Modelling mortality risks using hidden markov models with covariates

Seventeeth International Longevity Risk and Capital Markets Solutions Conference

Michelle Vhudzijena[1,2], Michael Sherris[1,2], Andrés Villegas[1,2] and Jonathan Ziveyi[1,2]

[1]ARC Centre of Excellence in Population Ageing Research
[2]School of Risk and Actuarial Studies, UNSW Sydney, Australia

Outline

- Introduction
- Data
- Methodology
- Results
- Conclusion

Do we have the same probability of dying?

- Mortality heterogeneity is the non-uniformity of each individual's susceptibility to death in a population

Why should we care?

- Understanding the drivers of mortality heterogeneity has a significant impact on the underwriting and pricing of life and health insurance products
- Individuals with a perceived shorter life expectancy due to inequities in education, health, wealth and income are less likely to purchase annuities because they are deemed expensive in relation to their mortality risk
- Correctly priced annuities that cover all risk levels have the potential to increase the demand of longevity risk products

Literature gaps

- Sherris and Wei (2020) show that disregarding health status can cause adverse selection from individuals with chronic conditions due to inaccurate pricing of mortality and morbidity risks.
- However, they fail to show the impact of individual risk factors such as body mass index and drinking status on health status which are known to affect mortality risk.
- We incorporate these predictors by fitting Hidden Markov models with health status as the response and covariates such as BMI, income and self reported–health.

Research questions

1. To what extent do the clusters developed from the multivariate-time series clustering of health trajectories using HMMs with covariates, provide well developed risk profiles that exhibit mortality heterogeneity?
2. Does clustering provide a better fit to empirical data when estimating transition rates and life expectancy in a multi–state model of health status and functional disability while controlling for age and gender?
3. To what extent do trajectories of health risk factors capture mortality differences so that there is a reduction in inequities in the pricing of life annuities and deferred life annuities?

Health Retirement Study

- We use data from the University of Michigan Health Retirement Study (HRS) which follows Americans aged 50 and above.
- We use data from waves 1 to 13 to estimate HMMs and clustering
- Multi–state models are fit using ADLs from wave 4 to 13 similar to previous studies using this dataset (Fong, Shao, and Sherris 2015; Li, Shao, and Sherris 2017; Shao, Sherris, and Fong 2017; Sherris and Wei 2020).
- We exclude individuals who provide inappropriate responses, who fail to respond in any wave and those who do not appear in consecutive interview dates.

Functional disability model

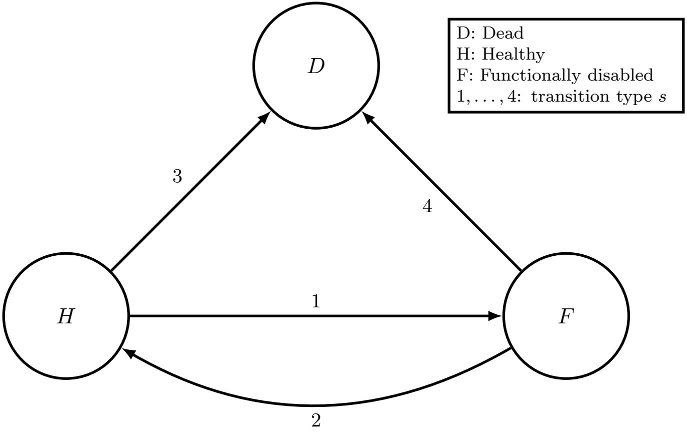

Figure 1: Three state functional disability model

Extending three state models

Following the proportional hazard specification in Li, Shao, and Sherris (2017); Sherris and Wei (2020), we model the transition intensity type s of type $s = 1, \dots, S$ for an individual k for $k = 1, \dots, K$ at time t years with

$$\lambda_{k,s}(t) = \exp(\beta_s + \gamma_s' w_k(t) + \alpha_s \psi(t)) H_{k,s}(t),$$

where β_s is the time invariant baseline log–intensity for transition type s, $w_k(t)$ is a vector of the observed predictors for each individual k, $\psi(t)$ is frailty which is a stochastic latent process, γ_s is a vector measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $w_k(t)$, α_s is a scalar measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $\psi(t)$ and $H_{k,s}(t) = 1$.

Multi–state models I

Static model

The transition rate $\lambda_{k,s}(t)$ is assumed to be dependent on age and sex only:

$$\begin{aligned} \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k \\ \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \gamma_s^{cluster} C_k \end{aligned} \tag{1}$$

where $x_k(t)$ is the k^{th} individual's age at time t, F_k is the binary variable indicating the gender for the individual k,C_k is the categorical variable indicating the cluster for the individual k, γ_s^{age} measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to age, $\gamma_s^{cluster}$ measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to cluster γ_s is a scalar measuring the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to sex.

Multi–state models II

Model with systematic trend

$$\begin{aligned} \ln\lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t \\ \ln\lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \gamma_s^{cluster} C_k \end{aligned} \quad (2)$$

where ϕ_s measures the sensitivity of $\ln\lambda_{k,s}(t)$ with respect to the time trend t.

Multi–state models III

Frailty model with systematic trend and uncertainty

$$\ln \lambda_{k,s} = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \alpha_s \psi_i$$
$$\ln \lambda_{k,s} = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \alpha_s \psi_i + \gamma_s^{cluster} C_k \tag{3}$$

where α_s measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to the latent factor ψ that is modelled as a random walk,

$$\psi_j = \psi_{j-1} + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma^2), \quad \psi_0 = 0 \quad \text{and} \quad \sigma^2 = t_j - t_{j-1}.$$

- We estimate the transition rates using code written by Fu, Sherris, and Xu (2021)

Hidden Markov models I

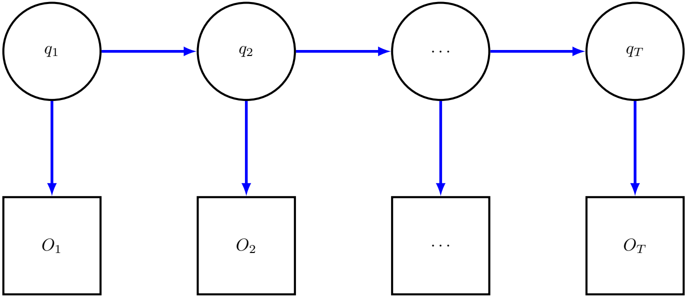

Figure 2: A graphical representation of a Hidden Markov Model

Hidden Markov models II

- Doubly embedded stochastic processes that are used to model an output based on an assumption that the system being modelled is Markovian and has unobservable states
- HMMs allow us to define a meaningful distance measure between trajectories through the use of probability densities that define trajectories.
- We use the Kullback–Leibler(KL) Divergence to calculate the distance between trajectories (Kullback and Leibler 1951).

Determining clusters

HMM estimation and K-medoids clustering

1. Preprocess the data sets so that we have the trajectories $T_i = (H_{it}; BMI_{it}; Income_{it}) \quad or \quad T_i = (H_{it}; BMI_{it}), \quad i = 1, \dots, N, \quad t = 1, \dots, n.$
2. This step actually consists of two tasks:
   a. Map each trajectory T_i into a HMM (λ) model $i = 1, \dots, N$
   b. Compute the likelihood $P(T_i|\lambda_j)$ of any trajectory T_i given the HMM λ_j Probability densities)
3. Compute the distance matrix $D(\lambda_i; \lambda_j)$ for $i; j = 1, \dots, N$
4. Perform k–medoids clustering
5. Summarise the characteristics of the individuals in each cluster by computing a number of features for each cluster, as shown by Ghassempour, Girosi, and Maeder (2014).

Cluster assessment

Table 1: Cluster quality assessment with BMI as the only covariate

Clusters	Sihlouette	Dunn	Davies Bouldin
2	0.7817	0.6336	0.4350
3	0.8353	1.2294	0.3946
4	0.8224	0.4662	0.4015
5	0.4405	0.0773	0.7798
6	0.4034	0.1542	0.9379
7	0.3515	0.1542	0.9493
8	0.3452	0.2242	0.7839
9	0.3618	0.2032	0.7989
10	0.3375	0.2032	0.7083

- Optimal cluster solution minimises Davies Bouldin Index and maximises both the Sihlouette and Dunn indices.

Table 2: Cluster quality assessment with covariates of BMI and income

Clusters	Sihlouette	Dunn	Davies Bouldin
2	0.3522	0.0149	0.9351
3	0.4486	0.0167	1.0049
4	0.6165	0.0135	0.6534
5	0.6439	0.0122	0.6857
6	0.7032	0.0179	0.5696
7	0.7455	0.0330	0.5172
8	0.7630	0.0232	0.4886
9	0.7842	0.0175	0.4078
10	0.6754	0.0175	0.5418

Multi–state model estimation results I

Table 3: Static model: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\widehat{\beta}_s$	-3.6100	***	-2.0366	***	-3.4748	***	-2.2359	***
	(0.0603)		(0.0879)		(0.0587)		(0.0756)	
$\widehat{\gamma}_s^{\text{age}}$	0.5081	***	-0.3927	***	0.6377	***	0.4537	***
	(0.0311)		(0.0442)		(0.0318)		(0.0332)	
$\widehat{\gamma}_s^{\text{female}}$	0.3147	***	0.0226		-0.2470	***	-0.2952	***
	(0.0567)		(0.0977)		(0.0548)		(0.0616)	
Log likelihood	-14,361							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Multi–state model estimation results II

Table 4: Static model with clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.2210	***	-2.1644	***	-3.0870	***	-2.0627	***
	(0.1265)		(0.2067)		(0.1265)		(0.1359)	
	(0.0312)		(0.0443)		(0.0319)		(0.0332)	
$\hat{\gamma}_s^{\text{age}}$	0.5076	***	-0.3917	***	0.6381	***	0.4532	***
	(0.0312)		(0.0438)		(0.0319)		(0.0332)	
$\hat{\gamma}_s^{\text{female}}$	0.3173	***	0.0379		-0.2319	***	-0.2977	***
	(0.0567)		(0.0977)		(0.0549)		(0.0617)	
$\hat{\gamma}_s^{\text{cluster}}$	-0.1477	***	0.0494		-0.1476	***	-0.0674	
	(0.0426)		(0.0720)		(0.0431)		(0.0442)	
Log likelihood	-14,348							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Multi–state model estimation results III

Table 5: Trend model: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\widehat{\beta}_s$	-3.7414	***	-1.4331	***	-5.1971	***	-3.1831	***
	(0.0709)		(0.1065)		(0.0880)		(0.1015)	
$\widehat{\gamma}_s^{\text{age}}$	0.4776	***	-0.2905	***	0.3114	***	0.3580	***
	(0.0322)		(0.0448)		(0.0328)		(0.0342)	
$\widehat{\gamma}_s^{\text{female}}$	0.3266	***	-0.0301		-0.1302	**	-0.2115	***
	(0.0567)		(0.0983)		(0.0546)		(0.0619)	
$\widehat{\gamma}_s^{\text{time}}$	0.2352	***	-0.9571	***	2.3323	***	1.0764	***
	(0.0643)		(0.1117)		(0.0674)		(0.0681)	
Log likelihood	-13,525							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Multi–state model estimation results IV

Table 6: Trend model with clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\widehat{\beta}_s$	-3.3352	***	-1.7948	***	-4.6454	***	-2.7374	***
	(0.1296)		(0.2141)		(0.1364)		(0.1425)	
$\widehat{\gamma}_s^{\text{age}}$	0.4758	***	-0.2813	***	0.3178	***	0.3537	***
	(0.0322)		(0.0451)		(0.0328)		(0.0343)	
$\widehat{\gamma}_s^{\text{female}}$	0.3210	***	-0.0345		-0.1346	**	-0.2159	***
	(0.0568)		(0.0983)		(0.0547)		(0.0620)	
$\widehat{\gamma}_s^{\text{cluster}}$	-0.1576	***	0.1471	**	-0.2172	***	-0.1888	***
	(0.0425)		(0.0746)		(0.0417)		(0.0433)	
$\widehat{\gamma}_s^{\text{time}}$	0.2492	***	-0.9906	***	2.3418	***	1.1200	***
	(0.0643)		(0.1131)		(0.0673)		(0.0687)	
Log likelihood	-13,495							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

Model performance

Table 7: Model selection using Akaike information criterion (AIC) and Bayesian information criterion (BIC)

	No Clustering		Clustering	
Model	AIC	BIC	AIC	BIC
Static	28745.19	28849.70	28727.58	28866.92
Trend	27082.43	27221.77	27029.51	27203.69
Frailty	27009.71	27183.89	26957.92	27166.93

Table 8: Likelihood tests for models with and without clustering

Null model	Alternative model	p–Value	Chi–square	Symbol[1]
Static	Static with Clustering	3.7837e-05	25.6148	***
Trend	Trend with Clustering	1.8611e-12	60.9171	***
Frailty	Frailty with Clustering	3.2100e-12	59.7907	***

[1] *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Future lifetime statistics I

Table 9: Comparison of future lifetime statistics for 65–year old healthy individuals

	Cluster 1[a]		Cluster 2[b]		Cluster 3[c]	
Change	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Static						
Total future lifetime	87.68%	86.24%	95.40%	94.51%	103.57%	103.06%
Healthy future lifetime	84.92%	84.63%	94.22%	93.89%	104.30%	103.78%
Disabled future lifetime	98.46%	96.47%	100.02%	98.47%	100.72%	98.55%
Age at onset of disability conditional on becoming disabled	98.17%	97.89%	99.30%	99.05%	100.60%	100.59%
Trend in 1998						
Total future lifetime	89.00%	88.90%	96.08%	95.74%	102.86%	102.61%
Healthy future lifetime	86.43%	87.21%	94.72%	94.96%	103.66%	103.11%
Disabled future lifetime	101.51%	101.69%	102.69%	101.64%	98.95%	98.82%
Age at onset of disability conditional on becoming disabled	98.80%	98.83%	99.63%	99.62%	100.25%	100.23%
Frailty in 1998						
Total future lifetime	89.23%	89.14%	95.97%	95.87%	102.80%	102.66%
Healthy future lifetime	86.47%	87.39%	94.77%	95.14%	103.46%	103.12%
Disabled future lifetime	102.35%	101.93%	101.67%	101.19%	99.68%	99.31%
Age at onset of disability conditional on becoming disabled	98.85%	98.90%	99.57%	99.58%	100.30%	100.30%

[a] Percentage difference in future lifetime betweeen Cluster 1 and model with no clusters of the same sex;

[b] Percentage difference in future lifetime betweeen Cluster 2 and model with no clusters of the same sex;

[c] Percentage difference in future lifetime betweeen Cluster 3 and model with no clusters of the same sex.

Future lifetime statistics II

Table 10: Comparison of future lifetime statistics for 75–year old healthy individuals

	Cluster 1[a]		Cluster 2[b]		Cluster 3[c]	
Change	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Static						
Total future lifetime	86.63%	85.67%	94.93%	94.33%	104.02%	103.95%
Healthy future lifetime	83.15%	83.40%	93.70%	93.63%	104.66%	104.09%
Disabled future lifetime	97.78%	97.94%	98.88%	98.09%	101.97%	103.18%
Age at onset of disability conditional on becoming disabled	98.80%	98.71%	99.55%	99.51%	100.37%	100.31%
Trend in 1998						
Total future lifetime	87.78%	87.16%	95.61%	95.10%	103.58%	103.07%
Healthy future lifetime	85.41%	86.01%	94.45%	94.25%	103.83%	103.58%
Disabled future lifetime	95.89%	93.24%	99.59%	99.61%	102.72%	100.39%
Age at onset of disability conditional on becoming disabled	99.09%	98.97%	99.67%	99.60%	100.30%	100.21%
Frailty in 1998						
Total future lifetime	88.00%	87.90%	95.55%	95.33%	103.28%	103.07%
Healthy future lifetime	85.32%	86.11%	94.40%	94.46%	103.80%	103.58%
Disabled future lifetime	96.90%	97.02%	99.35%	99.76%	101.54%	100.43%
Age at onset of disability conditional on becoming disabled	99.11%	99.06%	99.66%	99.62%	100.26%	100.25%

[a] Percentage difference in future lifetime betweeen Cluster 1 and model with no clusters of the same sex;

[b] Percentage difference in future lifetime betweeen Cluster 2 and model with no clusters of the same sex;

[c] Percentage difference in future lifetime betweeen Cluster 3 and model with no clusters of the same sex.

Summary

- Multivariate trajectories of BMI, income and self–reported health estimated using hidden markov models are able to capture mortality differences amongst different segments.
- We find that clustering provides a better fit to empirical data than results not based on clustering.
- Trend and frailty models had the best performance while static models had the worst fit.

Improvements

Limitations

- Small sample size used in this analysis which affects the estimation of transition rates.

Future work

- Pricing implications
- Larger sample
- 2 or 4 hidden states
- Describe hidden states

References I

Fong, Joelle H., Adam W. Shao, and Michael Sherris. 2015. "Multistate Actuarial Models of Functional Disability." *North American Actuarial Journal* 19 (1): 41–59. https://doi.org/10.1080/10920277.2014.978025.

Fu, Yu, Michael Sherris, and Mengyi Xu. 2021. "Functional disability with systematic trends and uncertainty: a comparison between China and the US." *Annals of Actuarial Science*, 1–30. https://doi.org/10.1017/S1748499521000233.

Kullback, S., and R. A. Leibler. 1951. "On Information and Sufficiency." *The Annals of Mathematical Statistics* 22 (1): 79–86. https://doi.org/10.1214/AOMS/1177729694.

References II

Li, Zixi, Adam W. Shao, and Michael Sherris. 2017. "The Impact of Systematic Trend and Uncertainty on Mortality and Disability in a Multistate Latent Factor Model for Transition Rates." *North American Actuarial Journal* 21 (4): 594–610. https://doi.org/10.1080/10920277.2017.1330157.

Shao, Adam W., Michael Sherris, and Joelle H. Fong. 2017. "Product pricing and solvency capital requirements for long-term care insurance." *Scandinavian Actuarial Journal* 2017 (2): 175–208. https://doi.org/10.1080/03461238.2015.1095793.

Sherris, Michael, and Pengyu Wei. 2020. "A Multi-state Model of Functional Disability and Health Status in the Presence of Systematic Trend and Uncertainty." *North American Actuarial Journal*, February, 1–23. https://doi.org/10.1080/10920277.2019.1708755.