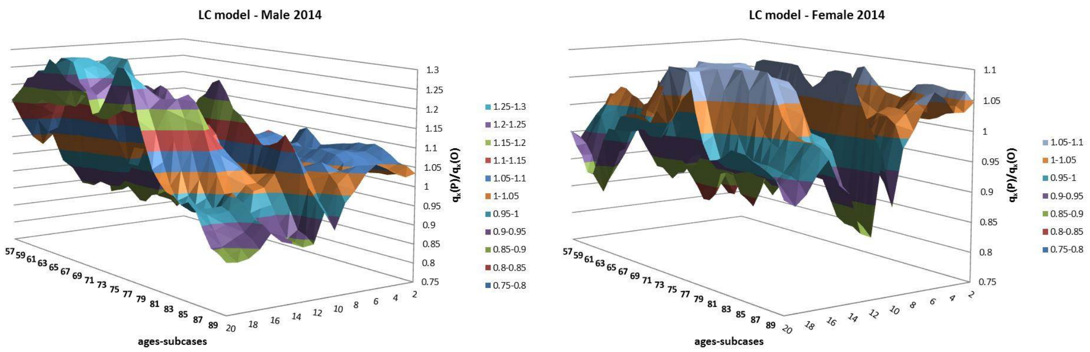
Note: The curves represent the average of the $\frac{q^P_x}{q^O_x}$ ratio over the five-years forecast horizon.

Hence, comparatively, we conclude that a good result for the performance index SSE [Table (4)] can hide some compensation for the forecast error in terms of age and time. Figure (11) above graphically shows the described scenario.

4.3 Rolling Fixed-Length Horizon Backtests

Finally, the analysis concludes with the study of the forecast convergence to the observed $q^O_{x,t}$ in the year[7] 2014. For this reason, we build a framework of nineteen groups of estimations and projection, rolling the database (fixed-length of 20 years) sequentially forward from[8] 1975 to 1993. Then, we compare the 2014 forecast obtained in each group with the realized mortality for that year. We observe that the comparison enhance the same critical issues analyzed in the previous paragraphs, with particular emphasis for two main aspects.
Firstly, scrolling the database over time year by year, gives rise to strong fluctuations in the performance of the prediction measured by the ratio of $q^P_{x,t}$ and $q^O_{x,t}$. These oscillations [Figures (12) and (13)] are evident for both sexes in the results of both the LC and the CBD models. Moreover, the trend is interrupted by a deep break in correspondence of the 1985-2004 database. In particular, the previous base (1984-2003) provided a strong overestimation of $q^O_{x,t}$ especially at old ages, the base 1985-2004 data has then reduced the size, the next one (1986-2005) moved closer to $q^O_{x,t}$.

Figure 12: LC Rolling Fixed-Length Horizon Backtests: $\frac{q^P_x}{q^O_x}$ ratio 2014.



It is objectively difficult to give an explanation for these results. However, they can be partially justified recalling that the year 2003 was characterized by a sharp rise in mortality, especially at old ages. Therefore, this historical event may have affected the estimated parameters. However, in the male case both models systematically underestimate $q^O_{x,t}$ when the age is lower than $x = 73$, and overestimate when is higher. This result is particularly evident when the "lookback" horizon is 1985-2004, and also for the following cases. In particular, CBD underestimates already when the database is referred to the period 1981-2000. However when for the period 1985-2004, the divergence becomes greater when compared to the LC model [Figures (12) and (13)]. As it is shown in this case, the choice of the database play a crucial role in forecasting mortality.

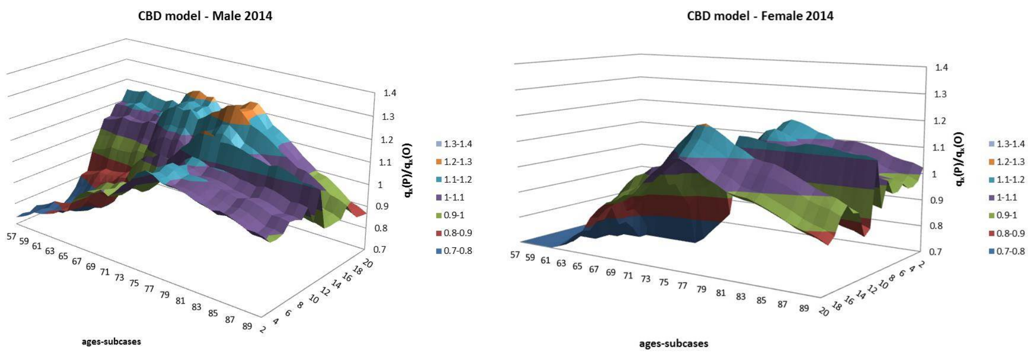
[7] The choice for the year 2014 was motivated by the observed regular mortality path. The 2015 mortality trend is expected to be increased, particularly at old ages [Istat, 2016].

[8] These represent the initial years of the 20-years-length database; i.e. 1975 refers to the estimation period 1975-1994, and so on so forth.

 doi:10.20944/preprints201701.0065.v1

Figure (12) shows the ratio between the projected and the observed death probabilities for the year 2014. Referring to the Table (3), the graph shows the projections obtained for that year on each pair of "Lookback" and "Lookforward" horizon.

Figure 13: CBD Rolling Fixed-Length Horizon Backtests: $\frac{q_x^P}{q_x^O}$ ratio 2014.



In particular, the-sub case index of the graph shows the position of that year on the projection horizon (i.e. 20 means that the year 2014 were the 20th year of projection, 19 means the 19th, and so on so forth.). Since the dataset is rolling over time decreasing the projection horizon, we decided to show the position of the year 2014 in order to take into account both the specific sub-case and the related length of the forecast horizon. Figure (13) shows the same for the CBD model with an inverted order sub-cases for males so to better show the shape of each curve.

Secondly, we detected substantial differences between the performances of the two models by analyzing female mortality. Figure (13) evidences how CBD model systematically underestimates real mortality until the age of 75 and then starts converging to $q_{x,t}^O$ after that "threshold" age. This result, which was already evident in the previous paragraphs, is likely to be linked to the combined effects on the CBD model (2) of the role of the mean age $\bar{x}$ (in our case $\bar{x} = 73.5$) of the age group, for which the forecast is made, and of the observed female mortality pattern. This results are also confirmed from the analysis of the confidence interval referred to the forecast. Figure (14) shows that in the female case at age 65 ($t = 2014$), $q_{x,t}^O$ is always outside the confidence interval, while at age 85 it is inside with central values almost converged to the real data in each sub-case. In the case of LC model, the initial underestimation of the $q_{x,t}^O$ is much less pronounced with respect to the previous case. Moreover, the "threshold" age with respect to which the forecast before underestimates and then overestimates $q_{x,t}^O$, increases as the database moves forward [Figure (12)].
Figure (14) shows the convergence of the projections to the obeserved data for the year 2014, at ages 65 and 85. The x-axis shows the position of the year on the forecast horizon as before. Figure (15) presents the same for the Lee-Carter model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

Figure 14: CBD Rolling Fixed-Length Horizon Backtests: convergence to real data (2014)

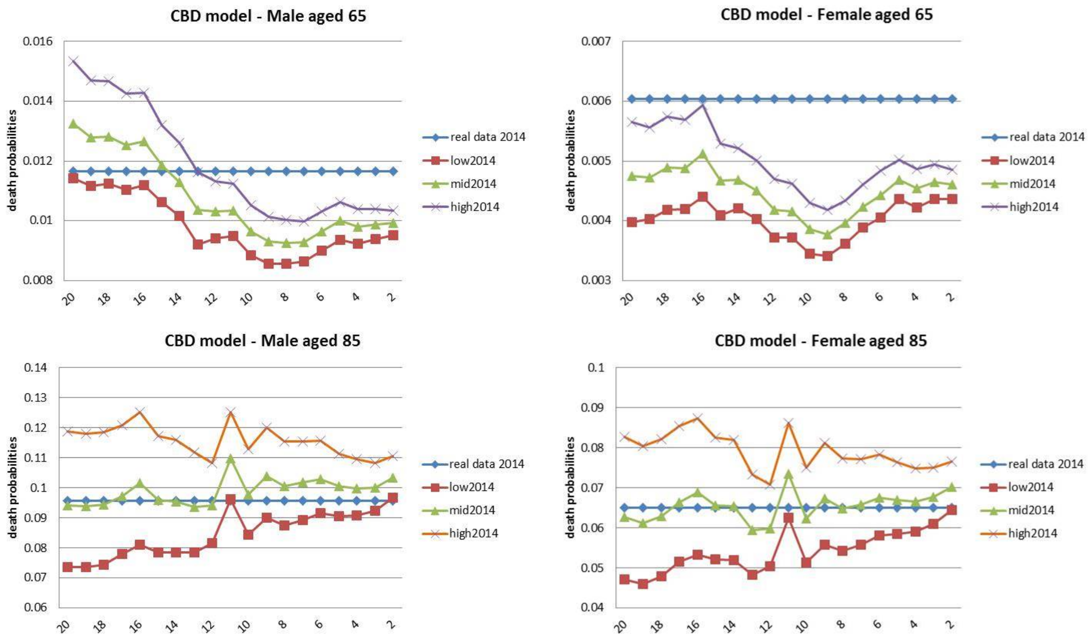
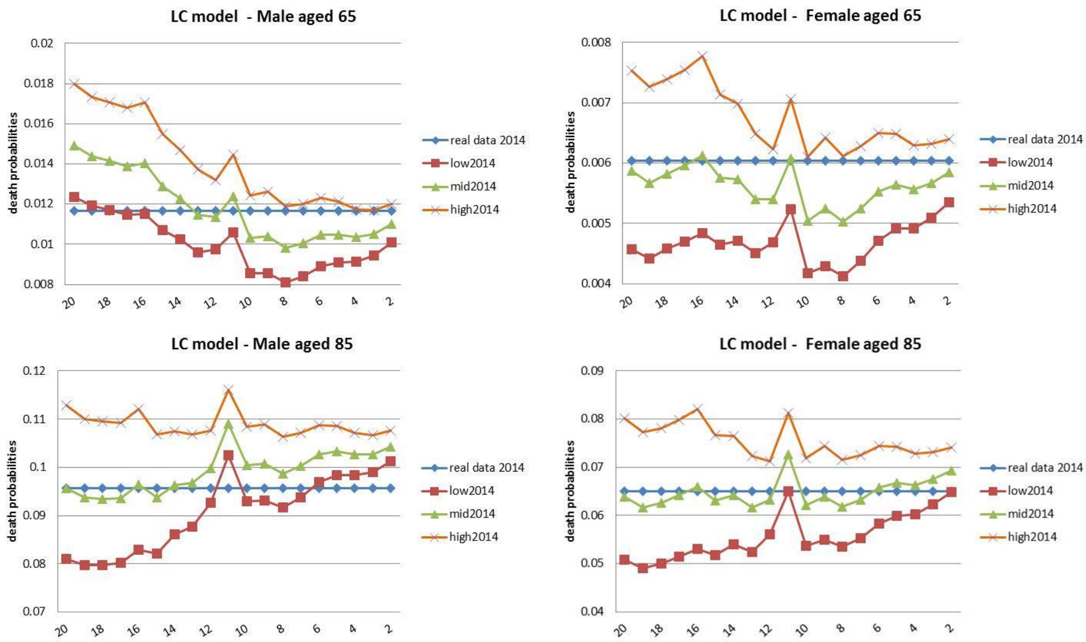


Figure 15: LC Rolling Fixed-Length Horizon Backtests: convergence to real data (2014)



5 Conclusions

The main focuses of this paper are to scrutinize the forecast for both sexes proposed by the models, and to analyze the long-term forecast with respect to short-term, observing qualitative differences in the estimation of the parameter accordingly to changes in the starting point of the database.
As regards to the former, we find that basically both models were not able to capture the shock - in terms of improvements - on the male mortality trend, with greater biases for the ages lower than $x = 75$, which were those more affected by the improvement. In this sense, CBD forecast for those ages are more biased than LC projections in terms of over-estimations. The limited capacity of the models in predicting male mortality is evident in all the three backtesting frameworks, table (4) numerically summarizes the difference - in terms of performances - between sexes for the first two backtesting approaches. Also the analysis of the forecast for the year 2014 that we provided with the third approach, confirms this result. Moreover, women's forecast are widely more accurate than men with small biases observed both in the short and in the long-term. However, in the female case CBD projections showed particularly deep and systematic under-estimations with respect to the ages lower than 75.
From the comparison between the short-term and the long-term forecast, we find that changes in the starting point of the database widely affect the estimation of the LC parameters, particularly for β_x with observable impacts on the projections. The female forecast are more influenced by those changes in β_x. The CBD model satisfy the "new-data-invariant" property for the estimation of the parameter $k_t^{(1)}$, while $k_t^{(2)}$ presents persistent changes for the same year as the dataset slides forward. This aspect is more evident on males than on females. In particular, the adjustment of the parameter $k_t^{(2)}$ (i.e. $x - \bar{x}$) affects mortality forecast with weights of opposite sign at the extremes of the considered age range. The weight is greater as large is the age range. This structural characteristic of the model - albeit simultaneously with $k_t^{(1)}$ - results in a systematic under-estimation of the $q_{x,t}^O$ for the ages lower that $\bar{x}$ that gradually decreases as x moves towards $\bar{x}$. Moreover, mortality forecast around $\bar{x}$ are almost explained exclusively by $k_t^{(1)}$, since $(x - \bar{x})$ is really close to 0 in that case. On contrary, as x gets closer to the upper limit of the age range, the weight of $(x - \bar{x})$ on mortality forecast changes with opposite sign with resulting over-estimation of the $q_{x,t}^O$. For these reasons, the risk in terms of application of the models is conspicuous, because it could potentially affect both the mortality risk and the longevity risk. Taking in consideration the variability of both the parameters β_x (LC) and of $k_t^{(2)}$ (CBD), it is difficult to judge *a priori* what of this two rigidities penalizes more the mortality forecast.

As far as CBD model is concerned, we find that projections are not reliable for describing mortality at ages before $x = 75$. For this reason, LC projections are preferable for describing Italian mortality in this particular framework of years and ages.

Finally, we would like to make clear that we examined the models in their original form, so we can not rule out the possibility that some extensions of the models might resolve the evidenced issues. Moreover, we should precise that we analyzed the performance of the models on this particular dataset, so we make clear that the work has no presumption of generality in judging the models.

Acknowledgments

The authors want to thank CeRP (Center for Research on Pensions and welfare Policies) for supporting this study. It is one of the contributes by CeRP to the hackUniTO research project about aging, promoted by the University of Turin.

6 Appendix A
Lee-Carter Estimation and Projection

Parameter Estimations

The parameter estimation was computed with respect to the *Ordinary Least Square* (OLS) estimation method in accordance to the original approach suggested by the authors.

The following constraints were used in order to find a unique solution for the parameters:

$$\sum_{x=x_1}^{x_m} \beta_x = 1 \qquad \textit{and} \qquad \sum_{t=t_1}^{t_n} k_t = 0 \tag{6}$$

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, as first order condition we get:

$$\hat{\alpha}_x = \frac{1}{t_n - t_1 + 1} \sum_t \ln m_x(t) \tag{7}$$

where the denominator simply represents the number of years considered in the dataset, and $x = x_1, ..., x_m$ is the considered range of ages.

As it is expressed by the equation (7), 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 A of elements $(\ln[m_{x_i}(t_j)] - \alpha_{x_i})$, with i as age index and j as time index (years considered in the data).

 doi:10.20944/preprints201701.0065.v1

At this point, the estimated parameters were recalibrated so that the differences between the actual and the estimated total deaths in each year were zero. This imply that the recalibrated $\hat{k}_t^*$ solve the equation[9]:

$$\sum_x d_{x,t} = \sum_x e^{(\hat{\alpha}_x^* + \hat{\beta}_x^* \hat{k}_t^*)} L_{x,t} \tag{8}$$

Finally, the estimated parameters were adjusted so to satisfy the constraint at (6) for the parameter $\hat{k}_t^*$. Then:

$$a_x^* = \hat{\alpha}_x + \hat{\beta}_x \bar{k} \tag{9}$$

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

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

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 $\left(\sum_{j=1}^{x_m - x_1 + 1} \hat{\beta}_{1j} \right)$ is simply the sum of all the estimated $\hat{\beta}$, which sum to 1. The fitted model is then used to estimate the median and the 95% prediction interval.

Parameter Projection

We projected the estimated parameters k_t^* of Lee-Carter model using a *Random Walk with Drift* equation:

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

where the drift d is estimated by the formula:

$$\hat{d} = \frac{(k_2^* - k_1^*) + (k_3^* - k_2^*) + \dots + (k_T^* - k_{T-1}^*)}{t_n - t_1} = \frac{(k_T^* - k_1^*)}{t_n - t_1}$$

with k_T^* and k_1^* respectively given by the first and the last elements of the vector $k_t^* = [k_1^*, \dots, k_T^*]$. The drift is simply the arithmetic mean of the differenced series of estimated parameters.

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

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

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

[9]The equation (8) has no explicit solution so it has to be solved numerically.

doi:10.20944/preprints201701.0065.v1

$$\hat{m}_{x,T+\Delta t} = e^{a_x^* + b_x^* \hat{k}_{T+\Delta t}}$$

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

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

7 Appendix B
Cairns-Blake-Dowd Estimation and Projection

According to the original formulation of the model proposed by Cairns et al. [2006] the equation (2) is the result of the logit transformation of the following model equation:

$$q_{x,t} = \frac{e^{k_t^{(1)} + k_t^{(2)}(x-\bar{x})}}{1 + e^{k_t^{(1)} + k_t^{(2)}(x-\bar{x})}}. \tag{13}$$

Fitted values for the stochastic processes $k_t^{(1)}$ and $k_t^{(2)}$ were obtained using least squares applied to (13). The fitted model is then used to estimate the median and the 95% prediction interval.

The parameters vector $\vec{k}_t = \left[k_t^{(1)}, k_t^{(2)}\right]'$ has been projected by considering the following equation of a two-dimensional random walk with drift:

$$\vec{k}_{t+1} = \vec{k}_t + \mu + CN(t+1) \tag{14}$$

where:

- μ is a constant 2 x 1 vector of drifts, computed as the arithmetic mean of the differenced series of estimated parameters;
- C is a constant 2 x 2 upper triangular matrix, derived by the unique Cholesky decomposition of the variance-covariance matrix $V = CC'$ of the parameters vector $\vec{k}_{t+1}$;
- $N(t+1)$ is a two-dimensional standard normal random variable.

The adopted forecast method treat the estimated parameters as if they were the true parameter values (parameters certainty). In particular, the presented projections were computed considering parameter certainty based on 5,000 simulation trials.

References

Paola Biffi and Gian Paolo Clemente. Selecting stochastic mortality models for the italian population. *Decisions in Economics and Finance*, 37(2):255–286, 2014.

Heather Booth, Rob Hyndman, Leonie Tickle, Piet De Jong, et al. Lee-carter mortality forecasting: A multi-country comparison of variants and extensions. 2006.

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.

Wai-Sum Chan, Johnny Siu-Hang Li, and Jackie Li. The cbd mortality indexes: Modeling and applications. *North American Actuarial Journal*, 18(1):38–58, 2014.

Valeria D'Amato, Gabriella Piscopo, and Maria Russolillo. The mortality of the italian population: Smoothing techniques on the lee—carter model. *The Annals of Applied Statistics*, pages 705–724, 2011.

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.

Kevin Dowd, Andrew JG Cairns, David Blake, Guy D Coughlan, David Epstein, and Marwa Khalaf-Allah. Backtesting stochastic mortality models: an ex post evaluation of multiperiod-ahead density forecasts. *North American Actuarial Journal*, 14(3):281–298, 2010a.

Kevin Dowd, Andrew JG Cairns, David Blake, Guy D Coughlan, David Epstein, and Marwa Khalaf-Allah. Evaluating the goodness of fit of stochastic mortality models. *Insurance: Mathematics and Economics*, 47(3):255–265, 2010b.

International Monetary Fund. Ch. 4 the financial impact of longevity risk. *Global Financial Stability Report*.

Gerry Hill. The entropy of the survival curve: An alternative measure. *Canadian Studies in Population*, 20(1):43–57, 1993.

Istat. La mortalità in italia nel periodo 1970-1992: evoluzione e geografia. *Istat - Istituto superiore di Sanità*, 1999.

Istat. Tavole di mortalità della popolazione italiana per provincia e regione di residenza. anno 1998. *Informazioni*, 2001.

Istat. Indicatori demografici. stime per l'anno 2015. *Statistiche Report*, 2016.

Nico Keilman. Ex-post errors in official population forecasts in industrialized countries. *Journal of Official Statistics Stockholm*, 13:245–278, 1997.

Nico Keilman. How accurate are the united nations world population projections? *Population and Development Review*, 24:15–41, 1998.

Nathan Keyfitz and Hal Caswell. *Applied mathematical demography*, volume 47. Springer, 2005.

Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

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

Ronald Lee and Timothy Miller. Evaluating the performance of the lee-carter method for forecasting mortality. *Demography*, 38(4):537–549, 2001.

Ronald D Lee. Modeling and forecasting the time series of us fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting*, 9(2):187–202, 1993.

Ronald D Lee. Probabilistic approaches to population forecasting. *Population and Development Review*, 24:156–190, 1998.

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

Han Li and Colin O'Hare. Mortality forecast: Local or global? *Available at SSRN 2612420*, 2015.

Jackie Li. Projections of new zealand mortality using the lee-carter model and its augmented common factor extension. *New Zealand Population Review*, 36:27–53, 2010.

Johnny Siu-Hang Li, Mary R Hardy, and Ken Seng Tan. Uncertainty in mortality forecasting: an extension to the classical lee-carter approach. *Astin Bulletin*, 39(01):137–164, 2009.

Nan Li and Ronald Lee. Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography*, 42(3):575–594, 2005.

Stéphane Loisel and Daniel Serant. In the core of longevity risk: hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps. 2007.

Carlo Maccheroni. Diverging tendencies by age in sex differentials in mortality in italy. *South East Journal of Political Science (SEEJPS)*, Vol. II(3):42–58, 2014.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2017 doi:10.20944/preprints201701.0065.v1

George Mavros, A Cairns, Torsten Kleinow, and George Streftaris. A parsimonious approach to stochastic mortality modelling with dependent residuals. Technical report, Citeseer, 2014.

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.

A Roger Thatcher, Väinö Kannisto, and James W Vaupel. The force of mortality at ages 80 to 120. 1998.

James W Vaupel. Biodemography of human ageing. *Nature*, 464(7288):536–542, 2010.

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.