UNSW Business School

Never Stand Still | Business School | Risk and Actuarial Studies

Regional Differences in Life Expectancy in Mainland China

William Li, Hazel Bateman and **Kevin Liu**

ASTIN, AFIR/ERM and IACA Colloquia, 26 August 2015

ARC Centre of Excellence in Population Ageing Research

UNSW Business School

Background and motivation

- Rapid population ageing – unprecedented challenge to China esp. the sustainability of its PAYG pension system
- Uneven socioeconomic development – large regional difference and urban-rural disparity
- Fragmented pension system – locally managed, different arrangements for urban and rural populations
- Mortality forecasting – predominantly at national level, failed to consider different regional mortality trends
- Lack of mortality projection at regional level & lack of understanding of the drivers of regional differences in mortality improvements – a major obstacle to pension reform policy formulation

Literature review and research gap

- Lee-Carter model (Lee and Carter, 1992) and its variants (e.g. Lee and Miller, 2001; Booth et al., 2002; De Jong and Tickle, 2006; Russolillo et al., 2011) require long term data
- Prior research has focused on developed countries and selected middle-income countries (Alho, 2008; Booth et al., 2006; Lee and Miller, 2001; Lee, 2000; Tuljapurkar et al., 2000)
- Limited studies on developing countries including China due to the lack of high quality long term data (Li et al., 2009)
- Li, Lee and Tuljapurkar (2004) developed an extention to the Lee-Carter model to address the limited data issue – national level forecasting, regional differences remain unknown
- Standard Lee-Carter model fails to capture socioeconomic factors which may have substantially impact on mortality (e.g. Hanewald, 2011)

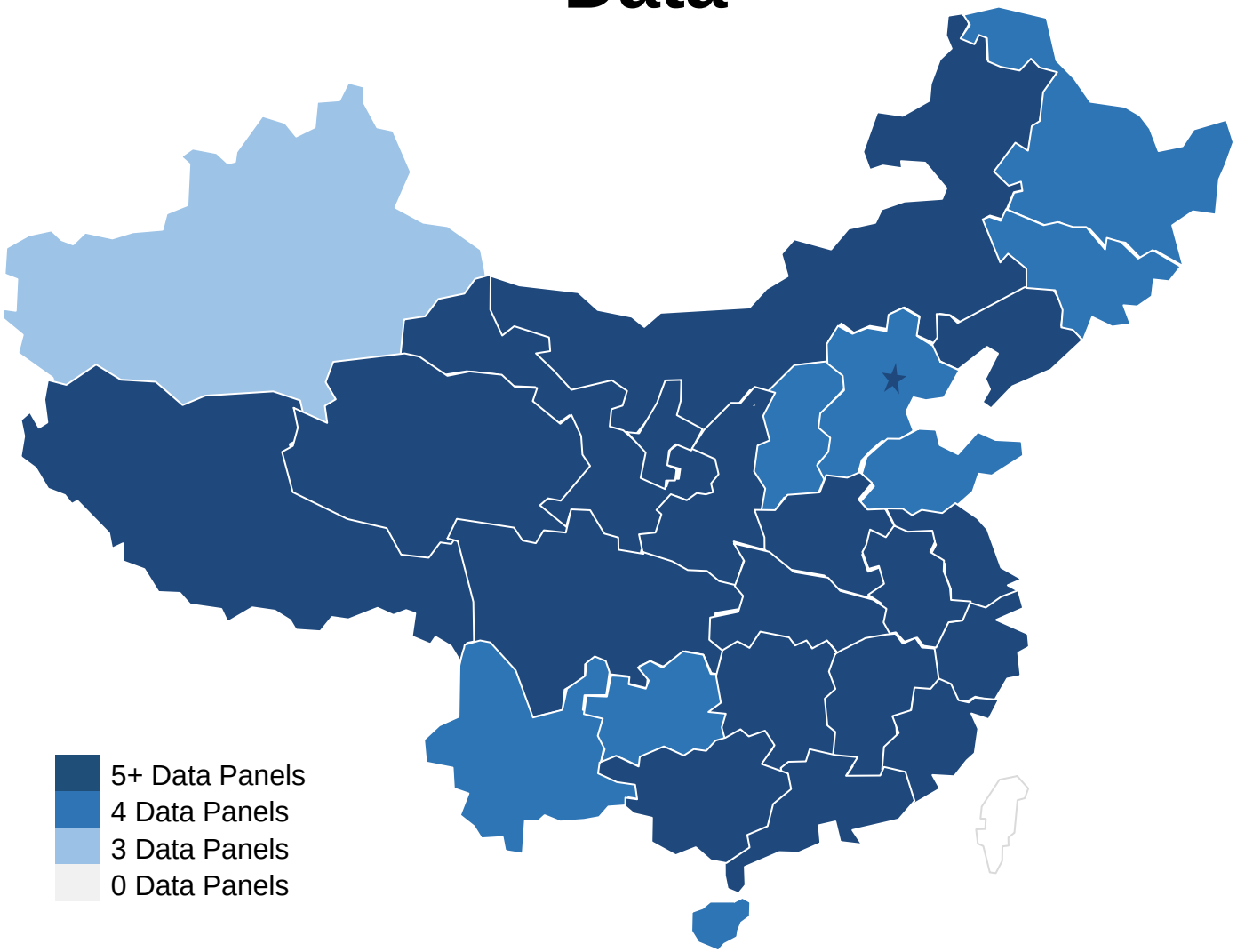
Research Objective

- Forecast the long-term mortality trends (and hence, the life expectancy) in China at both national and regional level (by province and urban-rural populations)
- Investigate the future changes in China's regional differences in life expectancy
- Examine the socioeconomic drivers of China's regional differences in life expectancy

Data

- Data sources:
  - Census (National/Provincial Reports: 1990, 2000, 2010)
  - National Population Survey (National/Provincial Reports: 1987, 1995, 2005)
- Databases:
  - National Bureau of Statistics of China
  - China National Knowledge Infrastructure
- Age-sex-specific mortality data:
  - Six panels (1987, 1990, 1995, 2000, 2005 and 2010); 30 provinces

Data



Data

- Data sources:
  - Census (National/Provincial Reports: 1990, 2000, 2010)
  - National Population Survey (National/Provincial Reports: 1987, 1995, 2005)
- Databases:
  - National Bureau of Statistics of China
  - China National Knowledge Infrastructure
- Age-sex-specific mortality data:
  - Six panels (1987, 1990, 1995, 2000, 2005 and 2010); 30 provinces
  - Death, exposure and death rates by age, sex and region at both national and provincial level for infants and standard five-year age groups up to age 100, and for age 100 and over
  - 800 regional life tables (by gender, province and urban/rural populations) constructed following Cai (2005)
- **A unique mortality dataset**

Methodology

Lee-Carter model (Lee and Carter, 1992):

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

The unique least square solution for the model can be obtained under two specific constraints:

$$\sum_{x=1}^{w} b_x = 1 \quad \text{and} \quad \sum_{t=1}^{w} k_t = 0$$

Under the constraints, a_x can be found as the average over time of $\ln(m_{x,t})$. Then though the application of singular value decomposition (SVD) method, b_x and k_t can be obtained as well.

Methodology

The extended time series methods developed by Li, Lee and Tuljapurkar (2004) are used to address the issue of limited data (unequal intervals) in China.

The extended model of the mortality index - k_t is modelled as follows:

$$k_{\mathrm{t},\varphi,\theta} = k_T + c\,(t-T) + \sigma \sum_{s=T+1}^{t} e(s)$$

where k_T represents the most recent intensity of mortality, and $\sum_{s=T+1}^{t} e(s)$ describes random changes in the future with each $e(s)$ following the standard normal distribution

Note: c and σ could be obtained as follows:

$$\hat{c} = \frac{\mathrm{k(u(T))} - \mathrm{k(u(0))}}{\mathrm{u(T)} - \mathrm{u(0)}},\quad \hat{\sigma} = \frac{\sum_{t=1}^{T}\{k(u(t)) - k(u(t-1)) - \hat{c}\,[u(t)-u(t-1)]\}^2}{u(T)-u(0)-\frac{\sum_{t=1}^{T}[u(t)-u(t-1)]^2}{u(T)-u(0)}}$$

where $\mathrm{u(T)}$ represents for the year of the last data point and $\mathrm{u(0)}$ represents for the year of the first data point

Methodology

Li, Lee and Tuljapurkar (2004): the forecasted 95% probability intervals of k_t can be obtained by taking $\theta = 0$. The wide and narrow bounds of the forecasted 95% probability intervals can be obtained by using $\theta = 1.96$ and $\theta = -1.96$ respectively

Then the prediction of mortality behaviour can be calculated using:

$$\ln(m_{x,t}) = \ln(m_{x,T}) + b_x\ [\,k_t - k_T\,]$$

where k_t is forecast from equation in the previous page, $m_{x,T}$ reflects the most recent observed age-specific central death

Methodology

Socioeconomic factors - following Hanewald (2011):

- Phillips-Perron test is applied to analyse the trend behaviour of the mortality index k_t and socioeconomic indicators for each region
- Pairwise Pearson correlation coefficients were evaluated between the stationary form of the mortality index and the socioeconomic factors
- Socioeconomic factors which have significant correlation with the intensity of mortality are then selected to be incorporated into modelling of k_t
  - For instance, if the first difference of k_t and logarithm of some crucial socioeconomic variables ($X_{t,i}$) are found to be correlated, then the regression model can be constructed as follows:

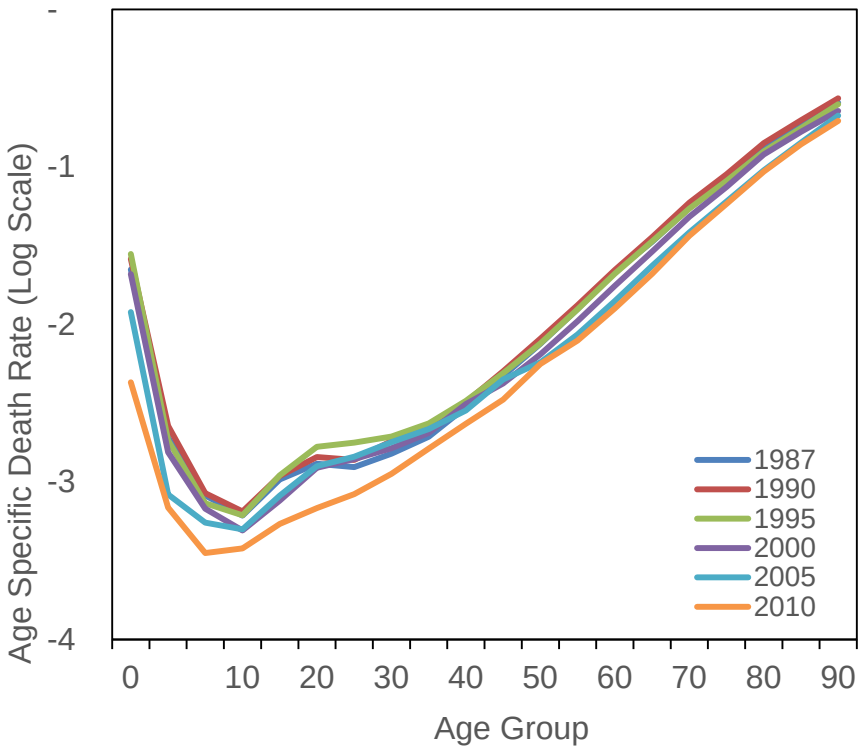
    $$\Delta\,(k_t) = \Theta + \sum_{i=1}^{n} \beta_i \Delta\,(\ln(X_{t,i})) + \varepsilon_{x,t}$$

    where $X_{t,i}$ represent selected socioeconomic factor at year t. Θ and β_i are the coefficients to be calculated. $\varepsilon_{x,t}$ is the error term with mean zero and variance σ_{ε}^2
  - The models developed in the previous sections will be compared based on the goodness-of-fit measure, R^2 (Cameron and Windmeijer, 1997)
  - The tractability and predictability of the models are taken into account in the comparison. The selected models will be used to forecast the future mortality rate

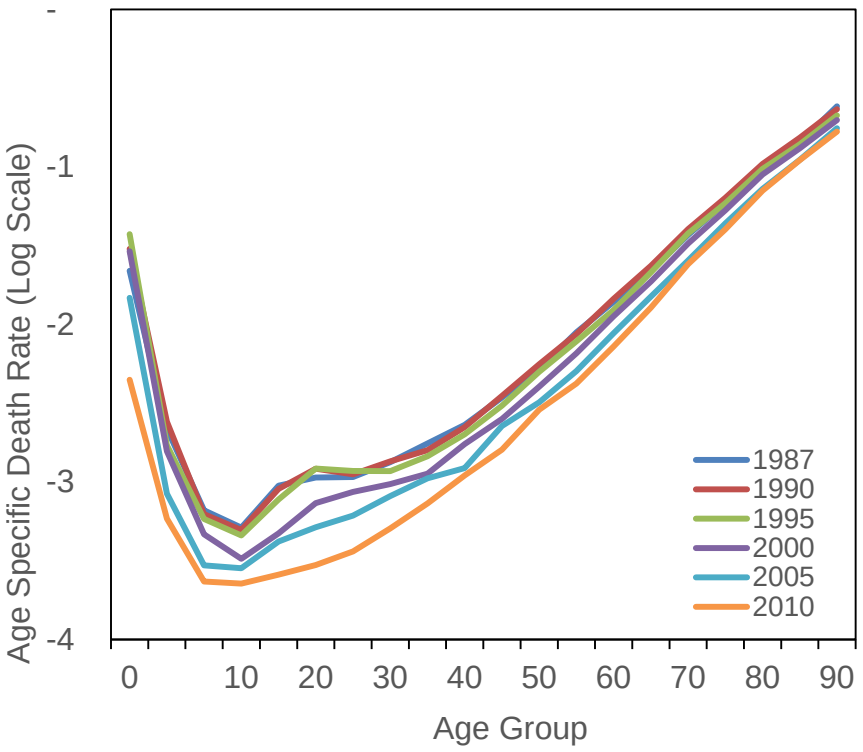
Results

Historical mortality rates (Total Population), 1987 - 2010

(a) Male



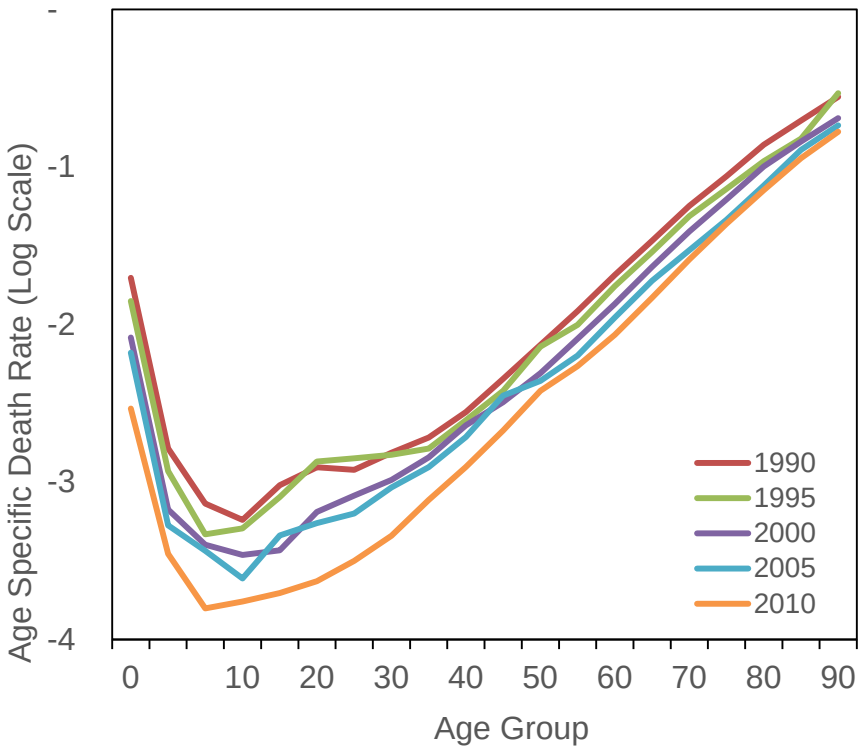
(b) Female



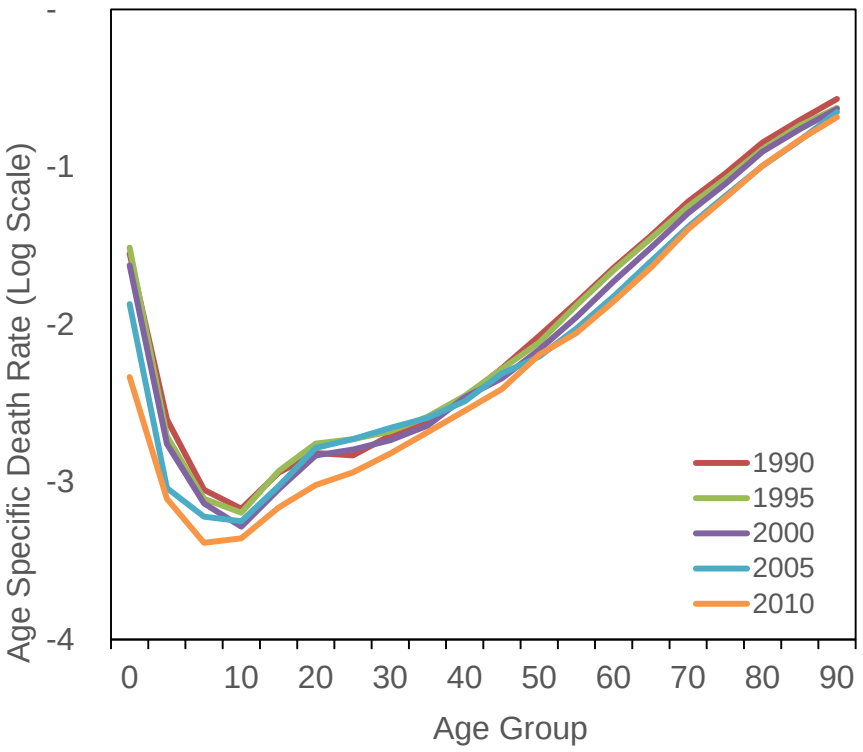
Results

Historical mortality rates (Male), 1990 - 2010

(a) Urban

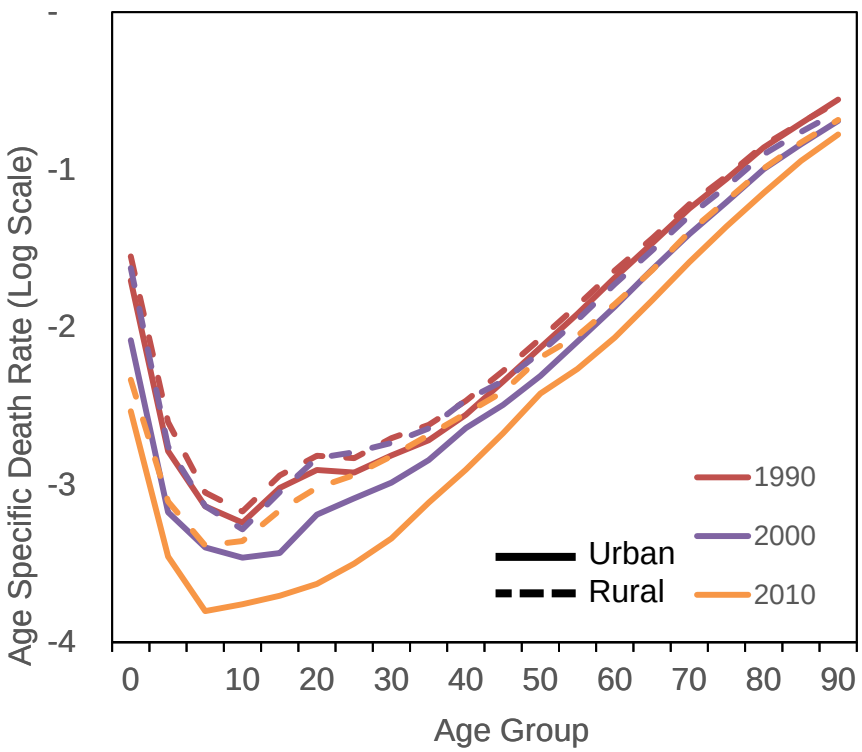


(b) Rural



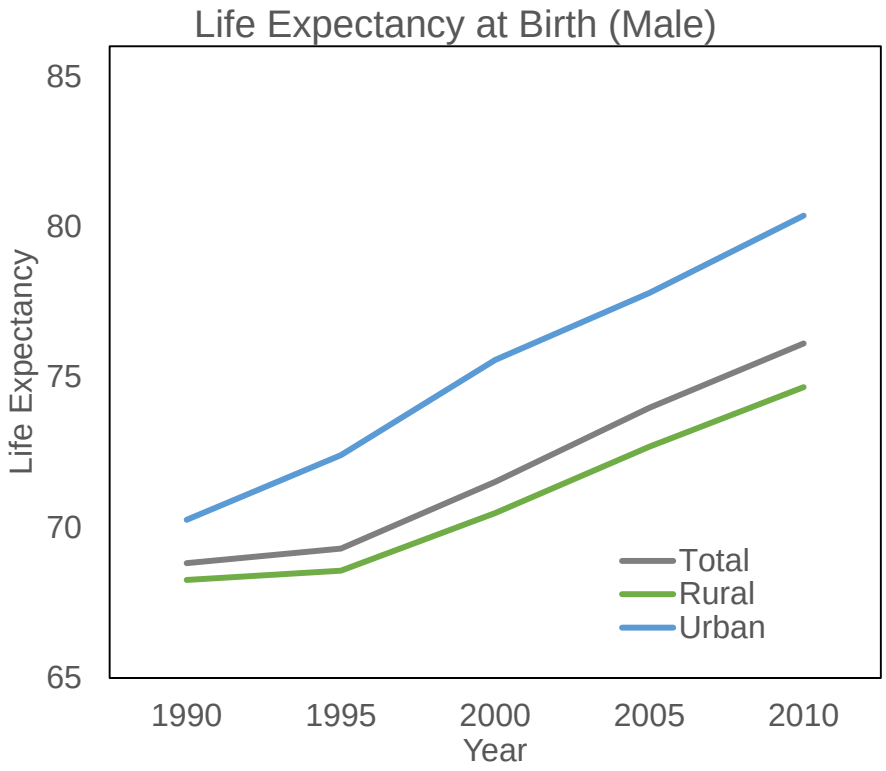
Results

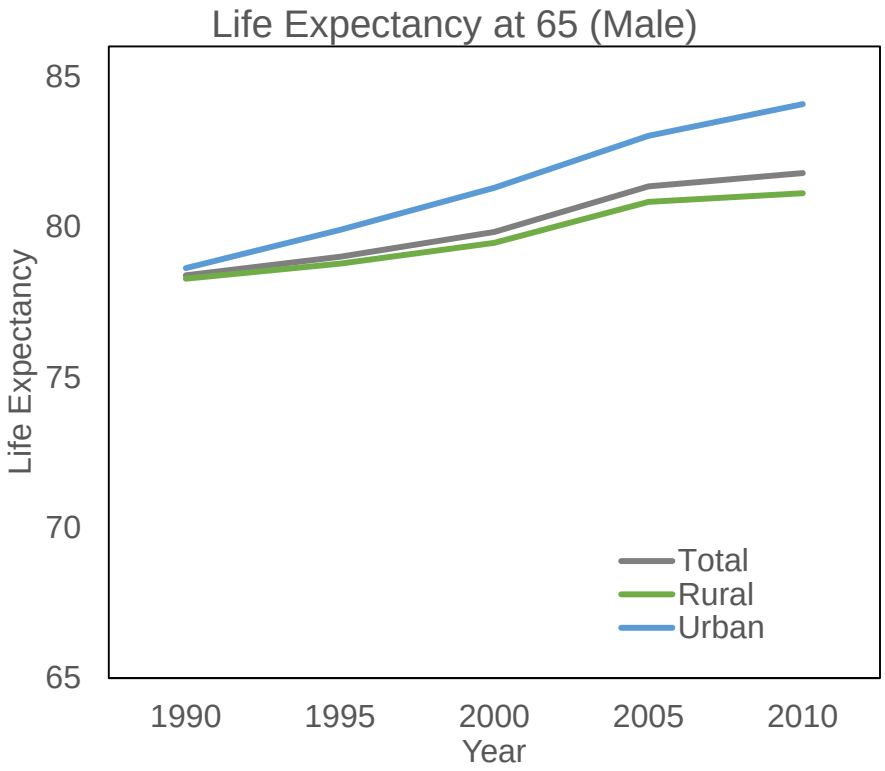
Historical mortality rates (Male), 1990 - 2010



Results

Life Expectancy (Male), 1990 - 2010

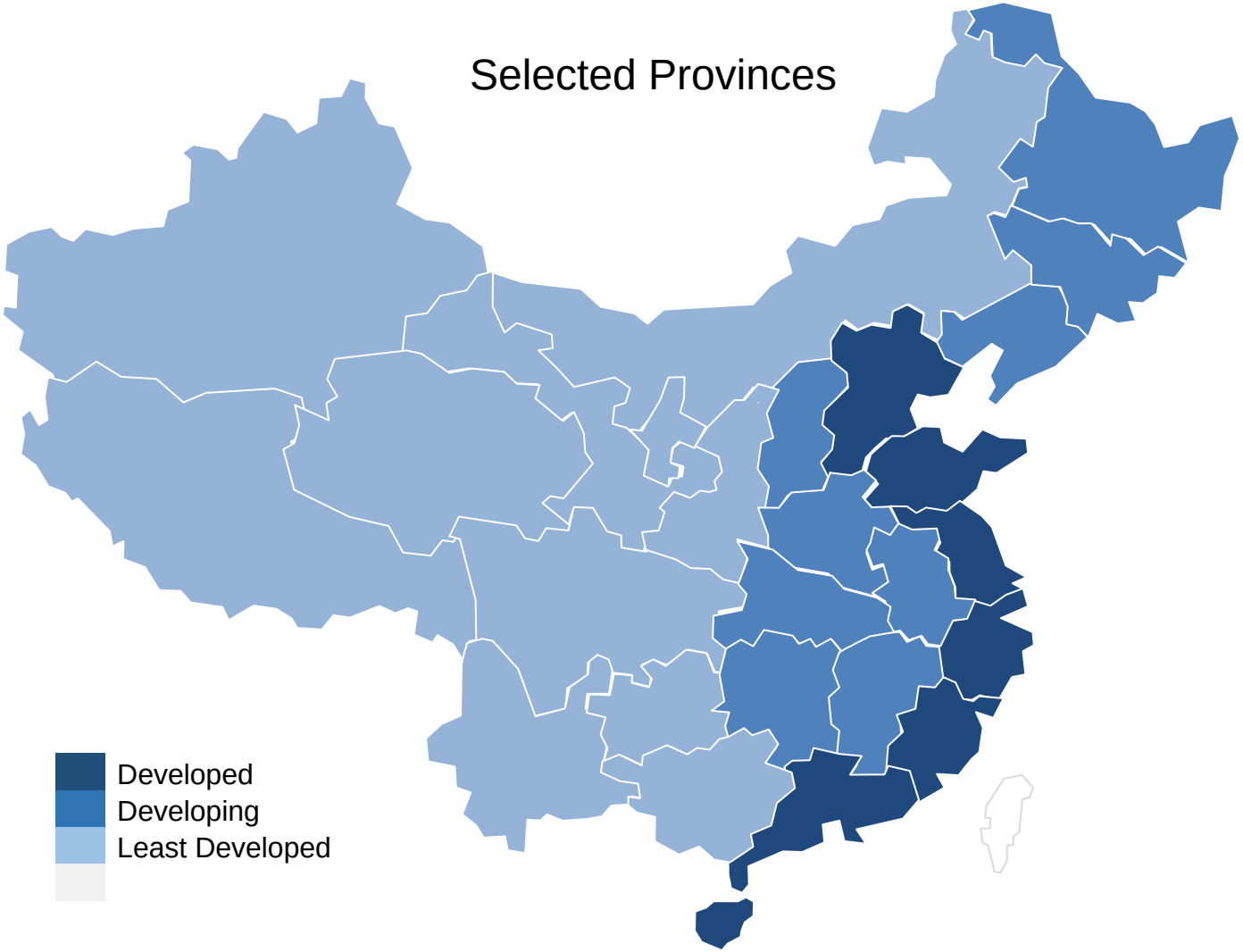




Results

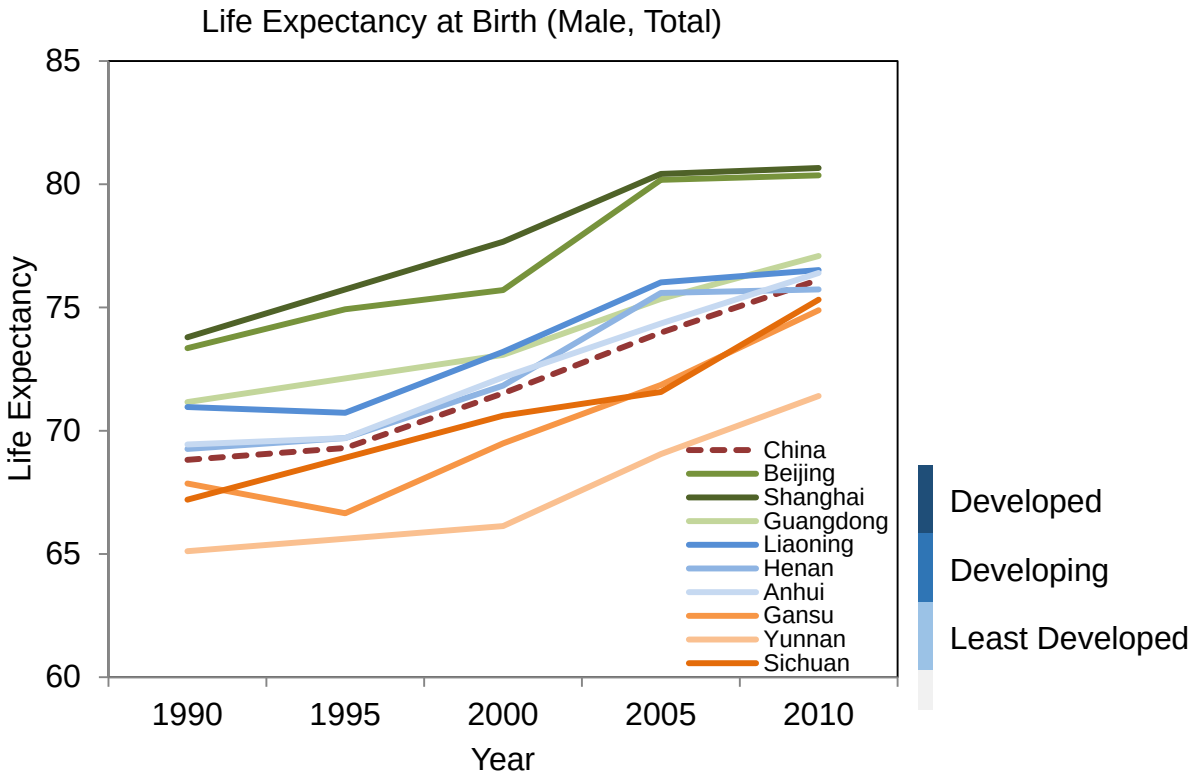
Selected Provinces

Results


Selected Provinces
Developed
Developing
Least Developed

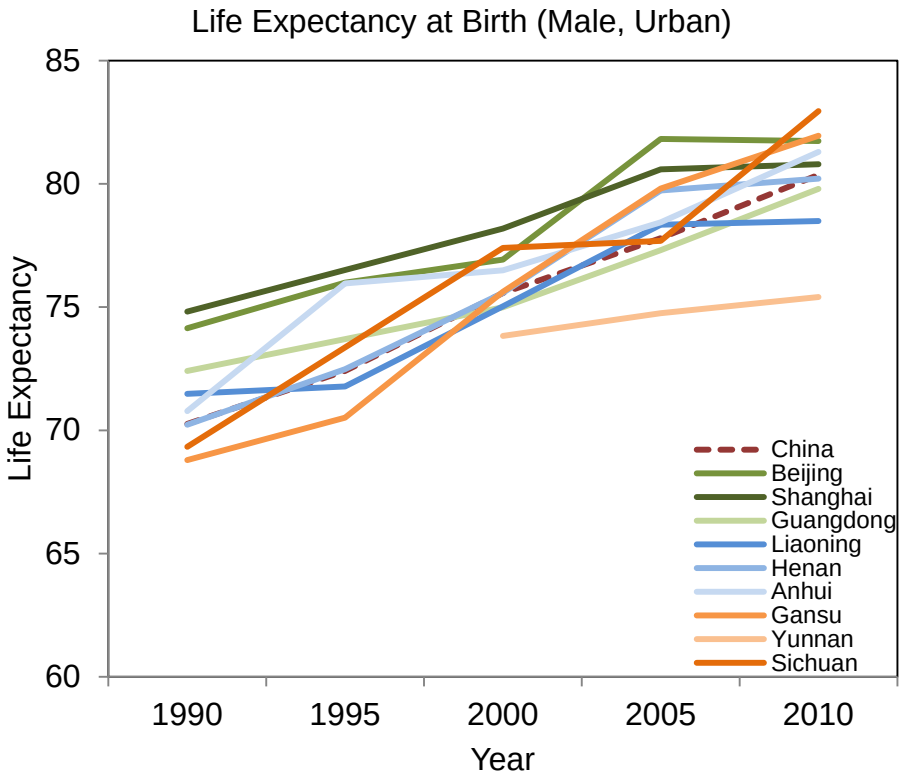
Results

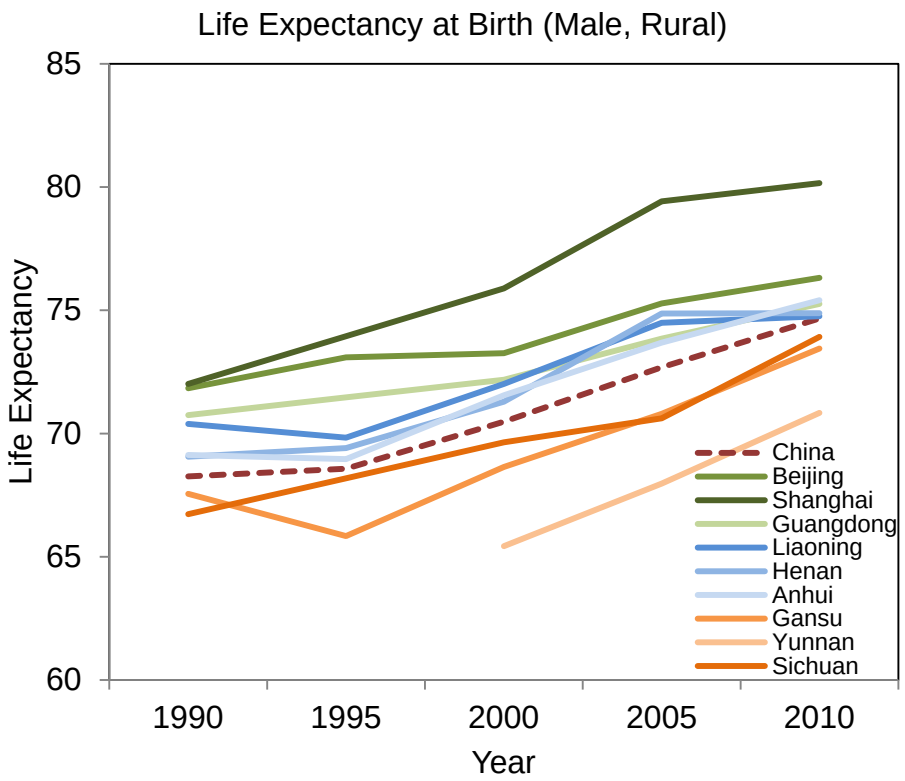
Life Expectancy (Male, Selected Provinces), 1990 - 2010



Results

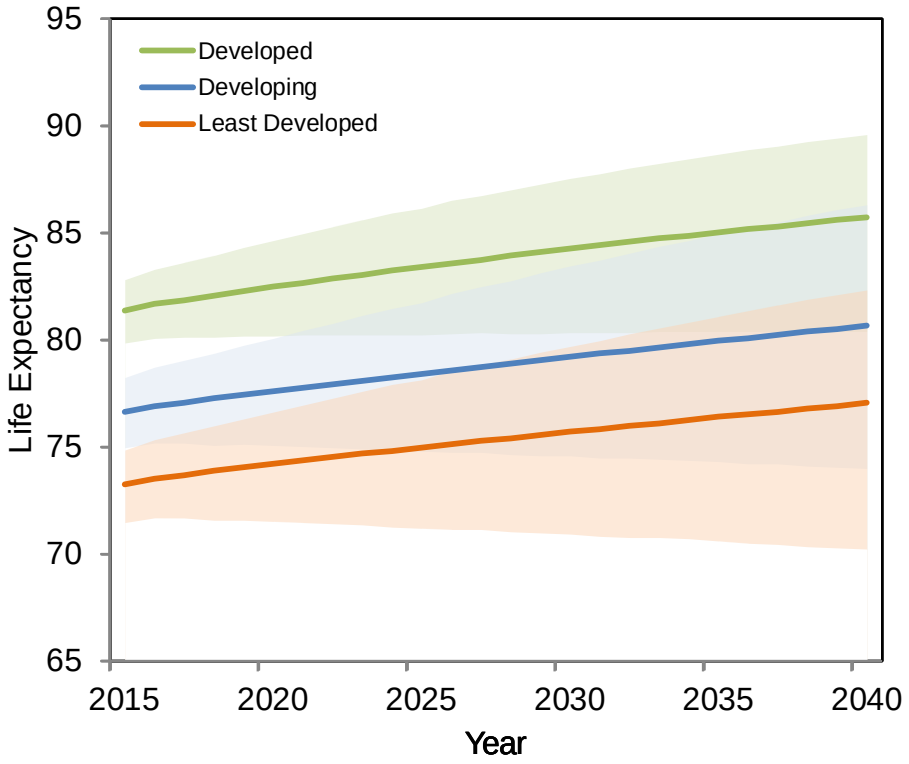
Life Expectancy (Male, Selected Provinces), 1990 - 2010





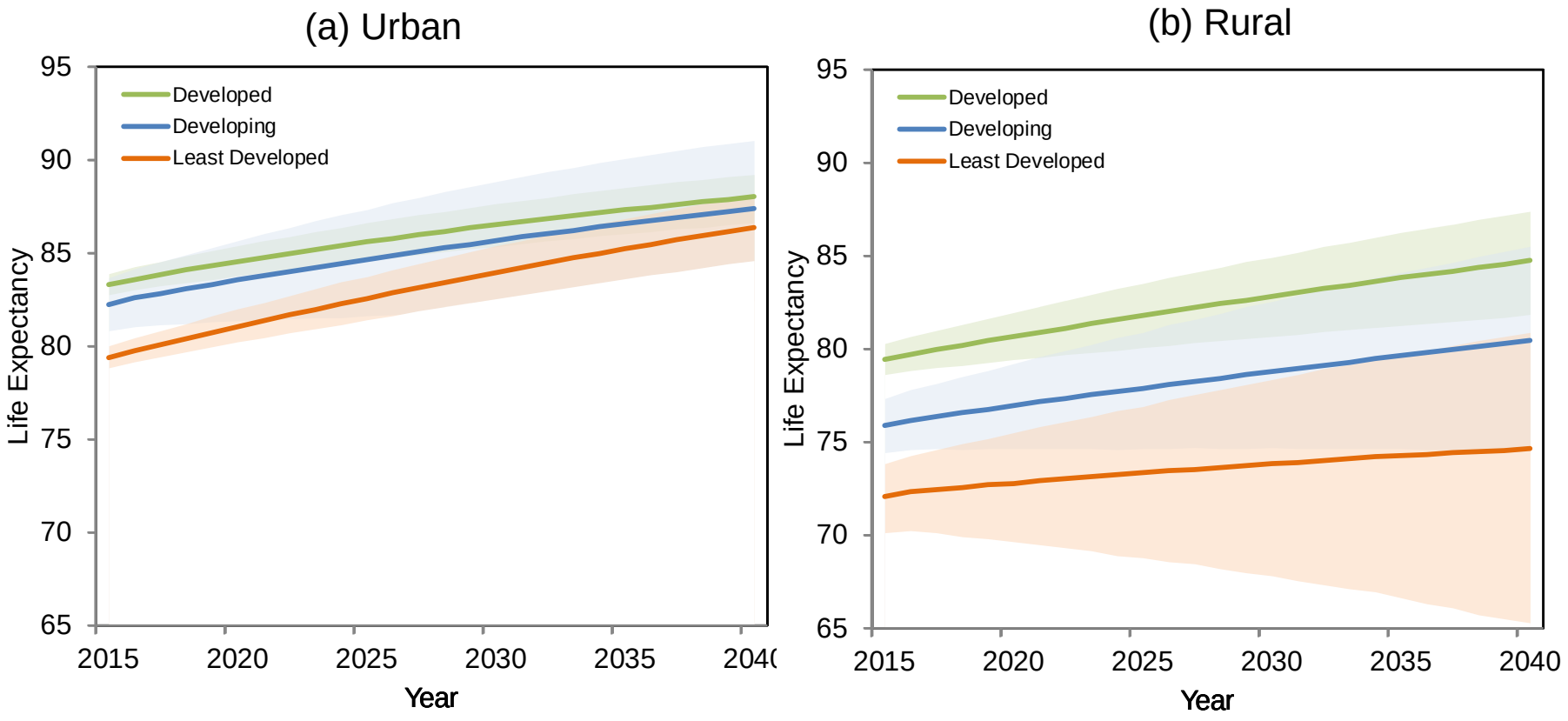
Results

Projected Life Expectancy (Male, Total), 2015 - 2040



Results

Projected Life Expectancy (Male), 2015 - 2040



Results

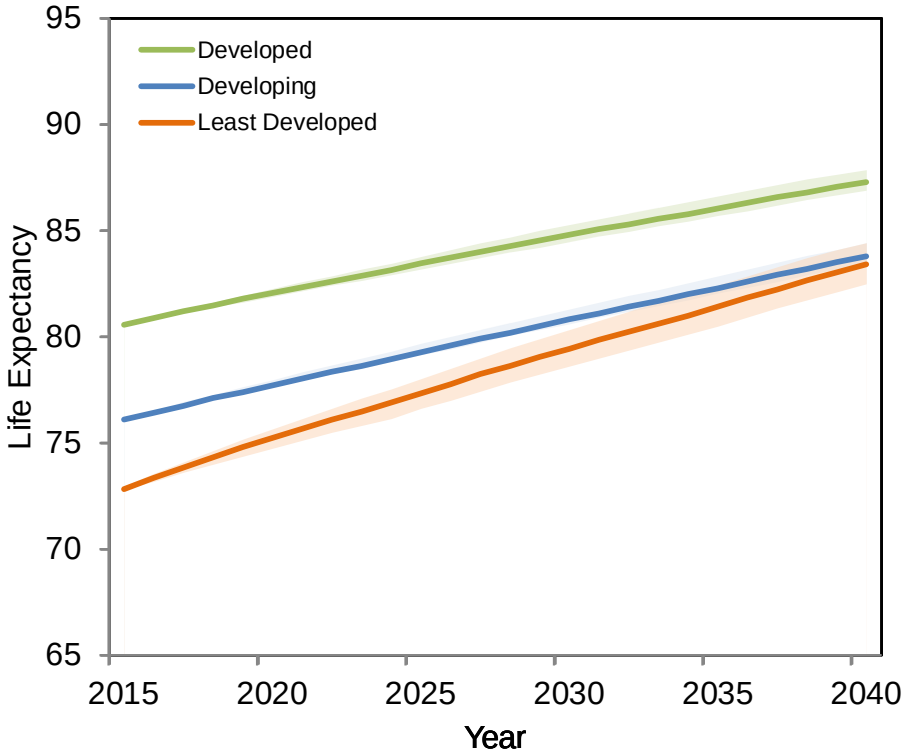
Socioeconomic factors

- Additional data for examining the impact of socioeconomic factors on mortality improvement:
  - Age-sex-specific mortality data: National China Population Statistics Yearbook (1994 – 2012)
  - Socioeconomic variables selected: GDP growth rate, GDP per capita, income per capita, unemployment rate, and change of number of doctor per 10,000 people (1994 – 2012)
- Following Hanewald's (2011) section criteria, GDP growth rate is found the have strong correlation with k_t (for both genders)
- Future regional growth rates of GDP were estimated according to Abeysinghe and Rajaguru (2004), Brandt and Rawski (2008), Yuan et al. (2007) and OECD (2012) which provided estimates of the growth rate of GDP for each region with varied level of development in China
- Base, optimistic and pessimistic scenarios were constructed for classified regions
- Then future mortality improvements were estimated and sensitivity analysis was conducted

Results

Projected Life Expectancy (Male, Total), 2015 – 2040

Incorporating socioeconomic factors



Conclusion

- Life expectancy – increase in all regions (rural/urban; provinces)
- Mortality improvements – concentrated in the young and middle age groups
- Large regional differences in life expectancy – increased significantly over the last 25 years
- The regional differences in life expectancy is predominately driven by the rural population (the LE difference in the urban population of different regions is relative small) → role of urbanisation / policy implications
- GDP growth rate – highly correlated with the mortality index, incorporating the impact of GDP growth in mortality projection → higher life expectancy, least developed regions is found to catch up with life expectancy in developing regions

UNSW Business School

Never Stand Still | Business School | Risk and Actuarial Studies

Regional Differences in Life Expectancy in Mainland China

William Li, Hazel Bateman and **Kevin Liu**

ASTIN, AFIR/ERM and IACA Colloquia, 26 August 2015

ARC Centre of Excellence in Population Ageing Research

UNSW Business School