Using Standardized Mortality to Estimate Life Expectancy of Small Areas

Sharon Yang and Jack Yue National Chengchi University Date: Sept. 8, 2023 Email: syang@nccu.edu.tw



Summary

- Life Expectancy of Small Areas
- Data and Methodology
- Estimating Longevity of Taiwan Counties
- Conclusion and Discussions

County-level Life Expectancy in Taiwan (2020)



Regional Life Expectancy in Taiwan

- Taiwan is not a big country but the difference of county-level life expectancy is not small.
- →The difference decreases (still large) after National Health Insurance was launched in 1995.
- Taiwan government decides not to construct township-level life tables (small samples).
- →Results of mortality graduation (including Lee-Carter model) would have noticeable fluctuation.

Life Expectancy in Taipei Cities



Note: These results are based on 1,000 bootstrap runs.

Estimating Longevity of Small Areas

- However, it is necessary to estimate longevity
 - of small areas due to regional discrepency.
- \rightarrow Q: Do we need to construct life tables?
- There are other mortality indices that can reflect regional characteristics.
- →For example, age-specific populations and deaths, Standardized Death Rate (SDR),
 Standardized Mortality Ratio (SMR).

SDR and SMR

SDR is often expressed in # of deaths per 100,000

$$SDR = 100000 imes rac{\sum_{x} N_{x} rac{d_{x}}{n_{x}}}{\sum_{x} N_{x}}$$

- → N_x & n_x are # of persons of age x of standard & target populations, D_x & d_x are # of deaths of standard & target populations.
- SMR is overall mortality to standard population

$$SMR = \frac{\sum_{x} d_{x}}{\sum_{x} n_{x} \frac{D_{x}}{N_{x}}}$$

SMR of Six Taiwan Major Cities (Female, 1998-2018)



Year

Boxplot of County-level SMR in Taiwan (1974-2018)



Estimation of Small Area Longevity

- SDR (& SMR) with Regression models
- →Lai et al. (1996, 2000): USA and China, SMR
- \rightarrow Skriver et al. (2018): Europe, log(SMR)
- \rightarrow Arias et al. (2018): USA, addition variables
- Other Estimation Methods
- →Sormin et al. (2019): HMD, Neural Network
- →Liou et al. (2020): UN, Multilevel Linear Regression
- →Raftery et al. (2013): UN, Bayesian Models

County-level SDR and Longevity (2000-2021)



SDR, SMR Regression Models

- Simple Linear Regression is a popular choice: (1) $Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it}$, $\varepsilon_{it} \sim \text{iid N}(0, \sigma^2)$ Y_{it} can be Life expectancy of age 0 (or 25, 65) and X_{it} is SDR or SMR.
- Extra Variables

(2) $Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \varepsilon_{it}$ such as the proportion of graduate education (%) X₂, the proportion of aborigines (%) X₃.

Table 1. Results of Regression Models (County-level)

	SDR Male	SDR Female	SMR Male	SMR Female
Correlation Coefficient	-0.981	-0.982	-0.975	-0.977
β_0	89.060	92.709	89.967	93.070
β_1	-0.022	-0.030	-14.013	-10.710
Adjusted R ²	0.963	0.964	0.950	0.954
KS test P-value	0.270	0.054	0.310	0.125

SDR and Life Expectancy at age 0 (County-level)



Using SDR & SMR to Predict Longevity (3 year data)



Table 2. Extra Variables Regression Models (County-level)

	Model 1		Model 2	
Gender	Male	Female	Male	Female
β_0 : Intercept	89.060*	92.710*	86.472*	90.716*
β_1 : SDR	-0.022*	-0.030*	-0.019*	-0.026*
β_2 : Graduate (%)			0.119*	0.178*
β_3 :Aborigines (%)			-0.019*	-0.009*
Adjusted R ²	0.963	0.964	0.970	0.971
KS test p-value	0.270	0.054	0.070	0.180

Results of SDR & SMR

- Regression with SDR (or SMR) is a feasible model for estimating life expectancy in Taiwan, with satisfactory results in residual diagnosis and backcasting.
- Adding extra variables is possible, similar to that of Arias et al. (2018, High School Degree and Latino percentage).
- \rightarrow Not a lot of improvement in estimation.

Applying SDR to International Data (Human Mortality Database)



https://www.scor.com/en/expert-views/scors-rd-teams-support-collection-global-mortality-population-data-human-mortality

Using SDR to International Data

• We continue to apply SDR to estimate life expectancy to other countries (HMD). \rightarrow Note: SDR is not linear vs. life expectancy. We consider Piecewise Linear Regression and check if residuals satisfy normality assumption. \rightarrow Dividing into 4 intervals (3 knots) is the best and it also has the smallest backcasting errors. Note: Taiwan data have only 3 intervals.

SDR vs. Life Expectancy at age 0 (HMD,1991-2020)



Piecewise Regression (Male Life Expectancy age 0)

Male



Piecewise Regression (Female Life Expectancy age 0)

Female



Backcast of Piecewise Regression (Male Life expectancy age 0)



Backcast of Piecewise Regression (Female Life expectancy age 0)



Bayesian Hierarchical Model

- Raftery et al. proposed Bayesian Hierarchical Model, using MCMC to construct model, to predict future life expectancy.
- Regression model vs. Bayes model
- →Predicting life expectancy of future 5 years for 38 HMD countries, based on data from 1950 to present (5 backcasting periods).
- \rightarrow Regression model has smallest MAPE.

Backcast of Piecewise and Bayes Regression (HMD)



Longevity of Piecewise & Bayes Regression (Taiwan)



Conclusion

Using SDR & SMR to estimate life expectancy is a feasible approach for small areas, with respect to estimation and backcast results. \rightarrow Backcast MAPE < 0.7% (0.5 year in Taiwan). SDR is more accurate than Bayes model. \rightarrow We need future SDR to predict life expectancy using regression model, but Bayes model does not require SDR.

Suggestions

Regression coefficients depend on SDR values. \rightarrow SDR Ranges are 585 & 366 for Taiwan men & women, 1545 & 676 HMD men & women. \rightarrow Need to select appropriate training data. Compare with the partial SMR approach (Wang et al., 2018) for population fewer than 50,000. Do results of SDR depend on population size? \rightarrow Simulation study under certain mortality model.

Thank you for your Attention!

