A note on stochastic survival probabilities and their calibration

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

has become.....

Non mean reverting affine processes for stochastic mortality

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

The mathematical tools: some basic ideas

The default time is often modelled as a "doubly stochastic stopping time".

The definition of doubly stochastic stopping times needs some preliminaries:

1) A counting process N_t is said to admit the stochastic intensity λ (where λ is 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) If a counting process N_t admits the intensity λ, then

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

In other words, the process λ gives information about the average number of jumps of the process under observation in a small period of future time.

3) A counting process with stochastic intensity is *doubly stochastic driven by the subfiltration* $\mathcal{G}_t \subset \mathcal{F}_t$ if, conditionally on the path of λ, the process $N_s - N_t$ has Poisson distribution with parameter $\int_t^s \lambda_u du$.

Stopping time

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 τ is the first jump time of a Poisson process with parameter λ, then τ has exponential distribution and

$$P(\tau > t) = e^{-\lambda t}$$

Similarly, an important result for doubly stochastic stopping times is that if τ is doubly stochastic with intensity λ, then:

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

If the stopping time τ is chosen to be the random time of death of an individual aged x at time s, T_x, then the probability $(\star)$ is the survival probability ${}_tp_x$.

The affine framework

The choice of the intensity process is crucial for the solution of ($\star$).

Classical results from the credit risk literature show that if the process chosen for the intensity is of the affine class, then the expectation ($\star$) turns out to be tractable.

A process λ_t is affine if it is a jump-diffusion process, i.e. 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 of affine processes in finance: Vasicek, CIR.

IMPORTANT RESULT

If λ is affine:

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

where the coefficients $\alpha(\cdot)$ and $\beta(\cdot)$ satisfy generalized Riccati ODEs, that can be solved at least numerically and in some cases analytically (Duffie Pan Singleton, 2000).

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).

The crucial point now becomes: how do we choose the process λ so that to apply the useful equation ($\star\star$)?

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), so we have started with them and have chosen three specifications for λ:

1. CIR process:

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

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 apply ($\star\star$) and express the survival function in closed form (these are standard results):

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

where:

1. CIR process:

$$\alpha(t) = -\frac{2k\gamma}{\sigma^2}\ln\left(\frac{c+de^{bt}}{b}\right) + \frac{k\gamma}{c}t \qquad \beta(t) = \frac{1-e^{bt}}{c+de^{bt}}$$

$$b = -\sqrt{k^2+2\sigma^2} \quad c = 0.5(b-k) \quad d = 0.5(b+k)$$

2. mean reverting with jumps process:

$$\alpha(t) = -\gamma(t+\beta(t)) - l\frac{\mu t - \ln(1-\mu\beta(t))}{\mu+k}$$

$$\beta(t) = \frac{e^{-kt}-1}{k}$$

3. Vasicek process:

$$\alpha(t) = -\frac{(\beta(t)+t)(k^2\gamma - \frac{\sigma^2}{2})}{k^2} - \frac{\sigma^2\beta(t)^2}{4k}$$

$$\beta(t) = \frac{e^{-kt}-1}{k}$$

The calibration

We have calibrated the three survival functions to the Italian population.

We notice that the process $\lambda_x(t)$ when t passes describes the future intensity of mortality for any age $x+t$ of an individual aged x at time 0: therefore the approach adopted here is a "diagonal" one. So, we must choose a generation mortality table and not a contemporaries one. The table selected is the projected mortality table RG48 (for generation born in 1948).

Assumptions for the calibration:

- initial age $x = 65$, for both males and females
- least squares method to the survival probabilities
- jump size $\mu < 0$ (to capture improvements in mortality rates)
- $\lambda_{65}(0) = -\ln(p_{65})$
- calibration error: minimized sum of the squared differences between the survival probabilities of the relevant table and the ones implied by the model

Results from the calibration: value of the optimal parameters

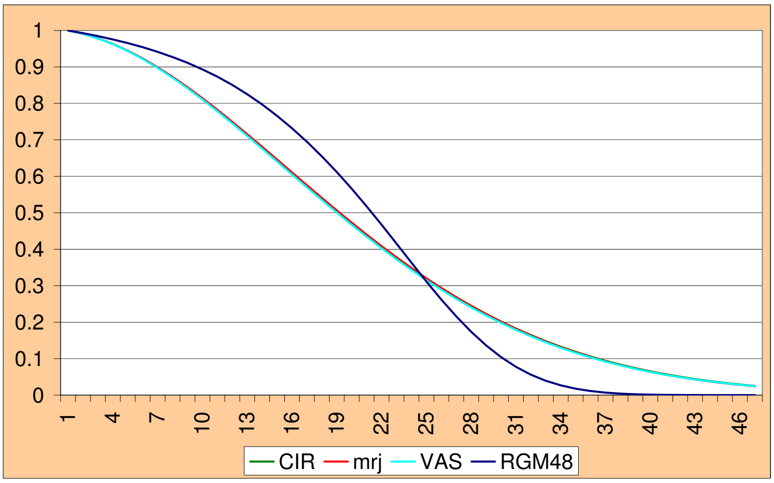
TABLE 1

	RG48M	RG48F
	$\lambda_{65}(0) = 0.007349$	$\lambda_{65}(0) = 0.002442$
CIR-k	0.0080658	0.0062041
CIR-γ	0.4486729	0.44
CIR-σ	0.0000045	0.0000486
CIR-error	0.277	0.698
mrj-k	0.006098	0.00606
mrj-γ	0.591	0.46
mrj-l	0.009895	0.009827
mrj-μ	-0.004916	-0.004862
mrj-error	0.271	0.699
Vasicek-k	0.009301	0.007538
Vasicek-γ	0.4	0.367
Vasicek-σ	0.00007229	0.00004828
Vasicek-error	0.282	0.708

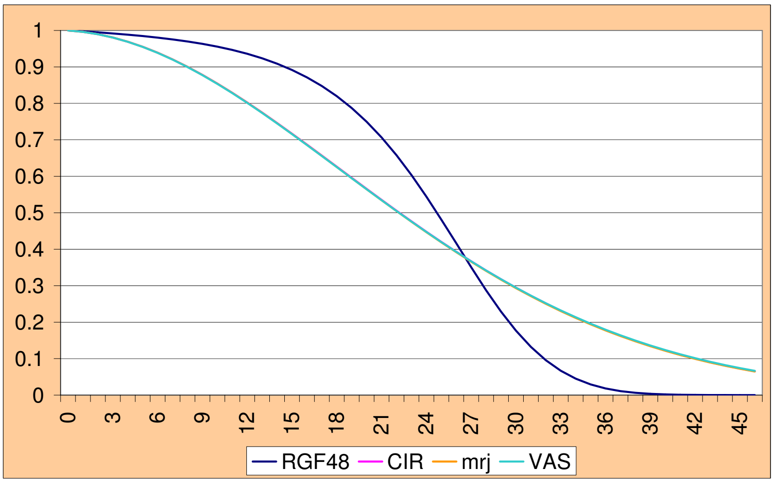
The calibration errors are quite high, and are higher for females than for males.

Results from the calibration: survival functions

MALES



FEMALES



Evidence:

- 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 RG48 tables
- the survival probability at very old ages (like 130-140) is positive

Conclusion: these models give a bad fit of the survival function.

QUESTION: is the bad fit due to the common feature of mean reversion of the three models?

SIMPLE IDEA: why not dropping the mean reversion term and try the calibration with a process that 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):

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

4. Feller process with jumps (FELj):

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

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

Survival function

The survival function in closed form is:

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

where:

1. OU process:

$$\alpha(t) = \frac{\sigma^2}{2a^2}t - \frac{\sigma^2}{a^3}e^{at} + \frac{\sigma^2}{4a^3}e^{2at} + \frac{3\sigma^2}{4a^3}$$

$$\beta(t) = \frac{1}{a}\left(1 - e^{at}\right)$$

Problem: the intensity λ can become negative. The probability of λ becoming negative is:

$$P(\lambda(t) \leq 0) = P\left(N \leq -\frac{\lambda(0)e^{at}}{\sigma\sqrt{\frac{e^{2at}-1}{2a}}}\right)$$

This probability in the applications is negligible.

2. OUj process:

$$\alpha(t) = (\frac{\sigma^2}{2a^2} + \frac{la}{a-\mu})t - \frac{\sigma^2}{a^3}e^{at} + \frac{\sigma^2}{4a^3}e^{2at} + \frac{3\sigma^2}{4a^3} +$$

$$+ \frac{l}{a-\mu}\ln(1 - \frac{\mu}{a} + \frac{\mu}{a}e^{at})$$

$$\beta(t) = \frac{1}{a}\left(1 - e^{at}\right)$$

3. FEL process:

$$\alpha(t) = 0 \qquad \beta(t) = \frac{1 - e^{bt}}{c + de^{bt}}$$

4. FELj process:

$$\alpha(t) = \frac{l\mu}{c-\mu}t - \frac{l\mu(c+d)}{b(d+\mu)(c-\mu)} \cdot$$

$$\cdot [\ln(\mu - c - (d+\mu)e^{bt}) - \ln(-c-d)]$$

$$\beta(t) = \frac{1 - e^{bt}}{c + de^{bt}}$$

The calibration of the non mean reverting processes

The calibration has been done for males and females, in two cases:

- for the Italian population: tables COH48 and RG48

- for the UK population: Standard Tables of Mortality: the "92" Series for immediate annuitants (IML92 and IFL92), generations 1935, 1945 and 1955

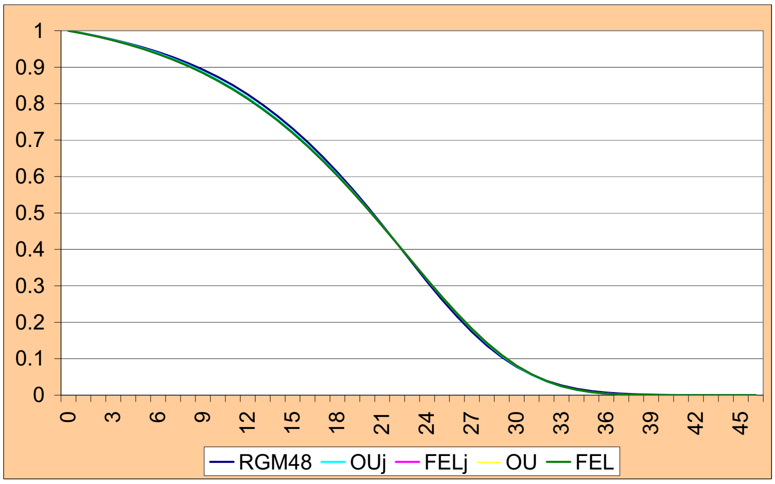
Results from calibration: optimal parameters

	RG48M	RG48F
	$\lambda_{65}(0)$=0.007349	$\lambda_{65}(0)$=0.002442
OU-a	0.126	0.155
OU-sigma	0.00003163	0.000009785
OU-error	0.001987	0.012
OUj-a	0.126843	0.158236
OUj-sigma	0	0
OUj-l	0.000357	0.002336
OUj-mu	-0.000015	-0.000015
OUj-error	0.001147	0.001924
FEL-a	0.1262899	0.1545647
FEL-sigma	0	0
FEL-error	0.00188	0.012
FELj-a	0.126291	0.15458
FELj-sigma	0	0.000203
FELj-l	0.000251	0.001004
FELj-mu	-0.00005	-0.0001
FELj-error	0.00188	0.01186

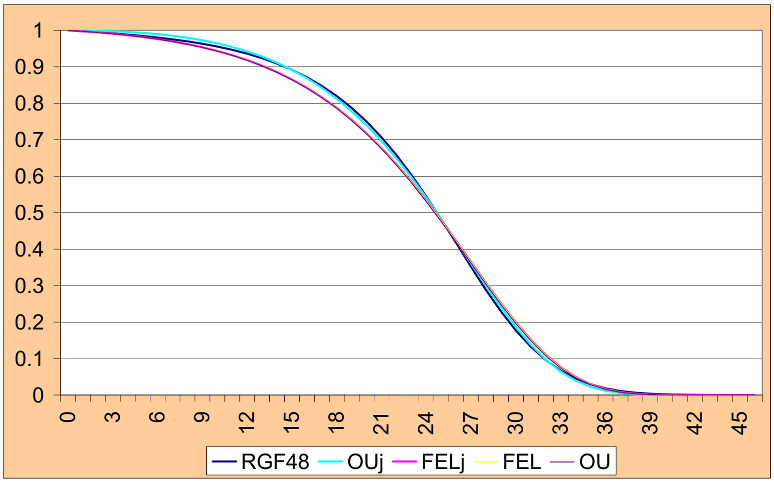
- The calibration errors are very small, compared to those of the previous models.
- The OUj model dominates the others. The models with jumps perform better than the corresponding models without.
- The value of σ is often 0.

Results from the calibration: survival functions

MALES



FEMALES

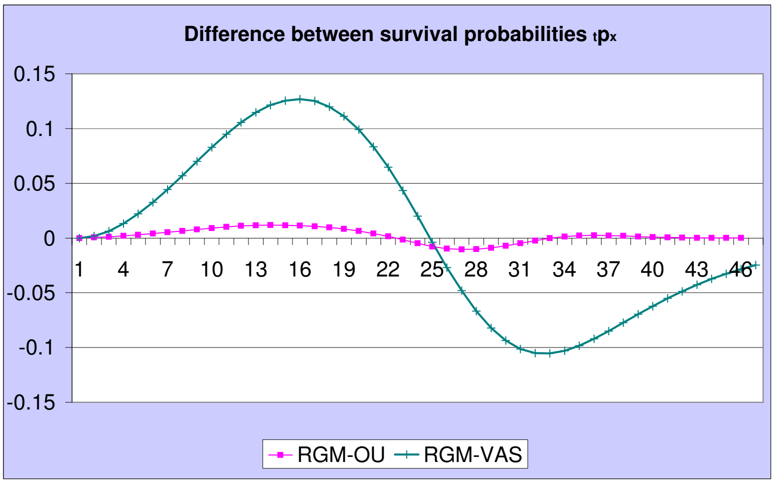


The fit is remarkable, especially in the males case.

Differences between survival function implied by the model and ${}_tp_x$ of the table (I).

To have a better idea of the goodness of the fit, we plot the differences between the survival function that has to be calibrated and the one implied by the four models are plotted.

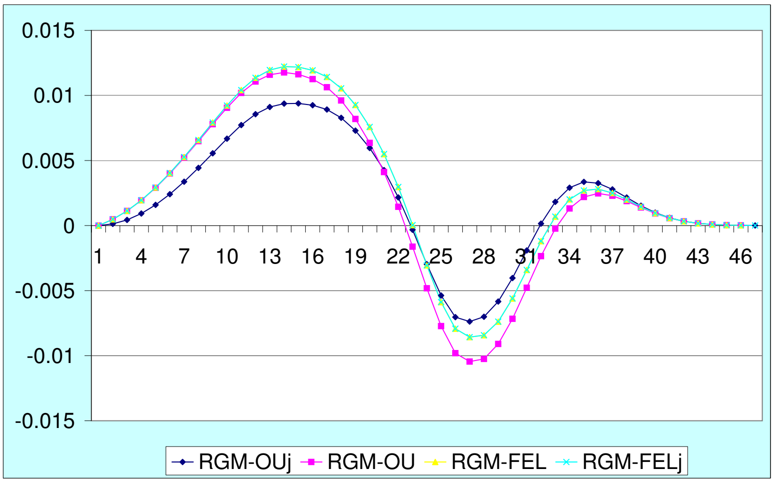
Next graph reports the comparison between one particular mean reverting process and the corresponding non mean reverting one.

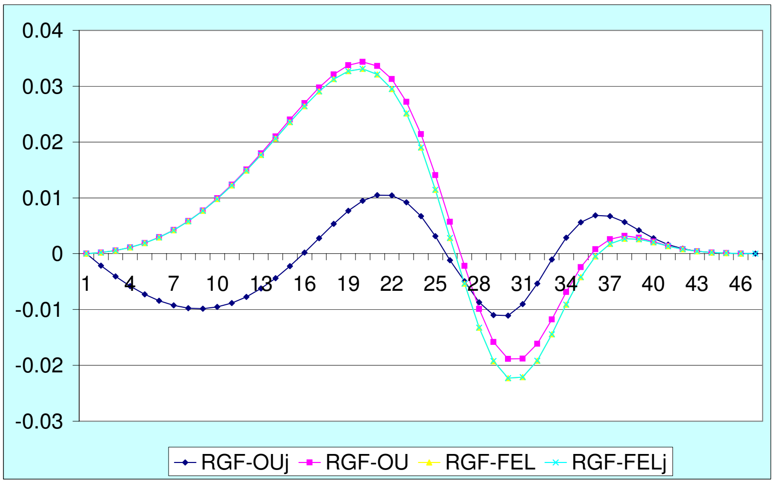


⇒ significant improvement in the fit when dropping the mean reversion term

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

Differences between survival function implied by the model and $_tp_x$ of the table (II).





Results from calibration: optimal parameters

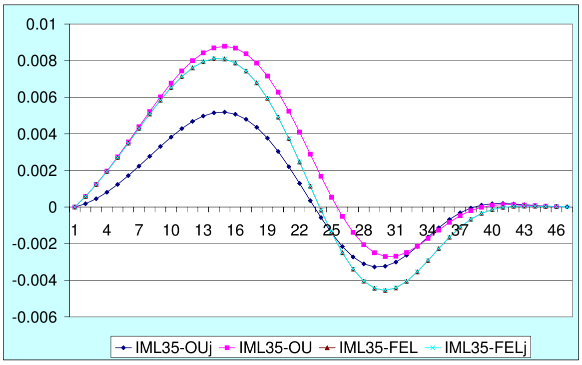
	IML92-1935	IML92-1945	IML92-1955
	$\lambda(0)$=0.011	$\lambda(0)$=0.00885	$\lambda(0)$=0.00706
OU-a	0.099	0.109	0.118
OU-σ	0.000098	0.0004756	0.0006785
OU-error	0.00085	0.0002706	0.0002075
OUj-a	0.099187	0.109188	0.118151
OUj-σ	0	0.000557	0.000677
OUj-l	0.000404	0.000326	0.000246
OUj-mu	-0.000015	-0.000015	-0.000015
OUj-error	0.00033	0.000096	0.00027
FEL-a	0.0986656	0.1081074	0.1164899
FEL-σ	0	0	0
FEL-error	0.00084	0.0002693	0.0006564
FELj-a	0.09867	0.10833	0.118645
FELj-σ	0	0.002733	0.007955
FELj-l	0.001008	0.001004	0.000996
FELj-mu	-0.000101	-0.0001	-0.0001
FELj-error	0.00084	0.00026	0.0001

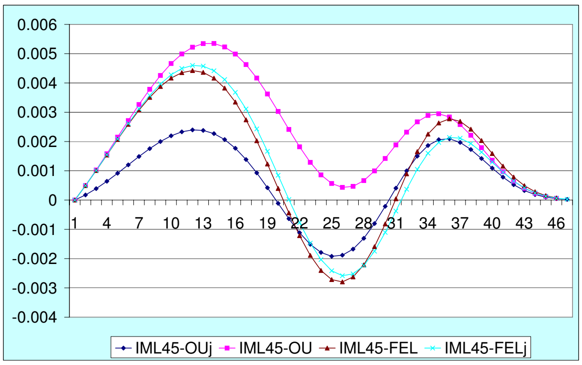
- The calibration errors are very small.

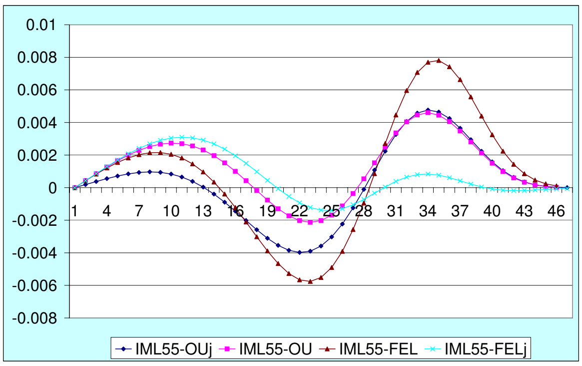
- The models with jumps perform better than the corresponding models without. Sometimes the OUj is better, sometimes the FELj.

- The value of σ is often 0.

Differences between survival probabilities.







Sensitivity analysis

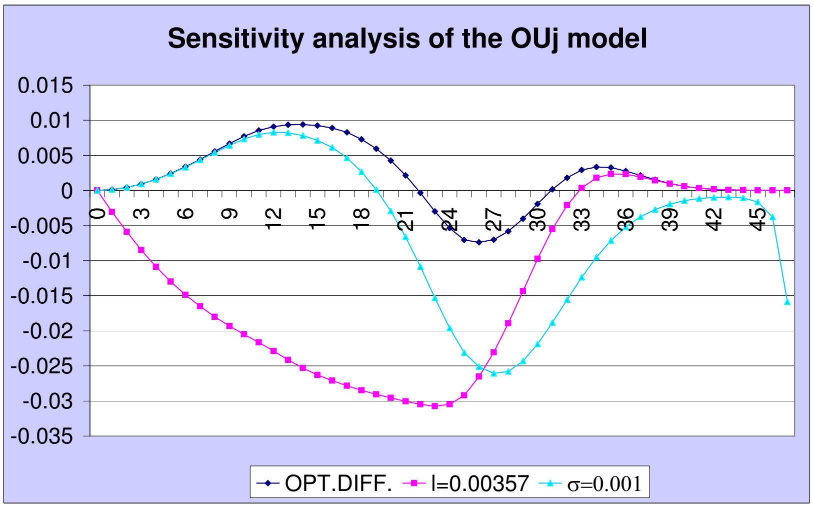
Evidence from the calibration of the non mean reverting processes:

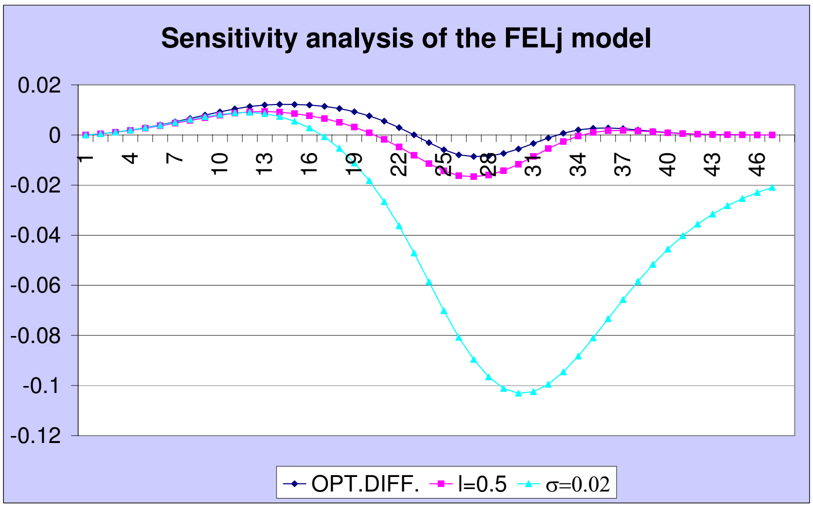
- low or null diffusion parameter (σ)
- improvements of fit when adding a jump component

Open question: is this due to the fact that the calibration is done on mortality tables that are *not* observed, but are constructed with deterministic algorithms?

Sensitivity analysis is possible: we have done stress tests on the parameters of the RG48 males, increasing the value of σ and l, considering the models with jumps OUj and FELj.

Differences between survival probabilities.





Main result from the stress tests:

Increasing the diffusion coefficient σ or the jump intensity l (the latter meaning a reduction in the expected arrival time of jumps), leads to differences becoming more negative.

$$\Rightarrow$$

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

This is confirmed by analytical results, showing, when possible, that the functions α and β are increasing functions of σ and l.

Conclusions & future research

- We have modelled the death time as a doubly stochastic stopping time, choosing affine processes for the intensity of mortality.
- Despite their large use and success in the financial context, mean reverting processes are not suitable.
- Non mean reverting processes with deterministic part that growths exponentially seem to be appropriate.
- In the stochastic part of the intensity process, negative jumps describe the mortality improvements better than the diffusion component.
- Increasing the randomness of the intensity process results in improvements in survivorship.
- The impact of these conclusions on insurance products pricing and reserving is in the agenda for future research.