Mortality Models and Longevity Risk for Small Populations

Jack C. Yue
National Chengchi Univ.
Date: Sept. 8, 2015
Email: csyue@nccu.edu.tw

Summary

- Small Populations and their Estimates
- Graduation and the Proposed Approach
- Computer Simulation
- Empirical Studies
- Conclusion and Discussions

Life Table Construction

- Smoothing the mortality rates (or graduation) is often necessary in constructing life tables.

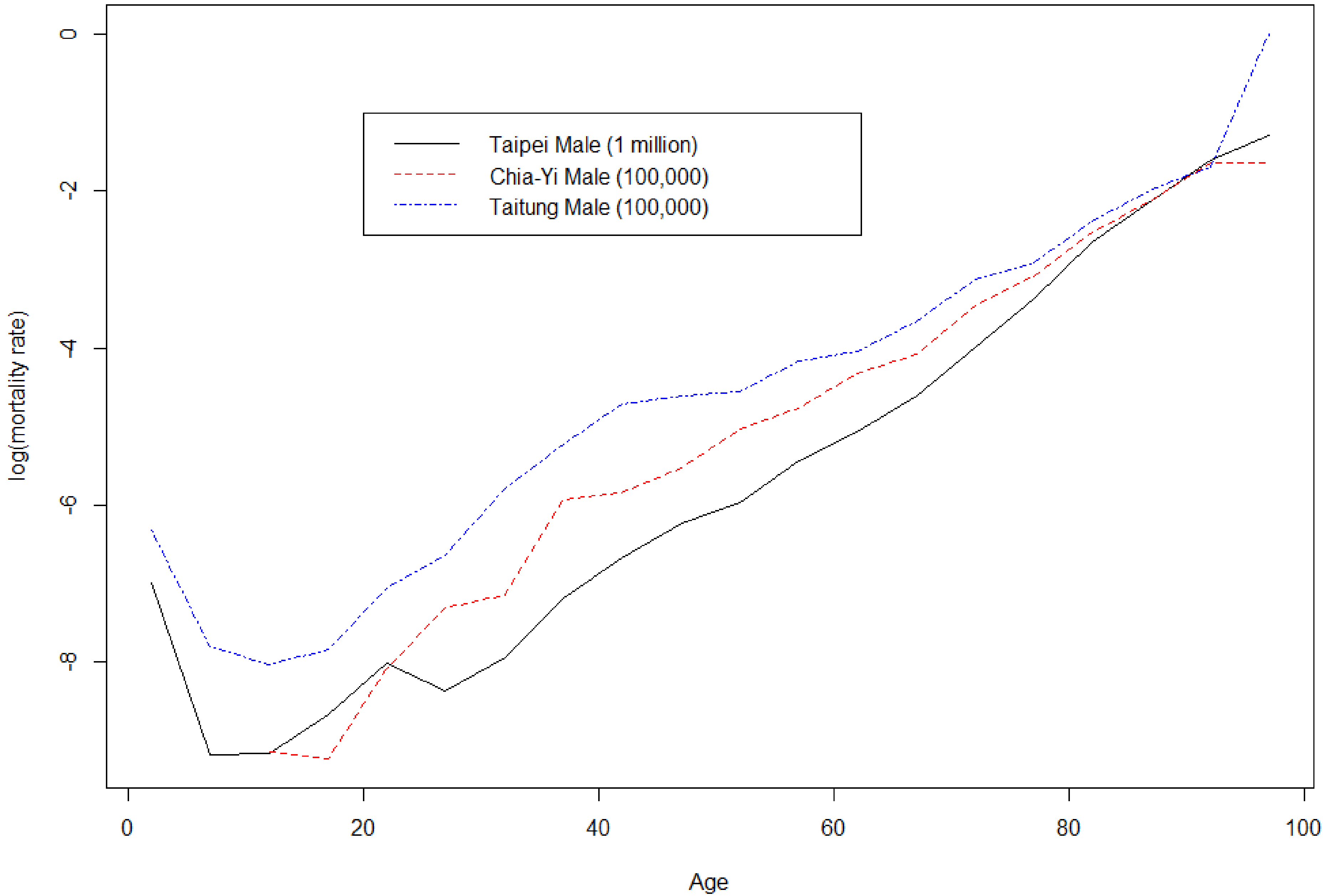
→Especially for younger ages and the elderly.

- Small areas need extra care!!

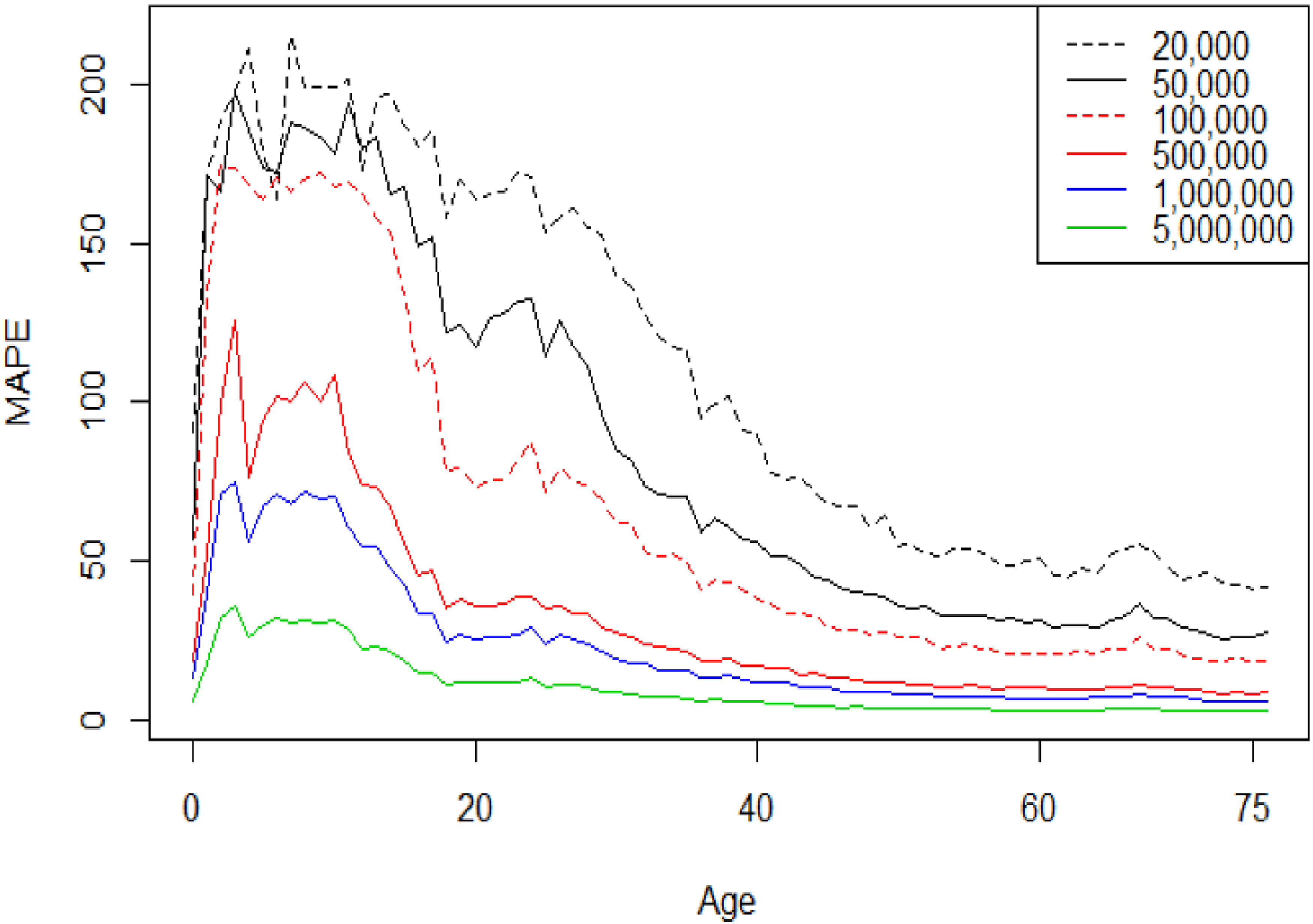
→Variance $\propto$ 1/(Sample size)

→The estimation can be unstable for small populations, even applying parametric models.

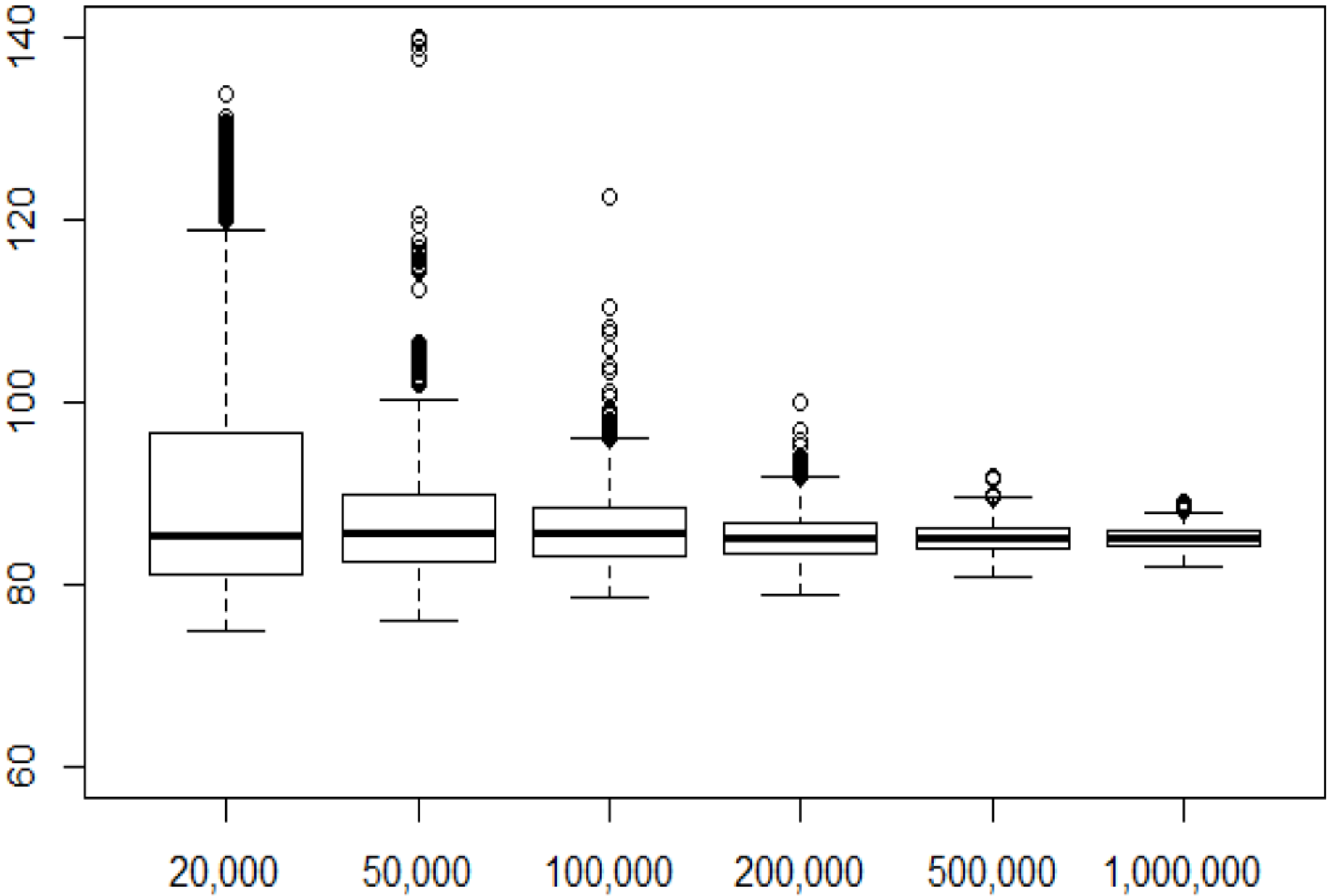
County-level Mortality Rates in Taiwan



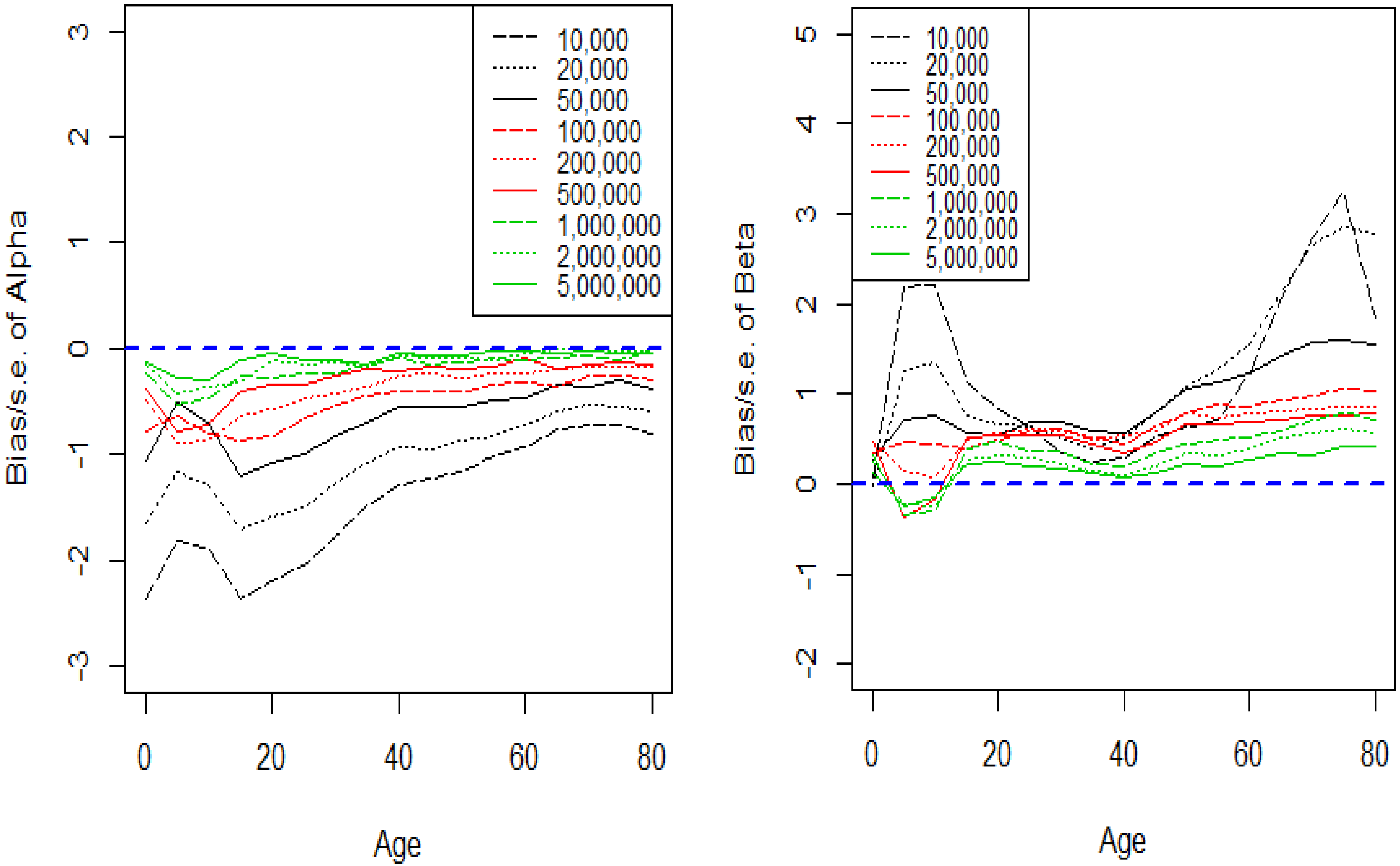
Estimation Error vs. Population Size (Taiwan Female)



Life Expectancy vs. Population Size (Taiwan Female)

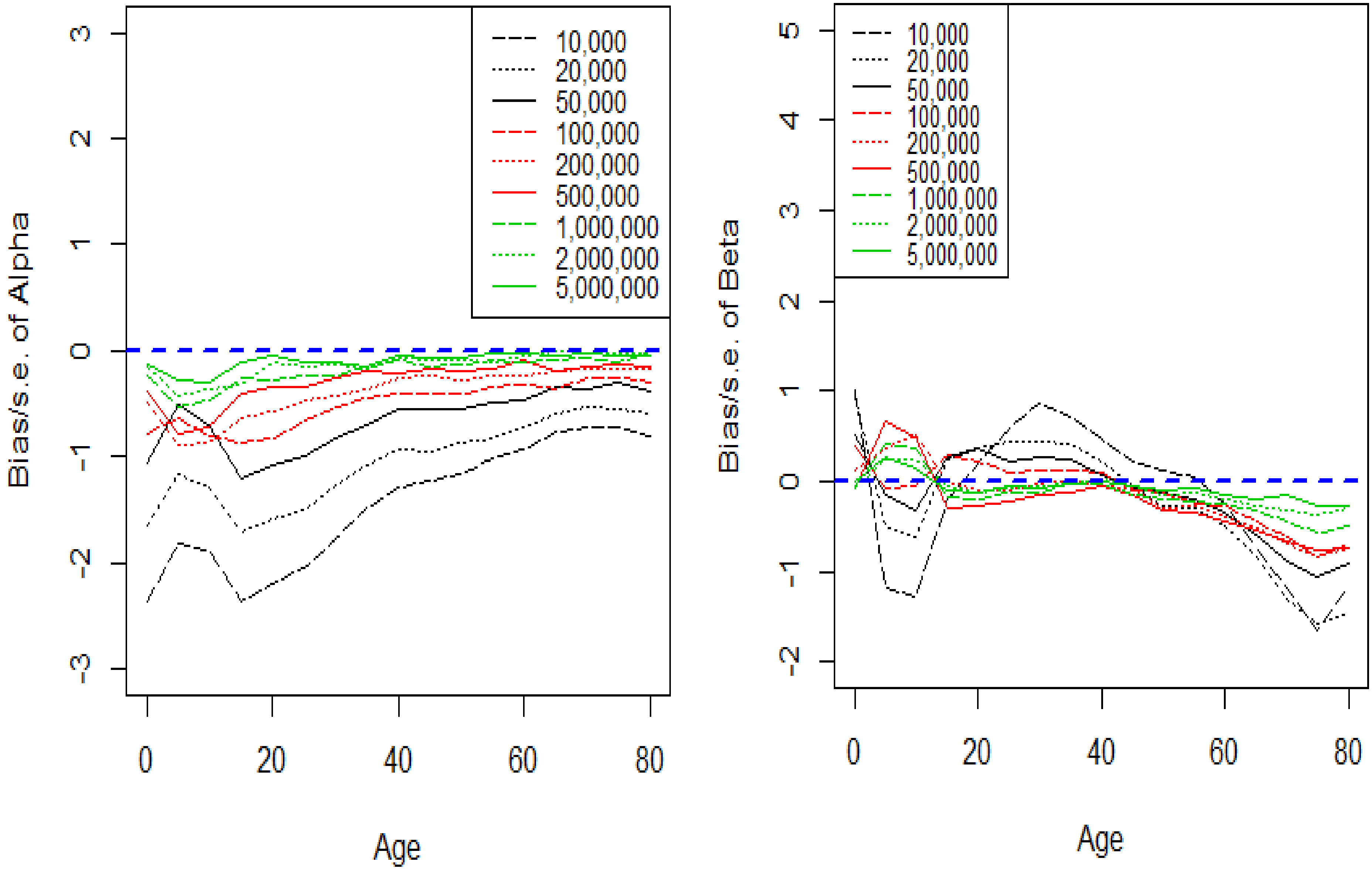


Lee-Carter Model (SVD)



"t-ratio" of $\hat{\alpha}_x$ & $\hat{\beta}_x$ estimates for Lee-Carter Model

Lee-Carter Model (Approximation)



"t-ratio" of $\hat{\alpha}_x$ & $\hat{\beta}_x$ estimates for Lee-Carter Model

Study Objective

- Develop SOP for graduating mortality rates of small areas, as well as their predictions.

→Suggest graduation methods according to the population size and mortality profile of the target area.

→Explore the limitations of parametric models and propose feasible modifications.

About the Graduation

- Increasing the sample size is the most intuitive way of stabilizing mortality estimates.

→Traditional graduation is to accumulate data with similar mortality attributes (e.g., same age for 3 or 5 consecutive years, ages x–1~x+1 or x–2~x+2 for single year).

→Combining data from populations with similar mortality profile is another possibility (e.g., Bayesian graduation).

The Proposed Approaches

- According to the data aggregation, we can classify the graduation methods into 4 groups, same area or not vs. one year or more.

→Traditional graduation methods usually are "same area & one year."

→Parametric models are of the type "same area & multiple years."

Note: We focus on (same area, multiple years) and (multiple areas, one year).

Lee-Carter Model & Graduation Methods

- Lee-Carter model (Lee & Carter, 1992) assumes that

$$\log(m_{x,t}) = \alpha_x + \beta_x \cdot \kappa_t + \varepsilon_{x,t}$$

where x is age, t is time, and α_x, β_x, κ_t are parameters. κ_t is a linear function of time.

- Greville's 9-term formula (1974) for single age:

$$q_x = \frac{1}{2431}(-99q'_{x-4} - 24q'_{x-3} + 288q'_{x-2} + 648q'_{x-1} + 805q'_x + 648q'_{x+1} + 288q'_{x+2} - 24q'_{x+3} - 99q'_{x+4})$$

□ Whittaker

→Minimizing the sum of Fit and Smoothness:

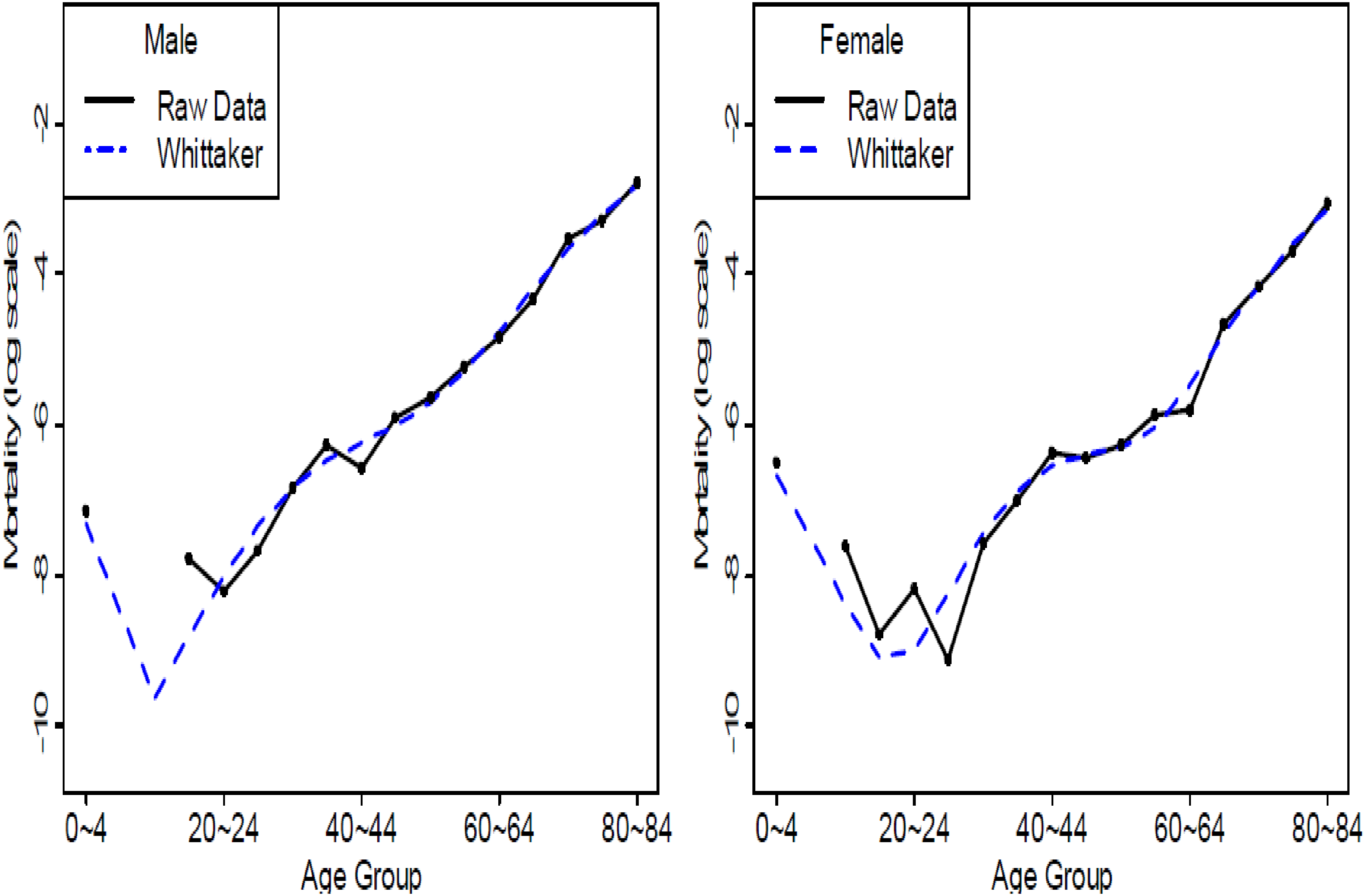
$$M = F + hS = \sum_{x=1}^{n} w_x (v_x - u_x)^2 + h \sum_{x=1}^{n-z} (\Delta^z v_x)^2$$

□ Partial SMR (Standard Mortality Ratio)

Lee (2003) proposed using the partial SMR (connection between large and small areas) to modify the mortality rates of small area:

$$v_x = u_x^* \times \exp\left(\frac{d_x \times \hat{h}^2 \times \log(d_x / e_x) + (1 - d_x / \sum d_x) \times \log(\mathrm{SMR})}{d_x \times \hat{h}^2 + (1 - d_x / \sum d_x)} \right)$$

$$\mathrm{SMR} = \frac{\sum_x d_x}{\sum_x n_x \cdot u_x^*} \qquad \hat{h}^2 = \max\left(\frac{\sum \left((d_x - e_x \times SMR)^2 - \sum d_x \right)}{SMR^2 \times \sum e_x^2} \text{ , } 0 \right)$$



Example of Whittaker Graduation (Population 230,000)

Simulation Setting

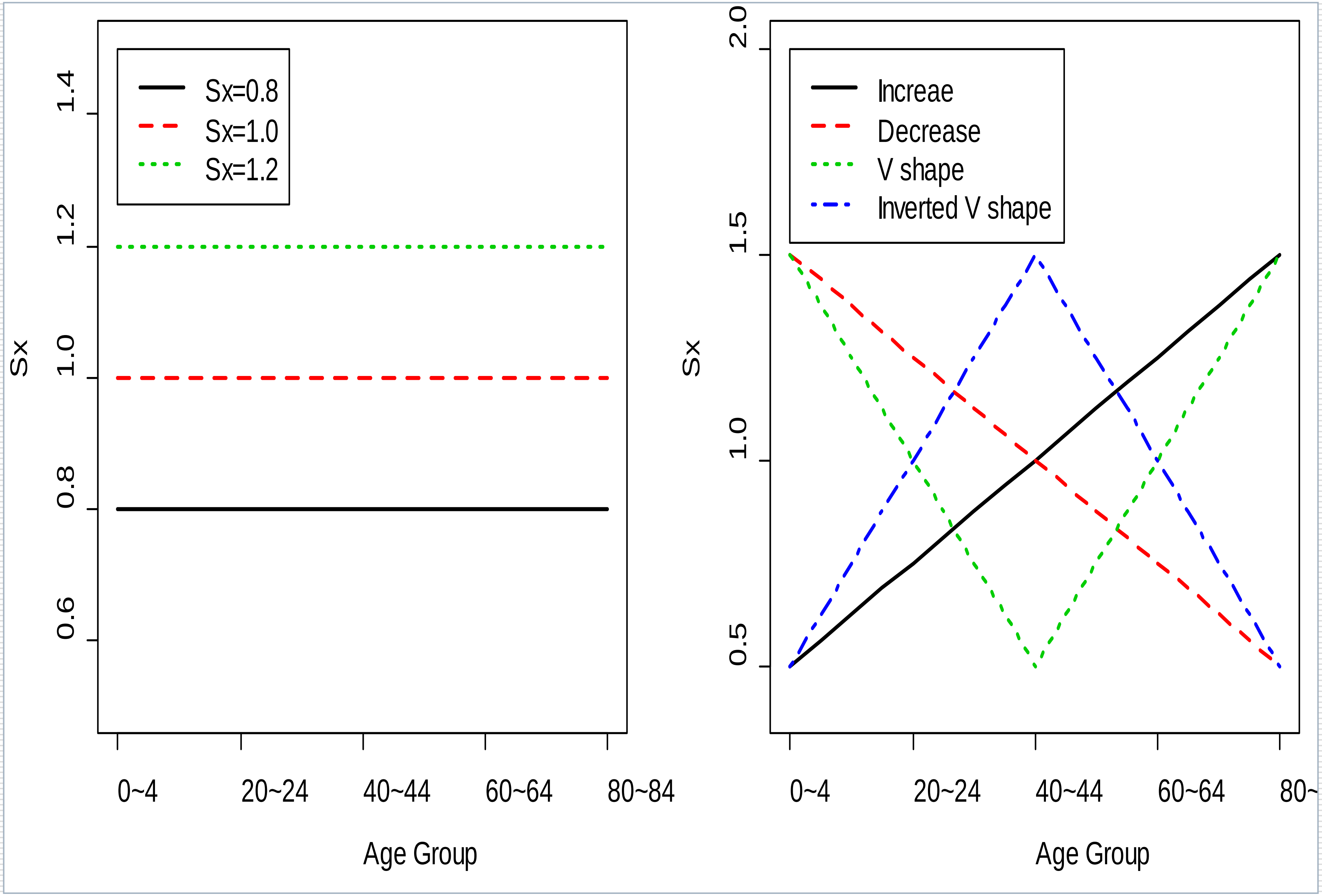
- The reference population is larger than the small population, and the mortality rates of reference population satisfy the LC model.
- The mortality rates of small population follow one of 7 mortality scenarios:

→Similar to the reference group (3 cases)

→Differ to the reference group (4 cases)

Note: We use mortality ratio $s_x = \dfrac{q_x}{q_x^*}$ to measure.

Seven Mortality Scenario



Simulation Setting (cont.)

- Taiwan is the reference population and counties in Taiwan are the small populations.

→5-age group (0-4, 5-9, …, 80-84)

→Training vs. Testing Periods

- Comparison criterion:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \times 100\%$$

Estimation Errors of Greville & Whittaker Methods

MAPE (%)

	10,000	20,000	50,000	100,000	200,000	500,000	1 mill.	2 mill.
Raw	125.56	101.45	73.01	54.89	39.40	24.60	17.45	12.32
Whittaker	89.41	68.06	45.75	33.44	24.92	17.61	14.14	11.75
Greville	87.15	66.36	43.55	30.83	21.85	13.92	9.96	7.20

Note: Target area is Taiwan 1-age male (1990-2009)

Multiple Areas & One-year Methods

- Enlarging the data of small area via a large population (various mortality scenarios).

→Use Partial SMR and Whittaker ratio (applying Whittaker method to the mortality ratio s_x.)

- We will only show the mortality scenarios of constant ($s_x = 1 + a$), increasing, and V-shape.

→ Population size of small area = 50,000 and 200,000.

“Multiple Areas & One Year” – Constant Scenario

MAPE (%)

(a) Population size = 50,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	30.1	28.7	27.6	26.5	25.3	24.7	23.3	23.0	22.4	21.8
Whittaker_R	13.2	12.8	12.5	12.3	11.9	11.7	11.5	11.4	11.1	10.9
PSMR	5.0	4.6	4.6	4.2	4.1	4.1	3.9	3.8	3.7	3.7

(b) Population size = 200,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	15.0	14.3	13.6	13.1	12.7	12.3	11.9	11.5	11.3	10.9
Whittaker_R	8.5	8.3	8.0	7.7	7.4	7.3	7.0	6.9	6.8	6.6
PSMR	2.5	2.4	2.2	2.2	2.1	2.0	2.0	1.9	1.9	1.8

"Multiple Areas & One Year" – Increasing Scenario

MAPE (%)

(a) Population size = 50,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	30.2	32.4	36.0	41.3	47.6	56.2	66.2	79.8	99.2	143.3
Whittaker_R	13.3	15.1	18.5	23.2	28.9	36.3	45.3	57.9	78.0	122.3
PSMR	4.9	7.5	12.7	19.4	27.4	37.0	48.8	64.9	88.8	140.9

(b) Population size = 200,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	14.9	17.2	22.2	27.8	35.5	44.0	54.7	68.3	90.0	132.9
Whittaker_R	8.4	10.4	14.6	19.4	25.8	32.8	42.0	53.5	72.5	112.8
PSMR	2.4	6.2	12.5	19.9	28.5	38.2	50.1	65.6	89.4	138.6

“Multiple Areas & One Year” – V-shape Scenario

MAPE (%)

(a) Population size = 50,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	30.2	31.0	33.6	38.0	43.7	51.4	60.1	72.7	90.9	130.6
Whittaker_R	13.3	14.2	17.1	21.5	27.1	33.7	41.4	52.4	68.7	105.8
PSMR	4.9	7.5	12.5	18.8	26.4	35.2	45.4	59.1	79.2	123.0

(b) Population size = 200,000

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	14.9	16.7	21.6	27.4	34.4	42.9	52.9	65.8	85.2	125.0
Whittaker_R	8.5	10.5	14.9	20.5	26.9	34.5	43.4	54.9	72.2	109.5
PSMR	2.4	6.2	12.4	19.5	27.5	36.5	47.0	60.5	80.3	121.0

Simulation Results (Multiple Areas)

- Partial SMR and Whittaker ratio still have smaller errors in the case of enlarging the data of small area via a large population.

→Partial SMR is better when the similarity level between different ages is higher.

- Using the partial SMR & treat the aggregation of historical data as the large population.

→We expect good mortality estimation unless the mortality pattern is not regular.

Est. Errors of “Same Area & One/multiple years”

MAPE (%)

	10,000	20,000	50,000	100,000	200,000	500,000	1 mill.	2 mill.
Raw	68.23	50.59	32.90	22.88	16.28	10.27	7.26	5.12
Whittaker	51.54	38.20	27.62	22.68	19.82	17.70	16.88	16.52
MA(3)	83.99	75.06	69.69	67.92	67.33	67.07	67.00	67.05
Lee-Carter	33.57	23.67	15.53	10.97	8.66	6.05	4.05	2.64
PSMR	14.31	11.75	9.68	8.70	8.09	7.50	7.03	6.48

Note: Target area is Taiwan 5-age male (1990-2009)

Est. Errors of "Same Area & One/multiple years"

MAPE (%)

	10,000	20,000	50,000	100,000	200,000	500,000	1 mill.	2 mill.
Raw	70.80	54.35	35.34	24.86	17.53	11.07	7.84	5.53
Whittaker	53.60	40.31	28.44	23.29	20.00	17.66	16.72	16.25
MA(3)	92.79	82.84	75.79	73.65	72.81	72.48	72.27	72.29
Lee-Carter	32.89	22.80	14.38	10.32	7.84	5.59	3.92	2.67
PSMR	28.13	26.20	24.65	23.79	22.97	21.51	19.90	17.76

Note: Target area is Pen-Hu 5-age male (1990-2009)

Conclusion

- The idea of increasing sample size can be used in small area estimations.

→Graduations of (same area, multiple years) and (multiple areas, one year) are recommended.

Note: (same area, multiple years) graduation can be treated an alternative approach to parametric mortality models (e.g. LC model).

→The proposed approach has smaller estimation errors for small areas.

Discussions and Future Study

- Modify the proposed approach and compare with the coherent Lee-Carter model.

→ From (same area, multiple years) to (multiple areas, multiple years)

- Simulation methods (e.g. Block Bootstrap) for the (same area, multiple years) graduation.
- Need to consider if the mortality improvement varies in different time periods.

Thank you
for your Attention!

Q & A