A Spatial Cluster Modification of the Lee-Carter Model

Authors: Yin-Yee Leong
and Jack C. Yue
National Chengchi University
Email: csyue@nccu.edu.tw

Outline

- Prolonging Life
- Stochastic Mortality Models
- Spatial Analysis & Mortality Models
- Empirical Studies
- Discussions and Conclusions

Prolonging Life Expectancy

- The life expectancies of all countries have being extended about 0.25 year annually since the end of World War II.

→Using static life tables to calculate annuity products would under-estimate the values and 30% of under-price is not too much.

→Stochastic mortality models (e.g., Lee-Carter model) are a popular choice for dealing with the longevity risk.

Impact of using different mortality bases *on pension projections*

mortality basis	value of deferred RPI-linked annuity relative to PMA80c2010-1
PMA80c2010-1	100%
No future improvements	104%
CMI projection	117%
Cohort projection	127%
Revised cohort projection	141%

Calculations are for a male, retirement age 65, payments monthly in advance, guaranteed 5 years, 3.5% expense loading, vesting year 2030, 3% real interest

資料來源：Richard Willets – “Mortality Update”

Lee-Carter Model

- Lee and Carter (1992) proposed the following mortality model for U.S.,

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

where

m_{xt} → Central Death Rate of age x, at time t

κ_t → Intensity of Mortality at time t (linear!)

α_x → Average Mortality of age x

β_x → Tendency of Mortality change for age x

Empirical Evidence for LC Model

- LC Model provides fairly accuracy forecasts for the countries such as the U.S. and Japan.

Fitting Error of the LC model (Ages 0-99)

MAPE	Japan	France	USA	Taiwan
Male	5.48%	5.47%	4.09%	7.73%
Female	7.34%	4.89%	3.26%	7.96%

Source: Table 2 from Yang et al. (2010)

Some Concerns in using LC Model

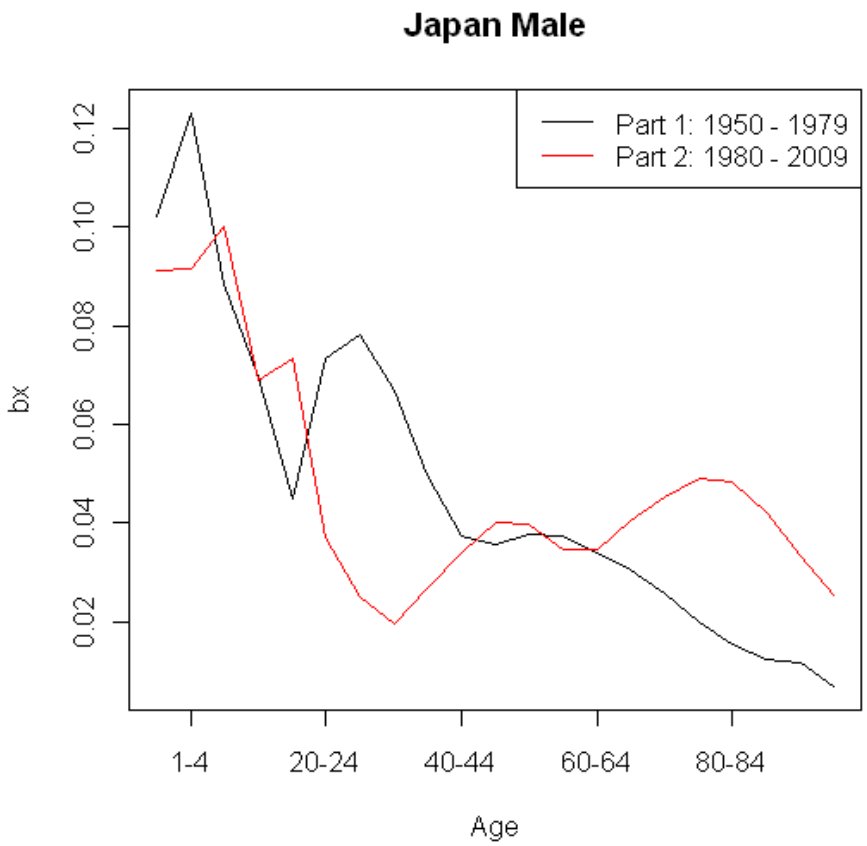
- However, the parameters α_x and β_x are found to change over time, and the parameter κ_t is also not linear in time.
- Several modifications have been proposed:

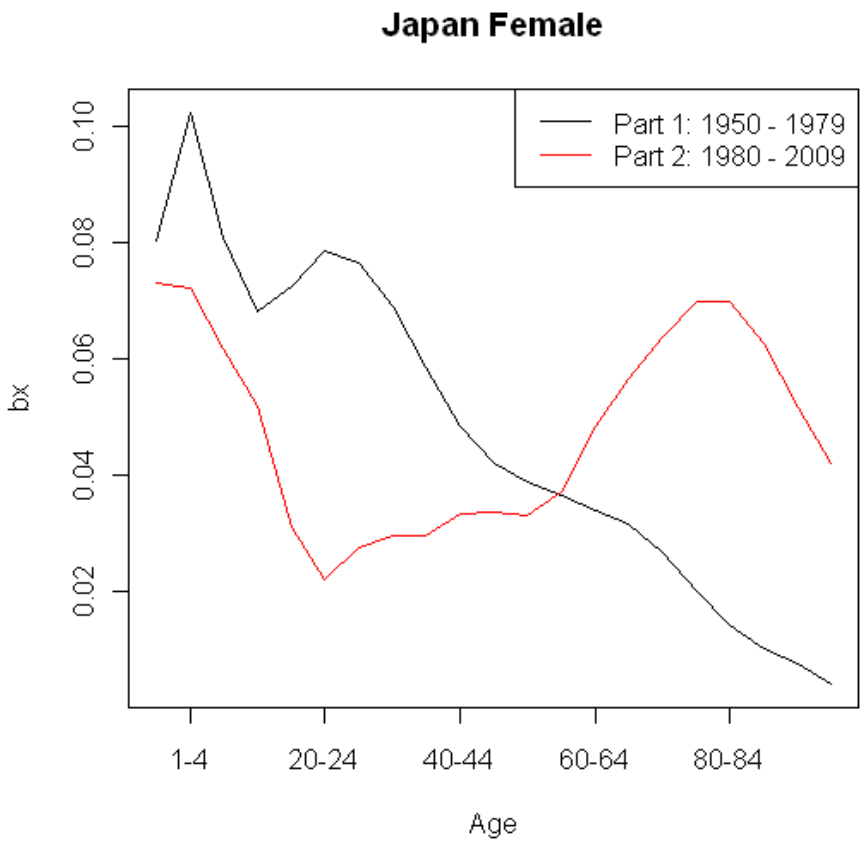
→2 or more period effects : Bell (1997), Yang et al. (2010)

→Cohort effect: Renshaw and Haberman (2006)

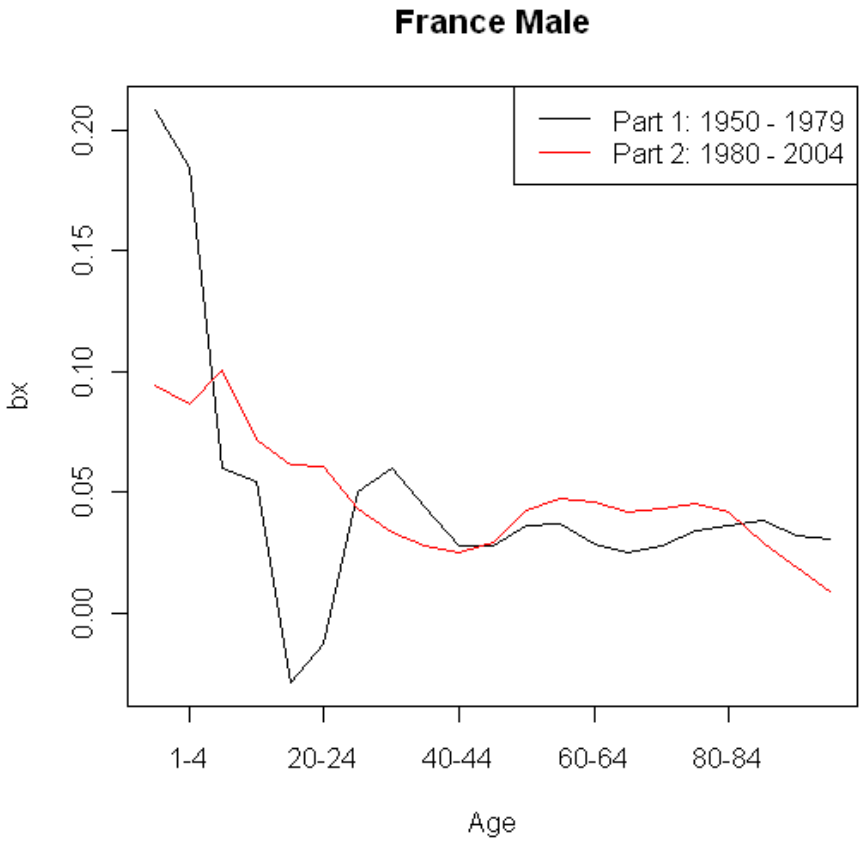
→Functional Analysis: Cairns et al. (2006), Hyndman et al. (2006).

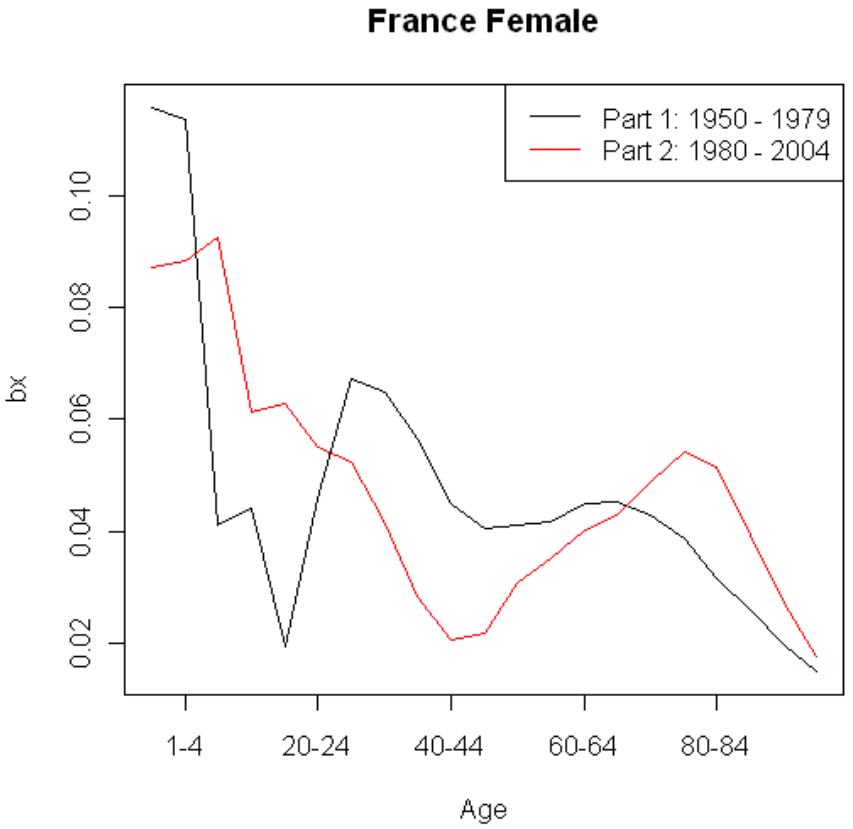
β_x Estimate in Japan





β_x Estimate in France





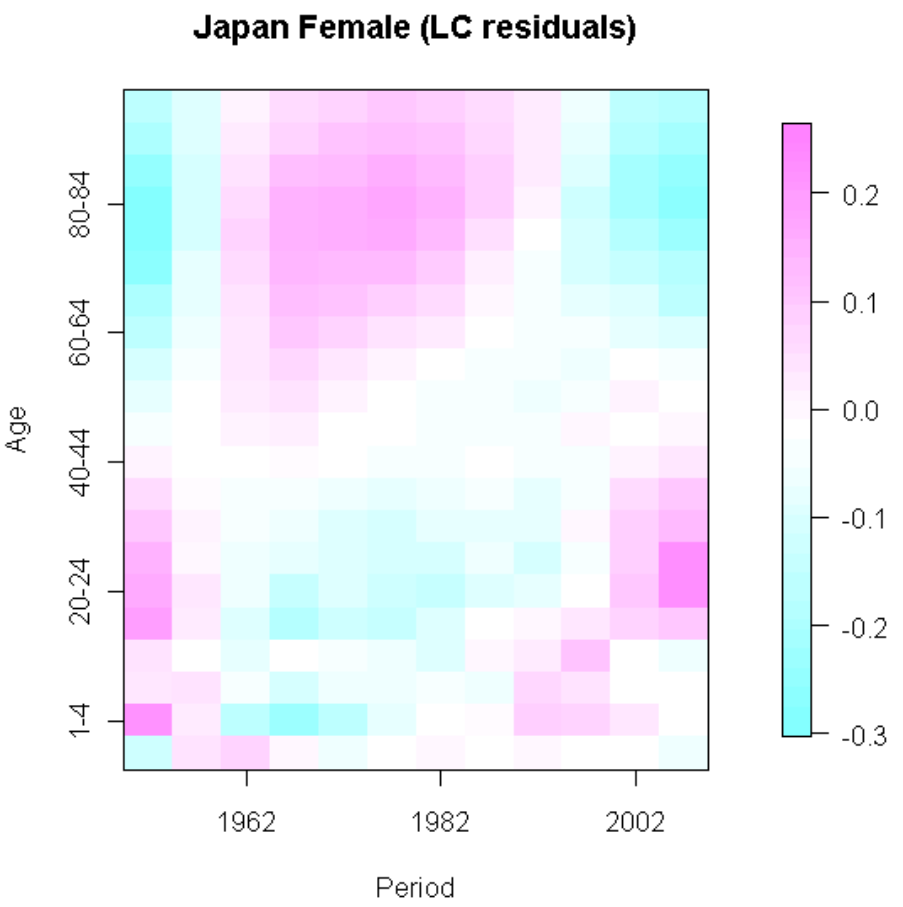
Residual Analysis of LC Model

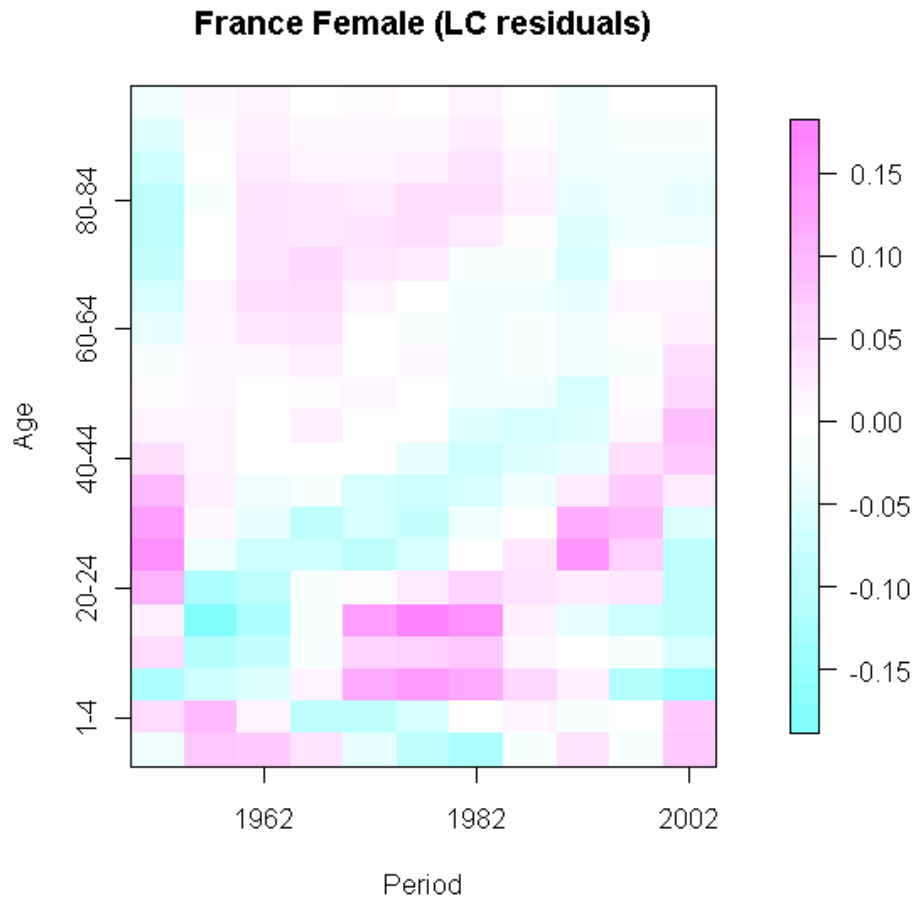
- The residual diagnoses of LC model show that the assumption of normality, independence, and constant variance of the errors are questionable.

→Debón et al. (2008) used geo-statistical analysis and found the errors are correlated spatially (i.e., variogram).

→We found that the errors are auto-correlated as well, and Moran's I and Geary's C both are significant.

LC model -- Residual Analysis





Note: There may exist cohort effects (diagonal).

Spatial Analysis with LC Model

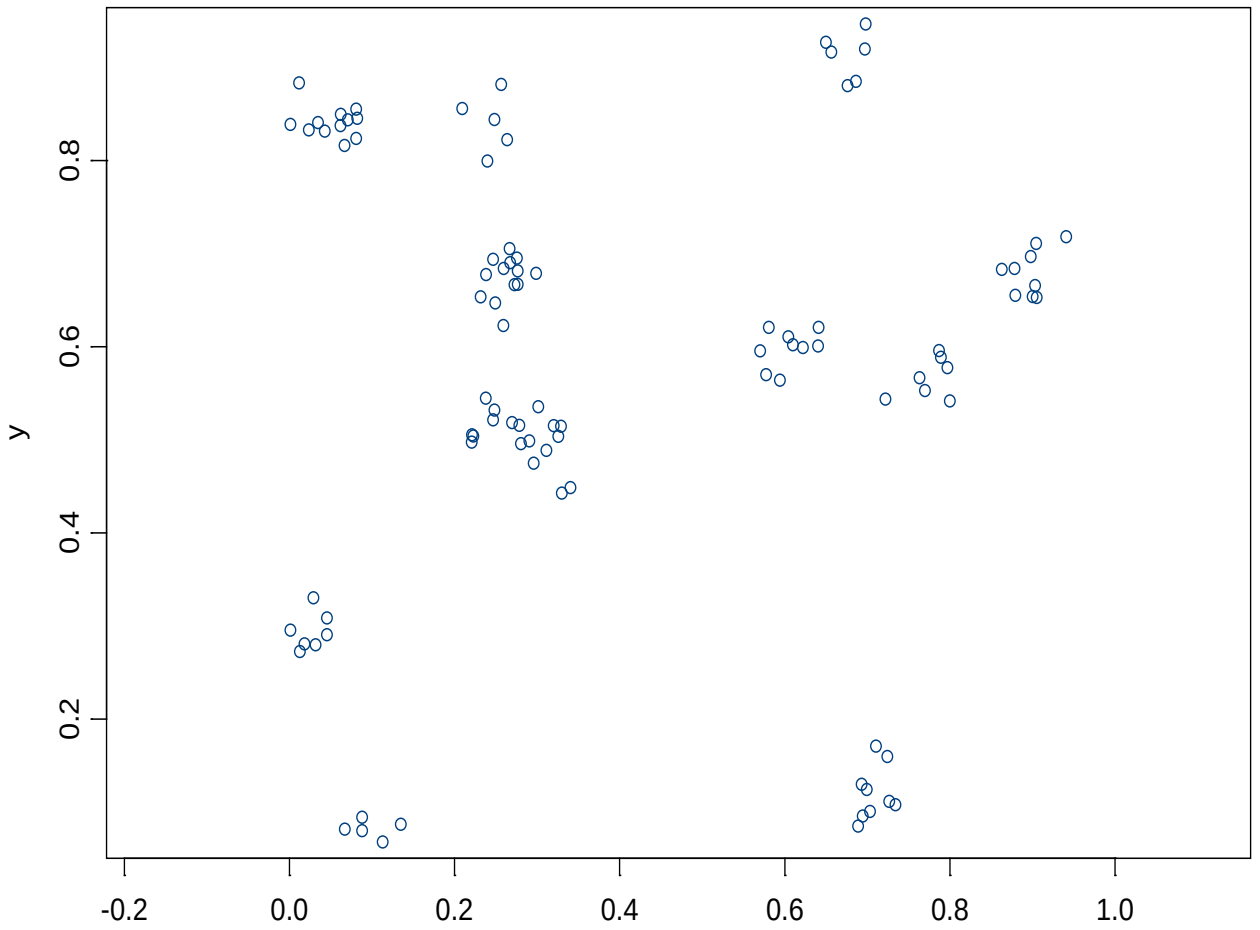
- Spatial in-homogeneity can be handled in two ways

→Mean sift for different “locations”

→Applying spatial covariance

- We shall adapt the first approach and apply the cluster detection technique. In other words, the cohort effect can be treated as “cluster.”

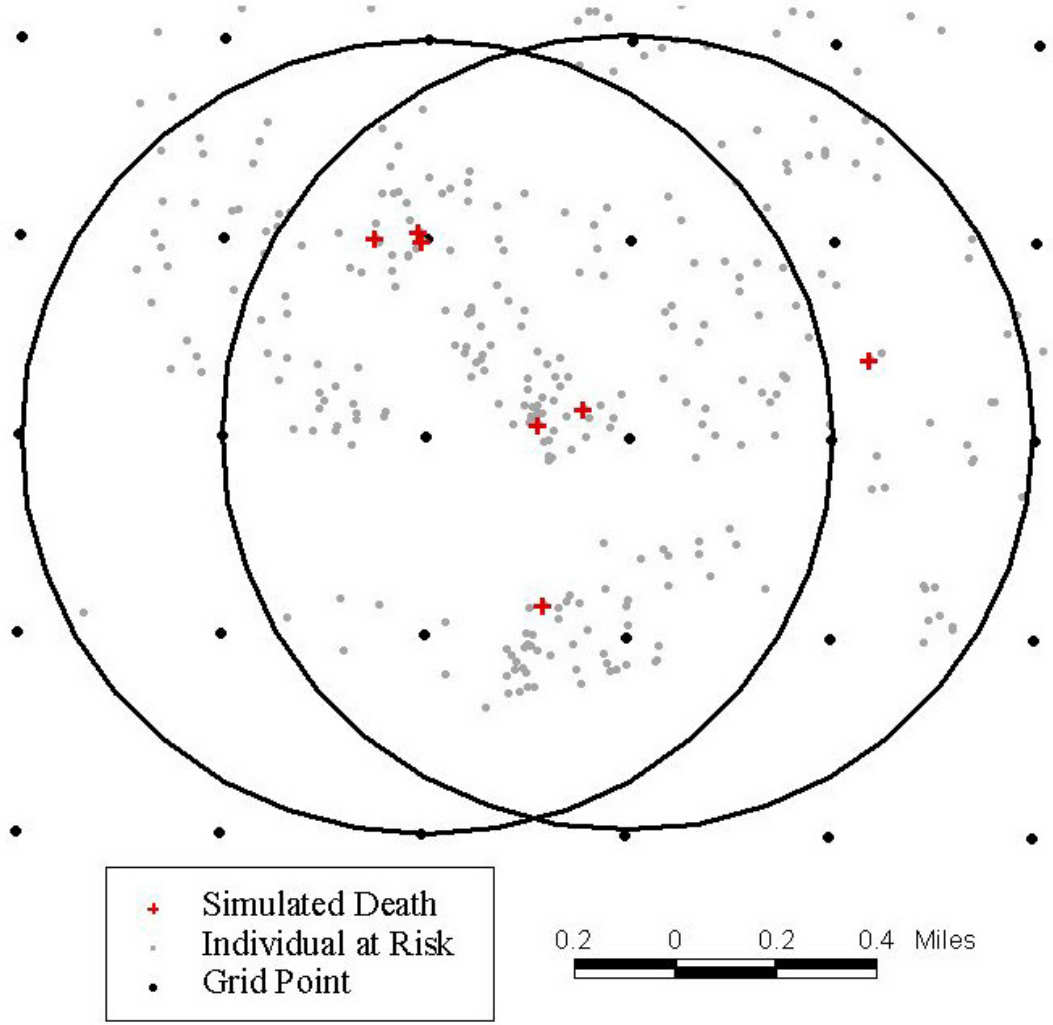
What are clusters?



- A bounded group of occurrences related to each other through some social or biological mechanism, or having a common relationship some other event or circumstance.

- SaTScan (Kulldorff & Nagarwalla, 1995)

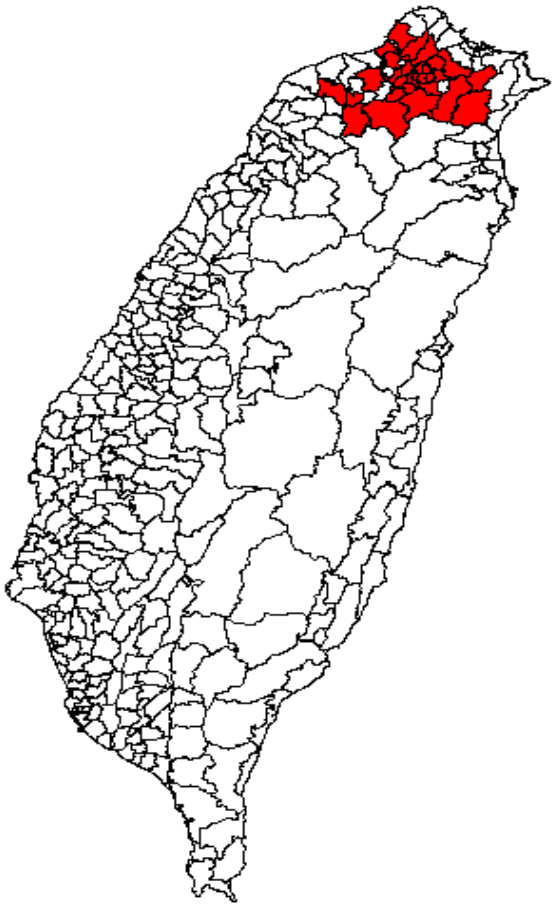
→recommended by National Cancer Institute and proved to be very effective



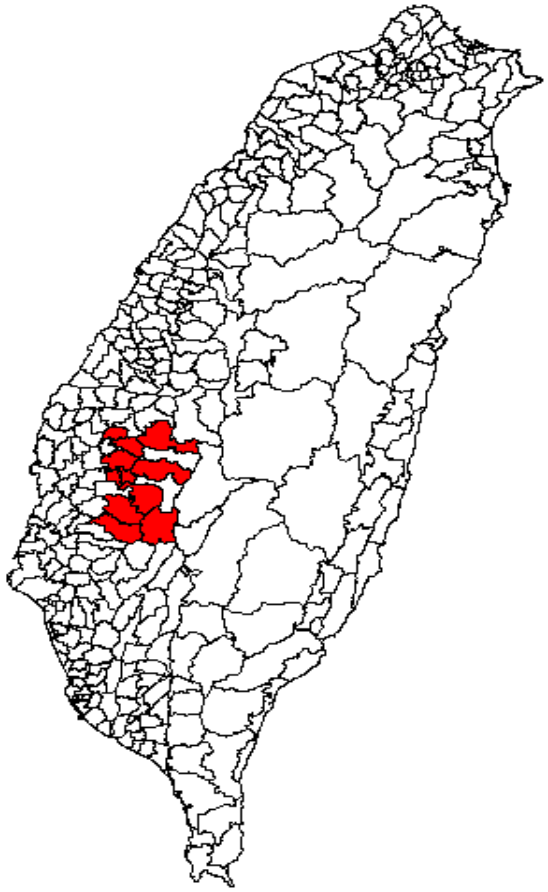
Application to National Health Insurance

- Number of out-patient visits in Taiwan

Significant increase



Significant decrease



Spatial LC Model

- Extend the LC model to incorporate cluster effect:

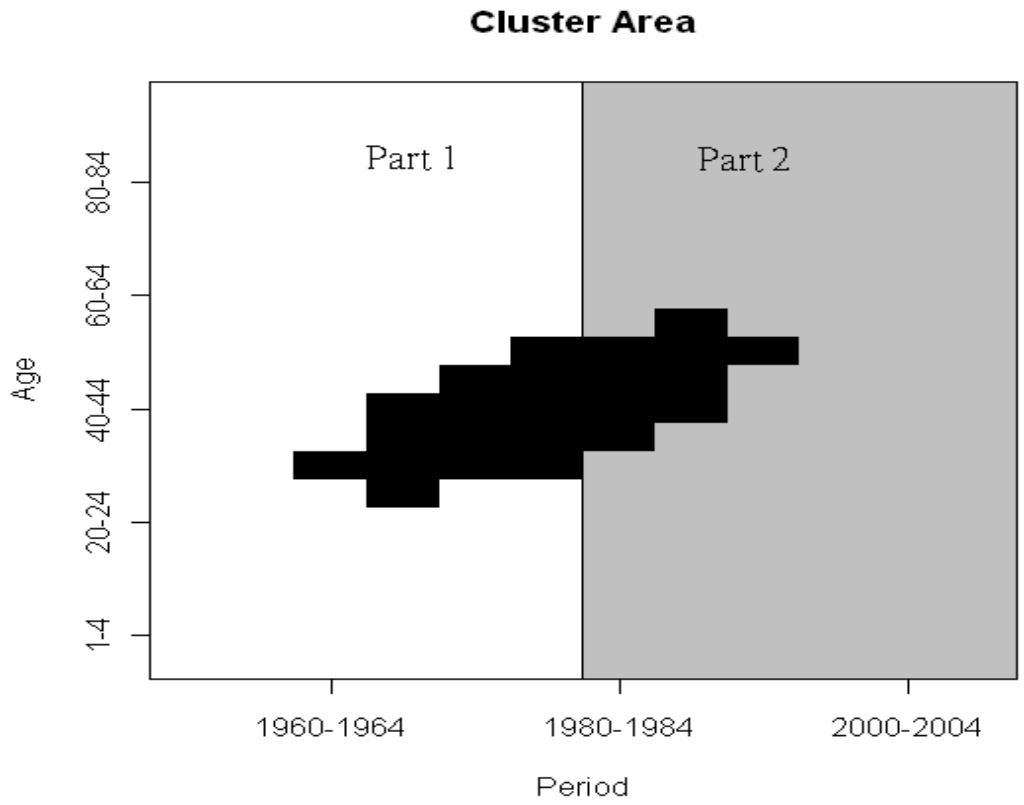
$$\log(m_{xt}) = \alpha_x + \beta_x k_t + \sum_{r=1}^{s} c_r d_{r,xt} + \varepsilon_{xt}$$

$$d_{r,xt} = \begin{cases} 1, & \text{if cluster } r \text{ is at location } (x,t) \\ 0, & \text{otherwise} \end{cases}$$

→In other words, we add a mean shift on particular locations.

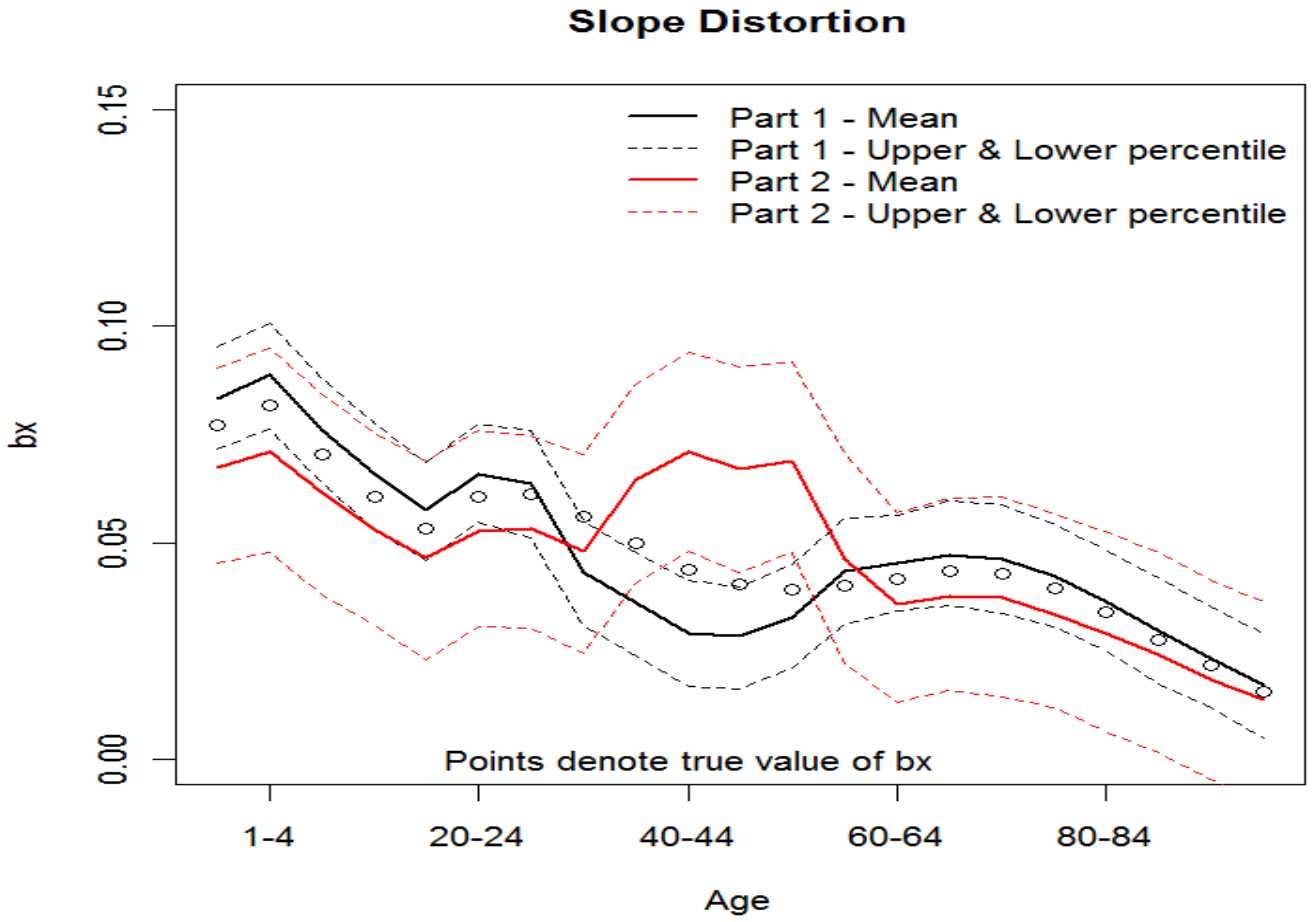
Simulation Study

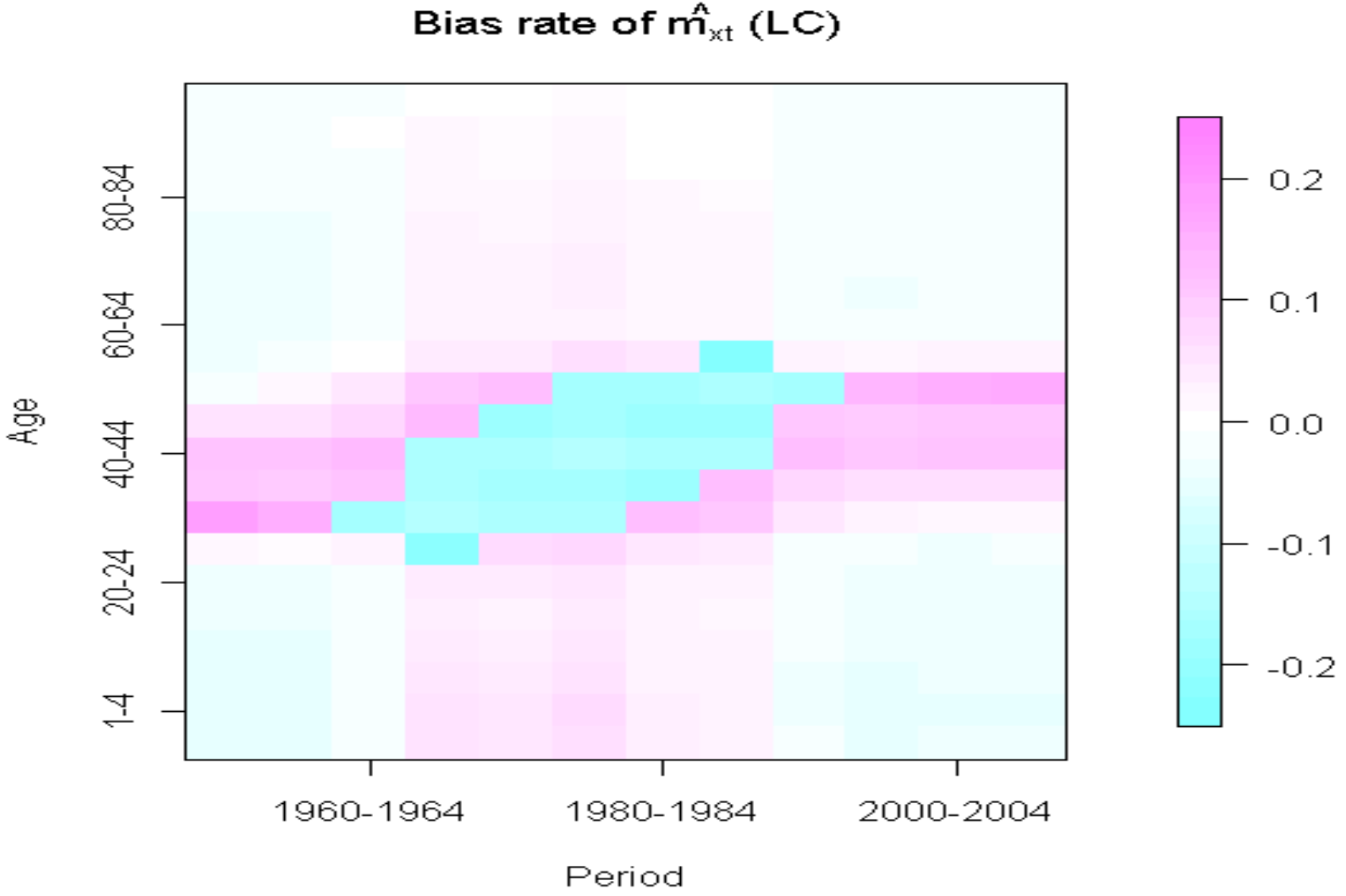
- We use simulation to evaluate the spatial LC model, assuming that there is a mean shift (i.e., cluster) or cohort effect.



Slope Distortion for Different Periods

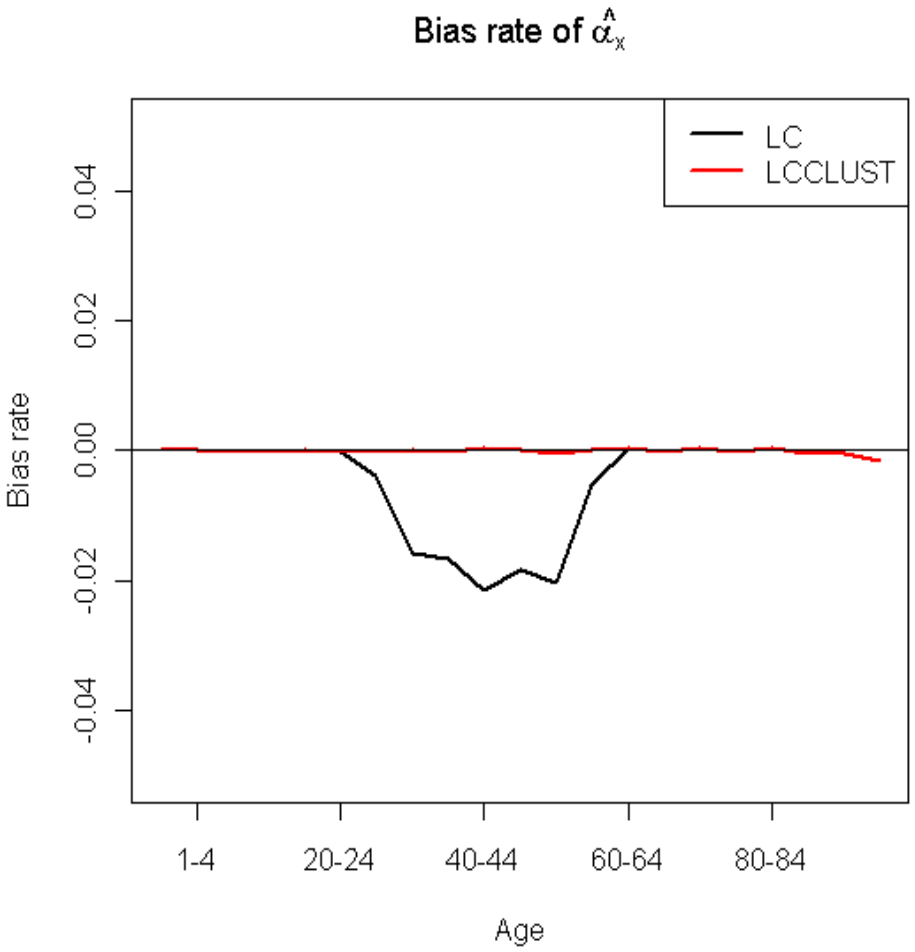
- The estimates of age parameter β_x will be quite different, if the data are separated into two periods.

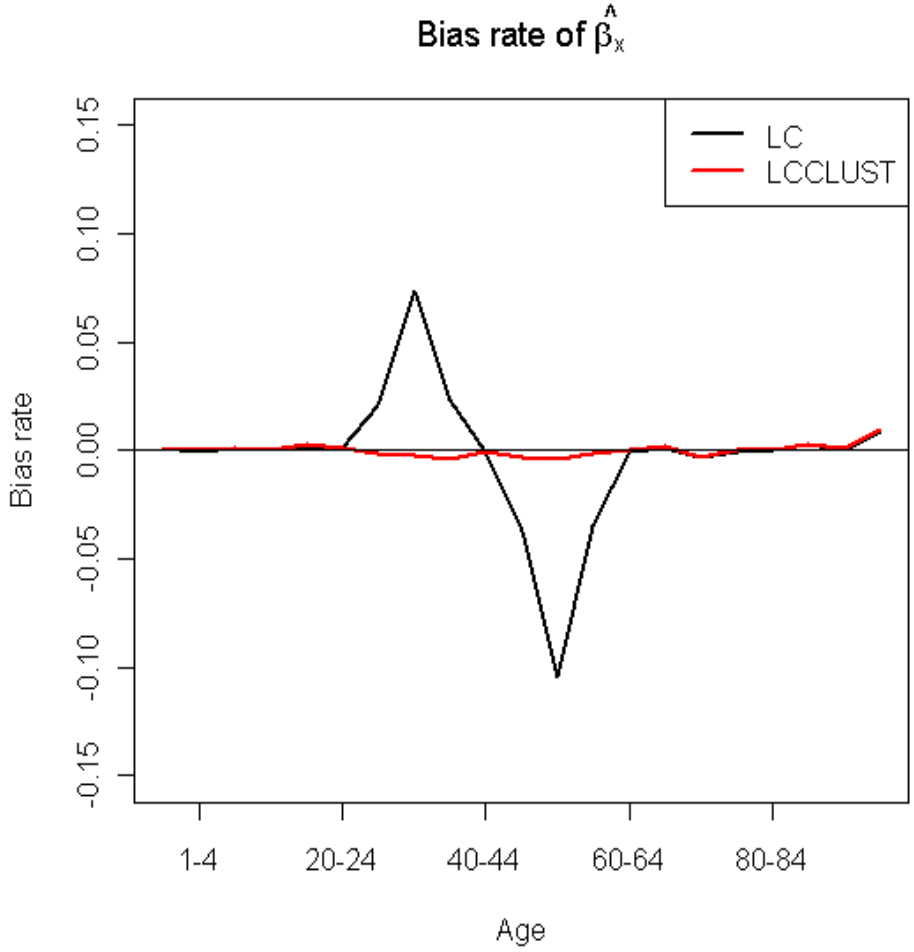




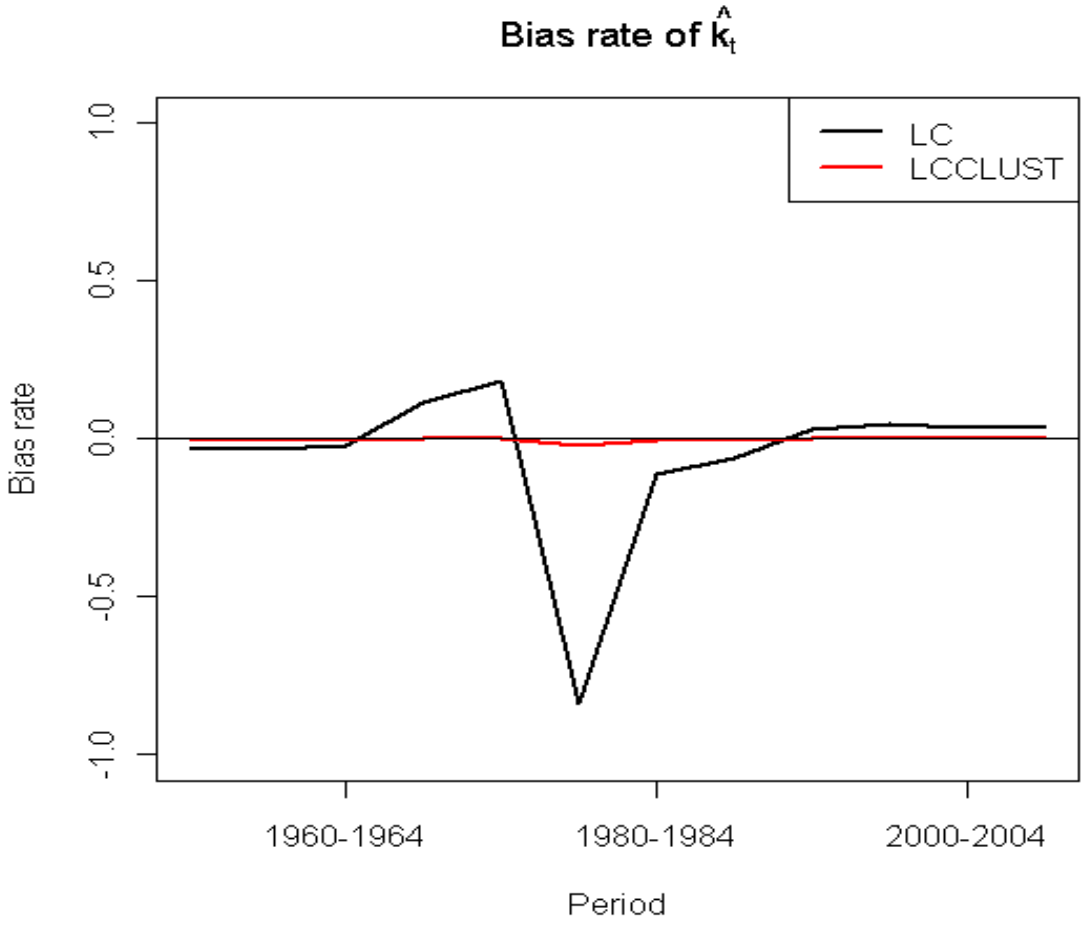
→Residuals show obvious sign of cluster!

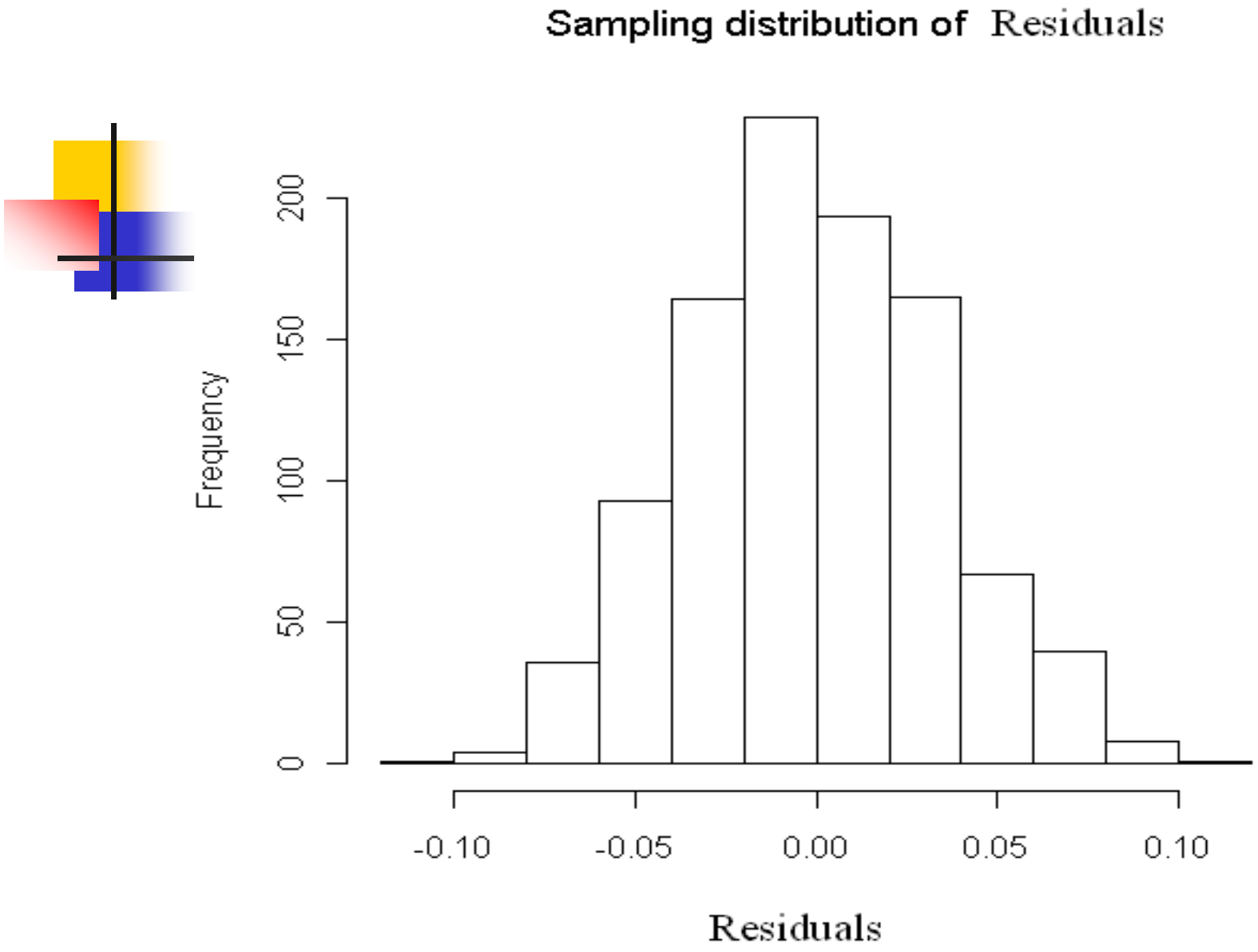
Better Estimate in Age Parameters α_x & β_x





Better Estimate in time Parameters κ_t as well!





→The normality test also supports the Spatial LC model.

Empirical Study

- We use the data from the Human Mortality Database to evaluate the proposed approach.

Country	Japan	France	USA	Taiwan
Data Period	1950~ 2009	1950~ 2004	1950~ 2004	1970~ 2009

Evaluation Criteria

- Two error criteria are used:

→Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|m_{xt} - \hat{m}_{xt}\right|}{m_{xt}} \times 100\%$$

→Bayesian Information Criterion (BIC)

$$BIC = l\left(\hat{\phi}_r\right) - \frac{1}{2}v_r \log(n)$$

Note: BIC is designed to avoid over-parameterization.

Empirical Study (conti.)

- The proposed approach has smaller fitting errors (MAPE).

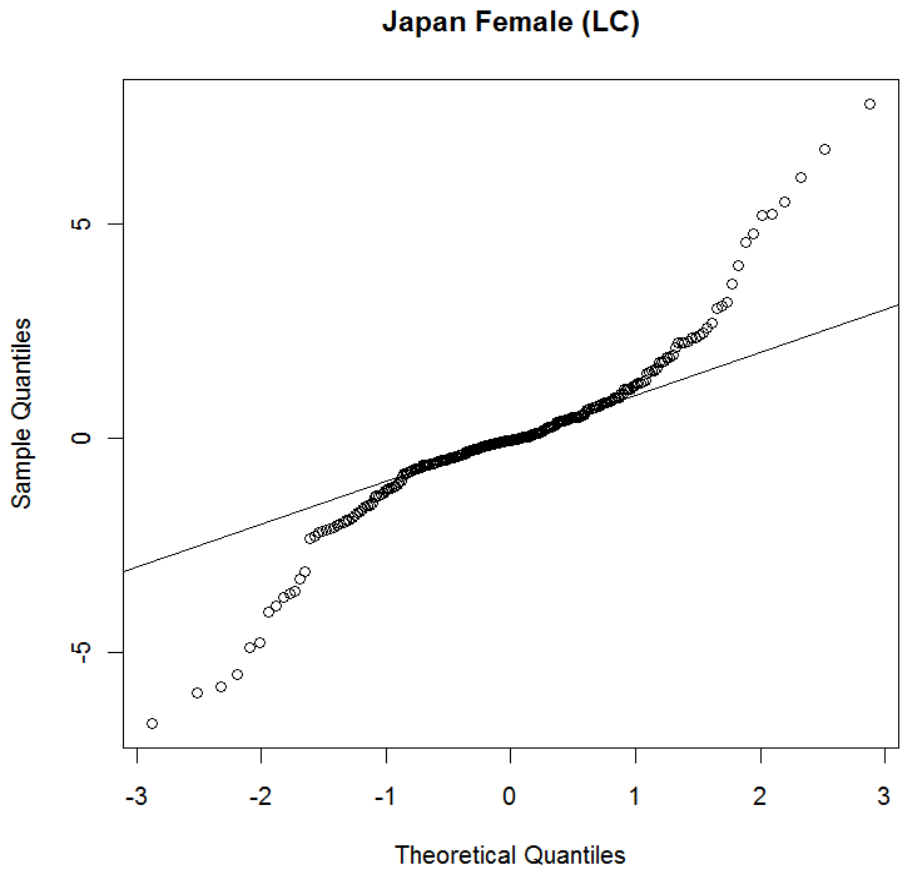
	Male		**Female**	
	LC	SLC	LC	SLC
Japan	5.70%	3.72%	8.35%	4.15%
France	5.12%	4.05%	4.42%	3.34%
USA	3.74%	3.22%	3.08%	2.31%
Taiwan	5.76%	5.20%	4.18%	3.64%

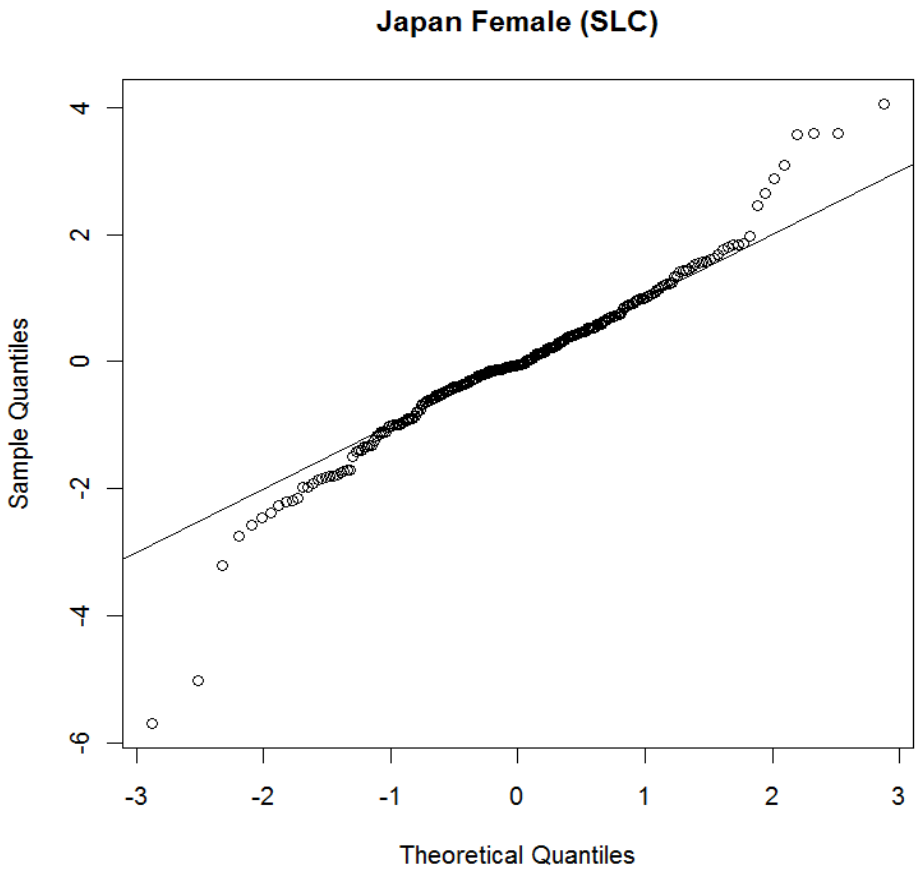
Empirical Study (conti.)

- The proposed approach also has larger (better) results with respect to.

	Male		**Female**	
	LC	SLC	LC	SLC
Japan	1370.59	1456.85	1388.01	1546.61
France	1211.51	1276.48	1389.56	1459.84
USA	1287.12	1330.19	1467.74	1522.20
Taiwan	847.08	868.16	974.65	1007.03

Normal Quantile Plot of Japan Mortality Data





Conclusions

- We propose a spatial modification of the Lee-Carter model. Simulation and empirical studies show positive evidence for the method.

→Cluster effect is another way to interpret the cohort effect.

→Spatial modification can be applied to other stochastic mortality models (e.g., PCA).

- We suggest checking the autocorrelation first and then apply the spatial modification.

Discussions

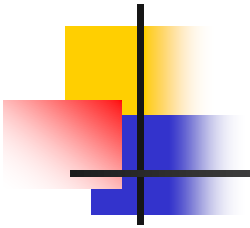
- Cluster detection techniques can be used to improve the estimation of mortality rates.

→They are not designed for forecasting mortality.

- Empirically, it is difficult to distinguish clusters and clustering (autocorrelation).
- SaTScan tends to have larger false positive error.

→We can choose other detection methods.

→Multiple testing!



Thank you!!