

Non mean reverting affine processes for stochastic mortality

Elisa Luciano, University of Turin and ICER
Elena Vigna, University of Turin

The mathematical tools

Survival time can be modelled as a “doubly stochastic stopping time”. Preliminaries:

1) A counting process N_t is said to admit the stochastic intensity λ (let λ be a nonnegative predictable process s.t. $E(\int_0^t \lambda_u du) < \infty$) if $M_t = N_t - \int_0^t \lambda(u) du$ is a martingale.

$$2) \Rightarrow E(N_{t+\Delta t} - N_t | \mathcal{F}_t) = \lambda_t \Delta t + o(\Delta t)$$

3) Conditionally on the path of λ until $s > t$, the process $N_s - N_t$ has Poisson distribution with parameter $\int_t^s \lambda_u du$.

4) \Rightarrow The stopping time of a doubly stochastic process is the analogous of the first jump time of a Poisson process, where the intensity is a stochastic process.

If the survival time τ is doubly stochastic with intensity λ , then:

$$P(\tau > t | \mathcal{F}_s) = E \left[e^{-\int_s^t \lambda(u) du} | \mathcal{F}_s \right] \quad (\star)$$

The affine framework

A process λ_t is affine if it can be described by the SDE :

$$d\lambda_t = \mu(\lambda_t)dt + \sigma(\lambda_t)dB_t + dJ_t$$

where J is a pure jump process and where the drift $\mu(\lambda_t)$, the covariance matrix $\sigma(\lambda_t)\sigma(\lambda_t)'$ and the jump measure associated with J have affine dependence on λ_t .

Examples: Vasicek, CIR.

IMPORTANT RESULT

If λ is affine:

$$E \left[e^{\int_t^T -\lambda(u)du} | \mathcal{G}_t \right] = e^{\alpha(T-t) + \beta(T-t)\lambda(t)} \quad (**)$$

where the coefficients $\alpha(\cdot)$ and $\beta(\cdot)$ satisfy generalized Riccati ODEs.

The actuarial application

We consider an individual aged x and model her random future lifetime T_x as a doubly stochastic stopping time with intensity λ_x .

According to (\star) the survival probability is:

$$S_x(t) = P(T_x > t | \mathcal{G}_0) = E \left[e^{-\int_0^t \lambda_x(u) du} | \mathcal{G}_0 \right]$$

Previous (recent) literature on this: Biffis (2004), Dahl (2004), Shrager (2004).

How do we choose the process λ_x ?

First application: mean reverting processes

In the credit risk literature, mean reverting processes work quite well to describe the intensity of default (Duffie and Singleton, 2003):

1. CIR process:

$$d\lambda_x(t) = k(\gamma - \lambda_x(t))dt + \sigma\sqrt{\lambda_x(t)}dW(t)$$

2. mean reverting with jumps (m.r.j.):

$$d\lambda_x(t) = k(\gamma - \lambda_x(t))dt + dJ(t)$$

3. VASICEK process:

$$d\lambda_x(t) = k(\gamma - \lambda_x(t))dt + \sigma dW(t)$$

with $k > 0, \gamma > 0, \sigma > 0, W(t)$ standard Brownian motion, $J(t)$ compound Poisson process with intensity l and jumps exponentially distributed with expected value μ .

Survival function

These processes are affine and we can express the survival function in closed form:

$$P(T_x > t | \mathcal{F}_0) = S_x(t) = e^{\alpha(t) + \beta(t)\lambda_x(0)}$$

where $\alpha()$ and $\beta()$ are known functions of the parameters k, γ, σ, l and μ .

The calibration to the UK population

Since $\lambda_x(t)$ is the intensity of mortality at age $x + t$ of an individual aged x at time 0, we choose a generation mortality table and not a contemporaries one.

Two observed generation tables (1880 and 1900, HMD data) and two projected mortality tables (1935, 1945).

Assumptions for the calibration:

- initial age $x = 65$, for both males and females
- jump size $\mu < 0$ (to capture improvements in mortality rates)
- $\lambda_{65}(0) = -\ln(p_{65})$

We minimize the sum of the squared differences between the survival probabilities of the relevant table and the ones implied by the model, and compute the calibration error.

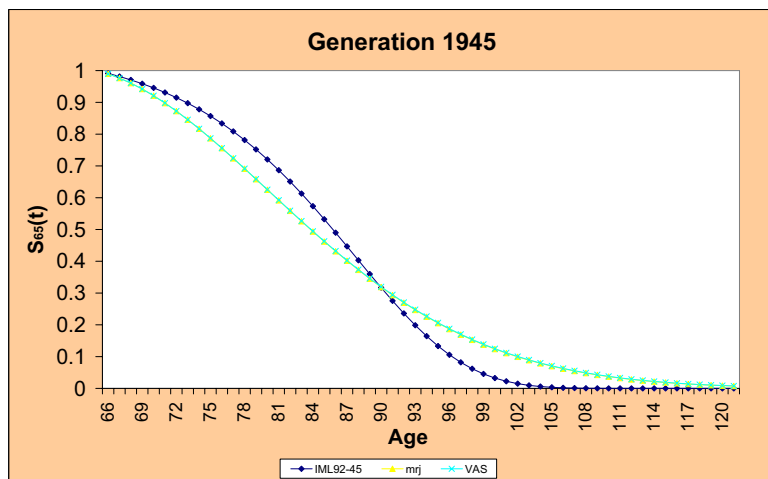
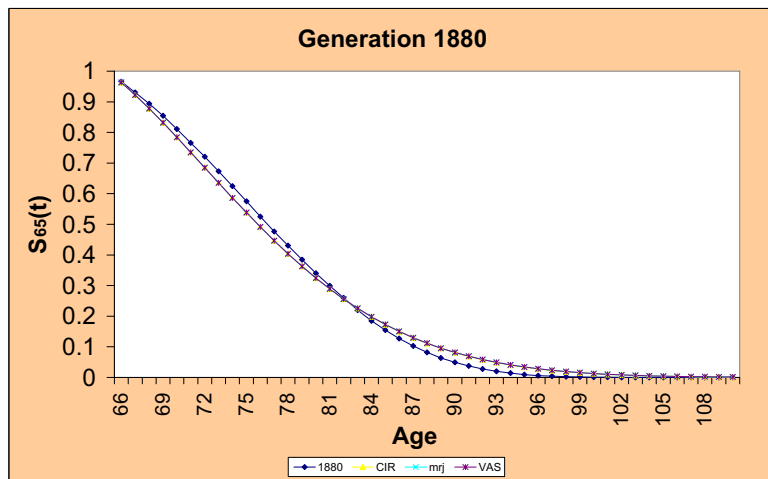
Results from the calibration: value of the optimal parameters

TABLE 1

| | 1880 | 1900 | 1935 | 1945 |
|-------------------|----------|----------|----------|----------|
| $\lambda_{65}(0)$ | 0.03515 | 0.03797 | 0.01145 | 0.00885 |
| CIR-error | 0.02182 | 0.01662 | 0.40945 | 0.20552 |
| CIR-k | 0.00448 | 0.01365 | 0.06494 | 0.0078 |
| CIR- σ | 0.00103 | 0.00298 | 0.00005 | 0 |
| CIR- γ | 1.24656 | 0.4301 | 0.07552 | 0.41711 |
| mrj-error | 0.02236 | 0.01327 | 0.15816 | 0.1965 |
| mrj-k | 0.00571 | 0.00392 | 0.005 | 0.00465 |
| mrj- μ | -0.00246 | -0.00227 | -0.00249 | -0.00492 |
| mrj-l | 0.00247 | 0.00234 | 0.00249 | 0.0099 |
| mrj- γ | 0.99382 | 1.31818 | 0.64908 | 0.67935 |
| VAS-error | 0.02247 | 0.01473 | 0.16191 | 0.1982 |
| VAS- σ | 0.00046 | 0.00048 | 0.00002 | 0.00002 |
| VAS-k | 0.00591 | 0.00835 | 0.00604 | 0.00526 |
| VAS- γ | 0.96029 | 0.65393 | 0.53278 | 0.59302 |

Result: the error more than decuplicates when passing from old observed tables to the projected tables for younger generations.

Results from the calibration: survival functions



Evidence:

- fit more satisfactory for the old generations
- for the younger generations, the rectangularization phenomenon is not captured
- the expansion feature is also not captured
- the survival probability at old ages is much higher and at lower ages much lower than in the tables
- the survival probability at very old ages (like 130-140) is positive

QUESTION: is the bad fit due to the common feature of mean reversion of the three models (see also Cairns, Blake and Dowd, 2004)?

OBSERVATION: the force of mortality shows no mean reversion, but rather an exponential increase

SIMPLE IDEA: why not dropping the mean reversion term and calibrate a process whose non stochastic part increases exponentially?

Second application: non mean reverting processes

We propose four models:

1. Ornstein Uhlenbeck process without jumps (OU):

$$d\lambda(t) = a\lambda(t)dt + \sigma dW(t)$$

2. Ornstein Uhlenbeck process with jumps (OUj):

$$d\lambda(t) = a\lambda(t)dt + \sigma dW(t) + dJ(t)$$

3. Feller process without jumps (FEL):

$$d\lambda(t) = a\lambda(t)dt + \sigma\sqrt{\lambda(t)}dW(t)$$

4. Feller process with jumps (FELj):

$$d\lambda(t) = a\lambda(t)dt + \sigma\sqrt{\lambda(t)}dW(t) + dJ(t)$$

with $a > 0$, $\sigma \geq 0$ and J a pure compound Poisson jump process, with Poisson arrival times of intensity $l > 0$ and exponentially distributed jump sizes with mean $\mu < 0$.

⇒ **Survival functions still in closed form**

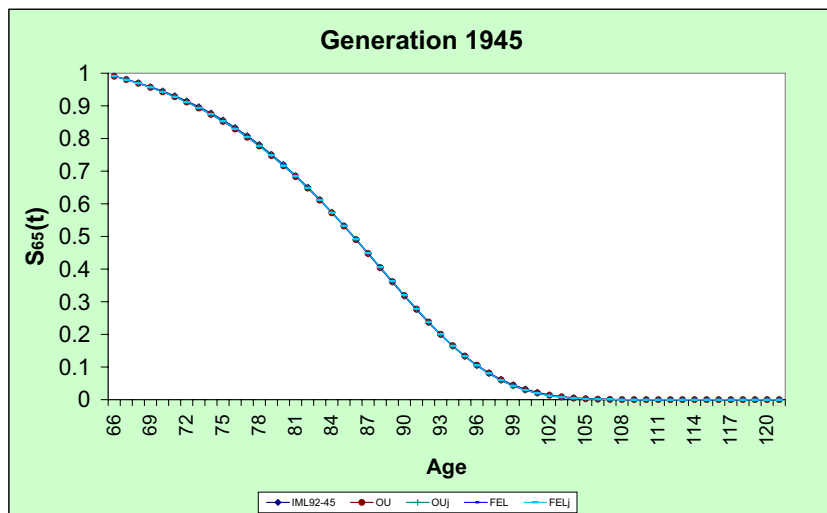
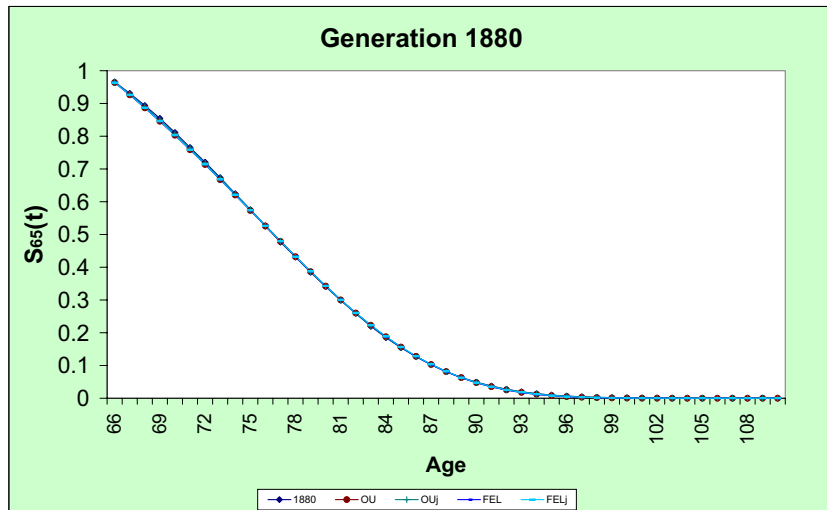
Calibration of non mean reverting processes

TABLE 2

| | 1880 | 1900 | 1935 | 1945 |
|-------------------|----------|----------|----------|----------|
| $\lambda_{65}(0)$ | 0.03515 | 0.03797 | 0.01145 | 0.00885 |
| OU-error | 0.00043 | 0.00012 | 0.00085 | 0.00027 |
| OU-a | 0.0861 | 0.07949 | 0.09856 | 0.10859 |
| OU- σ | 0.00183 | 0.00341 | 0.0001 | 0.00048 |
| OUj-error | 0.0001 | 0.00004 | 0.00002 | 0.00016 |
| OUj-a | 0.09101 | 0.08192 | 0.10014 | 0.10865 |
| OUj- σ | 0.00377 | 0.00414 | 0.0001 | 0.00011 |
| OUj-l | 0.00173 | 0.00088 | 0.00105 | 0.00036 |
| OUj- μ | -0.00003 | -0.00003 | -0.00003 | -0.00003 |
| FEL-error | 0.00044 | 0.00012 | 0.00084 | 0.00027 |
| FEL-a | 0.08553 | 0.07896 | 0.09867 | 0.10811 |
| FEL- σ | 0.00431 | 0.01348 | 0.00005 | 0.0001 |
| FELj-error | 0.00043 | 0.00012 | 0.00053 | 0.00027 |
| FELj-a | 0.0858 | 0.07897 | 0.10164 | 0.10811 |
| FELj- σ | 0.00735 | 0.01349 | 0 | 0.00001 |
| FELj-l | 0.001 | 0.001 | 0.1856 | 0.001 |
| FELj- μ | -0.0001 | -0.0001 | -0.00034 | -0.0001 |

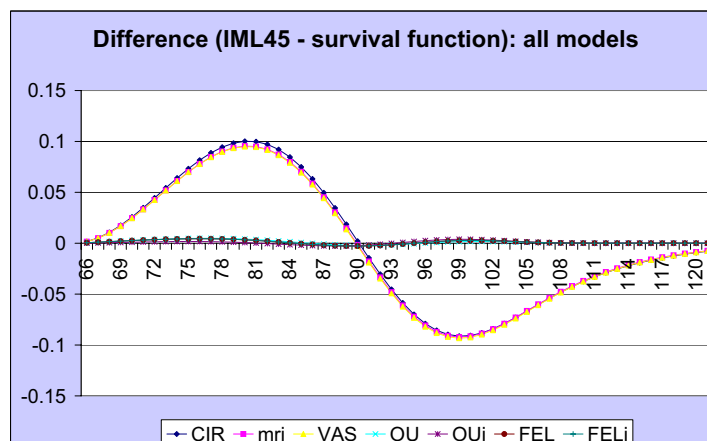
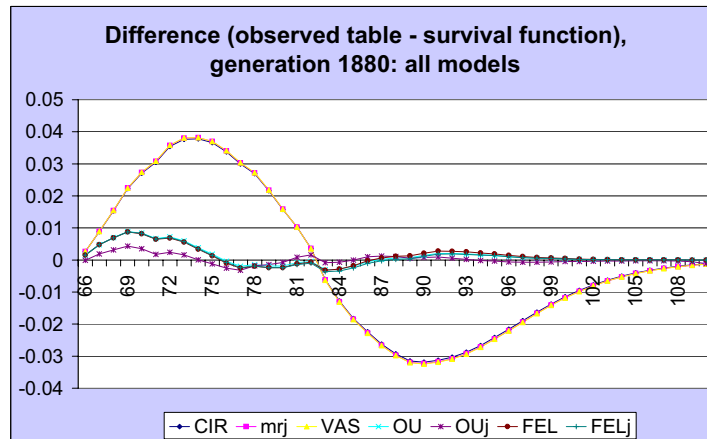
- The calibration errors are very small.
- The OUj model dominates the others. The models with jumps perform better than the corresponding models without.
- The value of σ is very low.

Results from the calibration: survival functions



The fit is remarkable, in all cases.

Differences between $S_x(t)$ and ${}_t p_x$



⇒ significant improvement in the fit when dropping the mean reversion term

⇒ **non mean reverting affine processes seem appropriate to describe the intensity of mortality**

Sensitivity analysis

Evidence from the calibration of the non mean reverting processes:

- low or null diffusion parameter (σ)
- improvements of fit when adding a jump component

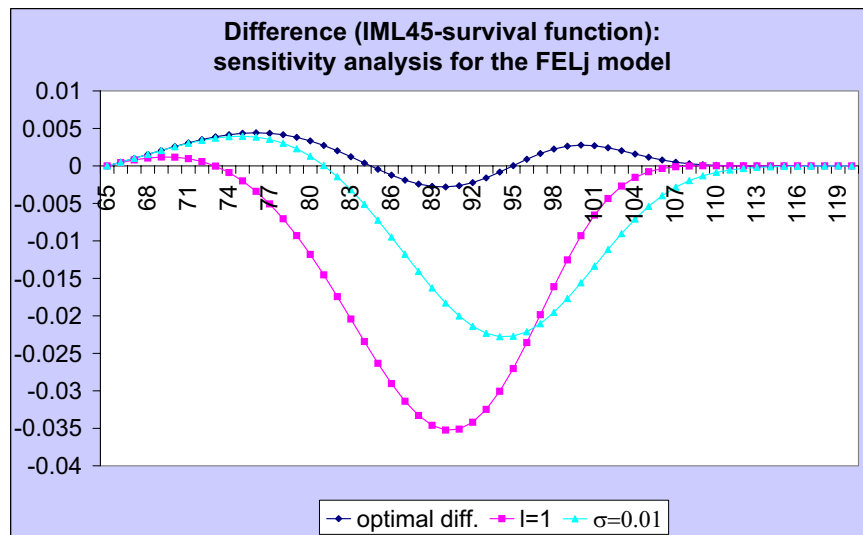
This does not need to be true for the future. So, what would be the effect of higher variability in the intensity λ on the survival probabilities?

Assessing the effect of higher variability

For the processes OU, OUj and FEL, we can answer this question with analytical results: when we increase the diffusion coefficient or the jump intensity (the latter meaning a reduction in the expected arrival time of jumps), these models predict a higher survivorship.

For the model FELj analytical results cannot be obtained and we provide a sensitivity or stress test analysis. Let us consider the differences between the survival probabilities of the table and those implied by the model

- under the optimal parameter values (optimal diff.)
- with a diffusion coefficient σ and an intensity l equal to a thousand times the optimal ones.



Increasing the diffusion coefficient σ or the jump intensity l , leads to differences becoming more negative.



Increasing the stochastic part of the intensity process implies improvement in the survival probability

The link with existing models for the force of mortality

What is the relation between the stochastic intensity of mortality and the deterministic force of mortality?

$$\mu_x = \lim_{h \rightarrow 0} \frac{P(x < T_0 \leq x + h | T_0 > x)}{h}$$

In our case, we have:

$$\mu_x = \lim_{h \rightarrow 0} \frac{1}{h} \left(1 - \frac{S(x+h)}{S(x)} \right) = -\alpha'(x) - \lambda_0(0)\beta'(x)$$

For example, in the OU model the force of mortality, becomes:

$$\mu_x = \lambda_0(0)e^{ax} - \frac{\sigma^2}{2a^2}(e^{ax} - 1)^2$$

If $\sigma = 0$ we have:

$$\mu_x = \lambda_0(0)e^{ax} = \lambda_0(x)$$

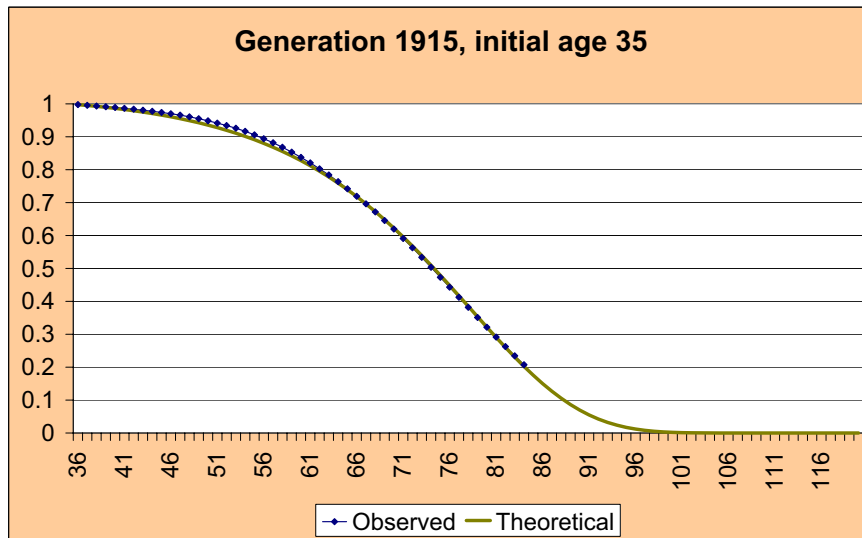
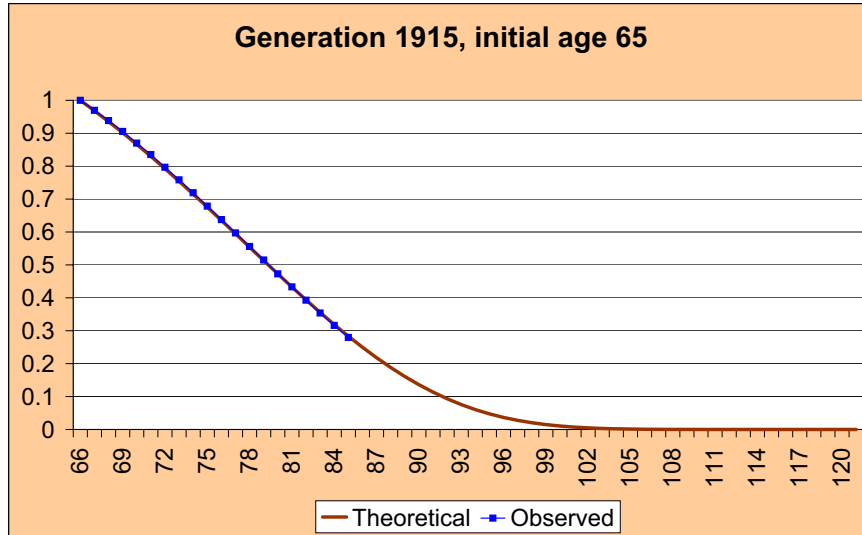
⇒ Gompertz type.

Also for the other three models, if the stochastic part is null ($\sigma = 0, l = 0$), we have the Gompertz model.

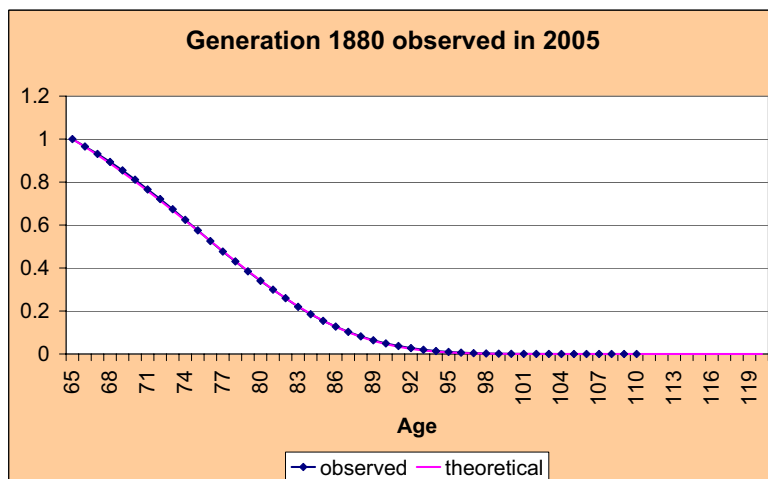
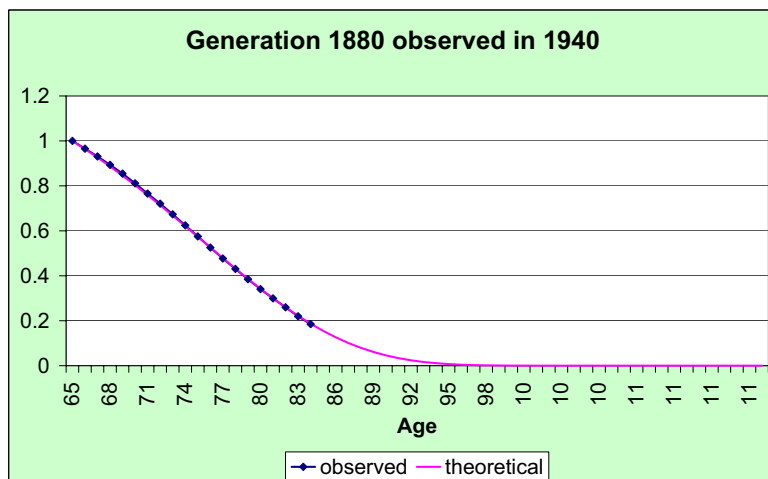
Forecasting mortality

One can use this model to see what is the future evolution of mortality for a given generation.

In the next two graphs, we report the mortality forecast for the generation 1915 for initial ages 35 and 65 (FELj model). The right tail of the “Theoretical” curve gives the forecast of the survival function beyond the observation date.



We have applied the same forecast procedure on the generation 1880, initial age 65, in order to compare the forecasted mortality with the experienced one.



Mortality trend

Let us consider the intensity of mortality for a given initial age x and different generations. A complete description of the intensity surface would be given by a two parameters-family $\lambda_{x,gen}$ (Biffis and Millosovich, 2005).

Here we focus only on the change of generation and omit the initial age x . We have a family of intensity processes:

$$d\lambda_{gen}(t) = f_{gen}(\lambda_{gen}(t))dt + g_{gen}(\lambda_{gen}(t))dW(t) + dJ_{gen}(t)$$

where the index gen refers to the year of birth (eg 1880, 1905).

The change in $\lambda_{gen}(0)$ and in the parameters that characterize f_{gen} and g_{gen} gives the description of the mortality trend in our setting.

Investigate the mortality trend

- calibration for the sixteen generations born in years 1900 to 1915
- initial age $x = 65$
- FELj model
- for each generation we calculate the value of $\lambda_{gen}(0)$ and find a set of optimal parameters:

$$a_{gen}, \sigma_{gen}, l_{gen}, \mu_{gen}, error_{gen}$$

Results

- decreasing trend of a and linearly decreasing trend of $\lambda_{65}(0)$ (R^2 of 0.912 wrt to calendar year)
- the calibration errors are very small.

TABLE 3

| <i>gen</i> | <i>a</i> | σ | <i>l</i> | μ | $\lambda_{65}(0)$ | <i>error</i> |
|------------|-----------|-----------|-----------|------------|-------------------|--------------|
| 1900 | 0.079247 | 0.0143406 | 0.0009996 | -0.0001999 | 0.0379719 | 0.0001177 |
| 1901 | 0.0790948 | 0.0150062 | 0.0009994 | -0.0001999 | 0.0374008 | 0.0000847 |
| 1902 | 0.0841481 | 0.0196919 | 0.0009989 | -0.0001998 | 0.0352956 | 0.0000298 |
| 1903 | 0.0781443 | 0.0140747 | 0.0009995 | -0.0001999 | 0.0362388 | 0.0000619 |
| 1904 | 0.0702891 | 0.0000099 | 0.0074853 | -0.0014807 | 0.0375461 | 0.0007773 |
| 1905 | 0.0726454 | 0.0001 | 0.001 | -0.0002 | 0.0362284 | 0.0001925 |
| 1906 | 0.0760964 | 0.0133211 | 0.001 | -0.0002 | 0.0348192 | 0.000033 |
| 1907 | 0.0714555 | 0.0000995 | 0.0010331 | -0.0002066 | 0.035192 | 0.000249 |
| 1908 | 0.0733996 | 0.0025116 | 0.001 | -0.0002 | 0.0341464 | 0.0001419 |
| 1909 | 0.0741234 | 0.0025277 | 0.001 | -0.0002 | 0.033443 | 0.0000641 |
| 1910 | 0.073784 | 0.00001 | 0.0046932 | -0.0009309 | 0.0331122 | 0.0001005 |
| 1911 | 0.0720482 | 0.0000099 | 0.0066647 | -0.0013194 | 0.0328228 | 0.0001631 |
| 1912 | 0.0756247 | 0.00001 | 0.0028953 | -0.0005766 | 0.0312121 | 0.0000266 |
| 1913 | 0.072663 | 0.0000243 | 0.0066073 | -0.0013085 | 0.0315629 | 0.0001378 |
| 1914 | 0.0717319 | 0.0000221 | 0.0071925 | -0.0014249 | 0.0315423 | 0.0003068 |
| 1915 | 0.0722413 | 0.0000332 | 0.0066523 | -0.0013185 | 0.0309851 | 0.0001773 |

To sum up

- We have described the time of death as a doubly stochastic stopping time (a jump time whose intensity is stochastic). The intensity has been described as an affine process, with mean reversion and without it. For both specifications, the survival probabilities have been provided in closed form.
- The intensity processes have been calibrated to the UK population, using observed mortality tables for old generations and projected tables for younger ones.
- Results seem to suggest that, in spite of their popularity in the financial context, mean reverting processes are not suitable for describing the death intensity of individuals.
- On the contrary, affine processes whose deterministic part increases exponentially seem to be appropriate for describing the intensity of mortality.
- We provide a procedure for mortality forecasting and mortality trend assessment: comparison of forecasted and experienced mortality for old generations gives very satisfactory results.