Spatial Smoothing and Small-Sample Bias in Mortality Forecasting

Jack C. Yue
National Chengchi University
Date: Sept. 9, 2026
Email: csyue@nccu.edu.tw

Global Life Prolonging

- Life expectancy continues to rise, occasionally affected by factors such as disease & war.

→COVID-19 reduced life expectancy by 1.5 years.

- Taiwan's Life expectancy continues to increase.

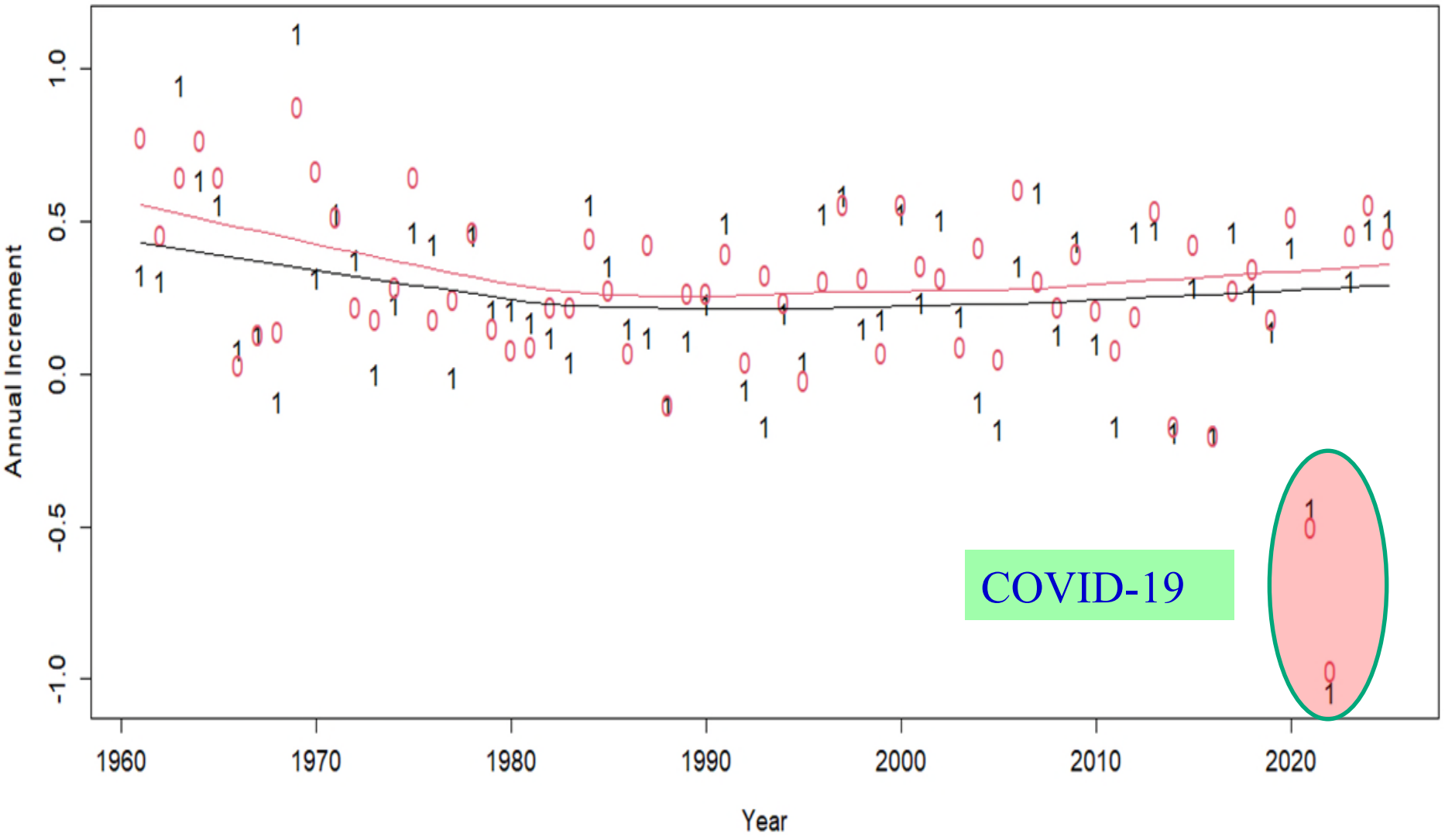
→Annual increment appears to remain unchanged.

→Mortality models are often used to predict life expectancy.

Note: Lee-Carter model is a fine choice.

https://s3.amazonaws.com/lowres.cartoonstock.com/death-life_expectancy-grim-grim_reaper-reaper-beggar-jcen178_low.jpg

Taiwan's Annual Increment Remains Unchanged



Lee-Carter Model

- Lee and Carter (1992) proposed

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

where α_x → Age-specific average mortality

β_x → Age-specific mortality improvement

κ_t → Time-varying mortality index

Note: SVD, Approximation, & Poisson (or MLE) Estimation.

Sample Size and Estimation

- Variance of estimates is inversely proportional to the sample size.

→Parameter estimation may also be biased for the small populations.

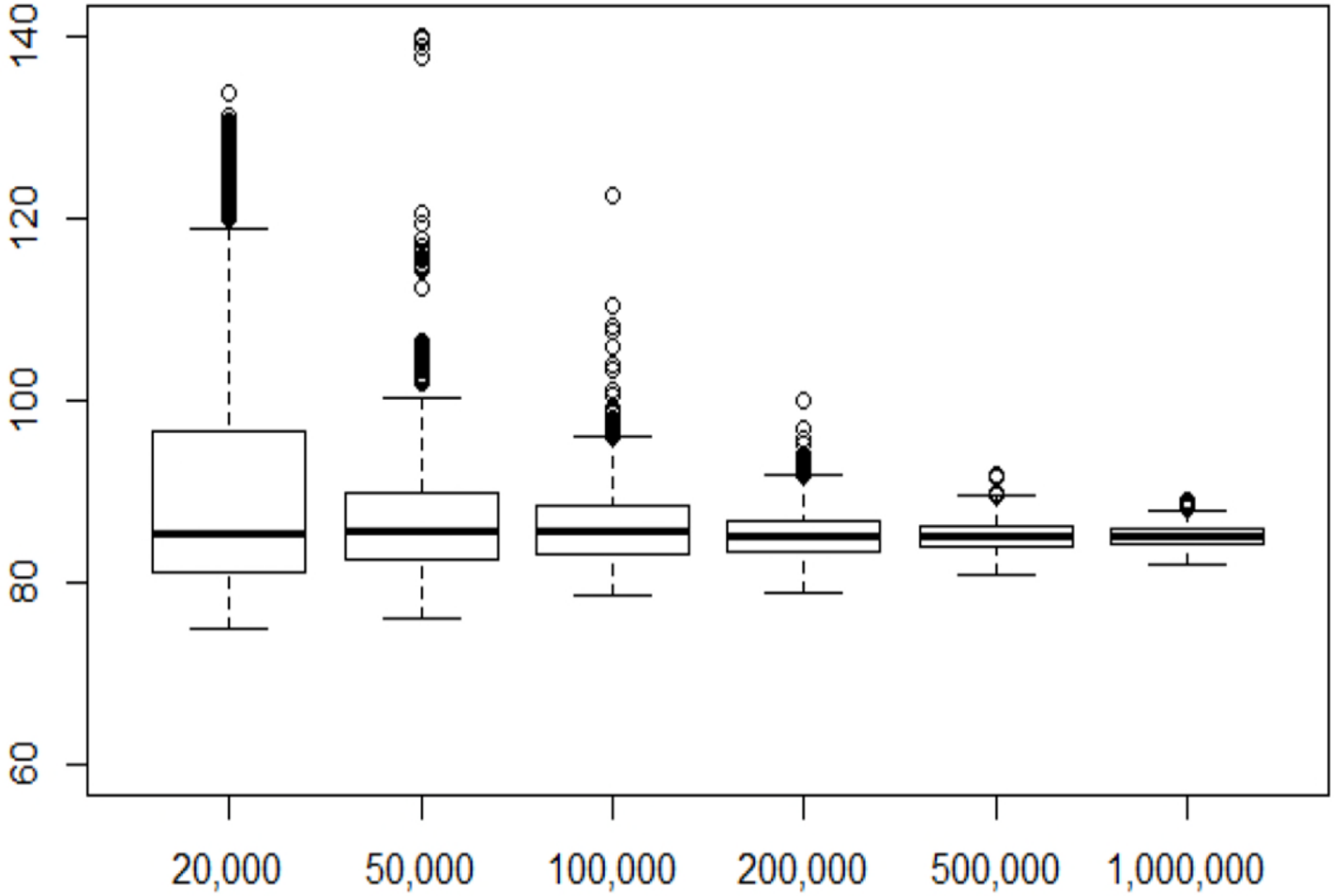
- Appling Lee-Carter model to small populations:

→α_x estimates are under-biased,

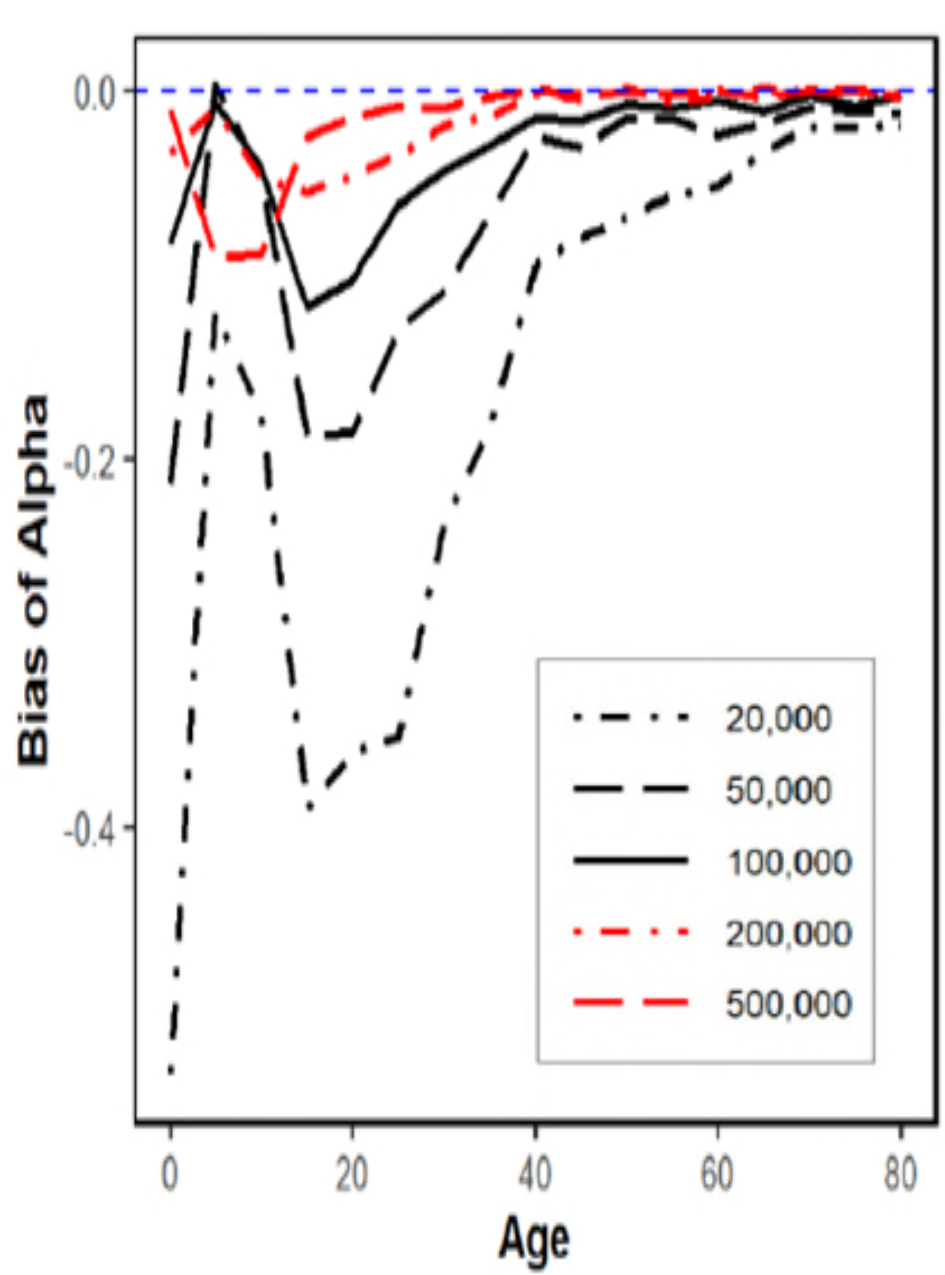
→β_x estimates can be over- or under-biased in younger ages.

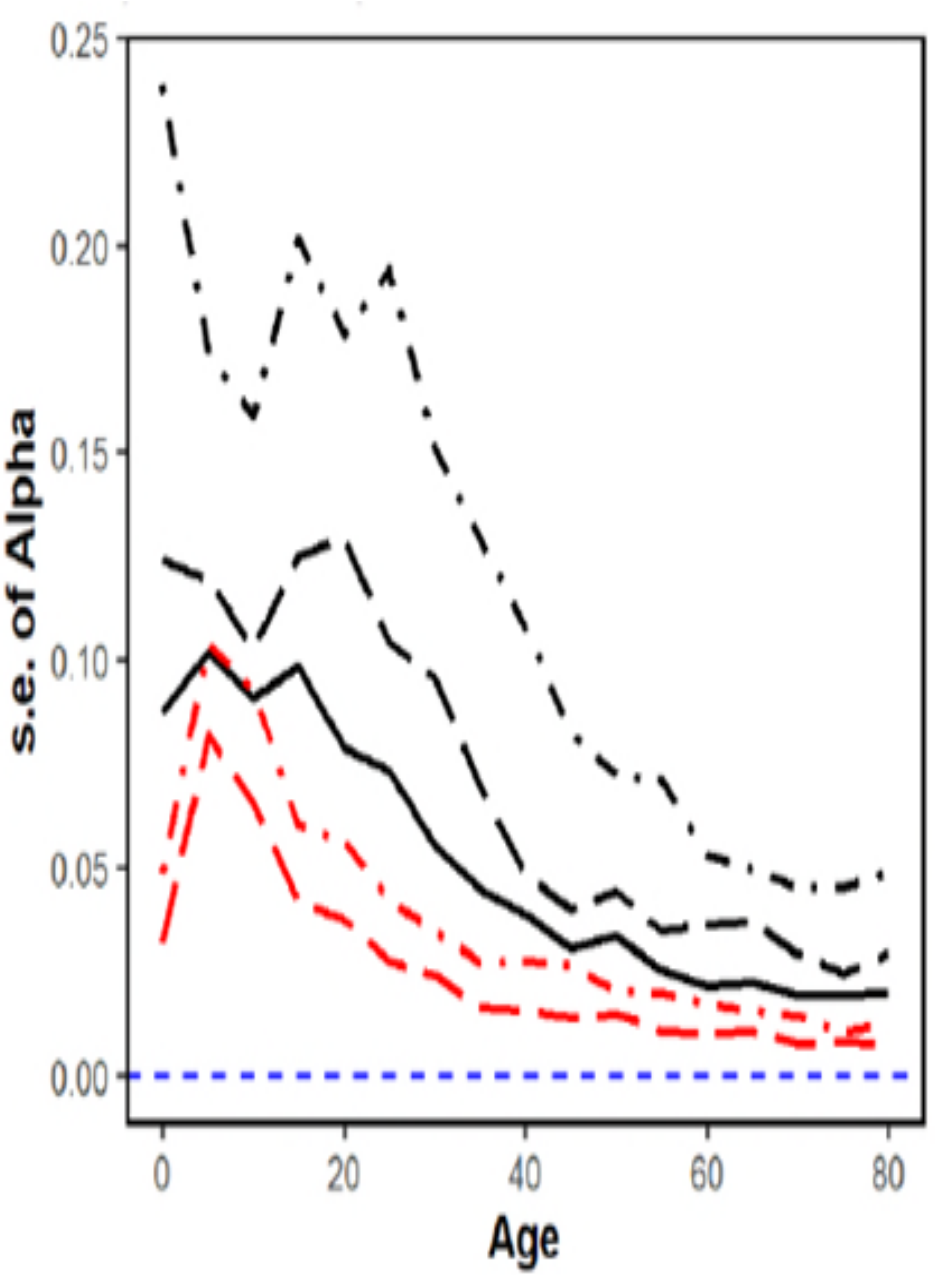
https://www.cartoonstock.com/directory/m/margin_of_error.asp?srsltid=AfmBOoqzbMzEds4AczV4gT6OfHstxYp8Xy_vEOxrVbTkQnVaHx63USVD

Variance decreases with sample size

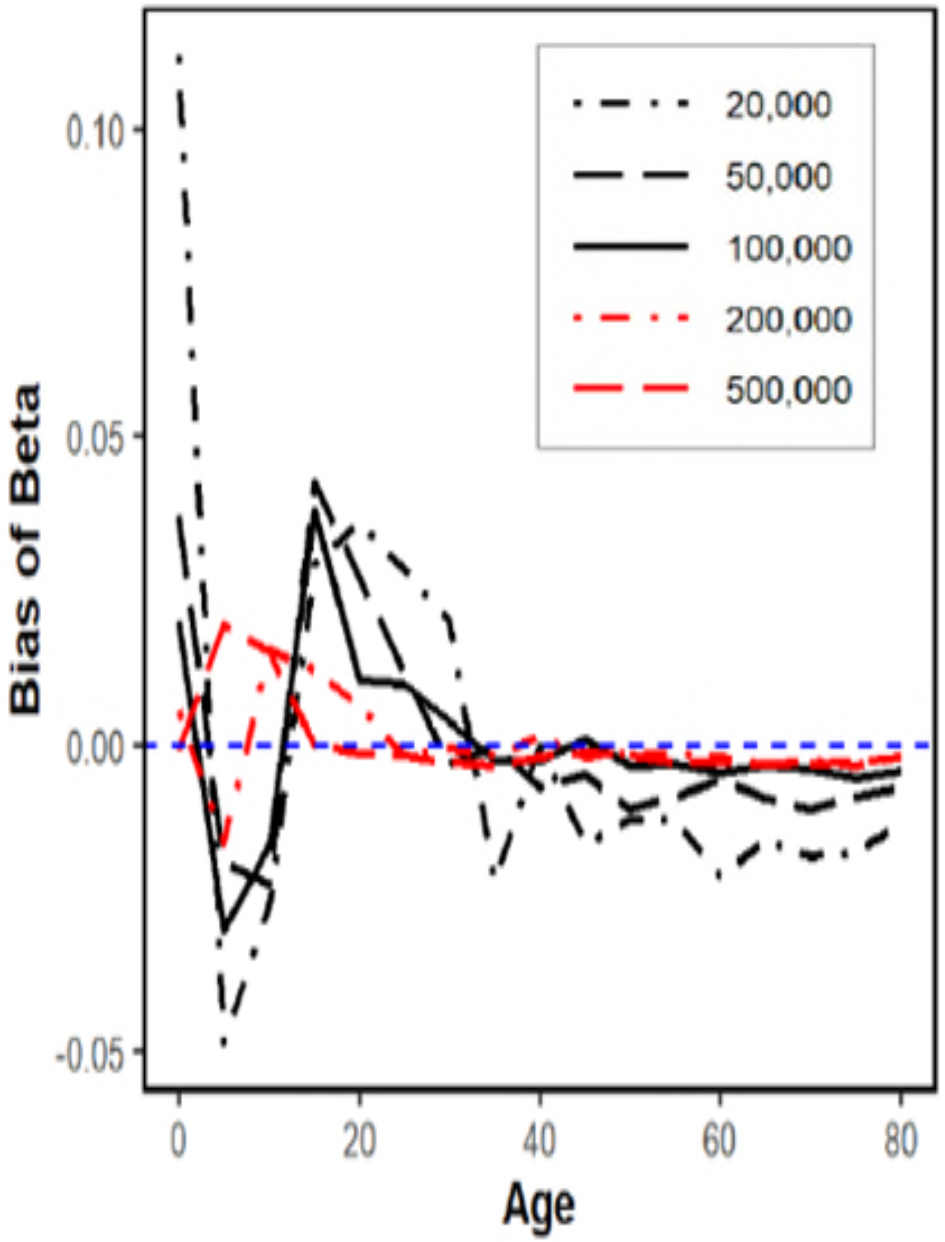


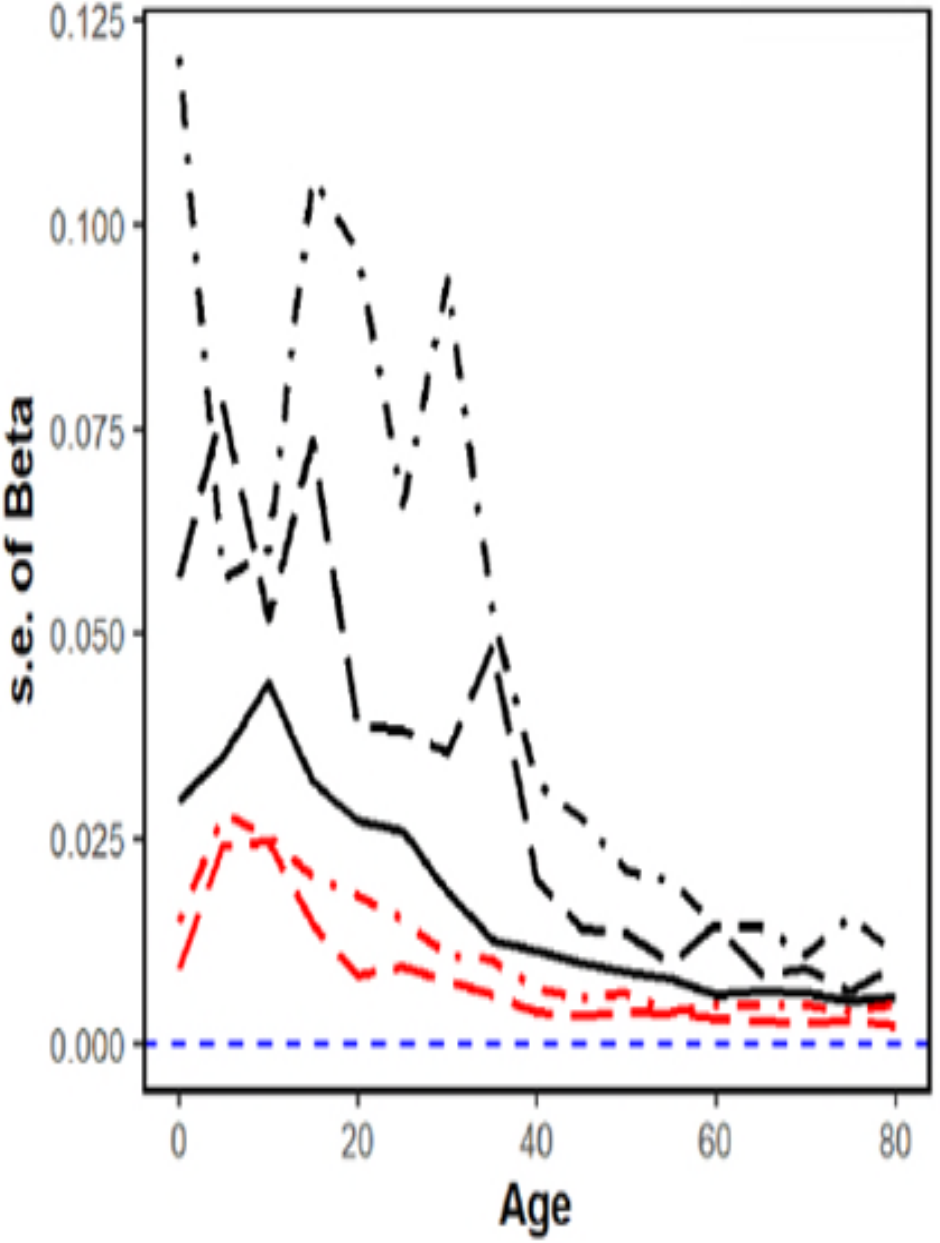
α_x Estimates for the Lee-Carter Model





β_x Estimates of the Lee-Carter Model





Bias Reduction in α_x & β_x

- The low number of deaths leads to high sampling variability in mortality rates.

→ α_x is the average of log(mortality) & the bias is most significant at about ages 0 & 10.

→ The bias of β_x is most obvious around ages 0, 10, & 20.

Note: The most effective way to reduce the bias of parameter estimates is to increase sample size.

Modification of Lee-Carter Model

- Modifications of Lee-Carter model are proposed to deal with small populations

→Poisson Lee-Carter model (Brouhns et al., 2002)

→Multiple populations (Li & Lee, 2005)

→Bayesian Lee-Carter (Czado et al., 2005)

→Graduation+Lee-Carter (Wang et al., 2018)

→ Spatial Lee-Carter (Leong & Yue, 2026)

Note: These methods can be seen as *increasing the sample size*.

Study Objective

- Mortality data in age and time format can be viewed as two-dimensional coordinates.

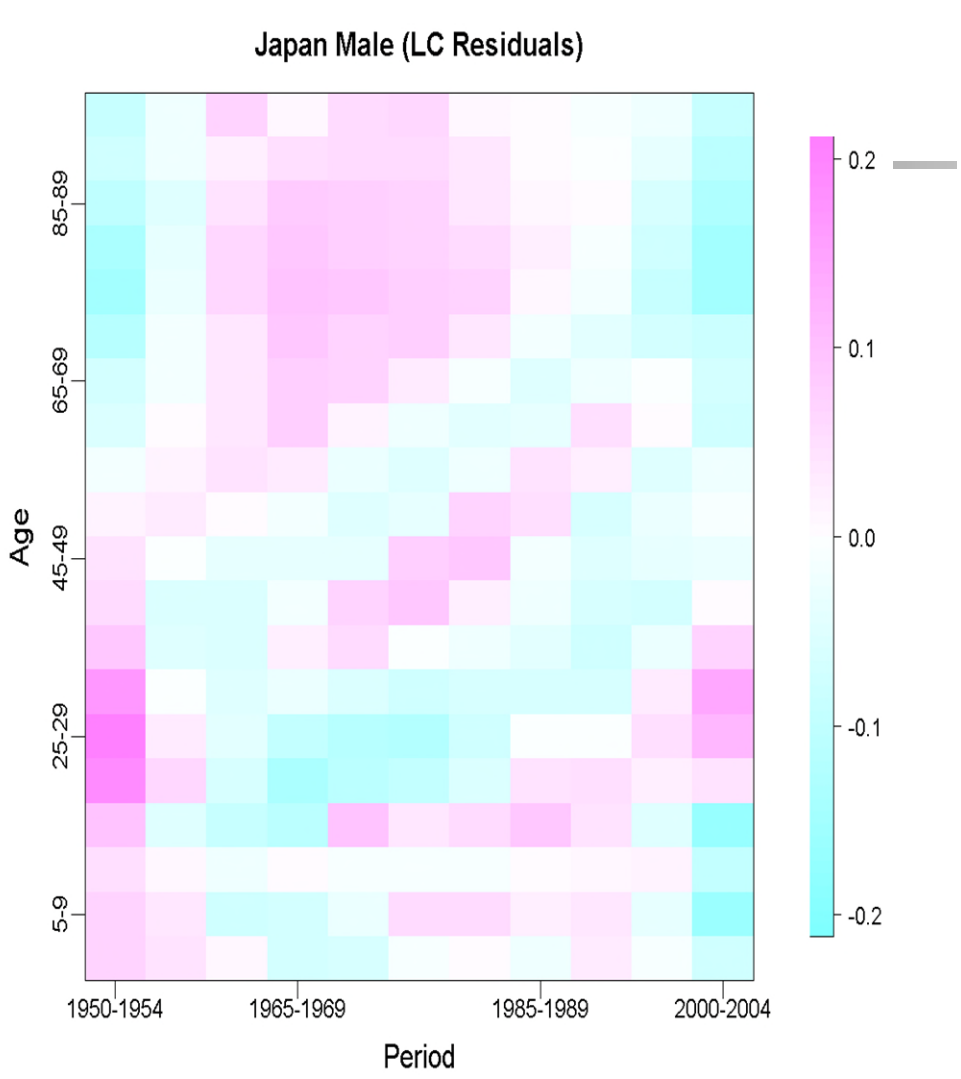
→Enlarge the sample size by including nearby cells.

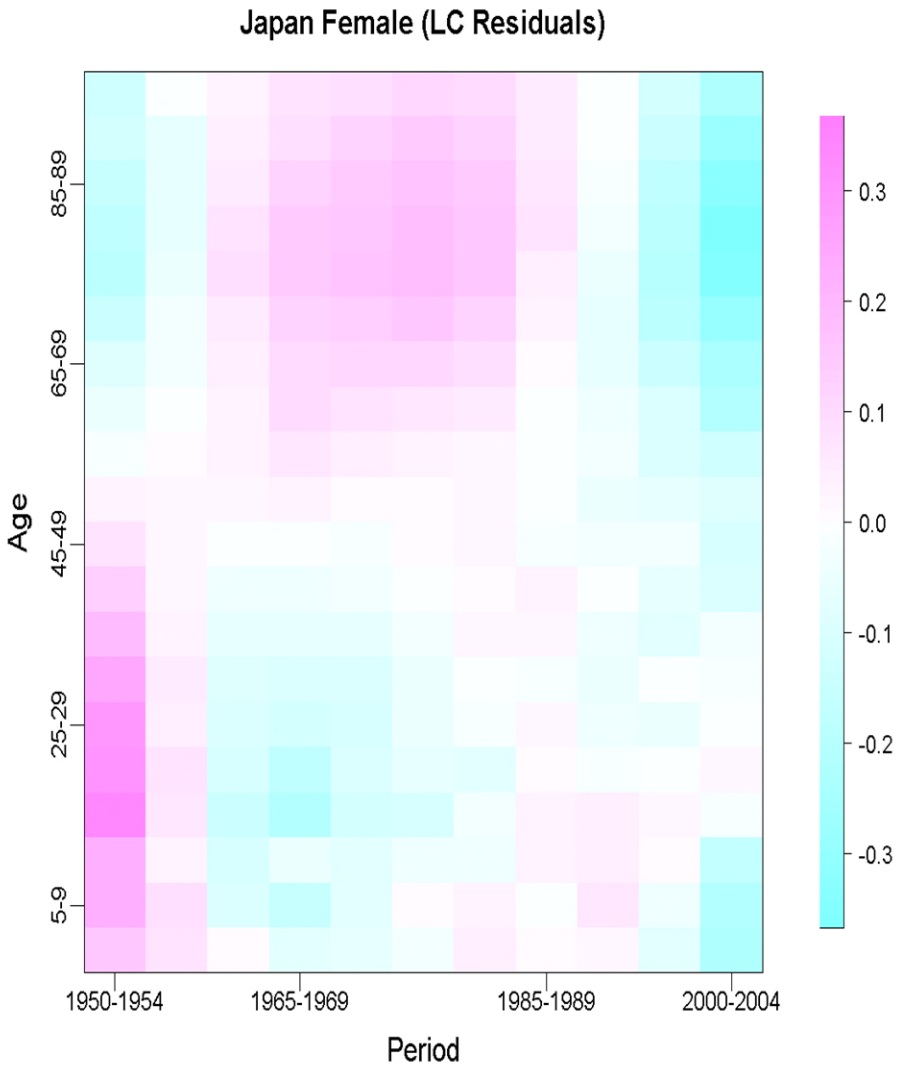
- Spatial modification of *Data Augmentation*.

→Partial SMR smooths mortality rates using the reference population in 1-dim.

→Treat age × time as 2-dim coordinates and apply GWR to increase the effective sample size.

Residuals of LC model (Japan)





Partial Standard Mortality Ratio

- Lee (2003) used SMR to smooth mortality rates based on a large reference area.

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

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

→ n_x & d_x: Numbers of people and deaths aged x

→ u_x^*:Mortality rates aged x in the large population。

GWR(Geographically Weighted Regression)

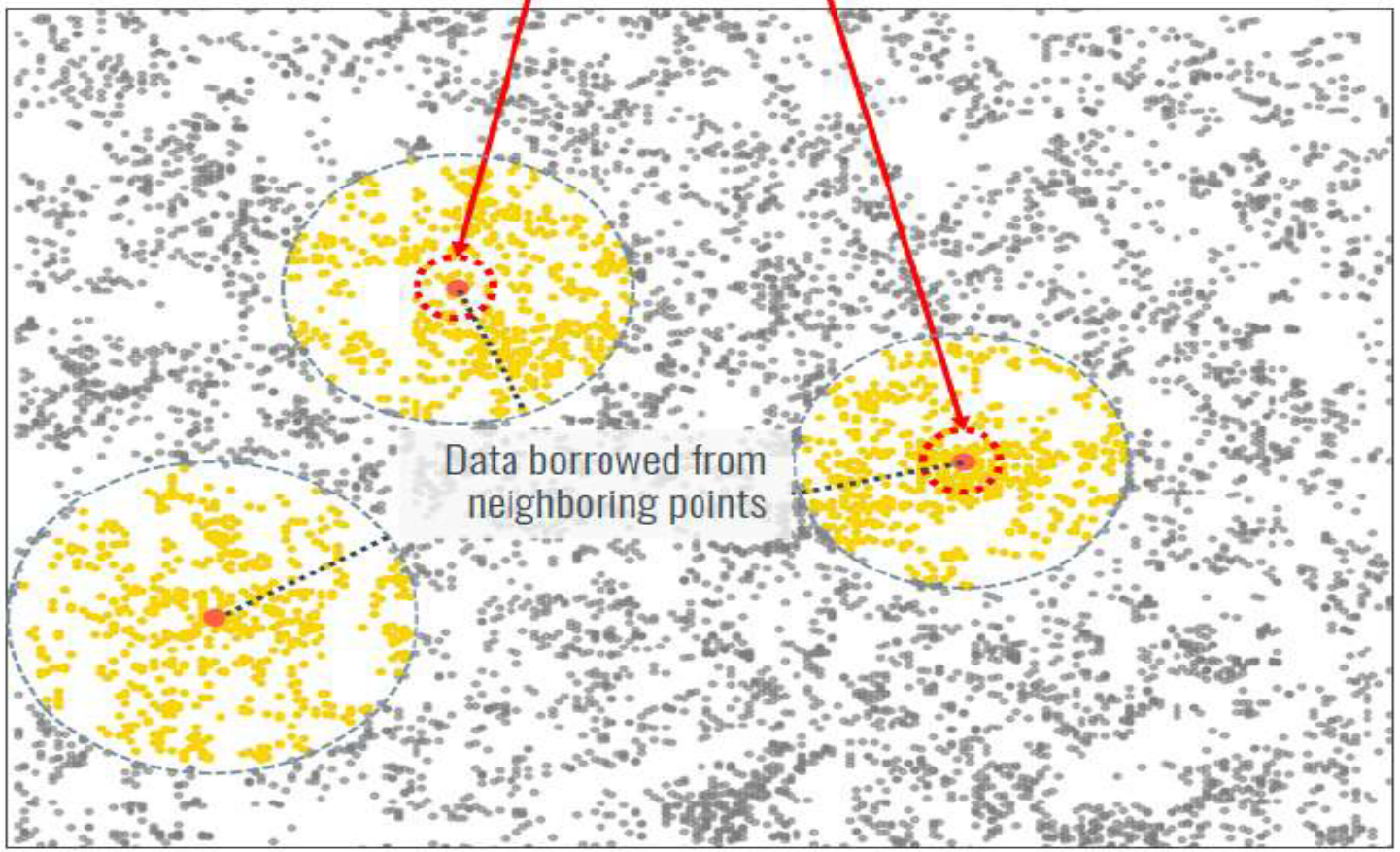
- (GWR) is a modelling technique designed to deal with spatial non-stationarity, i.e., the mean values vary by locations.

$$Y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i$$

Note: Determine how to implement GWR and select the optimal bandwidth.

Bandwidth Choice Alters GWR Estimates

$$y_i = \sum_j \beta_{ij}(u_i, v_i) X_{ij} + \varepsilon_i$$



https://mehak-sachdeva.github.io/MGWR_workshop_book/Html/Methodology

Computer Simulation Setting

- Mortality rates follow Lee–Carter model and parameters are estimated from Taiwan male mortality data, 1992–2020.

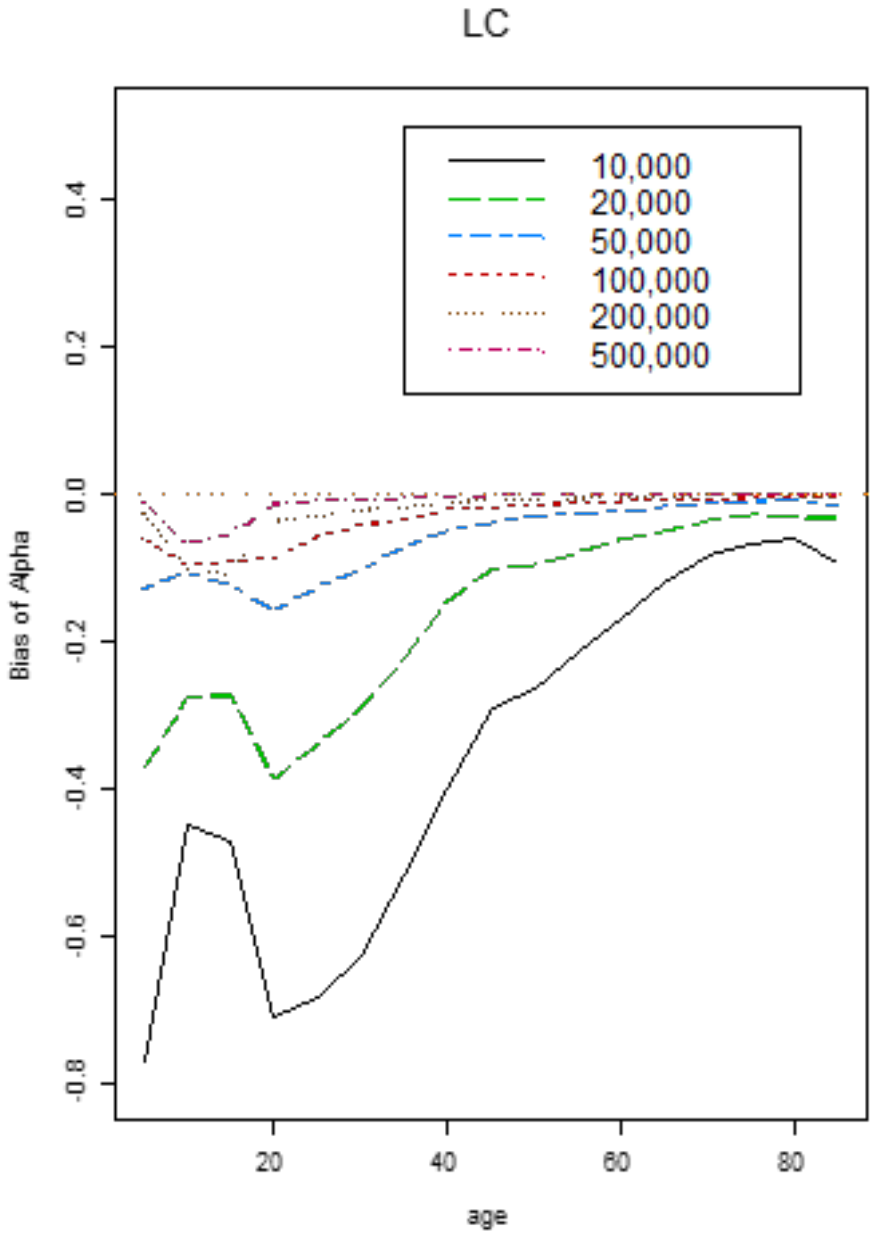
→Simulate no. of deaths using Poisson distribution.

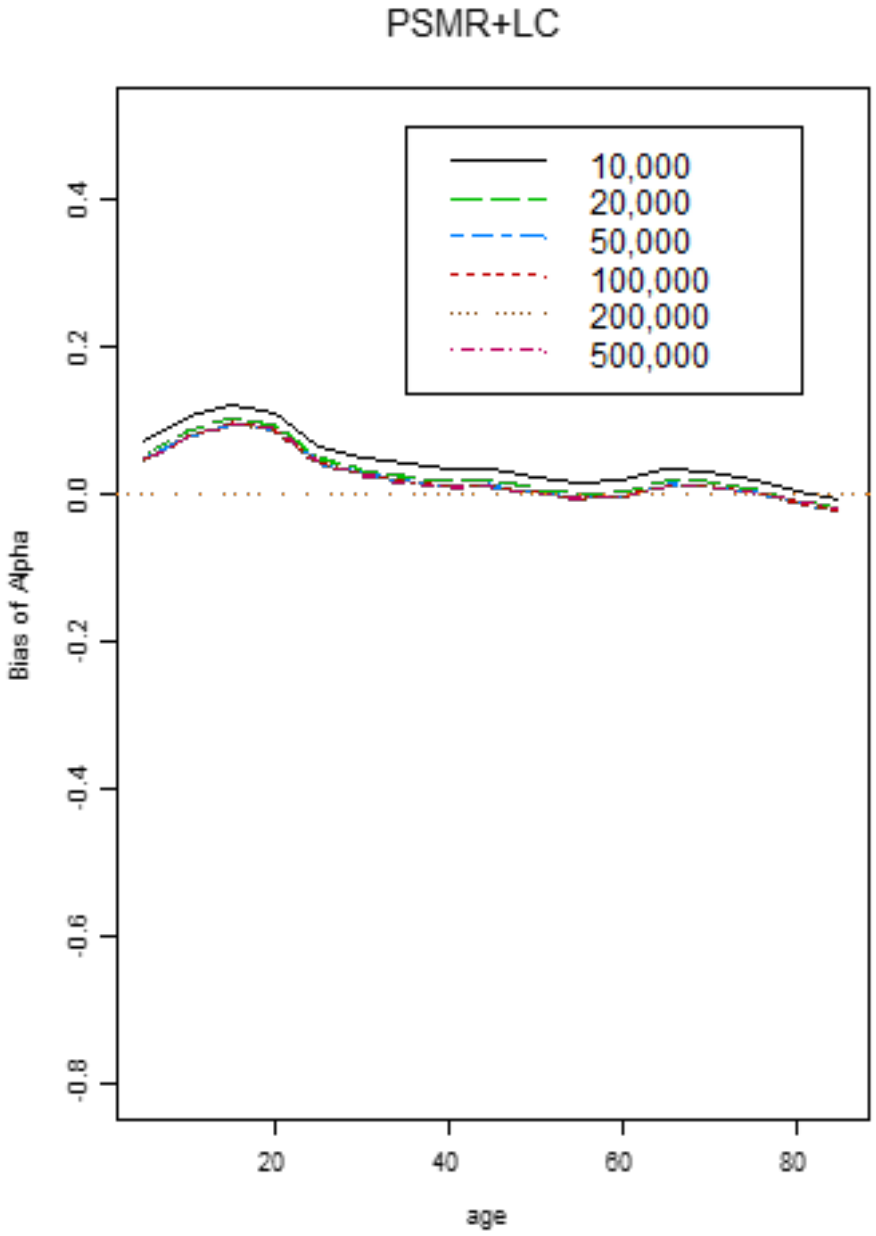
→Partial SMR reference: Pool data across all years within each age group (1992–2020).

→5-Age groups: 0–4, 5–9, 10–14, …, 80–84 (17).

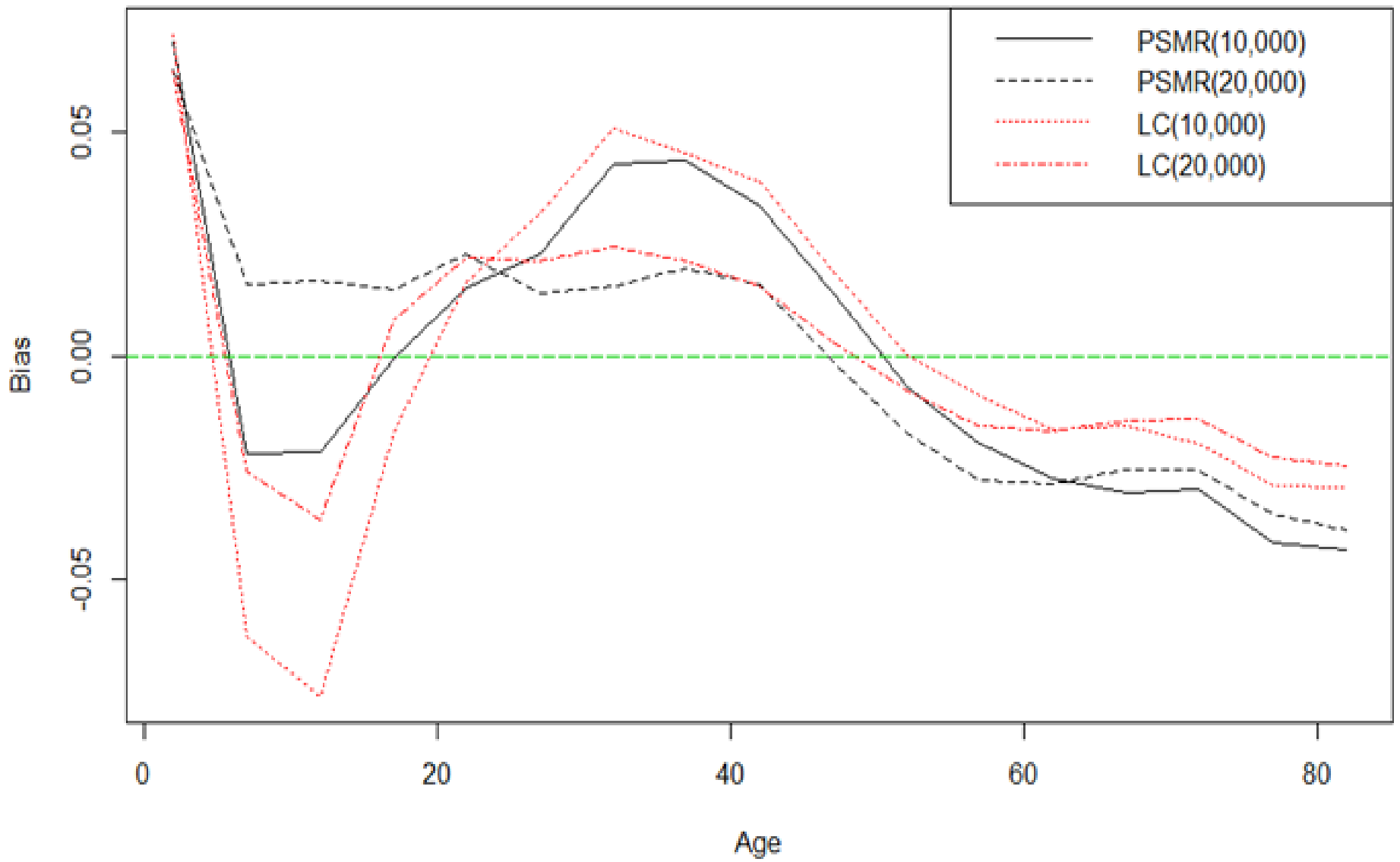
→Population sizes: 10,000, 20,000, 50,000,100,000, 200,000, and 500,000.

α_x Estimates of Partial SMR





β_x Estimates of Partial SMR



Partial SMR & Parameter Estimation

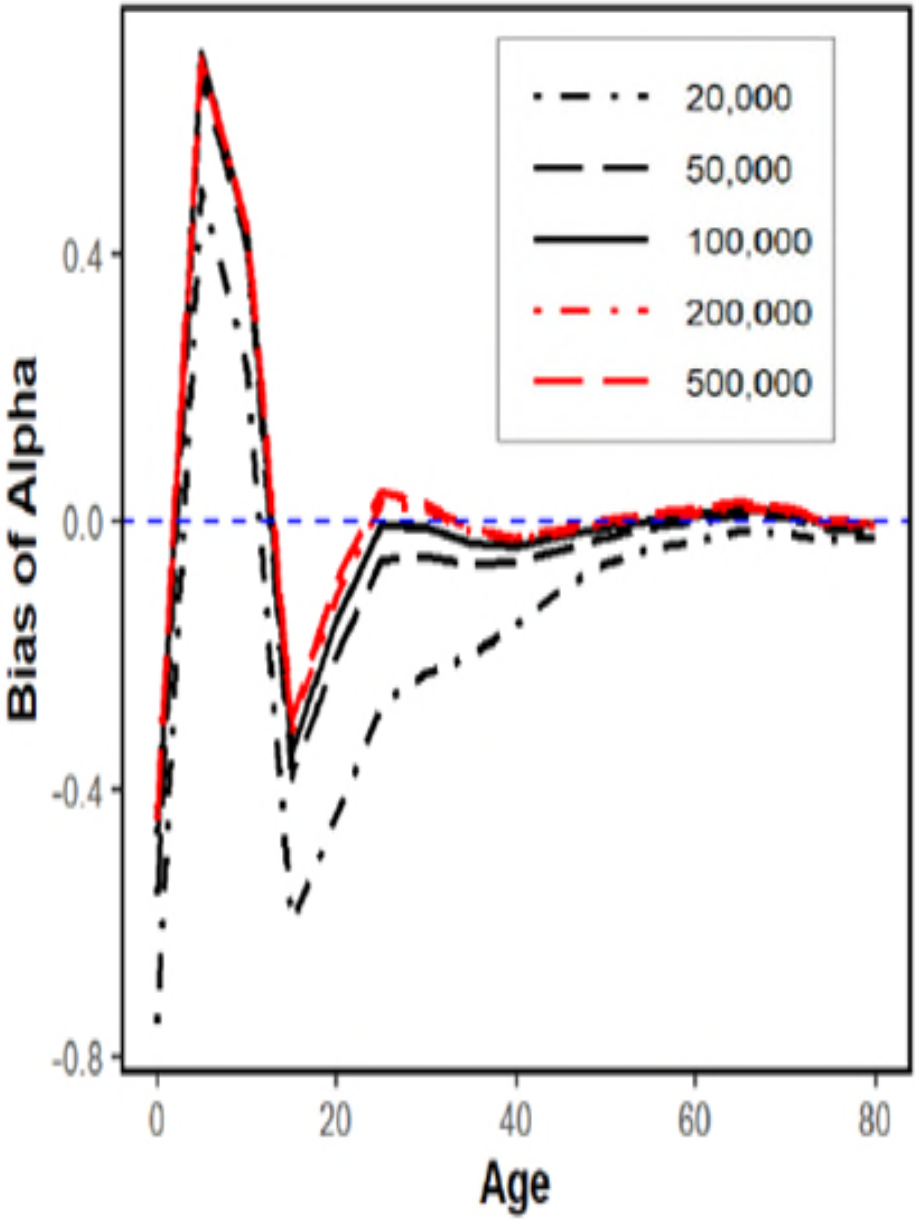
- Partial SMR shows improvement on estimating mortality rates when the population size is 20,000~500,000.

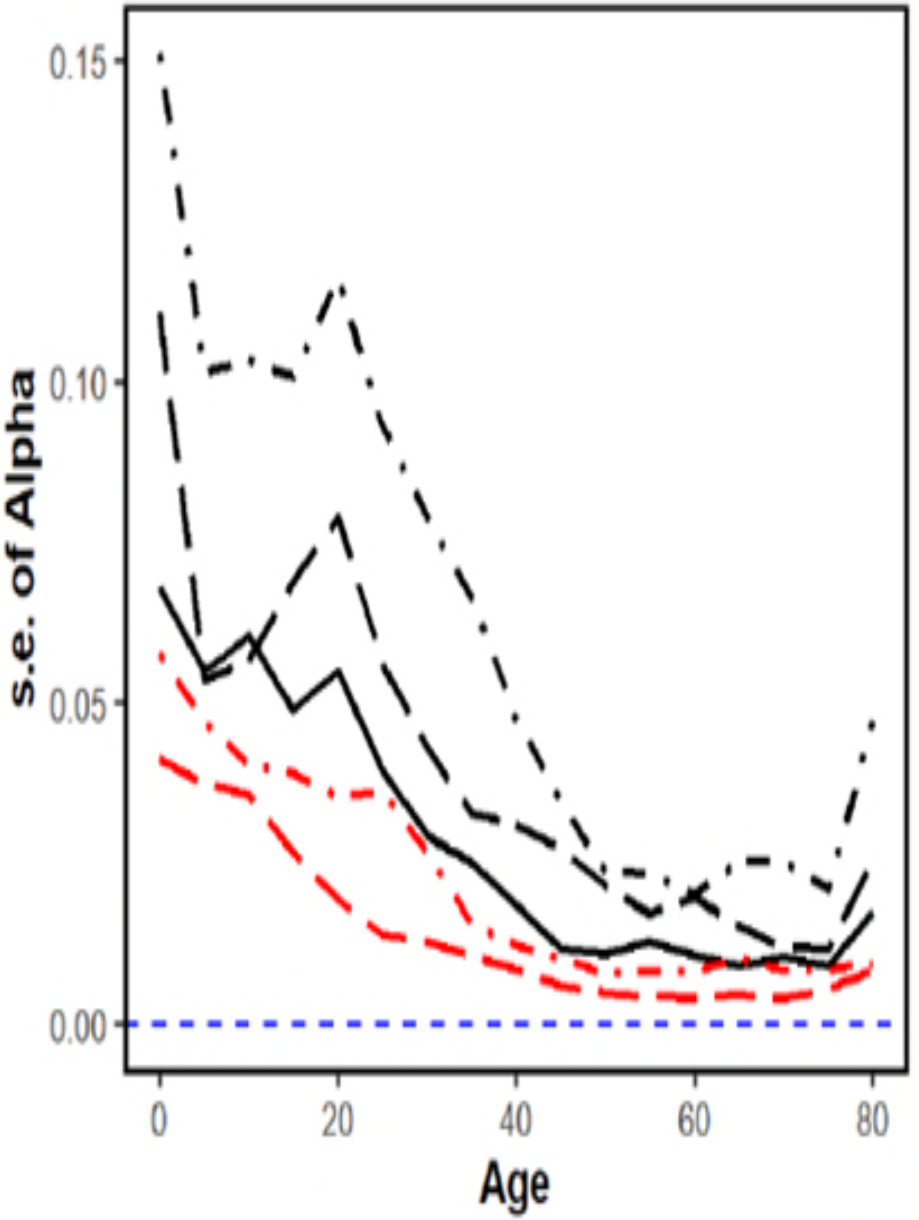
→Significant improvement on the bias of α_x estimate.

→Little improvement on the bias of β_x estimate.

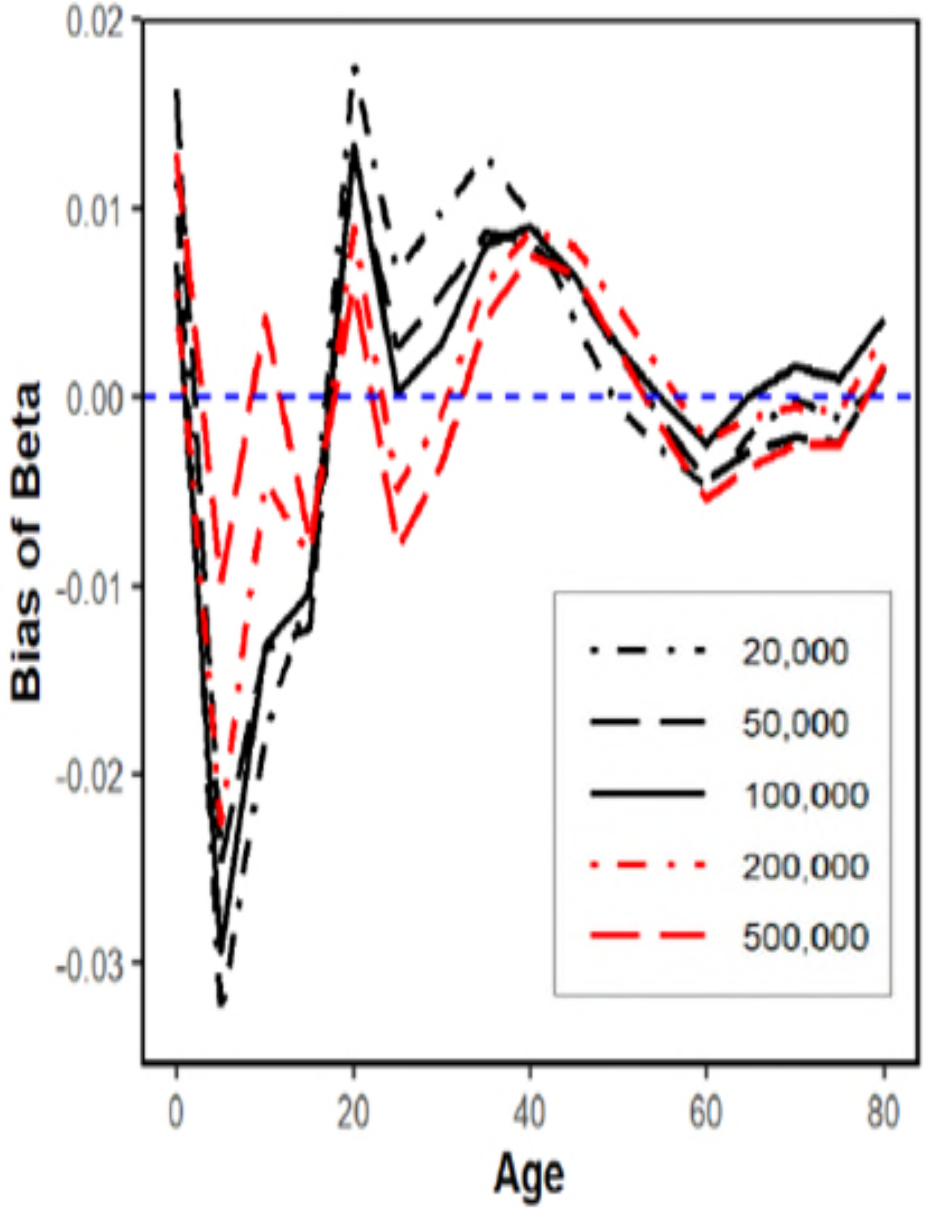
Note: Variances of parameter estimates are also smaller.

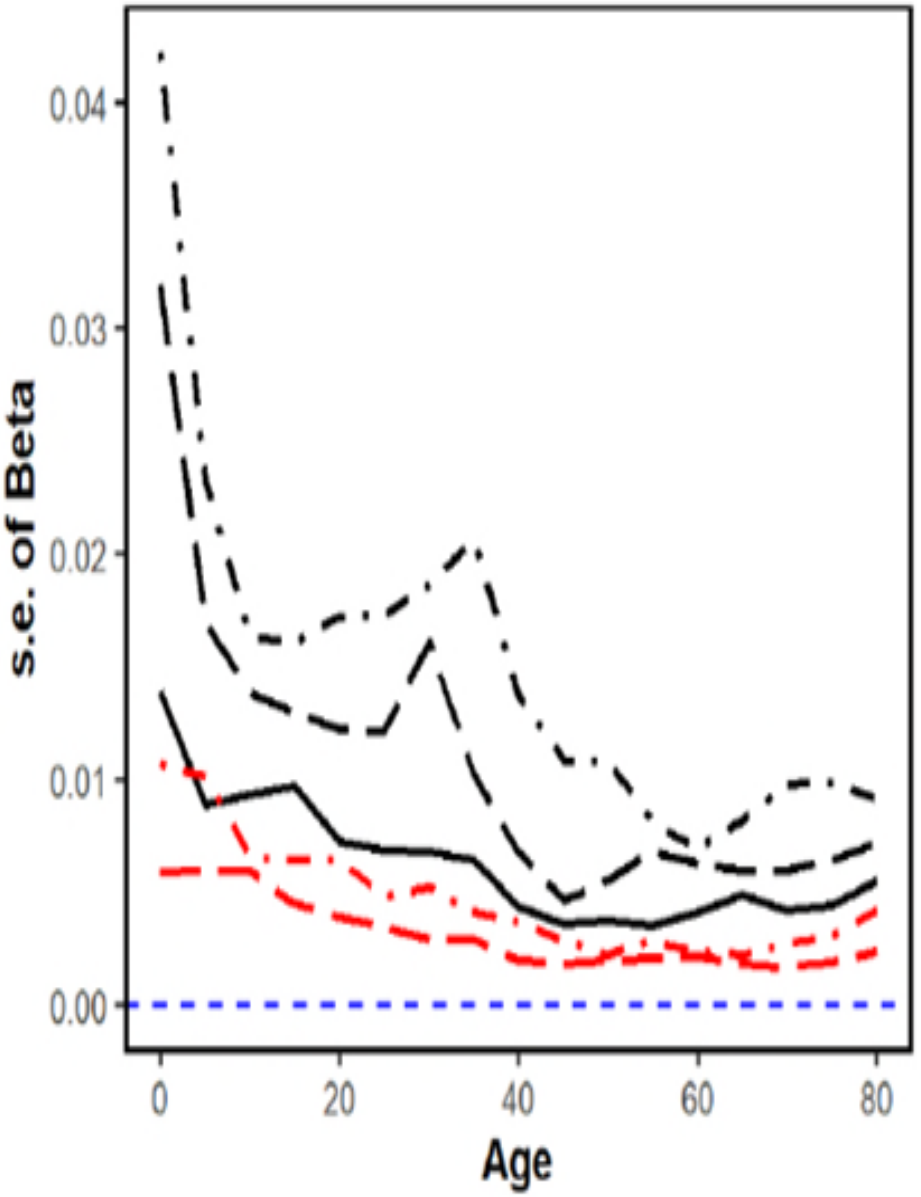
α_x Estimates of GWR





β_x Estimates of GWR





GWR & Parameter Estimation

- GWR has little improvement on mortality & parameter estimation, although the variances of α_x & β_x estimates are smaller.
- Why does GWR perform poorly?

→Does not increase the actual number of deaths

→May oversmooth real mortality patterns

→Sensitive to bandwidth selection

→Reduces variance more than bias

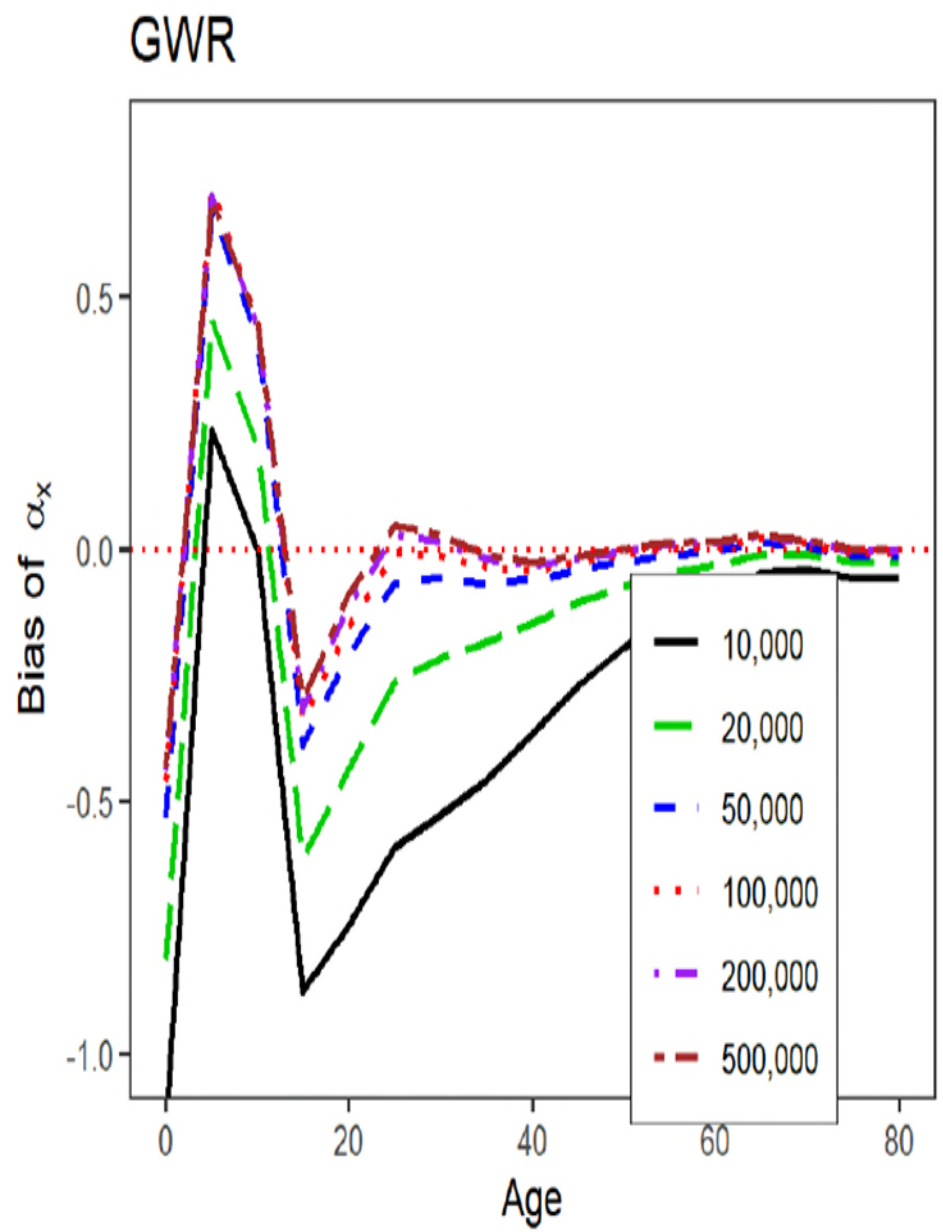
Proposed Spatial modification

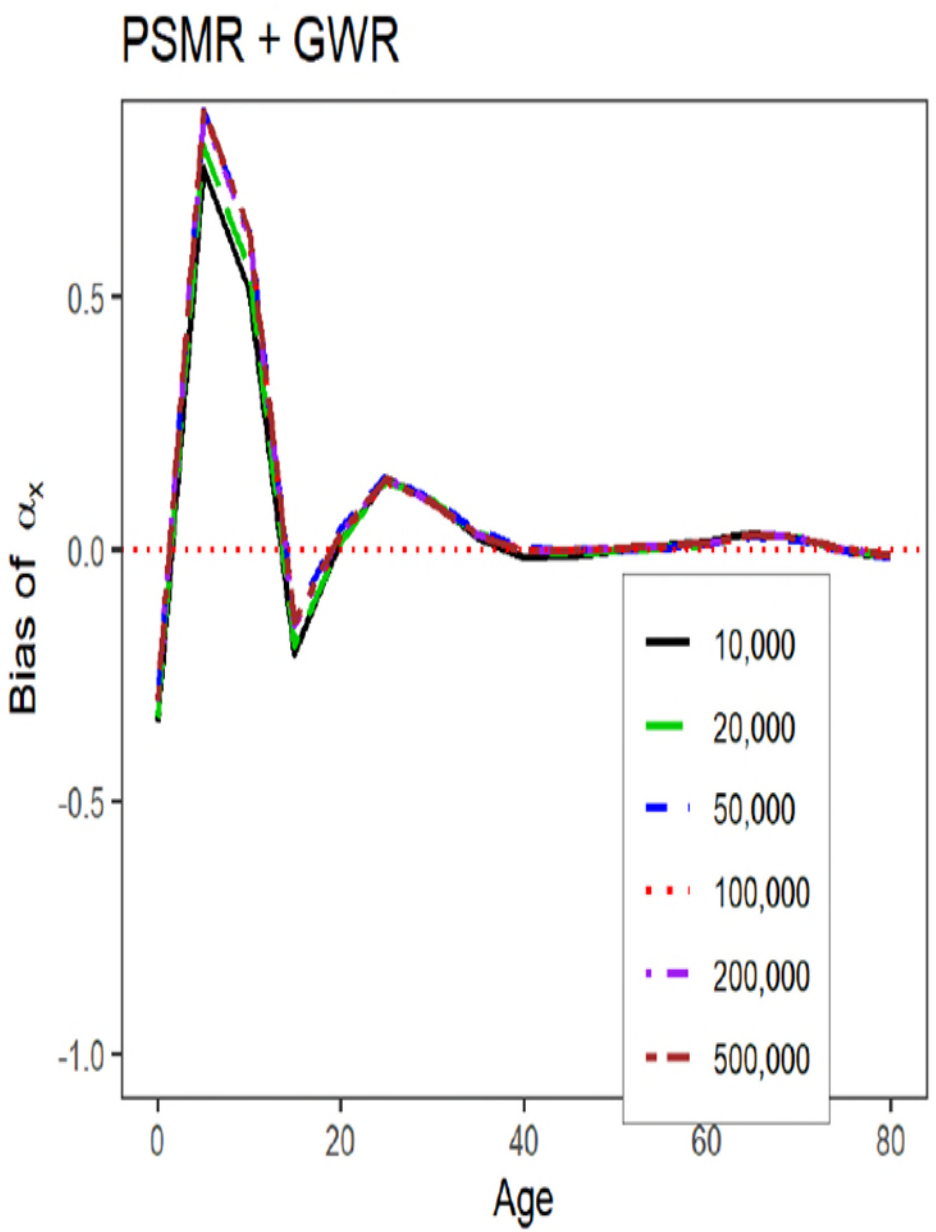
- Can GWR provide additional improvement after Partial SMR smoothing?

→Use the partial SMR to reduce small-sample bias and then apply GWR to further reduces the remaining local variability

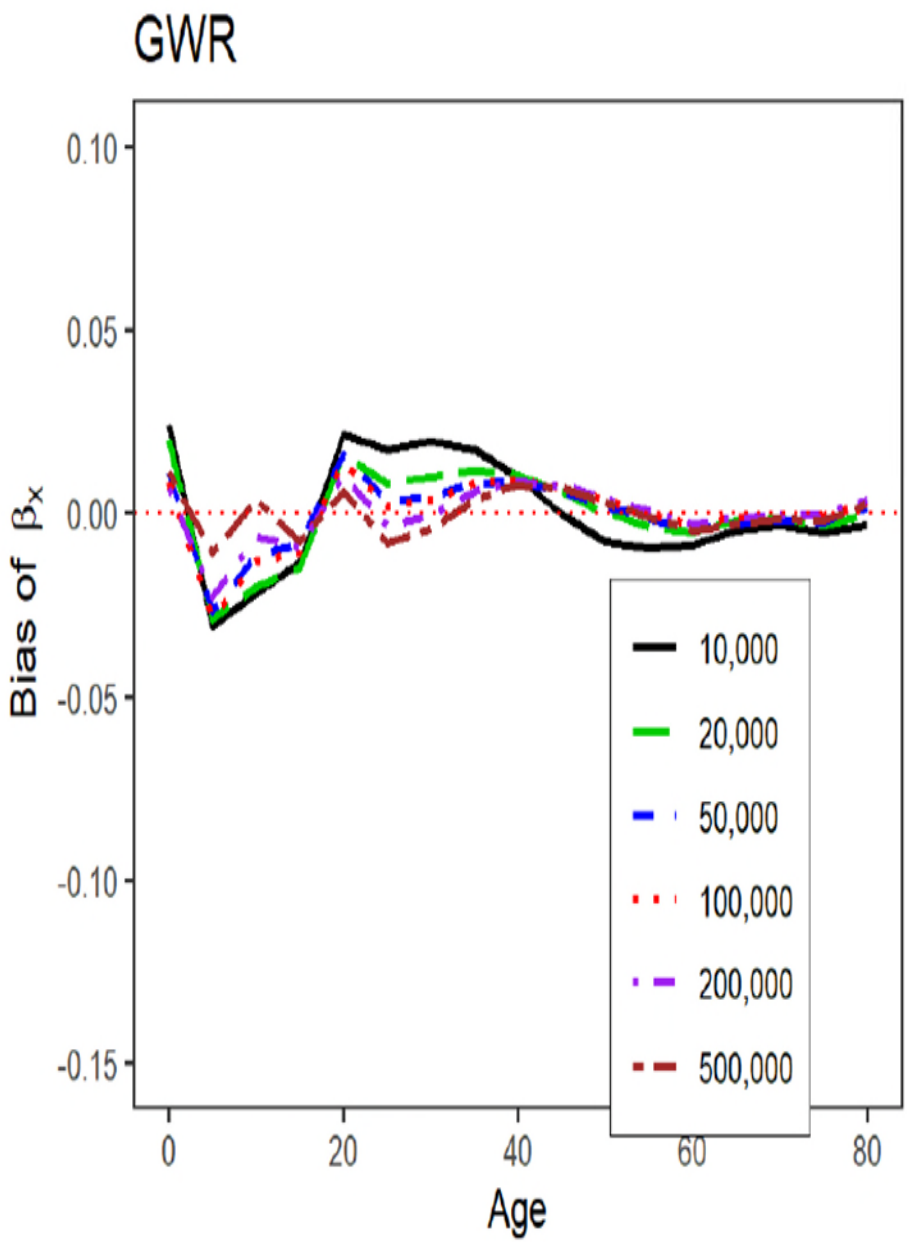
- The bias of α_x & β_x estimates of GWR+Partial SMR do not shown improvement.

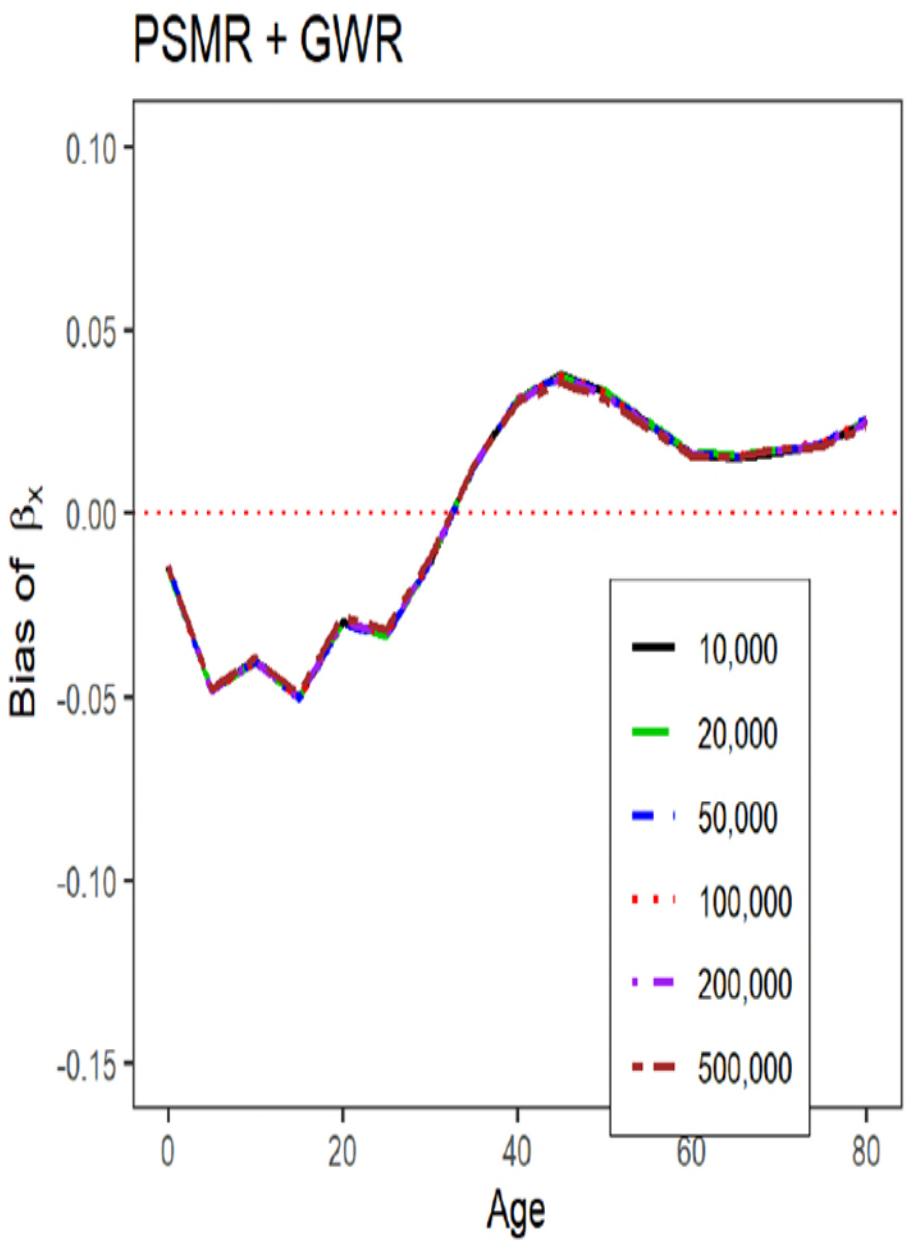
α_x Estimates of GWR & GWR+Partial SMR





β_x Estimates of GWR & GWR+Partial SMR





Adjustments for Proposed Approach

- Problem identified

→Partial SMR effectively reduces the bias of α_x.

→GWR may oversmooth local mortality patterns, with limited bias reduction.

→Performance of GWR is sensitive to bandwidth.

- Proposed adjustment:

→The key is to reduce bias first, then control local variability.

Conclusion

- Small populations induce substantial estimation bias

→The bias is particularly serious for α_x and the bias of β_x is more evident at ages 0, 10, & 20.

- GWR alone is not sufficient to solve the small-population bias problem.

→Partial SMR is effective for bias reduction and GWR mainly reduces variance, not bias.

Future Study

- Improved spatial modification

→Combine Partial SMR with other spatial model

- Optimal bandwidth selection

→Elliptical weighted window

→Age-time dependent weighting/flexible bandwidth

- Empirical application

→Apply the proposed method to Taiwan's small-area mortality data

Take-home Message

- Increasing the effective sample size can reduce small-population bias.
- Study notes:

  Partial SMR→ Effective for reducing bias

  GWR→ Effective mainly for reducing variance

Note: Future direction→ Develop a spatial method that can improve both bias and variance.

References

- Brouhns, Denuit, and Vermunt (2002), A Poisson Log-bilinear Approach to the Construction of Projected Lifetables, *Insurance: Mathematics and Economics*.
- Czado, Delwarde, and Denuit (2005), Bayesian Poisson log-bilinear mortality projections, *Insurance: Mathematics and Economics*.
- Leong and Yue (2026), A Spatial Approach to Model Mortality Rates, *Journal of Forecasting*.
- Li and Lee (2005), Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee–Carter Method, *Demography*.
- Wang, Yue, and Chong (2018), Mortality Models and Longevity Risk for Small Populations, *Insurance: Mathematics and Economics*.
- Wiśniowski, Smith, Bijak, Raymer, and Forster (2015), Bayesian Population Forecasting: Extending the Lee-Carter Method, *Demography*.
- Yue, Wang, and Wang (2021), Using Graduation to Modify the Estimation of Lee-Carter Model for Small Populations, *North American Actuarial Journal*.

Thank you for your Attention!