CEREMADE
UMR CNRS 7534

Dauphine | PSL
UNIVERSITÉ PARIS

BRIDGING THE LI-CARTER'S GAP

A LOCALLY COHERENT MORTALITY FORECAST APPROACH

Quentin Guibert

CEREMADE, Université Paris-Dauphine
Laboratoire SAF, Université Claude Bernard Lyon 1
Prim'Act
Email: guibert@ceremade.dauphine.fr

3rd HMD Conference
18 November 2020, Paris, France

Joint work with S. Loisel, O. Lopez and P. Piette

This presentation is based on the following working paper:

Guibert, Q., Lopez, O., and Piette, P. (2020). Bridging the Li-Carter's gap: a locally coherent mortality forecast approach. Working Paper. [Link HAL].

OUTLINE

Multipopulation modeling

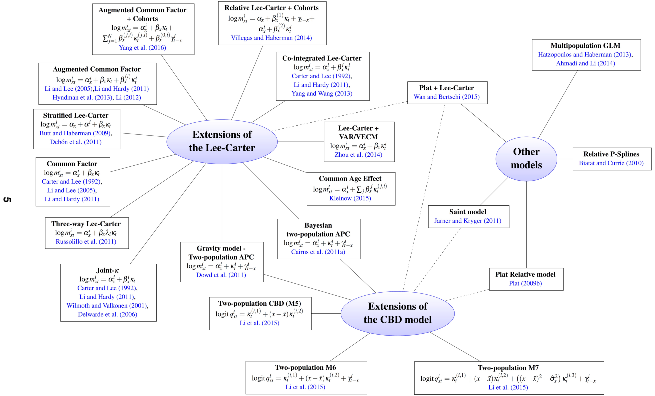
- Growing interest in mortality forecasting of multiple populations (e.g. Danesi *et al.*, 2015; Bergeron-Boucher *et al.*, 2018; Cairns *et al.*, 2019).
- Important issue for longevity risk assessment (government, pension fund, life insurance, ...).
- Several modeling challenges

Relationships between populations	**Heterogeneous exposure**
Correlation of the longevity trends	Lifestyle factors (alcohol, tobacco, obesity)
Evidence supporting coherence	Socioeconomic inequalities
	Health system
	Environmental factors

- Most of models for multiple populations are based on the coherence principle (Li and Lee, 2005), i.e. for populations i and j aged x

$$|\ln m_{x,t+h}^{(i)} - \ln m_{x,t+h}^{(j)}| \text{ do not diverge when } h \to \infty.$$

- This assumption is relevant for two-populations mortality models (Villegas *et al.*, 2017) when managing basis risk between 2 populations.



Source: Villegas *et al.* (2017)

AIMS OF THE STUDY

LIMITATIONS OF THE COHERENCE PRINCIPLE

- Only suitable for specific populations and over limited time windows (Li *et al.*, 2017).
- Not adapted for large heterogeneous longevity portfolio, e.g. global insurance or reinsurance company.
- Divergence between populations can exist $\rightarrow$ the coherence principle may distort the projections.

MAIN AIMS

- Introduce a framework for simultaneous modeling of several populations.
- Relax the mortality coherence principle $\rightarrow$ introduce a locally coherent assumption.
- Assess the impact in terms of simulated mortality dispersion for a large number of Western European populations.
- Improve risk assessment of the longevity risk SCR and longevity hedges basis risk.

SCOPE OF THE STUDY

Modeling simultaneously a large number of Western European populations.

DATA FROM THE HUMAN MORTALITY DATABASE (HMD, 2019)

- A collection $\mathcal{I}$ of $I = 16 \times 2$ populations (gender segregating).
- Austria, Belgium, Switzerland, West Germany, Denmark, Spain, Finland, France, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Sweden and England & Wales.

AGE–PERIOD TRAINING SET

- Age: 45-90.
- Period: 1960-2014.

Li-Lee vs. Lee-Carter for multiple populations

Consider the central mortality rates for the i-th population $m_{x,t}^{(i)} = \frac{D_{x,t}^{(i)}}{E_{x,t}^{(i)}}$.

Independent Lee and Carter (1992) model

Dynamic of the i-th population

$$\ln\ m_{x,t}^{(i)} = \alpha_x^{(i)} + \beta_x^{(i)} \kappa_t^{(i)} + \epsilon_{x,t}^{(i)}$$

- $\kappa_t^{(i)}$ are independent random walks with drifts.
- Estimation with SVD method for each population.

Coherent Li and Lee (2005) model

Dynamic with a common trend $B_x^{\mathcal{I}} K_t^{\mathcal{I}}$ for all populations

$$\ln\ m_{x,t}^{(i)} = \alpha_x^{(i)} + B_x^{\mathcal{I}} K_t^{\mathcal{I}} + \beta_x^{(i)} \kappa_t^{(i)} + \epsilon_{x,t}^{(i)}$$

- $K_t^{\mathcal{I}}$ is a random walk with drift.
- $\kappa_t^{(i)}$ are independent mean-reverting process (AR(1) models) $\rightarrow$ enforce coherence.

Coherence property

Gap between the Lee-Carter (LC) and the Li-Lee (LL) models

- LC model artificially may create some diversification in terms of longevity risk as no relationship between populations are taken into account.
- LL model imposes a strong coherence hypothesis for all the populations: mortality rates will not diverge in the long run, although $\kappa_t^{(i)}$ allow slight derivations.

Measuring dispersion

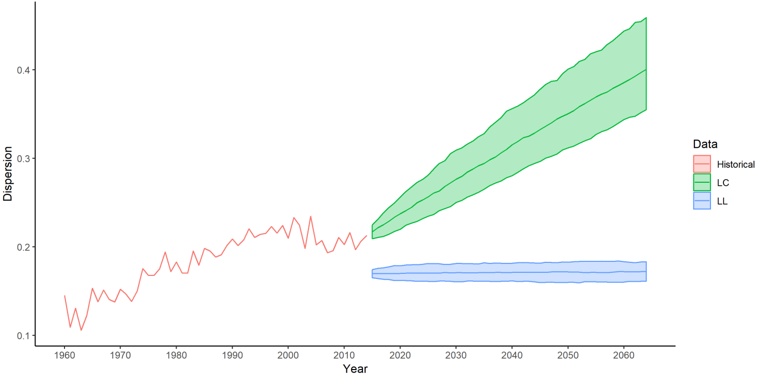
- Let introduce a metric to measure the divergence at age x of the mortality rates from a collection $\mathcal{I}$ of I populations

$$\delta_{x,t}^{\mathcal{I}} = \sqrt{\frac{1}{I-1} \sum_{i \in \mathcal{I}} \left(\ln m_{x,t}^{(i)} - \overline{\ln m_{x,t}} \right)^2}$$

where $\overline{\ln m_{x,t}} = \frac{1}{I} \sum_{i \in \mathcal{I}} \ln m_{x,t}^{(i)}$.

DISPERSION FORECASTING

Dispersion at age 85 in the western European populations ($I = 32$)



- Median projections by LC and LL models with the corresponding 95% prediction intervals (500 Monte-Carlo simulations).
- Both the LC and LL models present drawbacks:
  - Artificial diversification vs. high concentration.
  - Need to consider some intermediate scenarios $\rightarrow$ bridge the gap!

Local coherence

Key idea

- Populations are coherent by homogeneous sub-groups, and not all together at the same time.
- Local version in the populations space dimension of the coherence property.

Intermediate model

- Assume the existence of coherent sub-groups of populations.
- Denote $\mathcal{J}$ a partition of the populations collection $\mathcal{I}$ in J distinct sub-groups.
- Let $\phi : \mathcal{I} \to \mathcal{J}$ a function returning the label of the assigned sub-group.
- Dynamic of the i-th population

$$\ln \, m_{x,t}^{(i)} = \alpha_x^{(i)} + \underbrace{B_x^{\phi(i)} K_t^{\phi(i)}}_{\substack{\text{common trend of} \\ \text{the sub-group } \phi(i)}} + \underbrace{\beta_x^{(i)} \kappa_t^{(i)}}_{\substack{\text{independant AR(1)} \\ \text{short-term effet}}} + \epsilon_{x,t}^{(i)}$$

VAR MODEL FOR COMMON TRENDS

- Let $\boldsymbol{K}_t = \left(K_t^j\right)_{j \in \mathcal{J}}$ the vector of dominent trends related to sub-groups.
- Consider a VAR model with a lag p for capturing relationships between sub-groups

$$\Delta \boldsymbol{K}_t = \boldsymbol{C} + \sum_{k=1}^{p} \boldsymbol{A}_k \Delta \boldsymbol{K}_{t-k} + \boldsymbol{E}_t,$$

  where $\Delta K_t^j = K_t^j - K_{t-1}^j$ is the common mortality improvement of a cluster.
- $\boldsymbol{A}_k$, $k = 1, \ldots, p$, are $J \times J$-autoregressive matrices which capture the long–run relationships of mortality improvements between coherent sub-groups.
- $\boldsymbol{C}$ is a J–dimensional vector of drifts.
- $\boldsymbol{E}_t$ is a J–dimensional Gaussian white noise.

DYNAMIC OF SUB-GROUPS

BORDER CASES

The LC and LL models are included in this specification.

LC model	**LL model**
Single sub-groups: $\phi(i) = \{i\}$	Only one group: $\phi(i) = \mathcal{I}$
Lag $p = 0$	Lag $p = 0$
$V(\boldsymbol{E}_t)$ diagonal	Var-covar of the $\left(\kappa_t^{(i)}\right)_i$ diagonal

Dynamic of sub-groups

VAR Elastic–Net - A flexible estimation process

- The number of sub-groups can be large or small.
- Estimation based on VAR Elastic–Net specification (Guibert *et al.*, 2019).
- Consider T observations and minimize the criterion

$$L\left(\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_p\right) = \frac{1}{T-p} \sum_{p}^{T} \|\Delta \boldsymbol{K}_t - \boldsymbol{C} - \sum_{k=1}^{p} \boldsymbol{A}_k \Delta \boldsymbol{K}_{t-k}\|_2^2$$

$$- \alpha\lambda \sum_{k=1}^{p} \|\boldsymbol{A}_k\|_1 - \frac{(1-\alpha)\lambda}{2} \sum_{k=1}^{p} \|\boldsymbol{A}_k\|_2^2,$$

- $\lambda > 0$ is the strength of the penalization $\rightarrow$ 10-folds cross-validation method.
- $\alpha \in [0, 1]$ represents the mix between ridge ($\alpha = 0$) and LASSO ($\alpha = 1$) penalties.
- Hereafter, we fix $\alpha = 0.9$ and $p = 4$, which allow to have good fits.

How grouping populations?

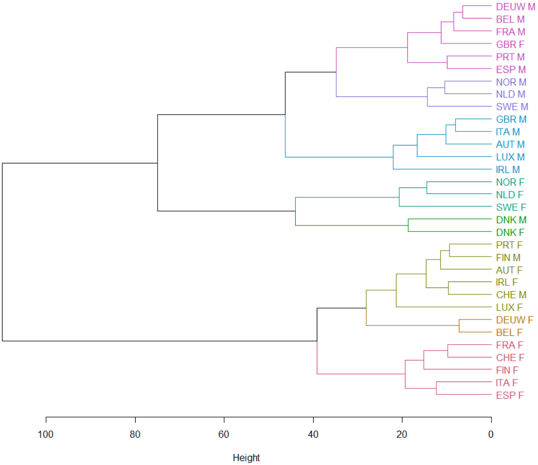
- Very difficult tasks based on 2 approaches: pure data-driven approach (see e.g. Hatzopoulos and Haberman, 2013) or expert judgments approach.
- Apply expert judgments (trends in data, economical, social, environmental, ... criteria).
- For instance, we can consider 16 sub-groups by grouping males and females of the same country $\rightarrow$ coherence by country.
- Clustering analysis based on times-series.

BASIC TIME SERIES CLUSTERING EXAMPLE

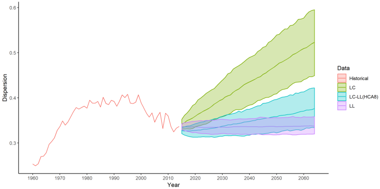
- Consider the time series $\left(\kappa_t^{(i)}\right)_{t>0}$ derived from the LC fitting.
- Apply an unsupervised hierarchical cluster analysis (HCA) method with the Euclidean metric and Ward's criterion.
- Gender indicator is one of the major splitting criteria (except Denmark).

Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Group 7**	**Group 8**
DEUW M	NOR M	GBR M	NOR F	DNK M	PRT F	DEUW F	FRA F
BEL M	NLD M	ITA M	NLD F	DNK F	FIN M	BEL F	CHE F
FRA M	SWE M	AUT M	SWE F		AUT F		FIN F
GBR F		LUX M			IRL F		ITA F
PRT M		IRL M			CHE M		ESP F
ESP M					LUX F		

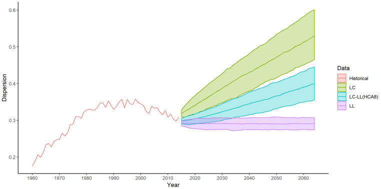
Dendrogram associated with the HCA applied on the LC's κ_t



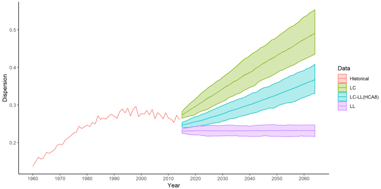
Dispersion in the western European populations with LC, LL and LC–LL(HCA8) models



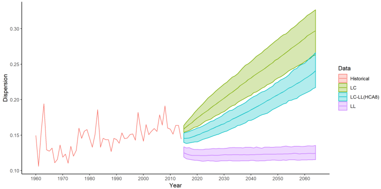
(A) Age 70



(B) Age 75



(C) Age 80



(D) Age 90

Application: Solvency II impact

Assessing the longevity risk of a global portfolio

- Global portfolio with pensions to be paid between ages 60 to 90 with 32 populations equally distributed.
- Annuity of 100 per year end of the year.
- One cohort aged 59 in 2014.
- Discount rate 1%.
- SCR calculated at the Value at Risk (VaR) at a 99.5% level of the best estimate provision (2,000 Monte-Carlo scenarios).
- Compare two clustering approaches: by gender (MF) and using HCA (HCA8).

	LC	LL	LC–LL(MF)	LC–LL(HCA8)
BE provisions (mean)	67,513	67,577	67,675	67,485
VaR 99.5 %	68,059	69,591	68,999	68,954
SCR (VaR - mean)	547	2,014	1,323	1,469

TREND SWITCHING MODEL

- Now, assume that the classification function $\phi_t : \mathcal{I} \to \mathcal{J}$ may change over time.
- Consider the following dynamics for the central mortality rates

$$\ln\ m_{x,t}^{(i)} = \alpha_x^{(i)} + B_x^{\phi_t(i)} K_t^{\phi_t(i)} + \beta_x^{(i)} \kappa_t^{(i)} + \mathrm{ad}_{x,t}^{(i)} + \epsilon_{x,t}^{(i)}$$

- $\mathrm{ad}_{t,x}^{(i)}$ are adjustment mortality levels to avoid abrupt jumps each time a population changes of dominant trend

$$\mathrm{ad}_{x,t}^{(i)} = \sum_{s=t_0+1}^{t} B_x^{\phi_{s-1}(i)} K_s^{\phi_{s-1}(i)} - B_x^{\phi_s(i)} K_s^{\phi_s(i)}$$

where t_0 is a time such as $ad_{x,t_0}^{(i)} = 0$ for all ages x.

- Consider the Longevity Divergence Index Value (LDIV) similar to the Swiss Re Kortis bond (Hunt and Blake, 2015).
- We construct an illustrative European LDIV for:
  - Longevity risk of the Swiss female population, aged between $x_1^{(\text{CHE F})} = 75$ and $x_2^{(\text{CHE F})} = 85$.
  - Mortality risk of the French female population, aged between $x_1^{(\text{FRA F})} = 55$ and $x_1^{(\text{FRA F})} = 65$.
  - A risk period n of 8 years, which ends at year $t = 2024$.
- At $t_0 = 2014$, the collection of populations follows the LC-LL(HCA8). FRA F and CHE F belong to Group 8.
- The LDIV at time t is obtained by

$$\text{LDIV}\,(t) = \text{Index}\,(t, \text{CHE F}) - \text{Index}\,(t, \text{FRA F})$$

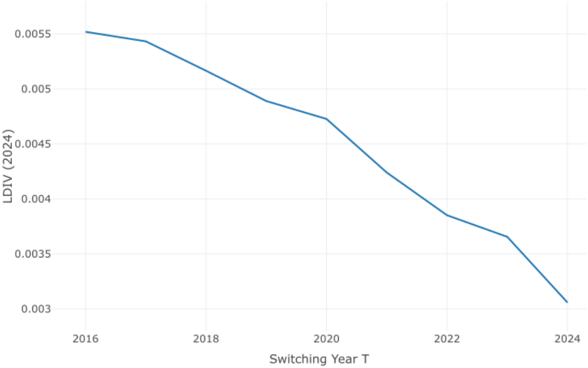
- The averaged improvement index is computed as

$$\text{Index}\,(t, i) = \frac{1}{1 + x_2^{(i)} - x_1^{(i)}} \sum_{x=x_1^{(i)}}^{x_2^{(i)}} 1 - \left[\frac{m_{x,t}^{(i)}}{m_{x,t-n}^{(i)}} \right]^{\frac{1}{n}}$$

- Suppose at a time $T > t_0$, the FRA F population (Group 8) switches to the trend of the BEL F population (Group 7), i.e for $t > t_0$,

$$\phi_t(i) = \begin{cases} \phi_{t_0}(i) & \text{if } i \neq \text{FRA F}, \\ \phi_{t_0}(\text{FRA F}) & \text{if } i = \text{FRA F and } t < T, \\ \phi_{t_0}(\text{BEL F}) & \text{if } i = \text{FRA F and } t \geq T. \end{cases}$$

Median European LDIV(2024) **according to the switching time** T



Conclusion and Future work

Main results

- Fully independent (LC model): artificial diversification.
- Fully coherent (LL model): exaggerated concentration.
- A locally coherent model to forecast populations with homogeneous mortality profiles.
- Relationship between dominant trend modeling through a VAR model without any coherence constraints.
- Allow intermediate situations over 32 European populations in terms of dispersion.
- Major impact in terms of SCR for the longevity risk and when pricing LDIV solution under a dynamic framework.

Conclusion and Future work

Future work

- Improve population clustering via time series clustering techniques.
- Future works are needed for empirically identifying coherent groups.
- The dynamic version of the model is not easy to calibrate $\rightarrow$ Backward analysis for detecting jumps in dominant trend are required.
- Assess the out-of-sample performances of our approach.

References I

Bergeron-Boucher, M.-P., Simonacci, V., Oeppen, J., and Gallo, M. (2018). Coherent Modeling and Forecasting of Mortality Patterns for Subpopulations Using Multiway Analysis of Compositions: An Application to Canadian Provinces and Territories. *North American Actuarial Journal*, **22**(1), 92–118.

Cairns, A. J. G., Kallestrup-Lamb, M., Rosenskjold, C., Blake, D., and Dowd, K. (2019). Modelling Socio-Economic Differences in Mortality Using a New Affluence Index. *ASTIN Bulletin: The Journal of the IAA*, **49**(3), 555–590.

Danesi, I. L., Haberman, S., and Millossovich, P. (2015). Forecasting mortality in subpopulations using Lee–Carter type models: A comparison. *Insurance: Mathematics and Economics*, **62**, 151–161.

Guibert, Q., Lopez, O., and Piette, P. (2019). Forecasting mortality rate improvements with a high-dimensional VAR. *Insurance: Mathematics and Economics*, **88**, 255–272.

Hatzopoulos, P. and Haberman, S. (2013). Common mortality modeling and coherent forecasts. An empirical analysis of worldwide mortality data. *Insurance: Mathematics and Economics*, **52**(2), 320–337.

HMD (2019). Human mortality database. university of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (data downloaded on 2019-07-08).

Hunt, A. and Blake, D. (2015). Modelling longevity bonds: Analysing the Swiss Re Kortis bond. *Insurance: Mathematics and Economics*, **63**, 12–29.

REFERENCES II

Lee, R. D. and Carter, L. R. (1992). Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, **87**(419), 659–671.

Li, J. S.-H., Chan, W.-S., and Zhou, R. (2017). Semicoherent Multipopulation Mortality Modeling: The Impact on Longevity Risk Securitization. *Journal of Risk and Insurance*, **84**(3), 1025–1065.

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

Villegas, A. M., Haberman, S., Kaishev, V. K., and Millossovich, P. (2017). A comparive study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin: The Journal of the IAA*, **47**(3), 631–679.