Modeling multi-state health transitions in China: A generalized linear model with time trends

Katja Hanewald, Han Li and Adam Shao

Australia-China Population Ageing Research Hub
ARC Centre of Excellence in Population Ageing Research (CEPAR)
UNSW Sydney

13th International Longevity Risk and Capital Markets Solutions Conference

21-22 September 2017
Australia-China Population Ageing Research Hub

- Based in the ARC Centre of Excellence in Population Ageing Research (CEPAR) at UNSW Sydney; funded by UNSW Sydney
- Research areas focusing on China:
  1. Aging trends
  2. Long-term care services and insurance
  3. Mature labor force participation
  4. Retirement incomes, financial products and housing
- Team:
  - Director: Prof John Piggott
  - Scientific Director: Prof Hanming Fang (University of Pennsylvania)
  - 4 full-time research fellows, 3 PhD students
Motivation

- Rapid population aging in China

- In 2015, 1 in 5 older persons (aged 65+) **globally** lived in China, while in 2050, 1 in 4 elderly (over 370 million people) will be Chinese (United Nations, 2015).

- China’s old age dependency ratio was 15% in 2015, will be close to 50% by mid-century (United Nations, 2015)

- Need for retirement planning, long-term care, and financial services for the elderly in China
Motivation

- Traditional family-based care under threat
  - Demographic changes, weakening of traditional values, greater geographic mobility, improved gender equality (see, e.g., Zhu, 2015; Lu et al., 2015).

- Current social security programs do not cover full nursing home cost; do not fund community-based services (Yang et al., 2013)

- Need for social security programs and/or private market solutions (e.g. LTC insurance, specialized home equity release products)

- Need to understand and model health transitions among Chinese elderly
Our paper

- We develop a generalized linear model (GLM) to estimate health transition intensities in a three-state Markov model
  - Builds on previous models developed by Renshaw and Haberman (1995) for UK data and Fong et al. (2015) for US data
  - Our model includes age effects, time trends and age-time interactions
- Provide first evidence on health transitions of Chinese elderly
Three-state time-inhomogeneous Markov process

- State **N**: non-disabled
- State **F**: functionally disabled
- State **D**: dead (absorbing)
Existing models for functional disability

- **Renshaw and Haberman (1995):**

  \[
  \log(\sigma_x) = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{1}
  \]
  \[
  \log(\varphi_{x,z}) = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 \sqrt{z} + \beta_4 xz + \beta_5 x \sqrt{z} \tag{2}
  \]
  \[
  \log(\nu_{x,z}) = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 (z - z_1)_+ + \beta_4 (z - z_2)_+ \tag{3}
  \]


- **Fong et al. (2015):**

  \[
  \eta_x = \sum_{s=0}^{k} \beta_s x^s \tag{4}
  \]

  where \( \eta_x = \log(\mu_x), \log(\sigma_x), \log(\varphi_x), \) or \( \log(\nu_x). \)


- **Li et al. (2017):**

  \[
  \ln(\lambda_{skx}(t)) = \beta_s + \gamma_{s,\text{female}} x_t + \gamma_{x,\text{female}} F + \phi_s t + \alpha_s \varphi(t) \tag{5}
  \]

  where \( t \) is the linear trend and \( \varphi(t) \) is the latent factor or frailty.

Stochastic mortality models

- Lee and Carter (1992):

  \[
  \log(m_{x,t}) = a_x + b_x \kappa_t, \tag{6}
  \]

  where \(a_x\) and \(b_x\) represent age effects and \(\kappa_t\) represents time effect.

- Cairns et al. (2006):

  \[
  \logit(q_{x,t}) = \kappa_t^1 + \kappa_t^2(x - \bar{x}), \tag{7}
  \]

  where \(\kappa_t^1\) and \(\kappa_t^2\) are time effects and are assumed to follow a bivariate random walk with drift process.

- Renshaw and Haberman (1996):

  \[
  \log(\mu_{x,t}) = \beta_0 + \sum_{j=1}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r} \sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i, \tag{8}
  \]

  where \(L_j\) is the \(j^{th}\) Legendre orthogonal polynomial.
A Generalized Linear Model

Link function: Adopt a log link function $g(\cdot)$:

$$g(\alpha_{x,t}) = \ln(\alpha_{x,t}) = \eta_{x,t},$$

(9)

for $\eta_{x,t} = \log(\mu_{x,t}), \log(\sigma_{x,t})$ or $\log(\nu_{x,t})$.

Linear predictor: Introduce a time trend and age-time interactions:

$$\eta_{x,t} = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 t + \beta_4 tx + \beta_5 tx^2$$

(10)

Probability distribution: Assume that the number of health transitions follows an independently distributed Poisson distribution.

Estimation and model selection: MLE, compare all possible model variants using BIC.
Our contribution

- **We combine** good model features and estimation techniques from multi-state models and mortality models.

- **We allow** for greater flexibility in the model and explore different functional forms.

- **We incorporate** a time trend in the transition intensities.

- **We compare** the distinct demographic differences between males and females in urban and rural areas in China.
Chinese Longitudinal Healthy Longevity Survey (CLHLS)

- Conducted by the Center for Healthy Aging and Family Studies (CHAFS) at the National School of Development at Peking University
- 22 of China’s 31 provincial regions
- Largest longitudinal survey of the “oldest old” (aged 80+) internationally
- Information on health status and quality of life of the elderly
Our sample

- Unbalanced panel, all individuals with 2+ consecutive observations
- Health transitions between 2 waves: 5 pairwise observations
- Focus on older ages 65–105
- Separate data for males/females and urban/rural
- We define the state “F” as having difficulties to perform 2+ Activities of Daily Living (ADL): bathing, dressing, eating, toileting, continence and transferring in and out of bed.
### Sample size

**Table: Number of transition counts.**

<table>
<thead>
<tr>
<th>Time</th>
<th>Males</th>
<th>Females</th>
<th>Males</th>
<th>Females</th>
<th>Males</th>
<th>Females</th>
<th>Males</th>
<th>Females</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>1998 - 00</td>
<td>99</td>
<td>153</td>
<td>175</td>
<td>292</td>
<td>277</td>
<td>604</td>
<td>362</td>
<td>793</td>
<td>141</td>
<td>240</td>
</tr>
<tr>
<td>2000 - 02</td>
<td>191</td>
<td>134</td>
<td>175</td>
<td>292</td>
<td>277</td>
<td>604</td>
<td>362</td>
<td>793</td>
<td>141</td>
<td>240</td>
</tr>
<tr>
<td>2002 - 05</td>
<td>168</td>
<td>134</td>
<td>175</td>
<td>292</td>
<td>277</td>
<td>604</td>
<td>362</td>
<td>793</td>
<td>141</td>
<td>240</td>
</tr>
<tr>
<td>2005 - 08</td>
<td>105</td>
<td>134</td>
<td>175</td>
<td>292</td>
<td>277</td>
<td>604</td>
<td>362</td>
<td>793</td>
<td>141</td>
<td>240</td>
</tr>
<tr>
<td>2008 - 11</td>
<td>214</td>
<td>134</td>
<td>175</td>
<td>292</td>
<td>277</td>
<td>604</td>
<td>362</td>
<td>793</td>
<td>141</td>
<td>240</td>
</tr>
<tr>
<td>Total</td>
<td>777</td>
<td>756</td>
<td>1,307</td>
<td>1,476</td>
<td>2,875</td>
<td>4,282</td>
<td>3,445</td>
<td>5,652</td>
<td>932</td>
<td>1,030</td>
</tr>
</tbody>
</table>

**Table: Number of exposure years.**

<table>
<thead>
<tr>
<th>Time</th>
<th>State N</th>
<th>State F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>1998 - 00</td>
<td>1,763</td>
<td>2937</td>
</tr>
<tr>
<td>2000 - 02</td>
<td>3,240</td>
<td>1,997</td>
</tr>
<tr>
<td>2002 - 05</td>
<td>5,570</td>
<td>7,516</td>
</tr>
<tr>
<td>2005 - 08</td>
<td>5,215</td>
<td>7,552</td>
</tr>
<tr>
<td>2008 - 11</td>
<td>4,946</td>
<td>8,627</td>
</tr>
<tr>
<td>Total</td>
<td>20,733</td>
<td>28,628</td>
</tr>
</tbody>
</table>
Plots of crude transition rates: urban females

(a) $\sigma: N \rightarrow F$

(b) $\mu: N \rightarrow D$

(c) $\nu: F \rightarrow D$
### Optimal model: parameter estimates

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>( \sigma: N \rightarrow F )</th>
<th>( \mu: N \rightarrow D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>-5.376***</td>
<td>-5.719***</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.122***</td>
<td>0.127***</td>
</tr>
<tr>
<td>( \beta_2 \times 10^2 )</td>
<td>-0.111***</td>
<td>-0.09**</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>-0.154***</td>
<td>-0.158***</td>
</tr>
<tr>
<td>( \beta_4 \times 10^2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_5 \times 10^3 )</td>
<td>-5.125***</td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>832.77</td>
<td>824.56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>( \nu: F \rightarrow D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>-2.267***</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.046***</td>
</tr>
<tr>
<td>( \beta_2 \times 10^2 )</td>
<td>-0.027***</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.047***</td>
</tr>
<tr>
<td>( \beta_4 \times 10^2 )</td>
<td>-0.029***</td>
</tr>
<tr>
<td>( \beta_5 \times 10^3 )</td>
<td>-1.622***</td>
</tr>
<tr>
<td>BIC</td>
<td>691.81</td>
</tr>
</tbody>
</table>

Note: Linear predictor: \( \eta_{x,t} = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 t + \beta_4 t x + \beta_5 t x^2. \)

\(* p < 0.05; ** p < 0.01.*
Estimation results

Example: urban females

\[ \log(\sigma_x) = -6.346 + 0.217x - 0.00259x^2 - 0.00154tx \] (disability rate)

\[ \log(\mu_x) = -4.684 + 0.137x - 0.00110x^2 \] (mortality rate from “N”)

\[ \log(\nu_{x,t}) = -2.619 + 0.053x - 0.026t \] (mortality rate from “F”)
Life expectancy and healthy life expectancy

- Use optimal models to compute LEs at age 65 and 75 conditional on initial health status and HLEs
- Results agree with Liu et al. (2009); Luo et al. (2016); Guo (2017)

**Table:** Healthy life expectancy at age 65 and 75.

<table>
<thead>
<tr>
<th>Year</th>
<th>Male Urban</th>
<th>Male Rural</th>
<th>Female Urban</th>
<th>Female Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>15.16</td>
<td>15.03</td>
<td>16.85</td>
<td>16.26</td>
</tr>
<tr>
<td>2011</td>
<td>15.16</td>
<td>15.17</td>
<td>17.36</td>
<td>16.68</td>
</tr>
<tr>
<td>2020</td>
<td>15.16</td>
<td>15.25</td>
<td>17.66</td>
<td>16.93</td>
</tr>
</tbody>
</table>

*Healthy life expectancy at 65*

<table>
<thead>
<tr>
<th>Year</th>
<th>Male Urban</th>
<th>Male Rural</th>
<th>Female Urban</th>
<th>Female Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>8.96</td>
<td>8.58</td>
<td>9.64</td>
<td>9.56</td>
</tr>
<tr>
<td>2011</td>
<td>8.96</td>
<td>8.76</td>
<td>10.21</td>
<td>10.04</td>
</tr>
<tr>
<td>2020</td>
<td>8.96</td>
<td>8.86</td>
<td>10.54</td>
<td>10.31</td>
</tr>
</tbody>
</table>

*Healthy life expectancy at 75*
Conclusion

- **Summary:** A new flexible approach to modeling health transitions at higher ages based on the GLM framework.
  - Model allows for time trends and age-time interactions
  - Results for Chinese aged 65-105 (males/females, urban/rural)

- **Results:**
  - Time trends and age-time interactions are important for modeling disability rates and disabled mortality rates
  - Estimated LEs and HLEs: persistent rural/urban health inequalities

- **Potential applications of the model:**
  - Estimate the demand for LTC services and insurance
  - Analyze other health conditions (chronic diseases, critical illnesses)
Thank you!

Any questions, comments or suggestions?

Contact email: k.hanewald@unsw.edu.au


