



A Spatial Analysis of Health and Longevity of Taiwan People

Jack C. Yue

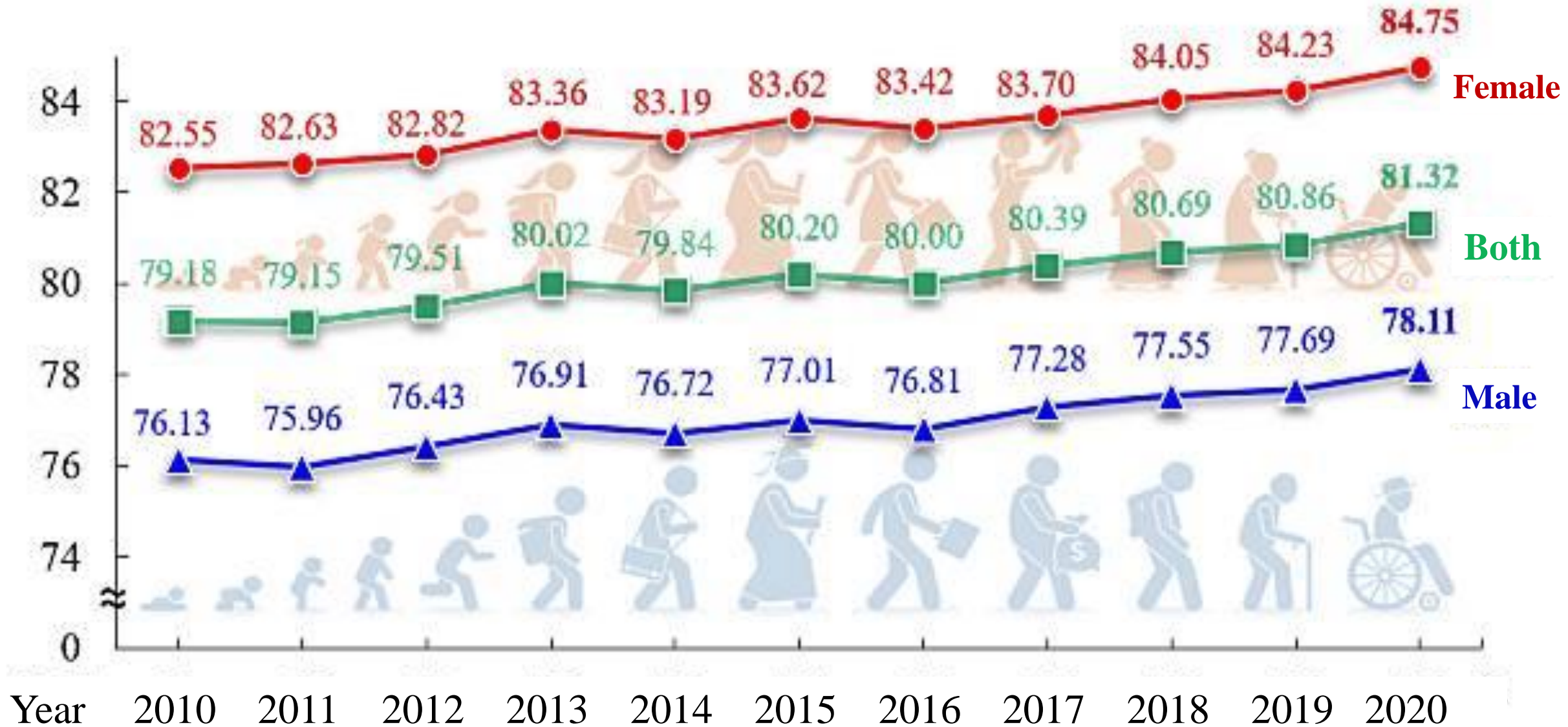
National Chengchi University

csyue@nccu.edu.tw

Sept. 13, 2022



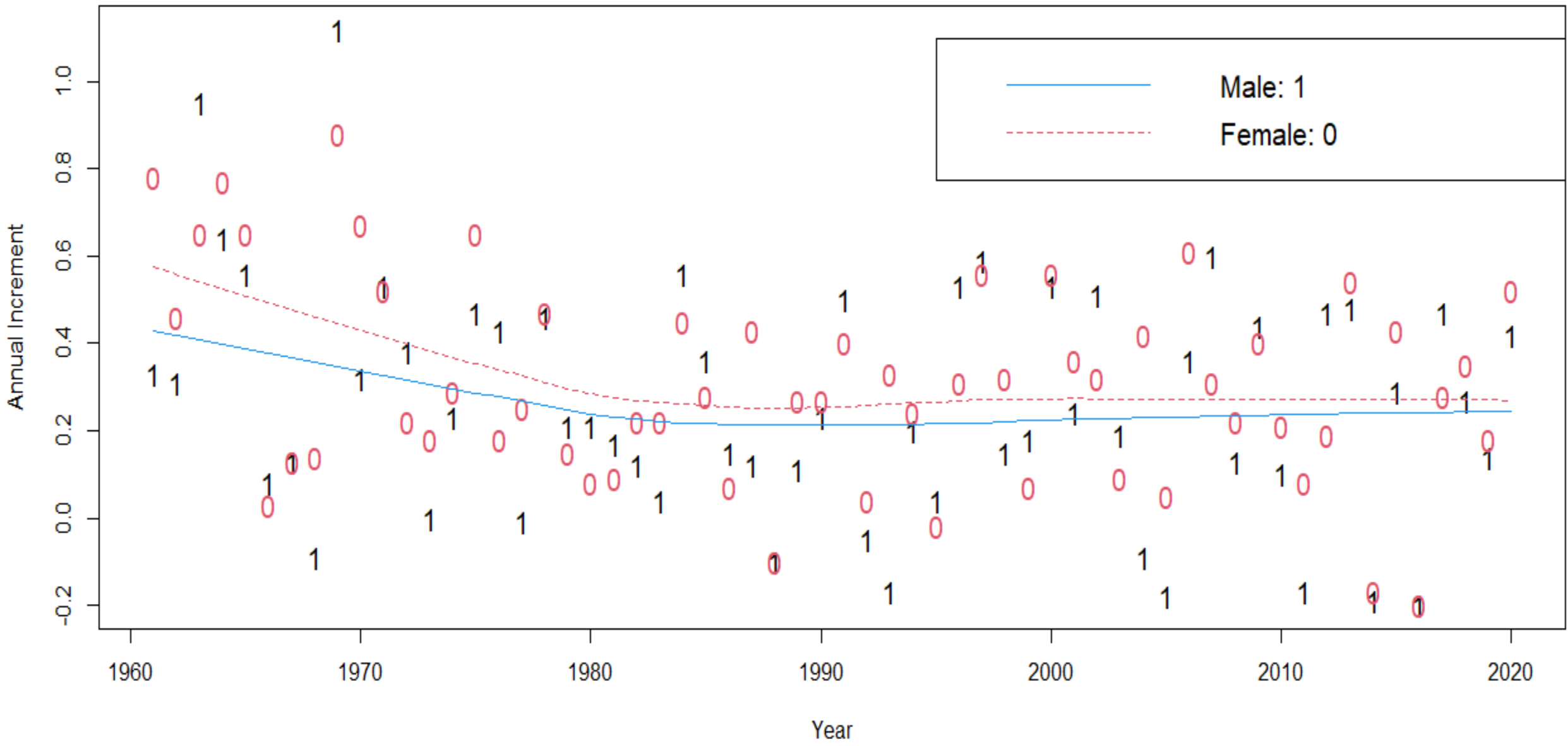
Life Prolonging Continues in Taiwan...



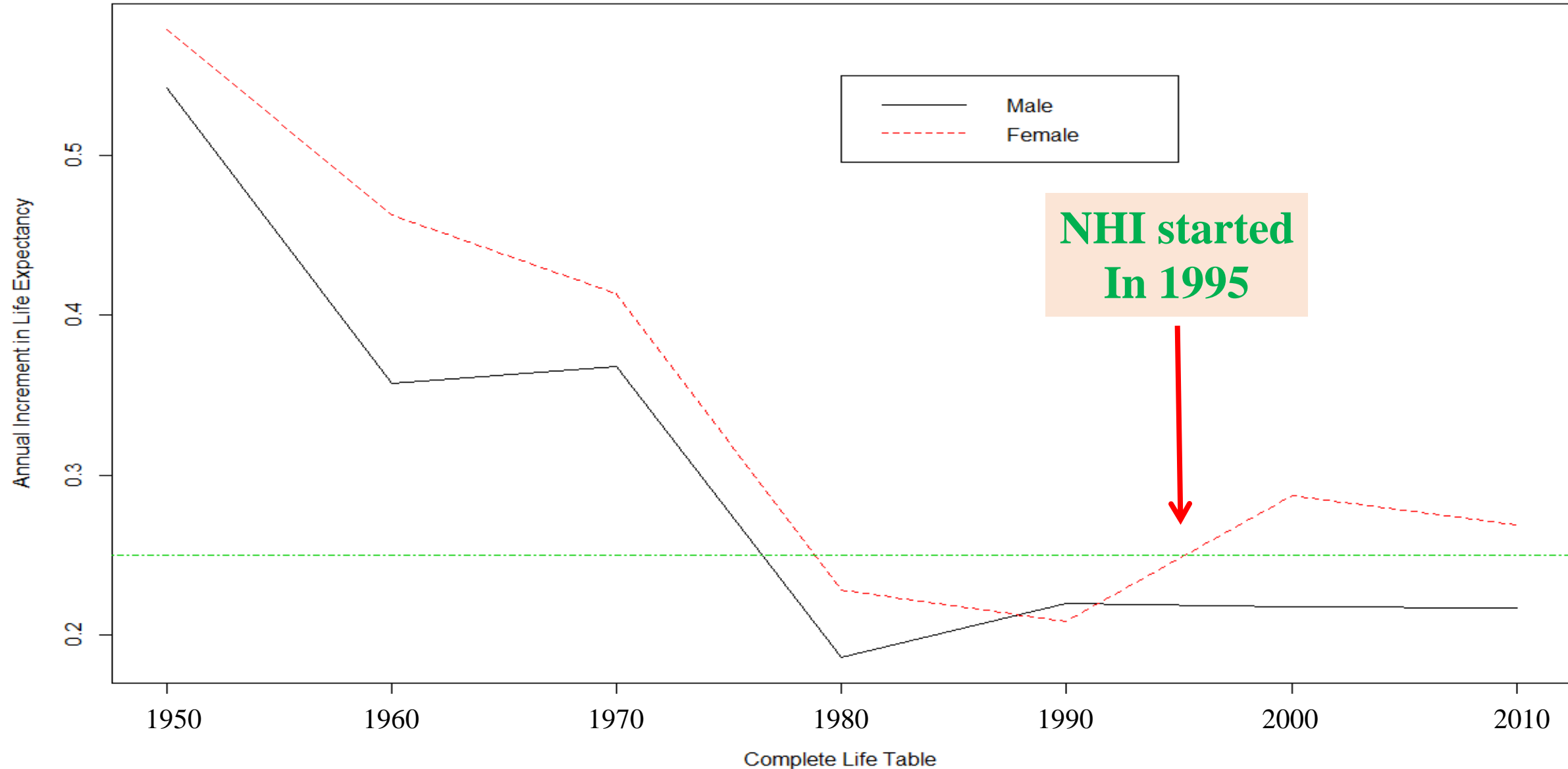
Life Prolonging in Taiwan

- ❑ Taiwan's prolonging life continues and life expectancy is likely to increase in the near future.
 - Significant decreases in 1960-80 & flattening trend after 1990's .
- ❑ Social insurance programs have a impact on the increment of life expectancy, e.g., National Health Insurance (NHI).
 - Wen et al. (2008) found that narrower health disparity after implementing NHI and the lower-ranked health classes (10 classes, 358 township) improved more in life expectancy.

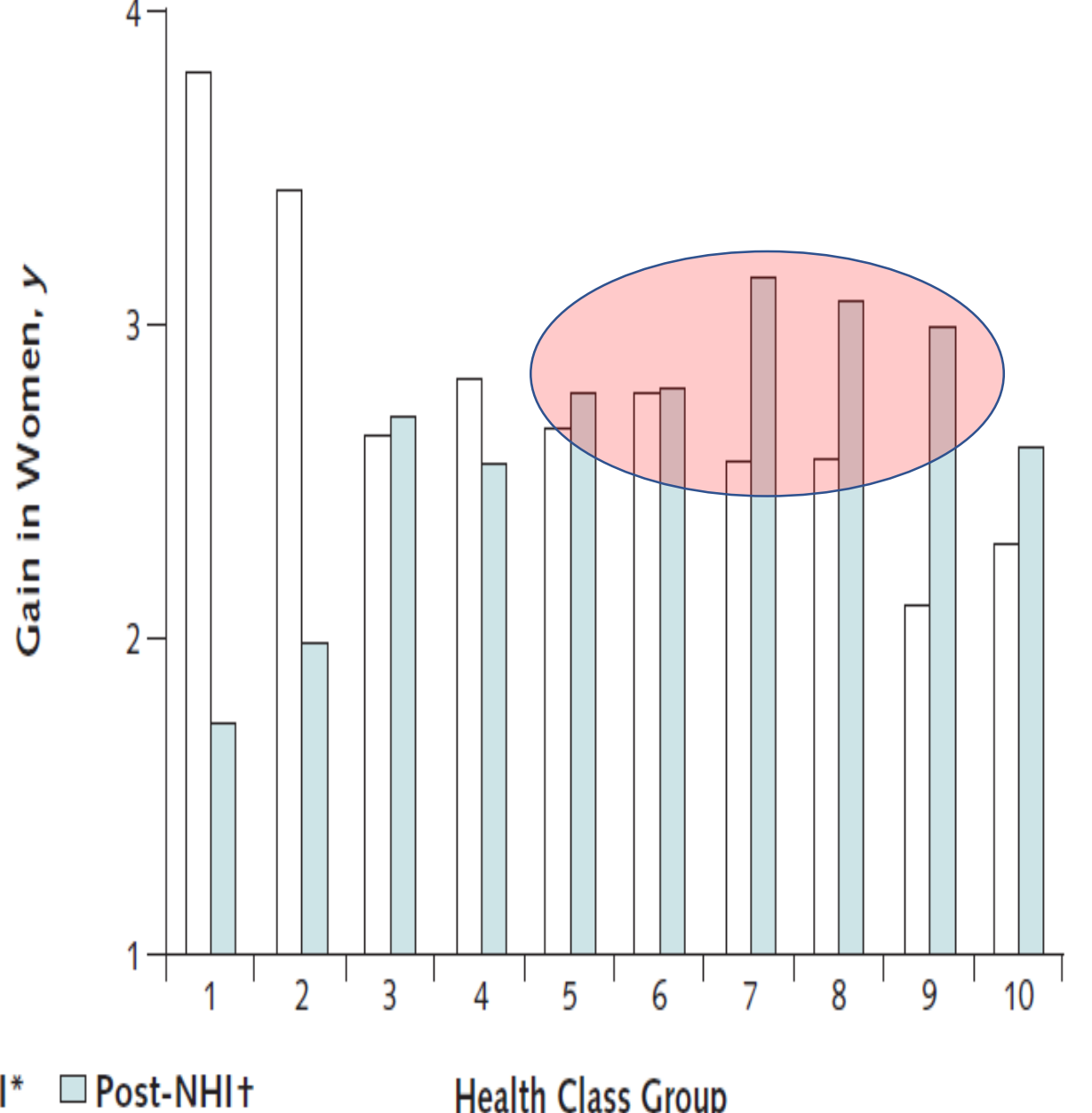
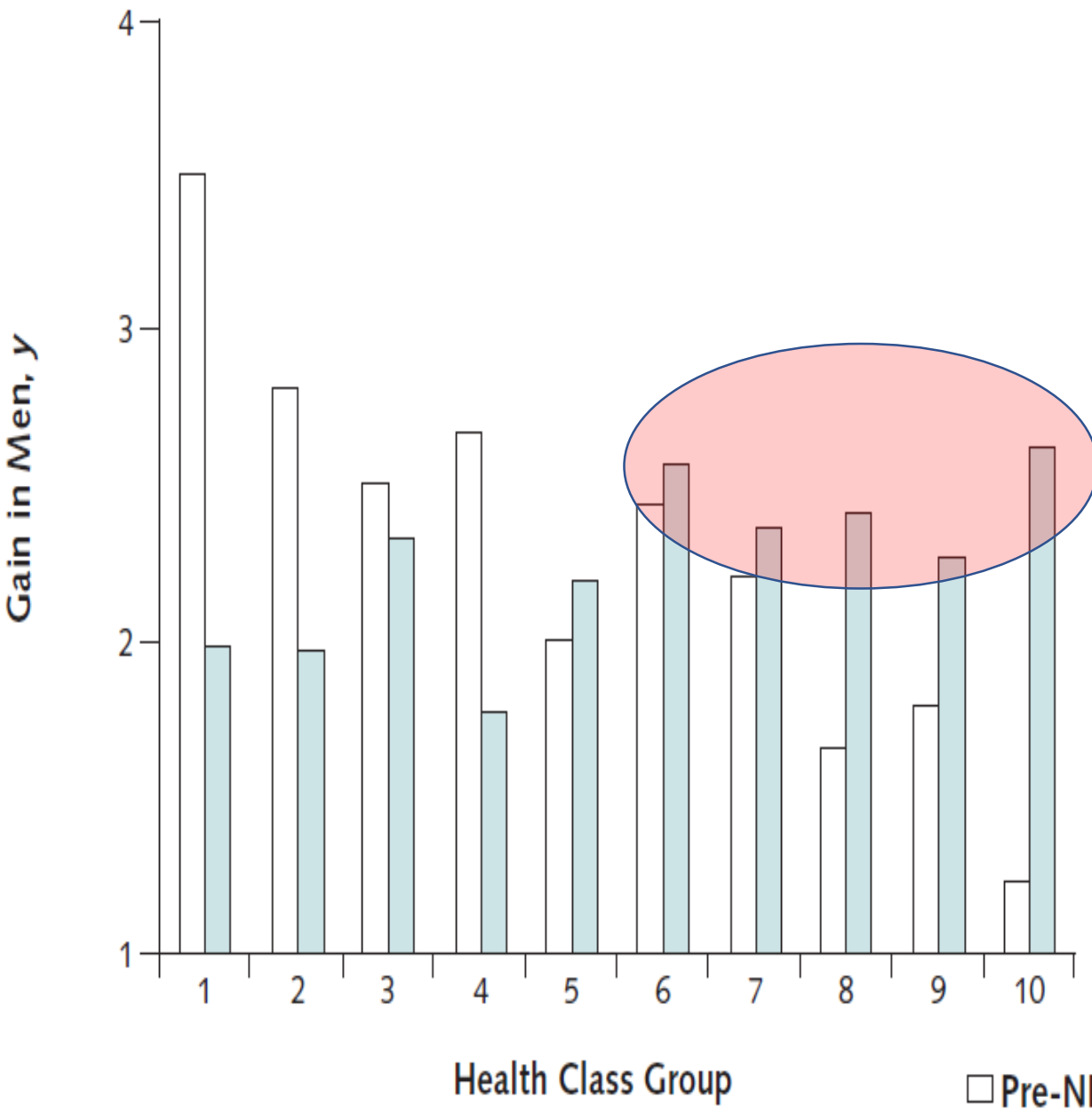
Flattening Life Expectancy Increment (Abridged Life Tables)



Life Expectancy Increment in Taiwan (Complete Life Tables)



Life Expectancy Gain before (1982-84 to 1992-94) & after (1992-94 to 2002-04) NHI



Regional Differences in Life Expectancy

- More medical sources and job opportunity in metropolitan areas, significant regional differences in Taiwan's lifespan, longer in metropolitan areas and shorter in mountain areas.
 - Largest gap of county-level lifespan is 7.5 years (2020).
- We want to explore the regional trend of life prolonging after implementing more social insurance programs.
 - Investigating the spatial attributes (**Spatial In-homogeneity!**)
 - Gap of lifespan between regions decrease/increase over time?

Study Goal

- Applying Spatial Analysis to explore health of Taiwan people at the county and township levels.
 - Check if there are urban-rural disparities in life expectancy and medical utilization.
 - Verify if the disparities change with time.
 - Introduce Geographically Weighted Regression (GWR) to study the relationship between mortality rates and medical utilization at township level.

Methodology

□ Standardized Mortality Ratio (SMR)

$$SMR = \frac{D^j}{\sum_x m_x^s \times P_x^j}$$

where D^j : number of deaths at area j , P_x^j : age x population at area j , and m_x^s : mortality rate of age x of standardized population.

□ Moran's I (y_i : SMR at region i & w_{ij} : weight between i and j)

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \times \frac{\sum_i \sum_j w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

Methodology (Cont.)

- Geographically Weighted Regression (**GWR**; Brunsdon et al., 2015)

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i$$

where x_{ij} is independent variable k and $\beta_0(u_i, v_i)$ & $\beta_k(u_i, v_i)$ are the parameters and ε is the error term at location (u_i, v_i) .

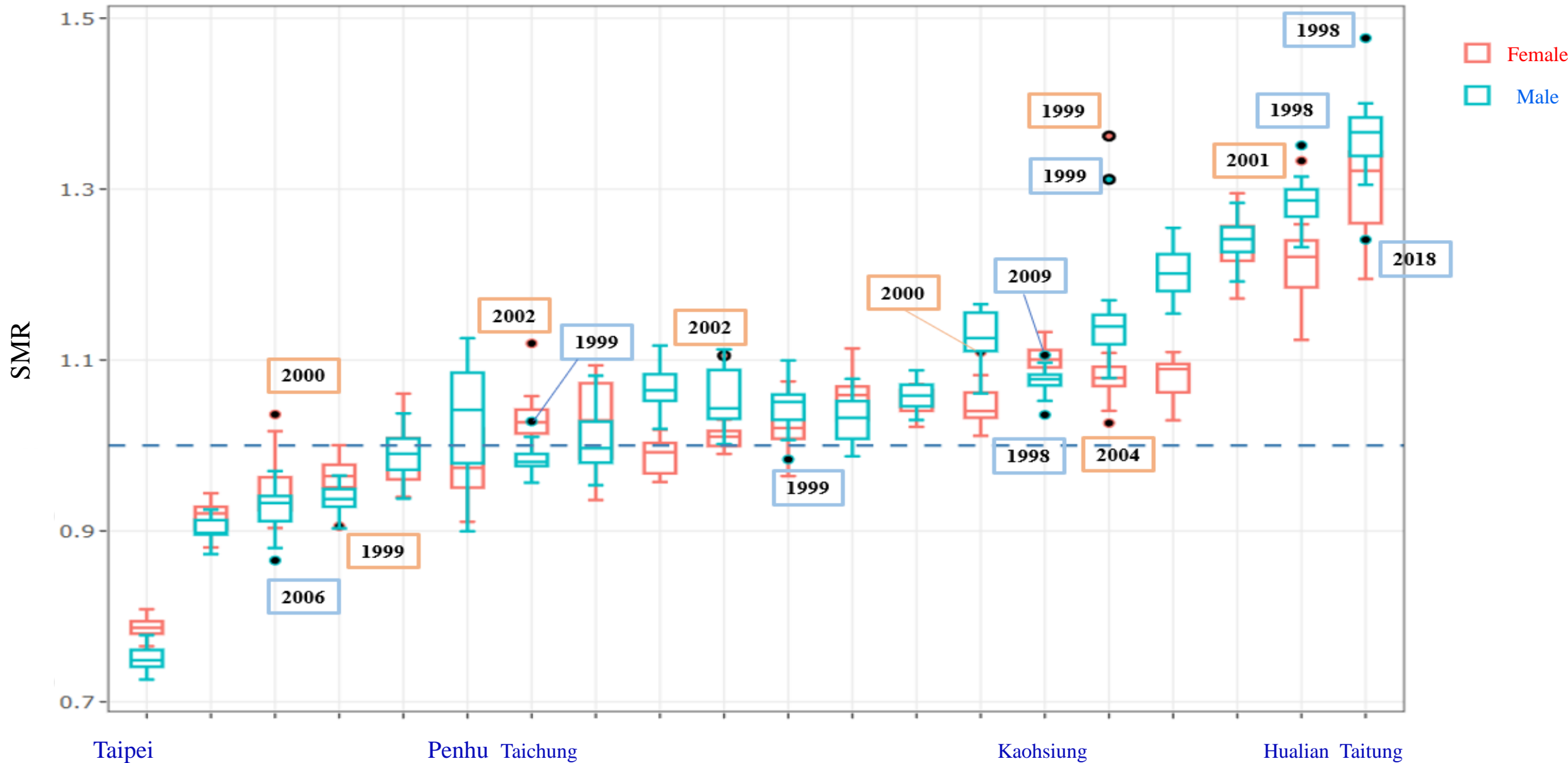
- Two modification of GWR:

→ Multiscale geographically weighted regression (**MGWR**; Fotheringham et al., 2017) and Conditional Geographically Weighted Regression (**CGWR**; Leong and Yue, 2017)

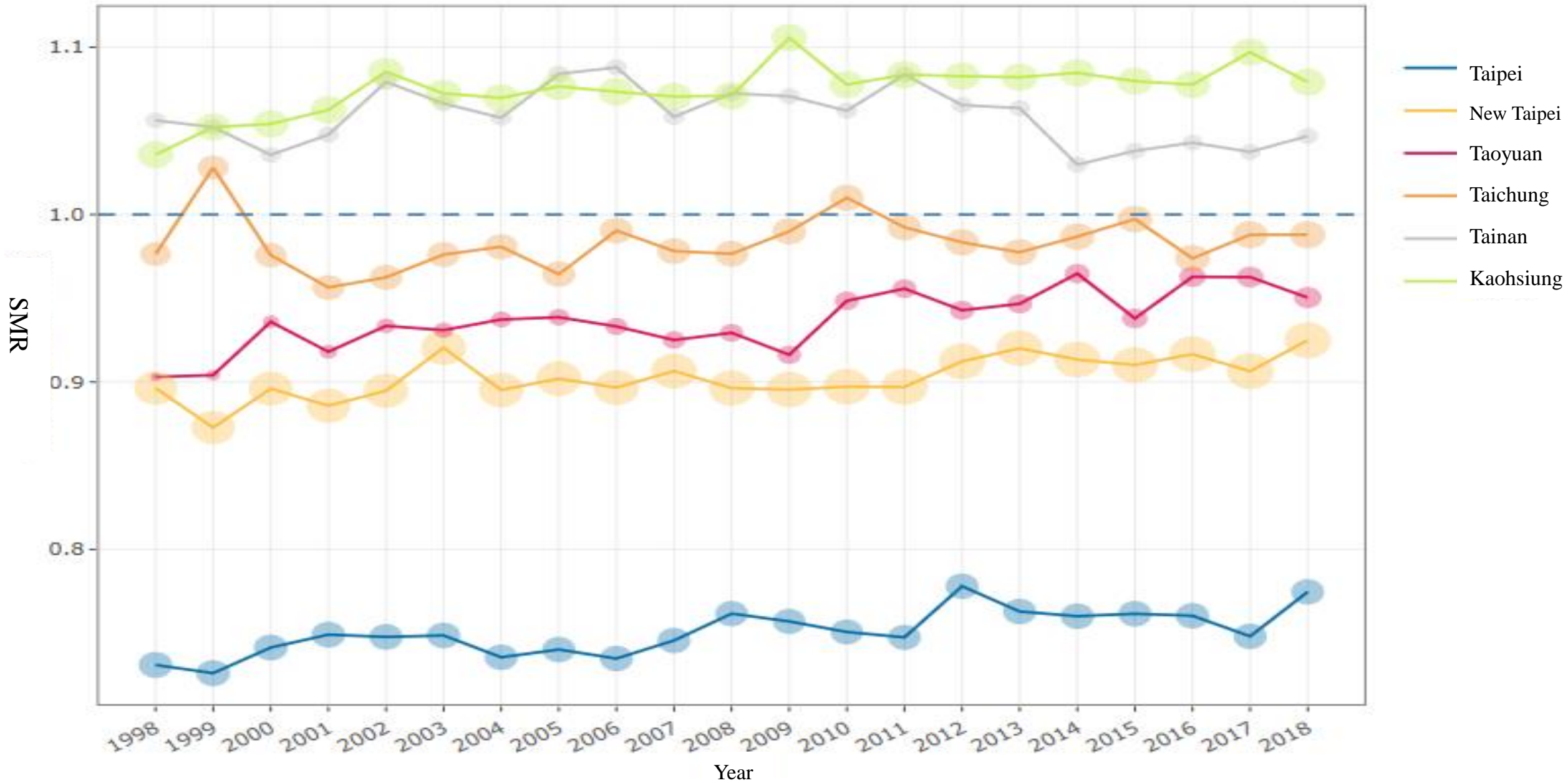
County-level Mortality Rates

- County-level SMR's show steady trend and strong signs of regional differences. (Note: Standardized Population is Taiwan every year.)
 - Larger fluctuations for counties with fewer populations.
- SMR's vs. Latitude/Population sizes
 - Northern cities have lower mortality rates and SMR's look unchanged in time.
 - More populations often are associated to lower mortality rates.

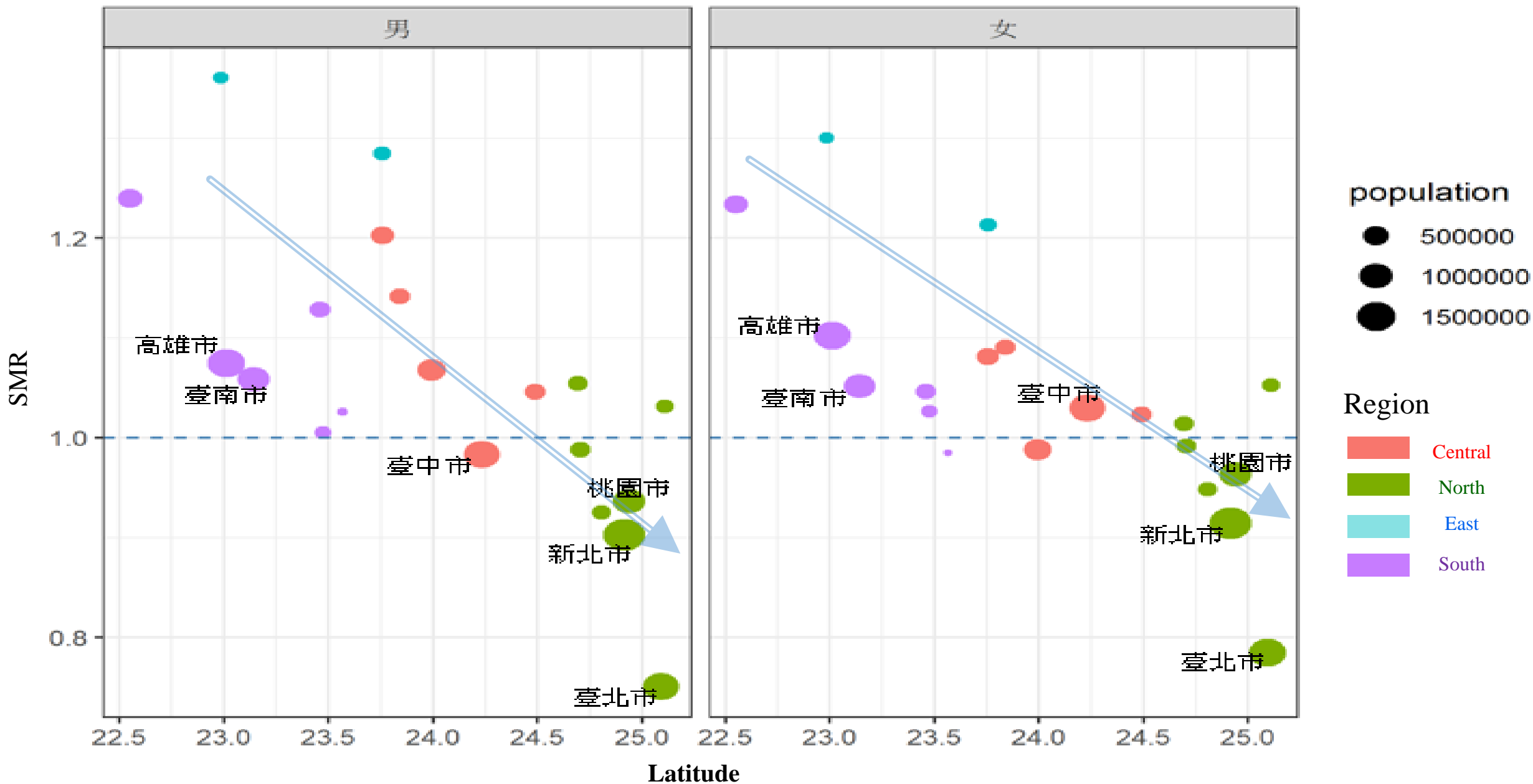
Boxplot of County-level SMR in Taiwan (1974-2018)



Male SMR of Six Taiwan Major Cities (1998-2018)



County-level SMR vs. Latitude (2020)



Spatial Analysis of Mortality Rates

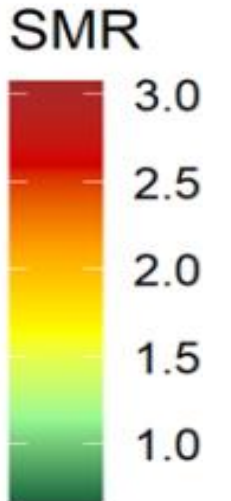
- All and major death causes SMR's (and standardized mortality rates) at township level show spatial in-homogeneity.

SMR	Moran's I	p-value
Cancer	0.379	< 0.05
Heart Disease	0.340	< 0.05
Accident	0.275	< 0.05
Cerebrovascular Disease	0.156	< 0.05
Pneumonia	0.117	< 0.05

Township Level Standardized Mortality Rates in Taiwan (2020)

Female

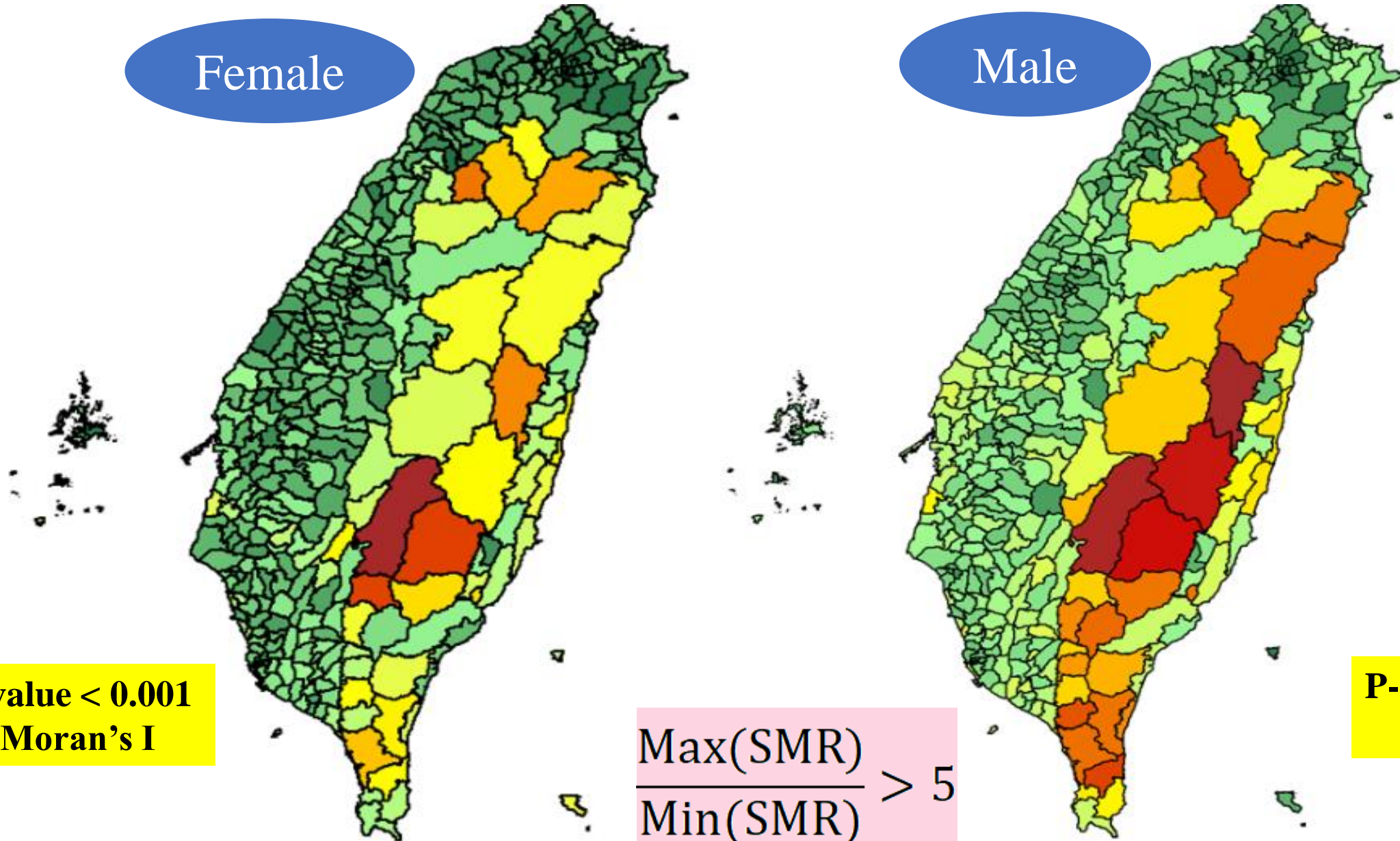
Male



P-value < 0.001
Moran's I

$$\frac{\text{Max(SMR)}}{\text{Min(SMR)}} > 5$$

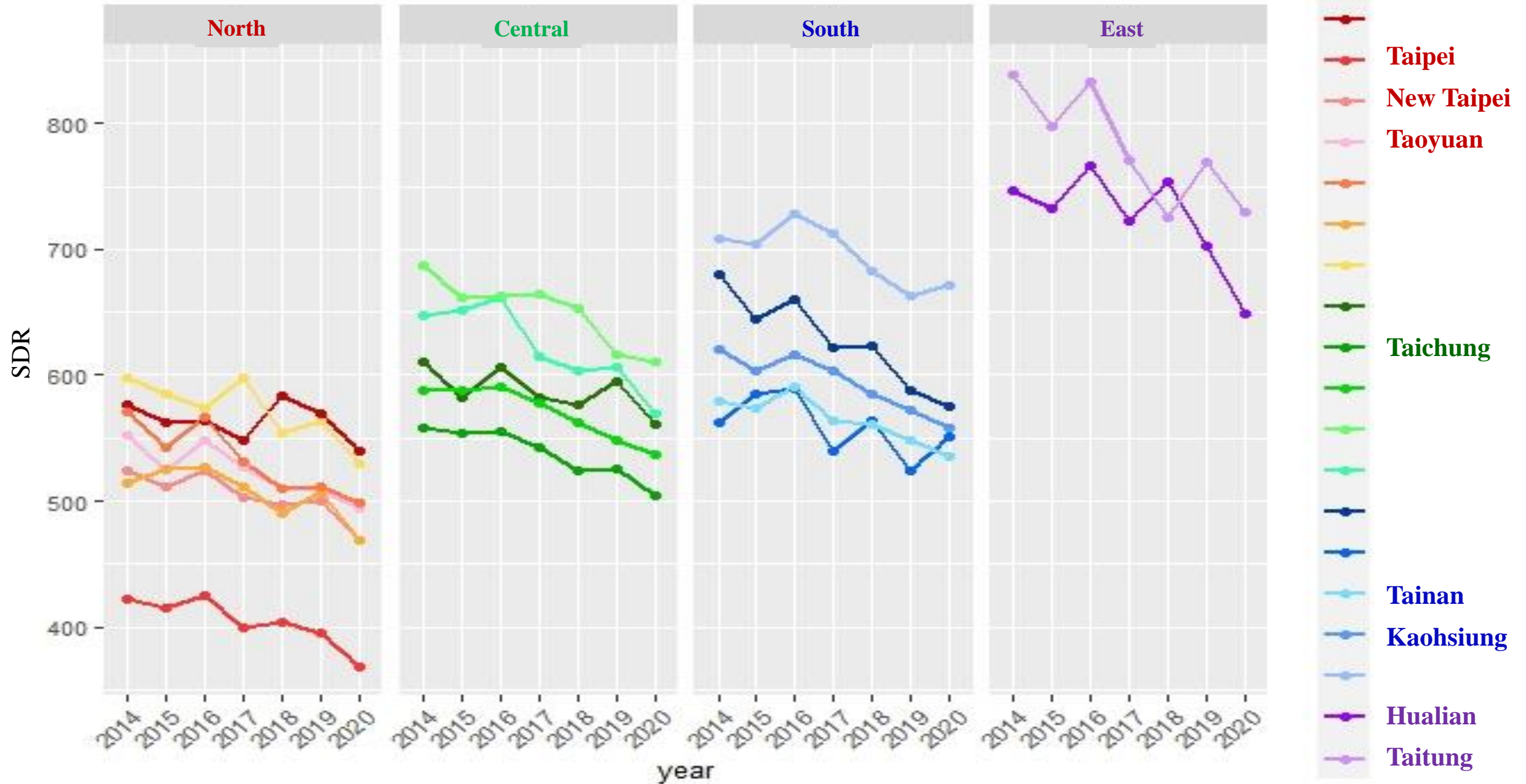
P-value < 0.001
Moran's I



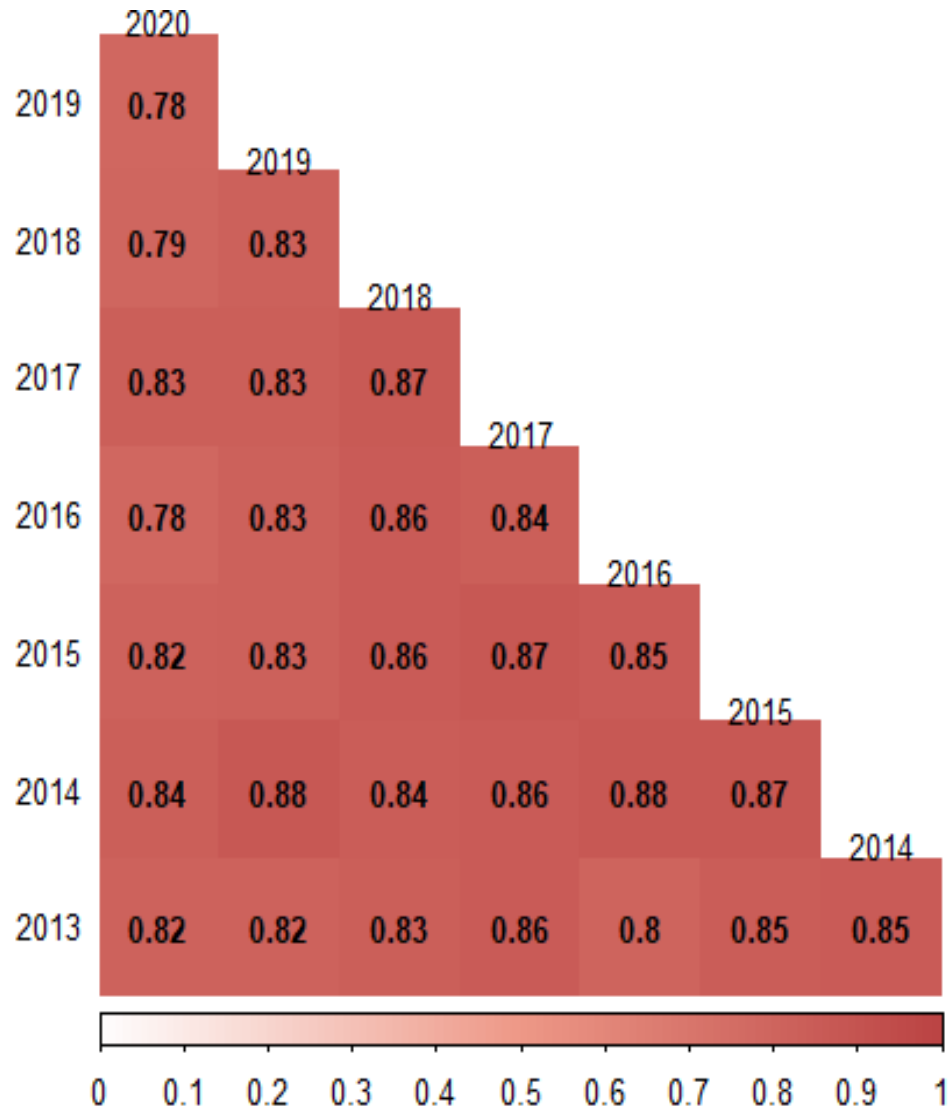
Time Trend of Regional Mortality Rates

- Regional differences of mortality rates appear to be fixed with time. (Mountain areas have higher mortality rates.)
 - The mortality rates of all counties appear to have same change rates and have high correlation between years.
- The year-to-year differences of mortality rates for all townships seem to be stable. (look like normal distribution!)
 - Applying Moran's I to these differences, we cannot reject the hypothesis if they satisfy spatial homogeneity.

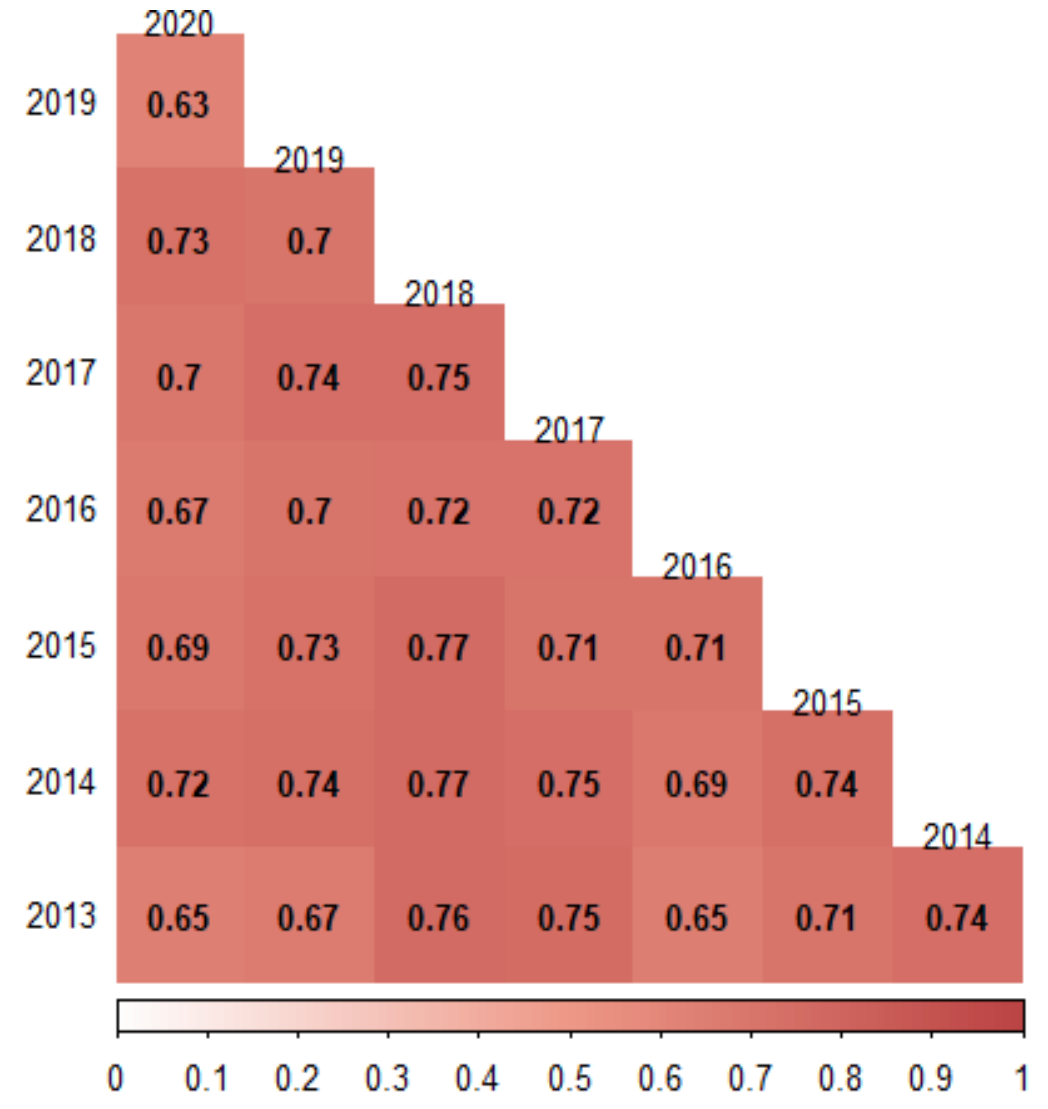
Time Trend of County-level Standardized Mortality Rates



Correlation of County-level Standardized Mortality Rates



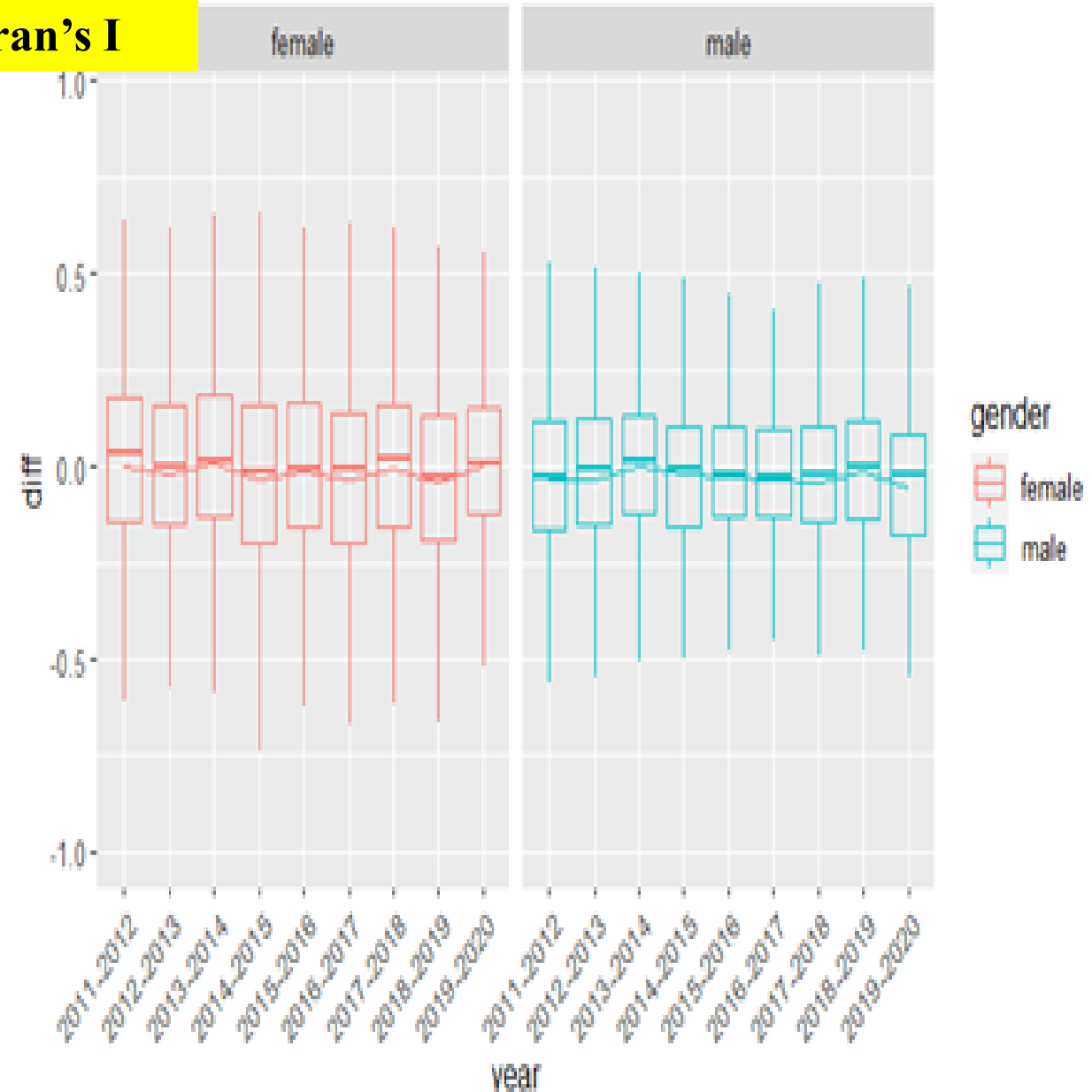
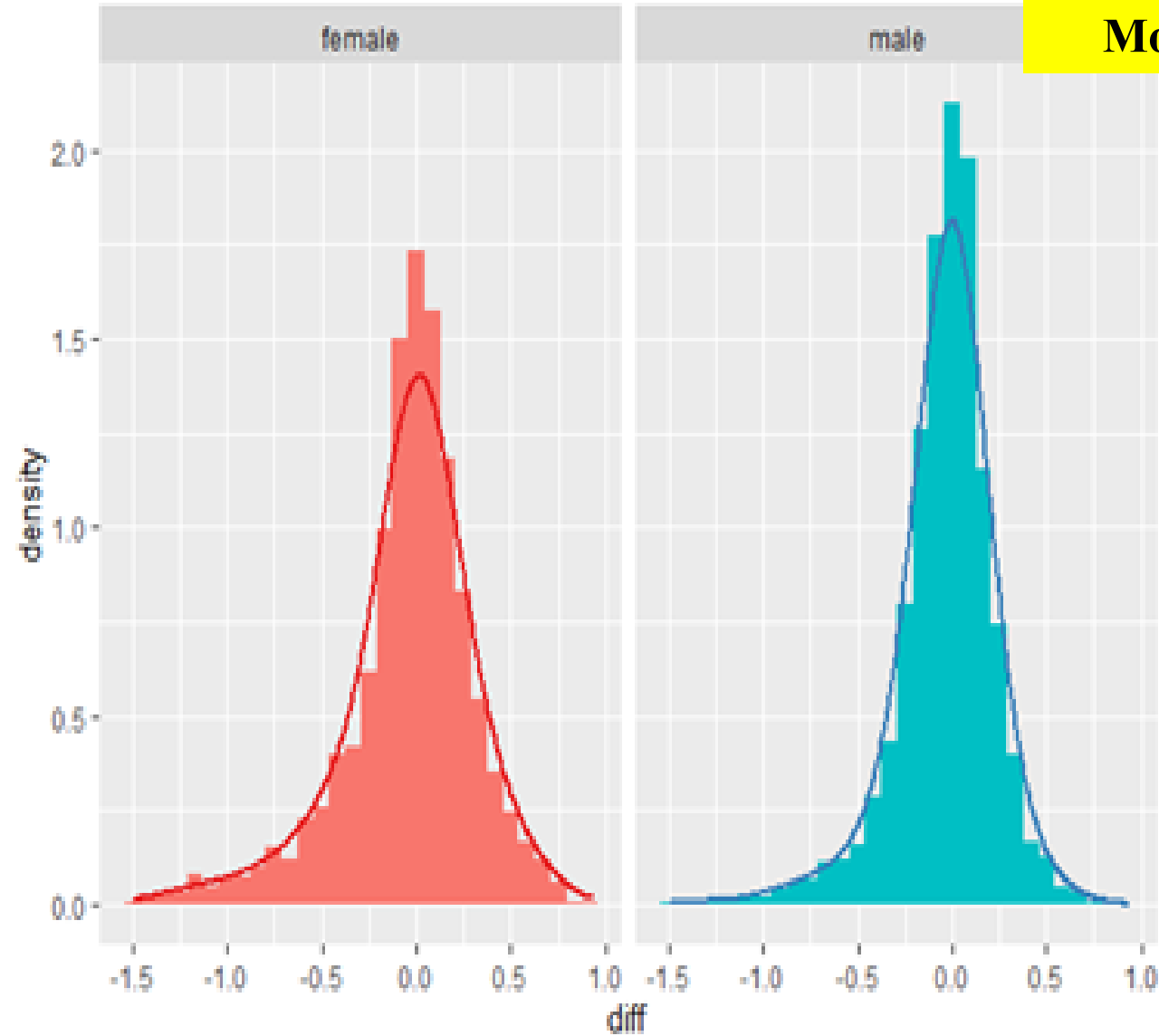
Correlation of Male mortality rates



Correlation of Female mortality rates

Differences of County-level Standardized Mortality Rates

P-value > 0.05
Moran's I



Exploring the Regional Medical Utilization

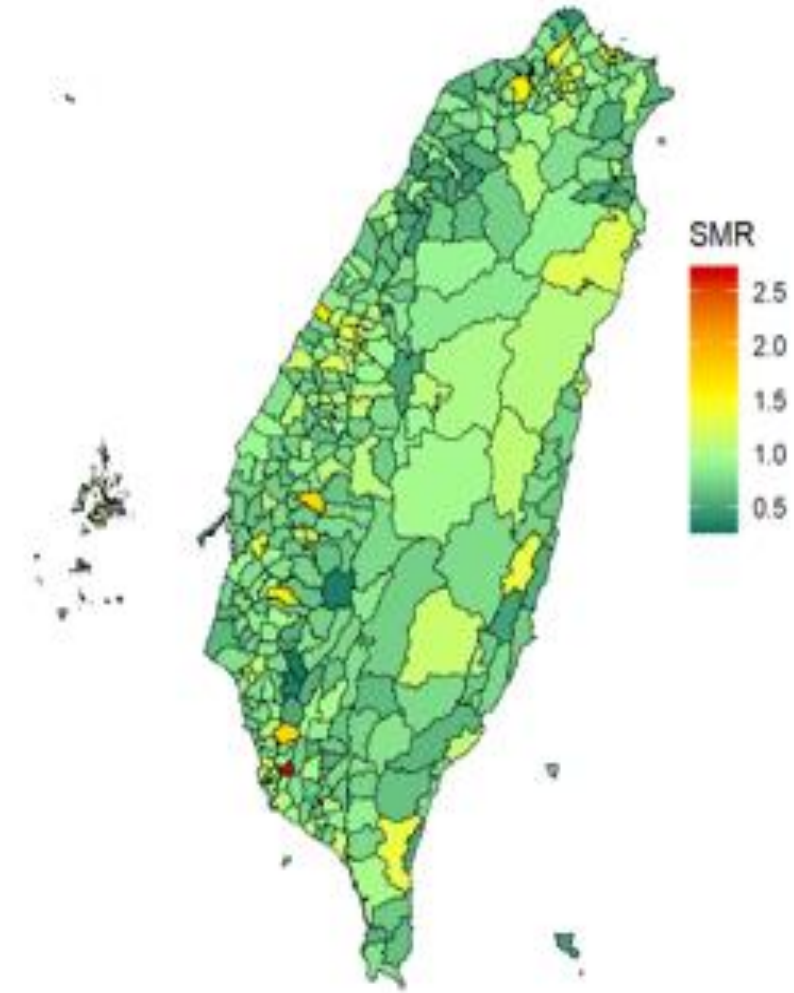
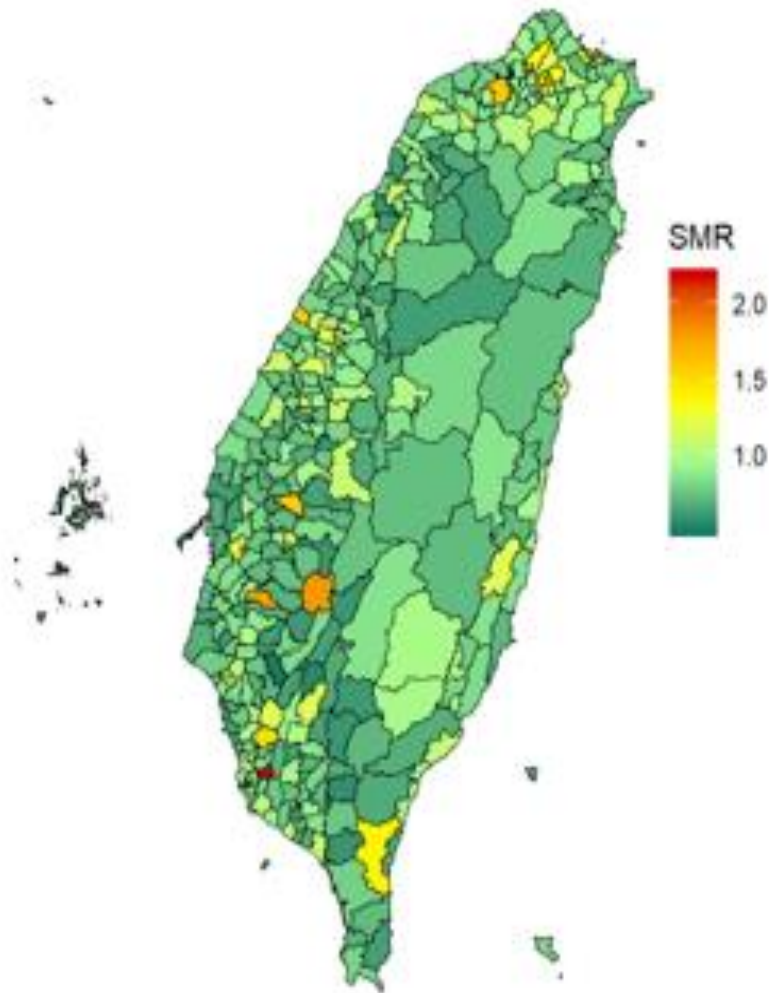
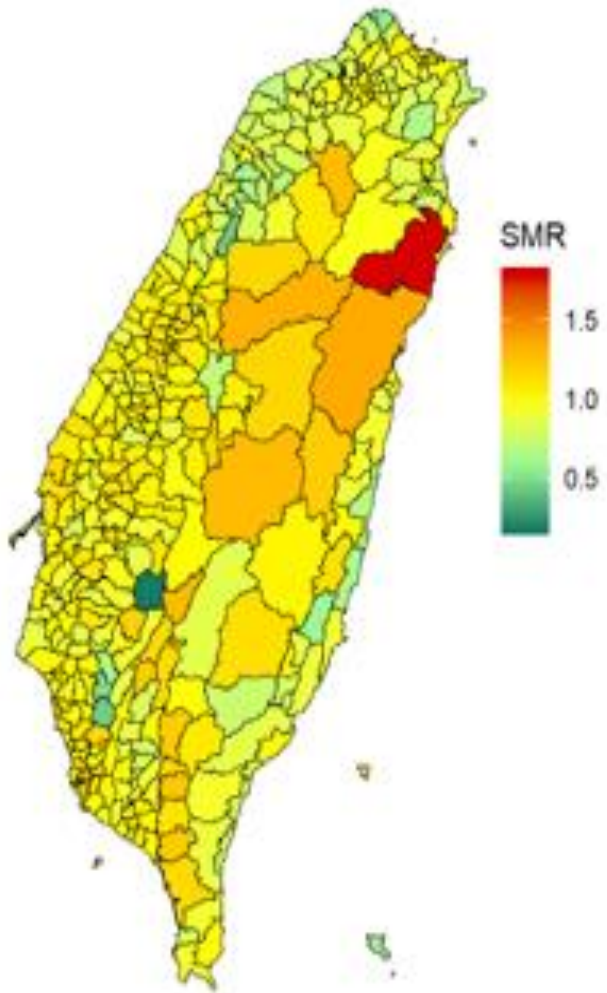
- ❑ Medical utilization do not have apparent spatial pattern.
 - Township-level number of outpatient visits, cost per visit, and total outpatient cost appear to be close to spatial homogeneity.
- ❑ Using the number of outpatient visits as an example, longer life is not associated with more/fewer medical visits.
 - Infants (& little kids) and older populations have more medical visits.
 - Fewest visits appear at about age 20, most at age 80, and the numbers of visits decrease after age 80.

Correlation of County-level Standardized Mortality Rates

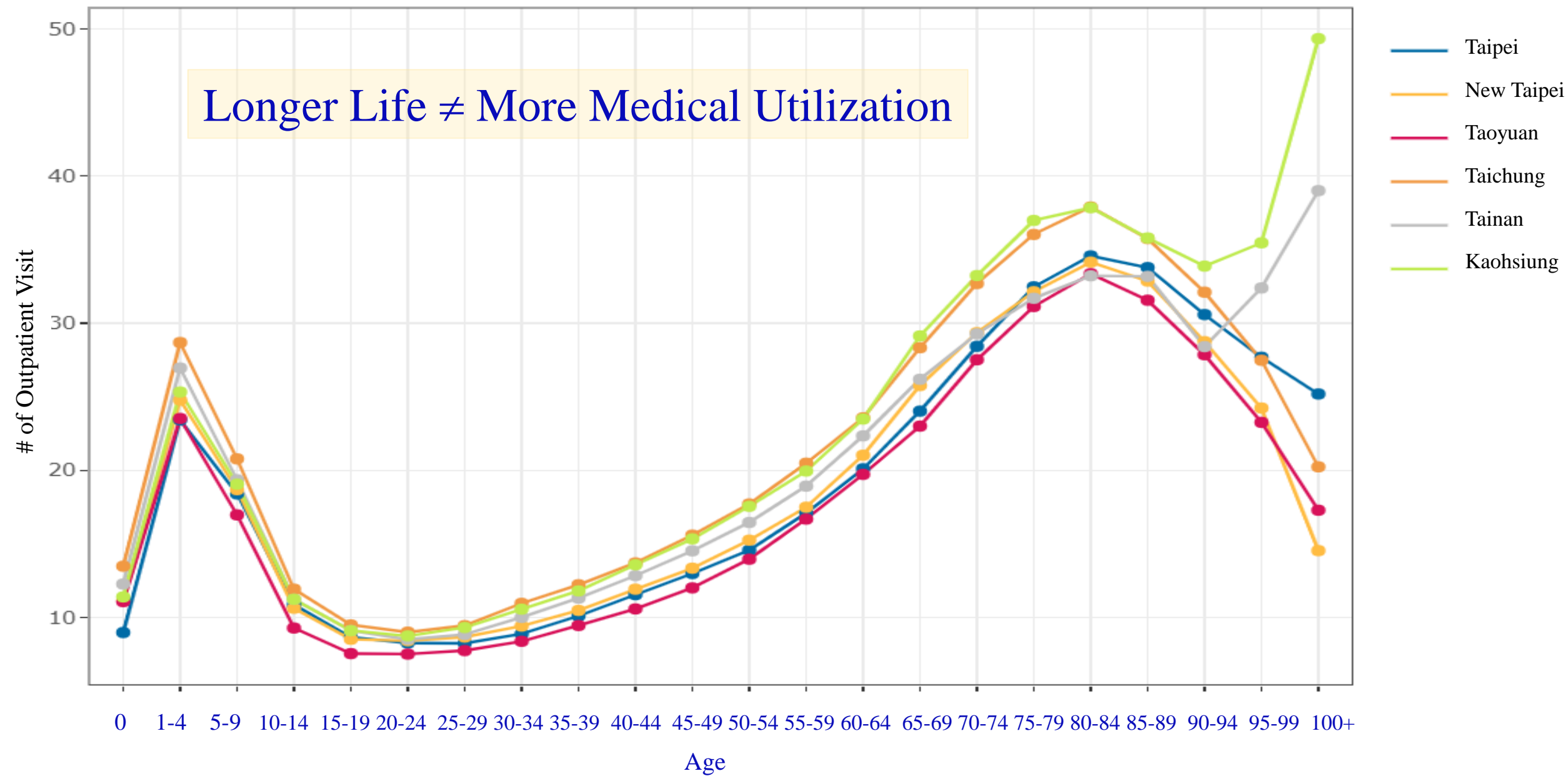
Outpatient Visits

Per Visit Cost

Total Cost



Outpatient Visits of Six Major Cities



Regional Mortality Rates and GWR

- Treat *log(township mortality rates)* as the dependent variable Y and *log(population size) & log(number of outpatient visits)*, or X_1 and X_2 , as the independent variables, i.e.,

$$y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \varepsilon_i$$

- Two modification versions of GWR have higher R^2 , 0.84, 0.94, and 0.94 for GWR, CGWR, MGWR, respectively.

→GWR have larger residuals.

Estimation Results of CGWR (2005 Taiwan Male)

Higher β_0 :
Mountains



© Mapbox © OSM

Population β_1 :
Mountains



© Mapbox © OSM

Visits β_2 :
Mountains

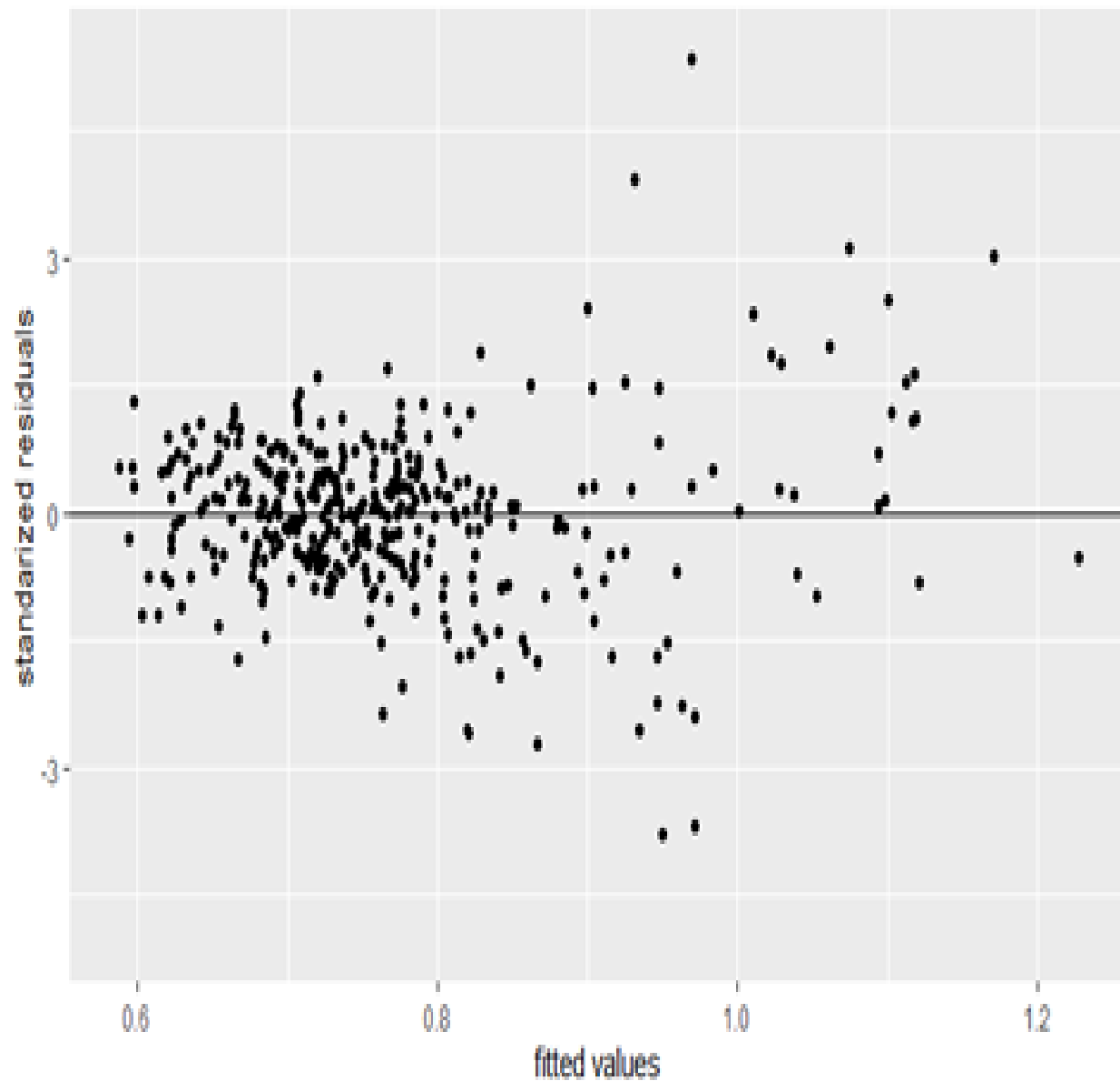


© Mapbox © OSM

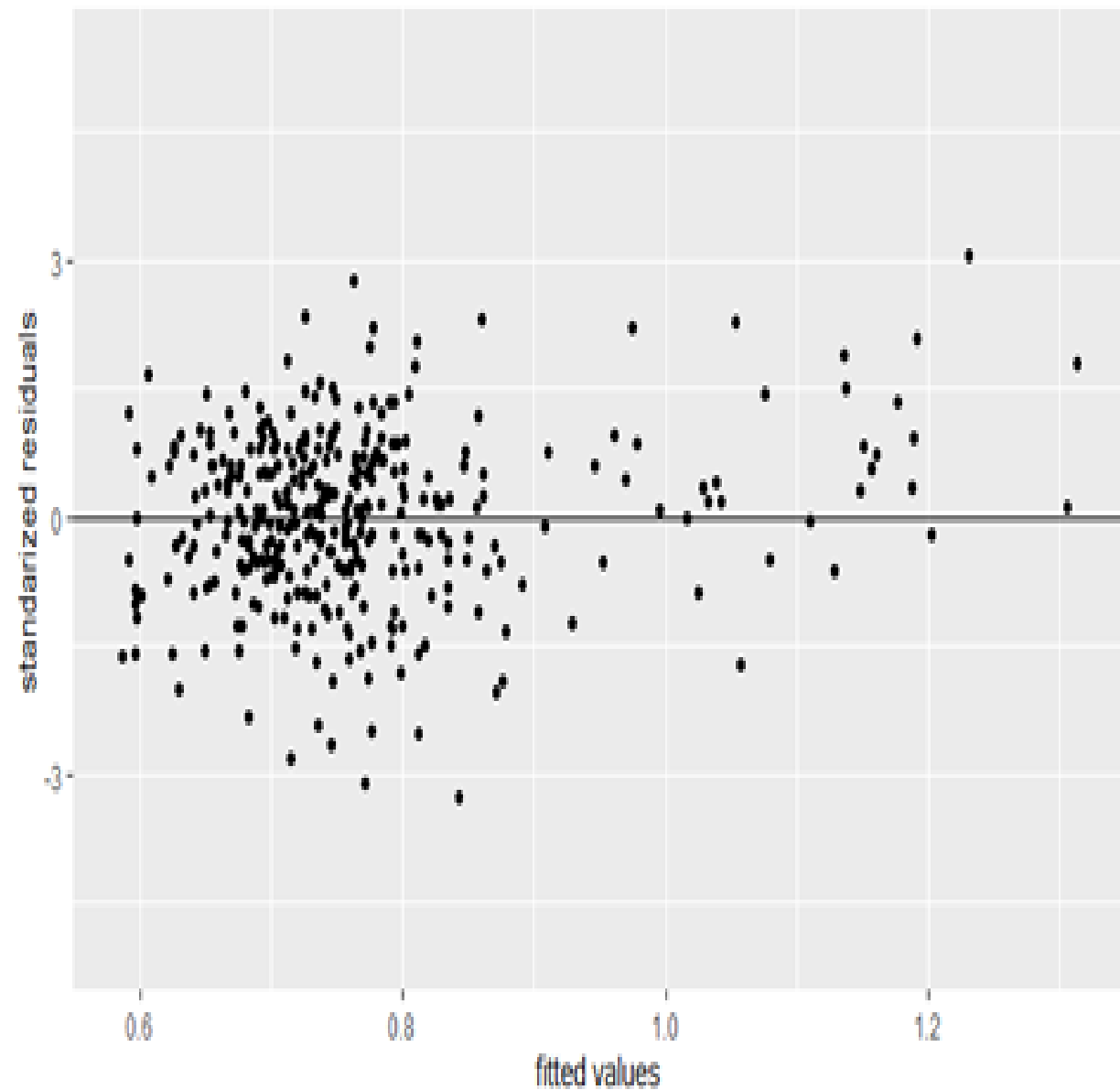


Residual Plots of GWR and CGWR

GWR residual plot



CGWR residual plot



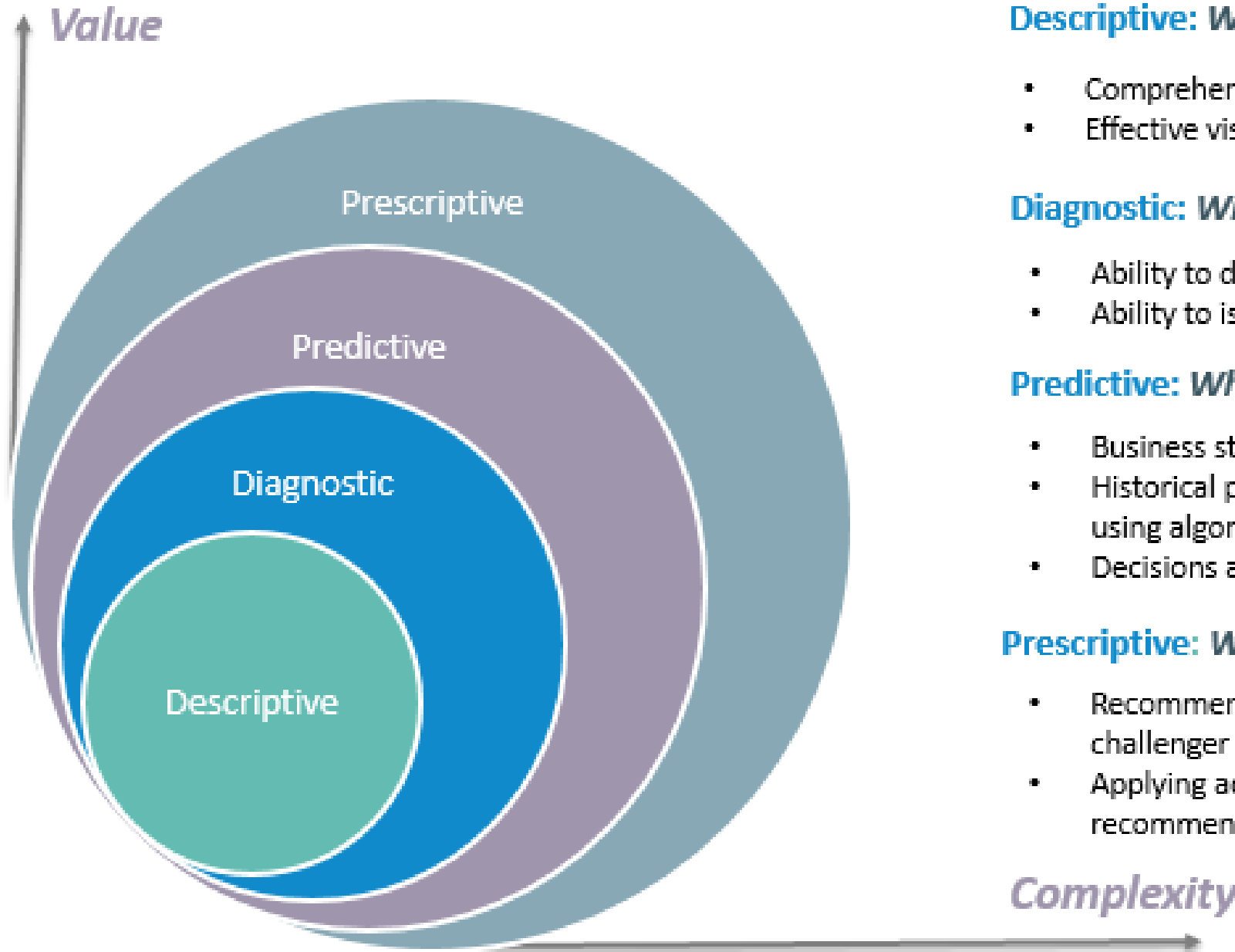
Conclusion

- ❑ Mortality rates and life expectancy show strong signs of spatial in-homogeneity. (But not medical utilization...)
 - People live northern areas and populous cities have lower mortality rates. (There may be clusters of higher mortality.)
 - Prolonging life seems to be at a similar pace in Taiwan and the gaps between areas are narrowing.
- ❑ Spatial methods can be useful in exploring regional properties.
 - Identifying hot spots (clusters) and measure the significance of independent variables.

Types of Data Analysis

- ❑ Statistical data analysis can be separated into:
 - **EDA (Exploratory Data Analysis)** applies visualization and summary statistics to get an overview of the data.
 - **CDA (Confirmatory Data Analysis)** evaluates the evidence using statistical tools, e.g., significance and inference.
- ❑ Other suggestions of types of data analysis:
 - **Descriptive Analytics** (*What has happened*), **Predictive Analytics** (*What would happen*), and **Prescriptive Analytics** (*What should we do*).

4 types of Data Analytics



What is the data telling you?

Descriptive: *What's happening in my business?*

- Comprehensive, accurate and live data
- Effective visualisation

Diagnostic: *Why is it happening?*

- Ability to drill down to the root-cause
- Ability to isolate all confounding information

Predictive: *What's likely to happen?*

- Business strategies have remained fairly consistent over time
- Historical patterns being used to predict specific outcomes using algorithms
- Decisions are automated using algorithms and technology

Prescriptive: *What do I need to do?*

- Recommended actions and strategies based on champion / challenger testing strategy outcomes
- Applying advanced analytical techniques to make specific recommendations

Discussions

□ Impact and further studies

→ Regional mortality differences on insurance products

→ Small areas estimation (Lee-Carter model): different α_x & same β_x

□ Data Issue (de jure population vs. registered population)

→ Linking NHI data to household registration records

→ More data (e.g., in-patient visits and cause of death)

□ Apply more spatial analysis methods

→ Cluster detection (e.g., *SaTScan*)

→ Spatial-temporal analysis (space & time)



THANK YOU!!

