ASTIN, AFIR/ERM and IACA Colloquia
Innovation & Invention
23-27 August 2015 | Amora Hotel Jamison Sydney

Actuaries Institute

Regional Differences in Life Expectancy in Mainland China

Prepared by Jichen Li, Hazel Bateman, Kevin Liu

Presented to the Actuaries Institute
ASTIN, AFIR/ERM and IACA Colloquia
23-27 August 2015
Sydney

Institute of Actuaries of Australia
ABN 69 000 423 656
Level 2, 50 Carrington Street, Sydney NSW Australia 2000
t +61 (0) 2 9233 3466 **f** +61 (0) 2 9233 3446
e actuaries@actuaries.asn.au **w** www.actuaries.asn.au

School of Risk and Actuarial Studies
Business School
UNSW AUSTRALIA

Regional Differences in Life Expectancy in Mainland China

Jichen Li

Under the supervision of:
Professor Hazel Bateman and Dr. Kevin Liu

A thesis proposal submitted in partial fulfilment of the requirements for the degree of Bachelor of Commerce (Honours in Actuarial Studies)

DECLARATION

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no materials previously published or written by anther person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

I also declare that this declaration is not my original composition.

Signature:

____

Date:

____

"Night coloured my dark eyes,

While I use them to look for light."

(Chen Gu 3.1980)

ABSTRACT

China's Central government is currently undertaking extensive pension reforms which requires accurate life expectancy forecasts at both national and regional levels. However, the lack of prior research in this area is an obstacle to policy formulation. The aim of this research is to investigate the trend of China's regional differences in life expectancy.

Three contributions of this thesis are as follows. First, we constructed 800 national and regional life tables (by province and for urban and rural populations) based on a unique hand collected mortality dataset. This extends investigations of life expectancy beyond just the national level. The analysis of the life tables reveals that significantly increasing life expectancy differences exist between urban and rural populations between 1990 and 2010. In addition, we find that across all regions the differences in life expectancy at birth and at age 65 were mainly driven by the differences in life expectancy of the rural population in each region. Second, estimations of future mortality improvements using constructed life tables and the extended Lee-Carted model which addresses the limited availability of mortality data shows a stable difference between the urban and rural population over next twenty-five years. Moreover, this thesis shows a narrowed regional difference for urban population and a widen regional difference for rural population in the future. Third, the growth rate of GDP, GDP per capita, income per capita and the change in the number of doctors per 10,000 people all show relatively strong correlations with the change of the index of k_t. Moreover, the sensitivity analysis displays that life expectancy at birth in the least developed regions will catch up with the developing regions by 2040.

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude for my supervisors, Professor Hazel Bateman and Dr. Kevin Liu. I would not have been able to survive over this stressful year without their dedication, kindness and guidance. They are like beacons driving out darkness and lighting up my way. I really appreciate having such two great mentors.

I would like to thank all the staff from the School of Risk and Actuarial Studies and ARC Centre of Excellence in Population Ageing Research (CEPAR). Thanks for sharing with me their invaluable knowledge and experience in research. I am sincerely grateful to God for giving me such a wonderful school and so many lovely people here. I am also honoured to be a recipient of the CEPAR Honours Scholarship, which has provided me both financial support and confidence in completing this thesis.

I am also deeply grateful to my fellow honours and PhD students, Clement Yeung, Estelle Liu, Shahin Elliin, Yolanda Li, Ann Vu, Adam Shao, David Bell, Mengyi Xu, Shang Wu, Vincent Tu, Xinda Yang and Yajing Xu. I really appreciate their generous support and enjoy the time with them, although they are not always to get my Chinese style of jokes.

Finally, I would like to thank all my family and friends for their love, understanding and support. I sincerely appreciate my parents, Haoshi Li and Lan Wu, for their patience and tolerance during this difficult year. Special thanks to my girlfriend, Sarah Xu, whose love, encouragement and tolerance give me the power to go through this unforgettable year.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Over the past few decades, China has experienced rapid population ageing which has been putting severe pressure on the sustainability of its public pension system. The Central government is currently undertaking extensive pension reforms in response to the demographic challenge of population ageing. This requires accurate pension liability estimation, and hence, accurate life expectancy forecasts at both national and regional levels. However, there have been few studies on life expectancy and mortality improvements in China due to a lack of high quality, long term data. Furthermore the vast regional differences in socioeconomic development raise questions about how severe the regional differences in life expectancy are in mainland China and the impact of socioeconomic factors on China's regional differences. The aim of this research is to forecast the future trend of the change in China's regional differences in life expectancy. The regional differences are evaluated in two ways: first, the difference between regions with different levels of economic and social development and second, the difference between rural and urban populations in each region.

National and regional life tables (by province and for urban and rural populations) are constructed based on a unique hand collected mortality dataset. This is also the first dataset at both national and regional level, by province and for urban and rural populations. These 800 life tables and analysis on historical mortality improvements in each administrative region provides a significant contribution as there are very few studies in this area due to the limited availability of relevant data in China.

Then, the Lee-Carter model (extended to address the issue of the limited availability of relevant data) is applied to fit the historical mortality data. Its estimations of mortality improvements in administrative regions (both in aggregate and by urban and rural populations in each region) allow analysis of the trend of the change in China's regional differences in life expectancy. This provides a better understanding of China's regional differences in life expectancy, which is crucial to better estimate the future sustainability of a pension system built around regional pools.

Finally, the study investigates the impacts of socioeconomic factors on the demographic experience in China. The sensitivity analysis through the incorporation of the growth rate of GDP into the standard Lee-Carter model enables a more accurate forecast of future mortality improvements by region. This also allows accurate estimation of the fiscal sustainability of China's pension system and provides a framework for future reform proposals regarding pensions and the implications of an ageing population more generally.

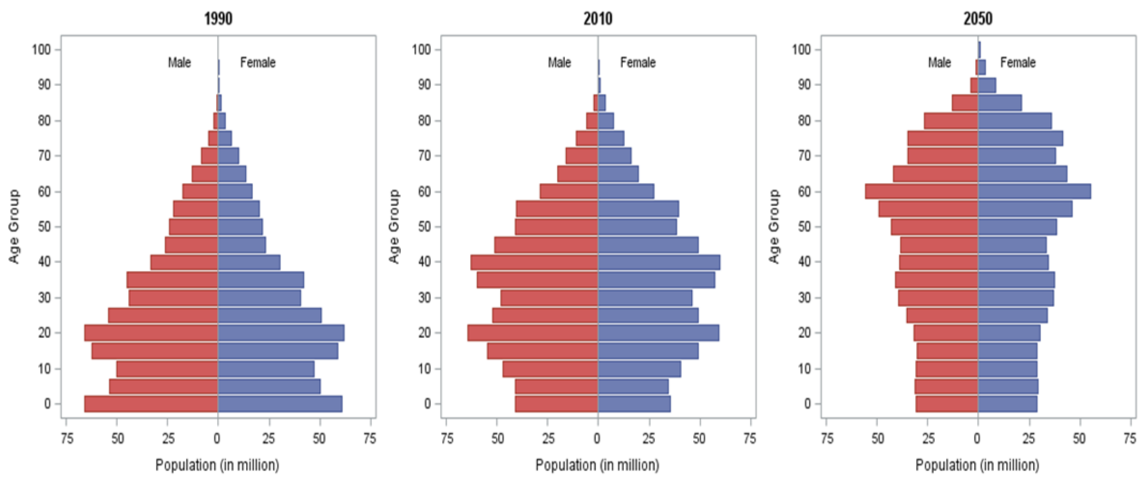
1.1 Background

China is facing an unpredictable challenge of population ageing which is putting heavy pressure on its existing pay-as-you-go (PAYG) pension system. Currently, low fertility rates and an increase in life expectancy as a result of improved medical care and living standards have led China to experience one of the fastest rates of population ageing in the world. There are about 200 million people (around 15% of the total population) over 60 years old and about 7.56 million Chinese move into old age every year. Moreover, the life expectancy in China as projected by the United Nations, for both males and females, is expected to increase from the current 73.8 years to 79.1 years by 2050 (United Nations, 2012), with substantial differences by region. Also, the United Nations forecast that China's aged population (those aged 65 and over) will reach 436 million by 2050.

Along with the increasing life expectancy, China's fertility rate has gradually decreased from around 6 births per woman in the 1960s to 1.6 births per woman as of 2013 (Pozen, 2013). This is largely due to the one-child policy implemented by the Chinese government since 1979. It has caused the well-known “4-2-1-1 problem” in which one working person has to provide for four grandparents, two parents, and one child, while at the same time needing to save for his or her own retirement (Salditt et al., 2008).

As a result, the increased longevity and decline in the fertility rate have started to erode the ‘demographic dividend’; that is, the positive demographics-related impact on growth of the Chinese economy. Furthermore, one implication of an ageing population will see a rapid reversal in the population pyramid (see Figure 1.1), and a dramatic decrease in the number of workers available to support the elderly from 6.1 in 2012 to 1.6 in 2050 (Bateman and Liu, 2013). These challenges are of a great concern to the affordability and sustainability of the age pension system in China which is designed as a PAYG system, whereby current contributions from employers and workers pay current pensions to the elderly. Only part of the contributions are invested to fund future pensions and most are generally invested in very low return assets.

Figure 1.1: Population Pyramid, China (1990, 2010, 2050)



Source: United Nations (2013) ’World Population Prospects: The 2012 Revision’.

On the other hand, benefiting from the ‘demographic dividend’, China has enjoyed phenomenal economic growth over the past twenty five years. Although China is now the second largest economy (by GDP) in the world, its GDP per capita ranks only 92nd worldwide (United Nations, 2013). In addition, China's disposable personal income in 2013 was only 12 per cent of that in the United States (Intl $4,330 compared $36,771 for the US) (Bateman and Liu, 2013). This suggests that there may be insufficient personal savings to finance individuals' retirement. As a result, the publicly provided pay-as-you-go (PAYG) pension system plays a very significant role in retirement income provision. However, the sustainability of the PAYG pension system as currently designed is heavily dependent on the difference in the growth rate of the working population and the old population (people aged 65 and over). Furthermore the PAYG pension system is locally managed – which means that local sustainability is related to local and regional rather than national mortality trends. As the benefits provided by the PAYG pension system are predetermined locally by a formula based on individuals' earnings history, tenure of service, age and a benefit multiple designed to reflect the life expectancy of the residents in the local area (Dickson et al., 2013). Hence, an accurate estimate of mortality improvement at the regional level is crucial to be able to monitor the sustainability of the public pension system.

In addition to the dramatic ageing process and inadequate personal savings, China's current pension system has a number of shortcomings itself. The pension system is immature and fragmented and there are inherent administrative hurdles in transferring individual accounts and accrued benefits across local borders. Although pension policies are determined according to the same rules set by the Central government, the pension pools in each of 2000 or so regions are still managed locally. This has resulted in a lack of transparency and uneven coverage among different regions (Alonso et al., 2011).

A summary of China's current pension system is presented in Table 1.1. This shows that there are huge differences in social assistance and income replacement on a scheme by scheme basis. For instance, the civil servants and public sector workers have a very generous PAYG pension system which does not require them to make contributions. Moreover, they enjoy around 100 per cent coverage and the most generous replacement rate (of all schemes) of around 90 per cent of pre-retirement earnings. This is very generous compared to employees in the other sectors, who not only need to contribute a certain percentage of their income but also obtain less generous pension benefits. For example, urban workers must make contributions and their expected replacement rate is only sixty per cent of their earnings. The situation is even worse for rural workers and urban residents not covered by the urban scheme as their basic pension income is only 55 RMB per month. Some other issues with the existing pension system, such as a lack of portability and narrow investment channels for pension assets, were pointed out in Salditt et al. (2008) as well. Later, Bateman and Liu (2013) utilized the data from China Household Finance survey to critically assess the practical implications of China's existing retirement income arrangements. Their study raised concerns about the inadequacy of pension income benefits, the vulnerability to political risks, the negative impact on inter-generational equity and economic efficiency and the limited contribution to capital accumulation and long term economic growth. According to recent research, the financing gap from 2002 to 2075 will be around 95 percent of China's GDP for 2001 (Dorfman et al., 2013). As all previous studies were conducted on a national level, an accurate estimate of the pension financing gap needs to be done on a regional basis (by province and for urban and rural populations) to account for differences in the mortality experience. There does need to be an understanding of 'regional differences' as well as good knowledge of expected improvements in mortality by region – to enable accurate forecasts of the sustainability of current age pension arrangements as well as possible pension reforms both nationally and by region.

Table 1.1: Current structure of the retirement income arrangements in China (2013)

	Urban	Rural	Urban Resident	Civil servants and public sector
Legal coverage	All urban workers (except those covered by the urban resident scheme)	Rural residents age 16+ (except students). Voluntary introduction by county	Urban residents age 16+ (except students). Voluntary introduction by county/city	All civil servants and public sector workers
Actual coverage	68% workers 90% retirees	22% workers 58% retirees	na	97% workers 99% retirees
Financing	Public PAYG (social pooling) plus publically managed individual account (IA)	Individual contribution plus subsidies from government/rural collectives	Individual contribution plus subsidies from government	General Revenue (PAYG)
Contributions	Public PAYG: 20% employer contribution. Publically managed IA: 8% employee contribution	Contributions over 5 levels (¥100-500 per year) Government subsidy (match) max ¥30 per year	Contributions over 10 levels (¥100-1,000 per year) Government subsidy (match) ¥30 per year	Nil
Pension age	Age 60 (men), 55 (female senior managers), 50 (female workers)[a]	Age 60 (men and women)	Age 60 (men and women)	Age 60 (men), 55 (female cadres), 50 (female workers)[b]
Benefits	Public PAYG: 1% of average local wage per year of service, indexed annually to wages/prices IA: accumulation/139 per month Expected replacement rate of 35% + 24.2% = 59.2% after 35 years.	Basic pension[c]: ¥55/month[d] Pension from voluntary contributions: accumulation/ 139 per month[e]	Basic pension[f]: ¥55/month Pension from voluntary contributions: accumulation/139 per month[g]	Around 90% earnings after 35 years.

Source: Alonso et al. (2011) , Bateman and Liu (2013), Dorfman et al. (2013)

[a] Age 55 (men), 45(women) if engaged in arduous work after 10 years continuous service.
[b] Civil Servant Act (2006), s14(87)
[c] Central government pays full basic pension in central and western areas, provides 50% subsidy to eastern areas.
[d] Note child contribution.
[e] Indexation according to economic development and changing prices.
[f] See footnote 2.
[g] See footnote 4.

1.2 Motivation

To address the risks associated with the combination of population ageing and economic development (that is of China getting old before it gets wealthy) (Jackson et al., 2009), a number of recommendations have been proposed to improve China's pension system. These include raising the retirement age from 60 (men) and 55 (women) to 65 and 60 respectively, centralizing the fragmented pension system from the current 2000 or so local pools (Pozen, 2013) and moving from a pure PAYG system to a notional defined contribution scheme (Oksanen, 2010).

It is likely that these strategies would be more effective if they were consistent with the average local life expectancy. However, the lack of prior research in this area is an obstacle to policy formulation. Although the standard model used to estimate mortality was developed by Lee and Carter in 1992, almost all extensions and applications of the Lee-Carter model have focused on developed countries and selected middle income countries with high quality, long term data, such as Argentina and Chile (e.g. Booth et al., 2006; Russolillo et al., 2011). There have been a very limited number of studies on China due to the lack of high quality, long term data. The standard Lee-Carter model demands age-specific death rates for at least the past twenty consecutive years. However China's mortality data at the national level are only available at the fixed points in time at which the Chinese government conducted the nationwide Censuses. These include 1974, 1981, 1987, 1990, 1995, 2000, 2005 and 2010.

On the other hand, there are vast regional differences in socioeconomic development in China (Demurger, 2001). This may cause significant regional differences in life expectancy according to the demonstrated strong correlations between socioeconomic vaiables and mortality behaviour (Singh and Siahpush, 2006). Nevertheless, there has been no study on China before which raises questions about how severe the regional differences in life expectancy are in mainland China and the impact of socioeconomic factors on China's regional differences. In addition, the PAYG pension system is still managed on a local basis in around 2000 pension pools. As a result, the pension benefits are determined according to the local wage and inflation indexes. Therefore, regional differences in life expectancy are vital issues to ensure the consistency and fairness of the entire age pension system, and to estimate future costs. Estimating and taking account of the regional differences in life expectancy are very important to enable policy makers to build a fair and sustainable pension system.

1.3 Research Objective

This research extends and applies the Lee-Carter model in order to estimate mortality at both the national and regional levels (that is, by province and for urban and rural populations). Li et al. (2004) developed an extension to the Lee-Carter model to forecast mortality for populations with limited data. However, they did not assess the cause of the regional difference of life expectancy. Due to the limited availability of regional mortality data, this research firstly applies the methodology in Li et al. (2004) to fit the mortality data from the National Bureau of Statistics of China. It then incorporates socioeconomic factors, such as growth rate of GDP, GDP per capita and income per capita and proxies for healthcare such as the change of number of doctor per 10,000 people, into the modelling of the mortality index k_t following Hanewald (2011). This approach is used to overcome a shortcoming of the Lee-Carter model which does not take account of the socioeconomic factors influencing mortality rates, such as the level of (and access to) medical care, real GDP and the unemployment rate. Ultimately, the research extends the application of the Lee-Carter model and provide forecasts of the future trend of the change in China's regional differences in life expectancy. This then allows more accurate estimates of the impact of population ageing on the sustainability of the current pension system (and other age-related policies and programs) and motivate possible reforms.

1.4 Expected Contribution

The contribution of this research is two fold. First it extends the existing Lee-Carter models to instances with limited data and second it incorporates socioeconomic variables in modeling China's regional difference in life expectancy both by province and for urban and rural populations. Incidentally, a set of regional life tables (by province and for urban and rural populations) are developed for the first time which is an additional contribution. This allows a better understanding of China's regional differences in life expectancy. In addition, the incorporation of socioeconomic indicators into the Lee-Carter model could provide a more accurate forecast which can help to better understand the liability of the pension system. This will enable more accurate estimates of its fiscal sustainability and provide a framework for future reform proposals regarding pensions and the implications of an ageing population more generally. As well, the extension to the Lee-Carter model will establish a solid base for future studies of the longevity risk in those developing countries without high quality, long term data.

1.5 Outline of Thesis

The remainder of the thesis is structured as follows. **Chapter 2** provides a literature review outlining the development and applications of the Lee-Carter Model. Two major shortcomings and modifications of the standard Lee-Carter model are then presented. Current literature gaps are also identified in the section. **Chapter 3** explains sources of data and the methodology applied to investigate China's regional differences in life expectancy. **Chapter 4** displays illustrative samples of constructed life tables and discusses a number of findings from China's historical demographic experience. In **Chapter 5**, predictions of future regional differences in life expectancy are investigated by using the models developed in **Chapter 3**. **Chapter 6** concludes by summarizing the main results and highlighting the practical applications of the research.

CHAPTER 2

LITERATURE REVIEW

In this chapter, the development of the Lee-Carter model for modelling mortality behaviour is first reviewed. Then the next two parts present the key extensions and applications of the Lee-Carter model which are relevant to the research questions addressed here. Later, two major shortcomings of Lee-Carter model are discussed. The first shortcoming of the Lee-Carter model is its incapacity to deal with populations with limited data. The second shortcoming is its failure to incorporate external factors which may have vital impacts on mortality. Possible solutions of those two shortcomings are then discussed. Finally, we identify the gaps in the literature that this thesis is going to fill.

2.1 The Lee-Carter Model

To accurately forecast future trends of life expectancy, the mortality rate has to be modelled. In the past, the mortality rate was traditionally modelled in a deterministic way. However such estimates might be significantly influenced by experts' personal subjective judgments (Lee and Miller, 2001). Alho (1990) pointed out that this use of experts' personal judgement would hinder rather than help the forecasts of human's mortality behaviour. In 1992, Lee and Carter developed a new demographic model of mortality, known as the Lee-Carter model (Lee and Carter, 1992). This is recognized as the first and most popular stochastic model of mortality (Cairns et al., 2008). The Lee-Carter method, as a new extrapolative approach, used standard time-series procedures to extrapolate the historical trends and future forecasts of age-specific death rates and has succeeded to reduce the influence of subjective judgement on the modelling (Li and Lee, 2005).

The Lee-Carter model was proposed as the following:

$$ln(m_{x,t}) = a_x + b_x \times k_t + \varepsilon_{x,t} \tag{2.1}$$

where $m_{x,t}$ is the central death rate for age x in year t; a_x and b_x are age-specific constants; k_t, as a time-varying index, is the index of the intensity of mortality; and $\varepsilon_{x,t}$ is the error term with mean zero and variance σ_ε^2.

This equation shows that the central death rate depends not only on age but also on time. An advantage of this model is the impossibility for death rates to be negative in the forecasts. The unique least squares solution for this model could be obtained under two specific constraints:

$$\sum_{x=1}^{w} b_x = 1 \qquad and \qquad \sum_{t=1}^{n} k_t = 0 \tag{2.2}$$

where w is the maximum age and n is the maximum year. Under the constraints, a_x could be found as the average over time of $ln(m_{x,t})$. Then through the application of the singular value decomposition method, b_x and k_t could be obtained as well. Alternatively, the estimation of k_t could be gained from the relationship:

$$D_t = \sum_{x=1}^{w} exp((a_x + b_x k_t) N_{x,t}) \tag{2.3}$$

where D_t is total deaths in year t, and $N_{x,t}$ is the population of age x in year t. Next, the model was used to forecast the mortality in the future through developing a proper stochastic time series model for the mortality index, k_t (Lee and Carter, 1992). The most commonly used model is a random walk with drift,

$$k_t = k_{t-1} + c + \varepsilon_t \tag{2.4}$$

where c is the drift term and ε_t follows the standard normal distribution. This allows the mortality improvement of each age group to be captured by its own age-specific rate, c. Ultimately the Lee-Carter model led to more certain forecasts of mortality and life expectancy.

As one of the most famous models, the Lee-Carter model had demonstrated its effectiveness in forecasting demographic variables, such as the mortality rate and fertility rate, during the previous two decades (Renshaw and Haberman, 2006). Nevertheless, there are still some limitations to the standard Lee-Carter model. For instance, the future gains in life expectancy are under predicted using the Lee-Carter model (Lee and Miller, 2001). This is probably because age patterns of mortality decline will be structurally changed due to advances in medicine and technology in the future. Also, the Lee-Carter model does not work well to forecast mortality for a group of populations (Li and Lee, 2005). Moreover, the homoscedasticity in the error term will be violated when applying the ordinary least squares (OLS) method to find parameters of the model. The two major shortcomings of the original Lee-Carter model, that is incapacity with limited data and failure to incorporate external factors which have severe impacts on mortality, will be discussed in section 2.4.

2.2 Variations and Improvements to the Lee-Carter Model

The Lee-Carter model opened the door for stochastic mortality modelling. Since then, a large number of extensions to the model have been developed. In 1993, to overcome the issues associated with zero death rates [1] and the shortcoming of original OLS method [2], Wilmoth (1993) conducted a new weighted least squares (WLS) method by constructing a set of new weighted parameters as follows:

$$\widetilde{a_x} = \frac{\sum_t d_{x,t}[ln(m_{x,t}) - \widetilde{b_x}\widetilde{k_t}]}{\sum_t d_{x,t}} \tag{2.5}$$

$$\widetilde{b_x} = \frac{\sum_t d_{x,t}\widetilde{k_t}[ln(m_{x,t}) - \widetilde{a_x}]}{\sum_t d_{x,t}\widetilde{k_t}^2} \tag{2.6}$$

[1]Since the standard Lee-Carter model needs to take logarithms of central death rates, taking logarithms of zero death rates definitely cause an issue.

[2]The homoscedasticity in the error term of the standard Lee-Carter model will be violated by using the OLS method.

$$\widetilde{k_t} = \frac{\sum_t d_{x,t}\widetilde{b_x}[ln(m_{x,t}) - \widetilde{a_x}]}{\sum_t d_{x,t}\widetilde{b_x}^2} \tag{2.7}$$

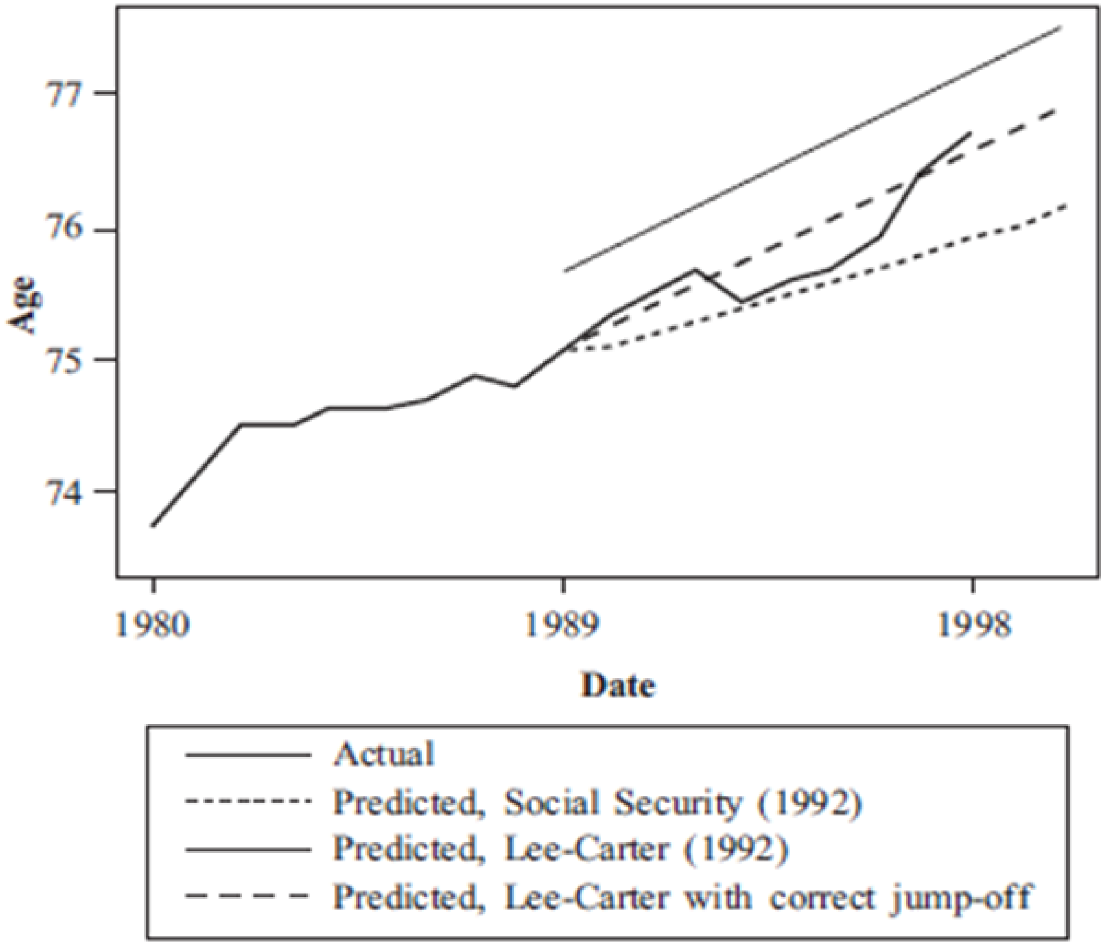
where $m_{x,t}$, a_x, b_x and k_t are interpreted the same as in the original Lee-Carter model, and $d_{x,t}$ represents the number of deaths for age x at time t.

Wilmoth (1993) then compared three different ways to estimate the parameters of the Lee-Carter model, including OLS, WLS and maximum likelihood estimation (MLE). In the paper, he found that WLS performed better. Later, Brouhns et al. (2002*b*) extended the original Lee-Carter method through the implementation of Wilmoth's recommendation. Moreover, they applied their new model to build projected life tables and critically assess the longevity risk (Brouhns et al., 2002*a*).

In 2000, Lee summarized various extensions and improvements to the Lee-Carter model. Instead of using random walk with drift to model k_t (Lee and Carter, 1992), Lee (2000) recommends that it is sometimes superior to add a moving average term or autoregressive term to the random walk model to help better describe the stochastic movement of k_t. In addition, Lee (2000) suggested that taking the most recently observed data as initial values can make forecasts of age-specific mortality rates smoother.

One year later, Lee and Miller (2001) found that the age pattern of mortality declining may change over time. Hence, they suggested that b_x should not be fixed and might vary over time. Furthermore, the jump-off error may be caused by the mismatch between fitted rates and actual rates for the last year of the fitting period (Lee and Miller, 2001). Hence, they developed the Lee-Miller variant which let k_t pass through zero in the jump-off year to avoid the jump-off bias. This improvement provides a better prediction of mortality behaviour than the standard Lee-Carter model (see Figure 2.1). Using American data, the figure shows that predictions of the extended Lee-Carter model has a much better goodness-of-fit to actual life expectancy than predictions made by the Social Security Administration and the standard Lee-Carter model. After then, more modifications of the Lee-Carter approach were proposed. Booth et al. (2002) developed the Booth-Maindonal-Smith variant to model the death process though the application of the Poisson distribution. The Booth-Maindonal-Smith variant is distinct from the Lee-Carter method in two ways. First, it fits the age-specific number of deaths, $D_{x,t}$, instead of the total number of deaths, D_t. Second, it applies the goodness-of-fit criteria to select the proper fitting period. They are all aimed at maximizing the goodness-of-fit of the overall model (Booth et al., 2002).

Figure 2.1: Life Expectancy Forecasts from 1989: Social Security Administration for the U.S., the original Lee-Carter, the Lee-Carter with correct jump-off (Lee-Miller variant)



Source: Lee and Miller (2001).

Later, Renshaw and Haberman (2003) realized that the single value decomposition method in the Lee-Carter model is potentially inflexible with respect to age after they applied the Lee-Carter methodology to forecast the male mortality rate in England and Wales, from 1950 to 1998. As a result, they proposed a study using the generalized linear modelling framework which was proven to have sufficient flexibility with respect to different age–specific features. Later, Renshaw and Haberman (2006) incorporated the cohort effect into the Lee-Carter model. They found that the age-period-cohort model is a great improvement over the fitted Lee-Carter model as it could remove the distinctive ripple effects better.

To extend the use of only one set of (b_x, k_t), Hyndman et al. (2007) proposed the Hyndman-Ullah functional data method as follows:

$$ln(m_{x,t}) = a_x + \sum_{n=1}^{N} b_{n,x}k_{n,t} + e_{x,t} + \sigma_{x,t}\varepsilon_{x,t} \tag{2.8}$$

where a_x is a measure of location of the smoothed mortality; the basis function, $b_{n,x}$ and the coefficient of time series, $k_{n,t}$; $e_{x,t}$ accounts for modelling error and $\sigma_{x,t}\varepsilon_{x,t}$ is observational error. This use of non-parametric smoothing methods to smooth death rates is able to reduce the observational noise. Also, the issue with outlying years due to wars or epidemics could be avoided under this application of robust methods. Furthermore, after fitting the extended model to French mortality data (1899-2001), it is found to be superior to the original Lee-Carter method.

In addition, a more generalized Lee-Carter model was developed through the incorporation of the state space framework (De Jong and Tickle, 2006). The proposed model is as follows:

$$ln(m_{x,t}) = Xa_x + Xb_xk_t + \varepsilon_{x,t} \tag{2.9}$$

where De Jong and Tickle (2006) used a known 'design' matrix (X) to multiply age-specific parameters a_x and b_x. This model is demonstrated to bring more robust results when fitted to the Australian women's mortality data for the period of 1921-2000.

Recently, in order to take account of multiple countries' mortality simultaneously, the Tucker3 model was applied as a variation of the Lee-Carter model. Russolillo et al. (2011) improved the standard Lee-Carter model by adding a new component γ_i which represented the country-specific factor. The extended model is shown as follows:

$$ln(m_{i,x,t}) = a_{i,x} + \gamma_i b_x k_t + \varepsilon_{x,t} \tag{2.10}$$

where $m_{i,x,t}$, b_x, k_t and $\varepsilon_{x,t}$ are interpreted the same as in the original Lee-Carter model. As a new model incorporating time, age and country, it is very useful to investigate the aggregated mortality performance of different countries. In addition, it is able to assist in the construction of concise life tables for the whole group of countries (Russolillo et al., 2011).

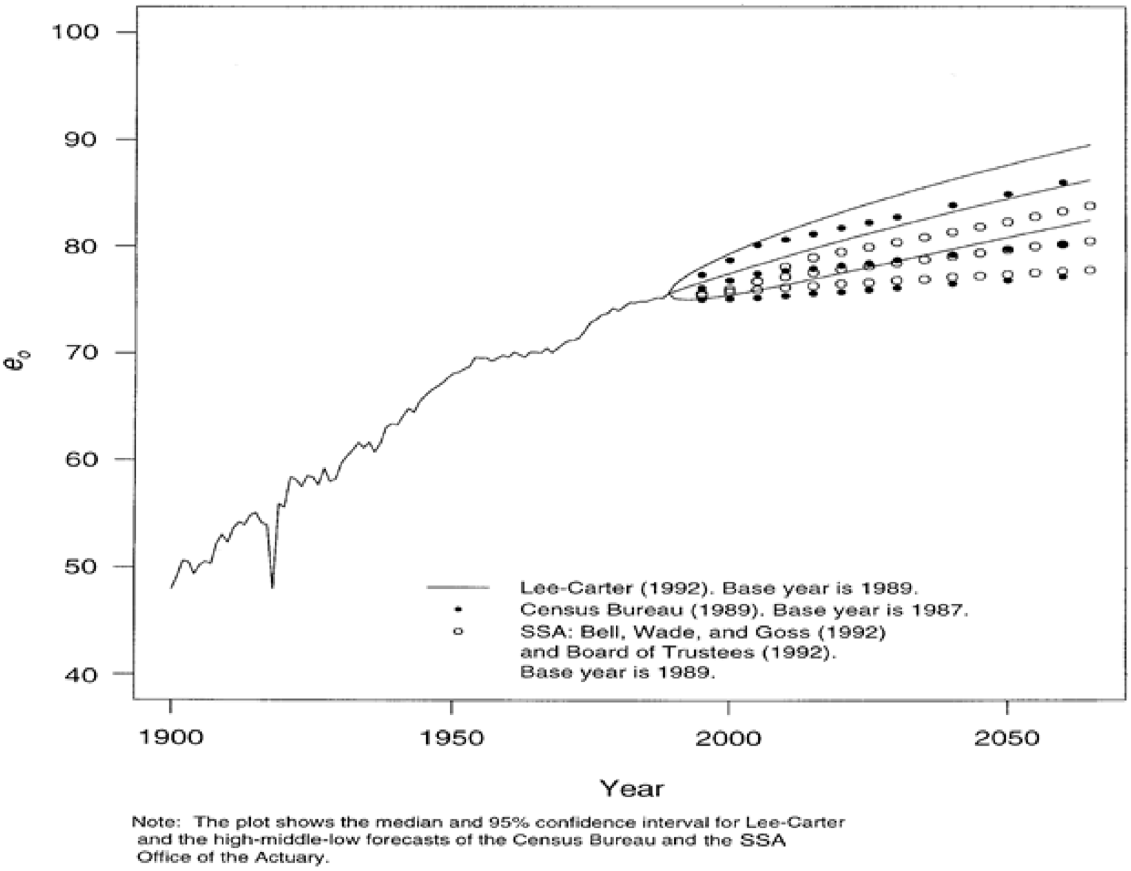
Another significant weakness of the original Lee-Carter model is its assumption of the linear improvement of the mortality index, k_t. To address the situation where the index is not linear, Li et al. (2011) proposed the importance of finding out the point of structural change of the mortality index. Advanced variations to deal with a nonlinear mortality index are found in Hainaut (2012) which suggested a way to incorporate a regime switching process to construct a multidimensional Lee-Carter model.

Many other variations and improvements to Lee-Carter model can be found in (Booth, 2006; Booth and Tickle, 2008; Li et al., 2013). The extensions in these papers were all developed to improve the goodness-of-fit and predictability of the original Lee-Carter mode, which provided a solid base for later studies in this area.

2.3 Applications of the Lee-Carter Model

Initially, the Lee-Carter model was applied to investigate the demographic behaviour of the U.S. population. The original Lee-Carter model was fitted to American mortality data for the period 1933 to 1987. It then forecast future mortality and life expectancy until 2065. The results of the life expectancy forecast brought a substantially higher value and greatly narrowed confidence bands than the forecast of the (US) Social Security Administration's Office of the Actuary. Later, with the development of stochastic mortality modelling, more variations of Lee-Carter model began to be applied to different countries. After Lee (2000) extended the Lee-Carter method to the U.S., Chile and Canada, he found similar results as in the original Lee and Carter (1992) paper. Figure 2.2 shows that the Lee-Carter model narrowed confidence bans and addressed the issue of underestimated life expectancy by Census Bureau and Social Security Administration. Lee (2000) concluded that the Lee-Carter model was a good baseline for the mortality and life expectancy forecasts.

Figure 2.2: Life Expectancy Forecasts with Intervals: Lee-Carter, Census Bureau, and Social Security Administration for the U.S., Sexes Combined, 1990–2065



Source: Lee (2000).

Later, more countries' demographic trends were investigated using the Lee-Carter approach. Tuljapurkar et al. (2000) applied the ordinary Lee-Carter model to the G7 countries and found that the Social Security projection systematically under predicted the life expectancy gain over time. Next, Lee and Miller (2001) applied the Lee-Miller variant to Japan, France, Canada and Sweden. Their results not only supported the argument in Lee and Carter (1992), Lee (2000) and Tuljapurkar et al. (2000), but also pointed out that the pattern of the mortality improvements had systematically changed in the second half of the twentieth century. This aroused interest in studying structural change of mortality improvements by using the Lee-Carter model.

Through numerous practical applications of the Lee-Carter method, it became clear that there existed severe geographic differences in mortality improvements (Lee and Nault, 1993). To solve this issue, Li and Lee (2005) fitted the standard Lee-Carter model to annual age-specific death rates (ASDR) for a group of fifteen developed and selected developing countries. The research found that the ASDRs of the countries with close connections and similar socioeconomic conditions tended to 'move up and down together over time' (Li and Lee, 2005). Furthermore, Li and Lee (2005) realized that adding a population-specific component to the model could substantially improve the fit for the countries. Hence, Li and Lee (2005) decided to add a pair of common factors, B_xK_t, to the original Lee-Carter model. This is shown as follows:

$$ln(m_{x,t,i}) = Xa_{x,i} + B_xK_t + b_{x,i}k_{t,i} + \varepsilon_{x,t,i} \tag{2.11}$$

where B_x and K_t can be obtained by fitting the ordinary Lee-Carter model to the weighted average ASDR of those fifteen countries. This common factor Lee-Carter model has the advantage of being able to statistically assess whether one population is following the demographic behavior of a typical group of countries. Nevertheless, this approach will fail if such a common coherent trend over time is not real for the populations (Li and Lee, 2005).

To further investigate the behaviour of mortality under similar socioeconomic circumstances, Härdle and Myšičková (2009) assessed the change of mortality in East Germany after it was merged with West Germany. The study demonstrates that the mortality for different populations will converge when their socioeconomic conditions also gradually converge. Moreover, this process will speed up if the two populations have very close connections and frequent migration. The later point was substantiated in Alho (2008) who did a similar investigation for France, Austria, Switzerland, Germany, Ireland, Portugal and other European countries.

In addition, Booth et al. (2006) compared the ordinary Lee-Carter method (Lee and Carter, 1992), the Lee-Miller variant (Lee and Miller, 2001), the Booth-Maindonald-Smith variant (Booth et al., 2002), the Hyndman-Ullah functional data method (Hyndman et al., 2007) and the de Jong-Tickle Lee-Carter model (De Jong and Tickle, 2006) by fitting them to the ASDRs of Australia, Canada, Denmark, England and Wales, Finland, France, Italy, Norway, Sweden, and Switzerland. Their assessment, based on the mean absolute error, two-way ANOVA model, displays that the forecasts of log death rates using the Hyndman-

Ullah and de Jong-Tckle variations are more accurate. On the other hand, the Booth-Maindonald-Smith variant performs better in forecasting life expectancy. However, further investigations are still required as the accuracy of various methods are still significantly influenced by the particular fitting period (Keyfitz, 1944; Murphy, 1995).

2.4 Shortcomings of the Lee-Carter Model

2.4.1 Limited Data

The Lee-Carter model has two major shortcomings. The first shortcoming of the Lee-Carter model is its incapacity to deal with populations with limited data. The Lee-Carter model needs at least consecutively around the past thirty years' mortality data to provide a good fit. However most developing countries only started to collect their mortality data around twenty years ago. As a result, although plenty of prior studies have been implemented for developed countries, there are few studies for developing countries, such as China, due to their lack of high quality long term data. The most complete mortality data for China has been collected in their nationwide Censuses (Banister and Hill, 2004). However, there have only been six Censuses over the past half century. This has meant that it has been difficult to study the mortality behaviour in China. Banister and Hill (2004) tried to re-estimate mortality levels from China's incomplete life table for the period 1964-2000. They applied the General Growth Balance method using China's Census data for 1964, 1982, 1990 and 2000 and a Cancer Epidemiology Survey. Similar attempts had been made by Jiang et al. (2013). They undertook a forecast of age specific death rate until 2030 based on the available Chinese Census data for 1982, 1990 and 2000. Their prediction of life expectancy at birth by 2020 was about 3.87 years lower than the projection by the Ministry of Health of China. Application of their prediction into pension benefit design resulted in an underestimation of future pension liabilities which threatened the sustainability of the pension system. Due to a lack of long term and accurate data for China, there had been substantial discrepancies in the trend of future demographic change amongst the different studies. Furthermore, due to limited data, most mortality forecasts for China used deterministic projections and have relied on experts' subjective judgement (Li et al., 2009). Some studies also applied scenario analysis to construct a band for mortality forecasts Jiang et al. (2013). However, the application of scenarios always resulted in under predicted results (Keilman, 2001). Hence, the limited data has been an obstacle for scholars wishing to study the trend of China's demographic change.

To solve the issue of limited data, some studies have suggested borrowing information from countries with similar socioeconomic conditions (Lee, 1998). Alternatively, other studies have recommended restructuring the Lee-Carter model to account for limited data (Li et al., 2004; Zhao et al., 2013). Li et al. (2004) proposed a new approach to model the index of the level of mortality, k_t, as follows:

$$k(u(t)) - k(u(t-1)) = c[u(t) - u(t-1)] + \sigma[e(u(t-1)+1) + ... + e(u(t))] \tag{2.12}$$

where $u(t)$ represents the year of the t-th data point. $k(u(t))$ represents the index of the intensity of mortality at $u(t)$ and $e(u(t))$ is the error term. c and σ could be obtained as:

$$\widetilde{c} = \frac{k(u(T)) - k(u(0))}{u(T) - u(0)} \tag{2.13}$$

$$\widetilde{\sigma} = \frac{\sum_{t=1}^{T} k(u(t)) - k(u(t-1)) - c[u(t) - u(t-1)]^2}{u(T) - u(0) - \frac{\sum_{t=1}^{T}[u(t) - u(t-1)]^2}{u(T) - u(0)}} \tag{2.14}$$

where $u(T)$ represents the year of the last data point and $u(0)$ represents the year of the first data point.

Using this approach, Li et al. (2004) fitted China's mortality data for 1974, 1981 and 1990. Based on those only three data points, they undertook a full forecast of life expectancy for China to 2040 without borrowing any information from other countries. This brought a better probabilistic interpretation of the future trend of China's demographic situation than the previously used scenario analysis which had been based on experts' judgment of low and high ranges (Li et al., 2004). In addition, this approach is typically useful for data with widely unequal intervals of time collection. After few years, Zhao et al. (2013) proposed another modified Lee-Carter model in which it was assumed that the occurrences of death followed a binomial distribution with probability $p_{x,t}$ age x in year t. The extended model is as follows:

$$ln(\frac{p_{x,t}}{1 - p_{x,t}}) = a_x + b_x k_t \tag{2.15}$$

where a_x, b_x and k_t were interpreted the same as in the original Lee-Carter model and were estimated using the MLE method. Zhao et al. (2013) applied this logistic regression model to fit China's mortality data from 2001 to 2009. In their study, they found that the new method would bring smoother predictions of the death rate over the years. However, the confidence range for the estimation is still quite wide, which is an incentive to undertake further research on how to deal with the limited data issue.

2.4.2 Socioeconomic Variables

The second major shortcoming of the original Lee-Carter model is its failure to incorporate external factors which may have vital impacts on mortality. It has been evident that the change of mortality has been substantially affected by the level of economic growth, the development of a health system and the improvements in the standard of living and education (Banister and Hill, 2004). Investigations have also shown that economic crises has led to an increase in infant mortality because of the reduction in pregnant women's expenditure on health care under economic downturn. For example, Alexander et al. (2011) found that the infant mortality increased by about 2 per cent on average during the U.S. economic crisis of 2007-08. On the other hand, an increase in public expenditure on health could reduce such a detrimental impact of a recession (Alexander et al., 2011).

Interestingly, good economic conditions might raise the mortality rate due to typical causes as well. Ruhm (2007) found that coronary heart disease (CHD) mortality increased during the early period of economic recovery. Moreover, the study revealed that each percentage reduction in the unemployment rate would correspondingly cause CHD mortality to rise by 0.75% in the United States (Ruhm, 2007). Granados (2005) formally tested the correlation between mortality and socioeconomic factors. In the study, he selected the four observable indicators as GDP, the unemployment rate, manufacturing hours and industrial production. A positive relationship is found between economic activity and mortality. When the economy expanded, the work pace, work time, and level of alcohol and tobacco consumption will increase which will eventually contribute to the rise of mortality levels (Granados, 2005). Later, a further investigation on Spain, Japan and Sweden provided similar conclusions that economic growth definitely brought a negative impact on mortality in the short run (Granados, 2005, 2008; Tapia Granados and Ionides, 2008). However, the study on Sweden also demonstrated a positive correlation between health progress and economic development (Tapia Granados, 2008).

In order to further investigate the factors influencing the difference in the mortality rate and life expectancy, Hanewald (2011) analyzed the correlation between the mortality index k_t and macroeconomic factors using data for six OECD countries, including Australia, Canada, Japan, Netherlands, United Kingdom and United States. In the study, she investigated the correlation between macroeconomic indicators and the index of intensity of mortality, k_t, in the Lee-Carter model. Through the application of the Phillips-Person test and cointegration analysis, she found that k_t has significant correlation with real GDP growth rates in Canada, Australia and the United States, and with unemployment rate changes in Japan. In addition, Hanewald (2011) incorporated the real GDP growth rate into the linear regression modelling k_t as:

$$\Delta(k_t) = \theta + \beta\Delta(ln(x_t)) + \varepsilon_{x,t} \tag{2.16}$$

where X_t represents real GDP at year t. θ and β are the coefficients to be estimated. $\varepsilon_{x,t}$ is the error term. Then, Hanewald (2011) extended the model to

$$\Delta(k_t) = \theta + \beta\Delta(ln(x_t)) + \alpha_1\Delta(k_{t-1}) + \alpha_2\Delta(k_{t-2}) + \varepsilon_{x,t} \tag{2.17}$$

The practical implication of macroeconomic fluctuations on the mortality index, k_t, was further investigated in Hanewald et al. (2011). Here the authors critically assessed the impact on the financial stability of life insurance companies after incorporation of the correlation between macroeconomic indicators and the mortality index. Through simulation, the insolvency probabilities of the insurance companies are found to increase when the dependencies between the mortality index and economic variables were taken into account. Furthermore, Hanewald et al. (2011) is aware that the link between the mortality and the macro economy has gradually changed, which might have an impact on the forecasting of future trends of demographic change. Therefore, it is important to incorporate macroeconomic factors into the stochastic modelling of mortality (Niu and Melenberg, 2012). In China, similar findings were presented in Banister and Hill (2004) as well. Moreover, Banister and Hill (2004) found that the decline of mortality for men had slowed due to occupational and lifestyle factors in China, such as smoking, drinking alcohol and work-related diseases. Additionally further studies could be proposed to assess other specific causal variables apart from macroeconomic factors, such as environmental factors, medical advances and biotechnology advances (Hanewald et al., 2011).

Hanewald also finds that there are considerable regional differences in demographic variables, like mortality and life expectancy, due to different socioeconomic conditions. It is evident that differences in real GDP and unemployment significantly influence the difference in life expectancies for various countries (Hanewald, 2011). Later, Cairns et al. (2011) implemented the Bayesian Markov chain Monte Carlo approach to model the mortality rate and life expectancy for two populations with different socio-economic backgrounds. Their paper developed a mean-reverting stochastic spread model which could explain the difference between the two populations' death rates quite well.

Similar research has been conducted for China as well. Zhao et al. (2013) applied the modified Lee-Carter model to the sex-age-specific mortality data of people from four different groups (country, city, town and county). In the study, they found that the expected mortality rate decreased for males in cities over the period 2000 to 2008, but the rate did not change for males aged 13 to 36 in towns and even increased for males aged 13 to 43 in counties. Zhao et al. (2013) also realized that the life expectancies increased the most in cities. The increase in life expectancies was about 4.5 years in cities, compared to only 3 years in towns. Also the development of the demographic structure across provinces was found to be highly correlated with the aggregate savings rate in China (Zhang et al., 2013).

2.5 Literature Gaps

Since the initial proposal of Lee-Carter model in 1992, various extensions have been developed to improve the robustness and adaptability of the model under different complex environments, such as the Lee-Miller variant (Lee and Miller, 2001), the Booth-Maindonal-Smith variant (Booth et al., 2002), the Hyndman-Ullah functional data method (Hyndman et al., 2007), the de Jong-Tickle Lee-Carter model (De Jong and Tickle, 2006) and the incorporation of the Tucker3 model into the standard Lee-Carter model (Russolillo et al., 2011). Meanwhile, the Lee-Carter model has been applied to investigate the demographic behaviour in developed countries, such as Australia, Canada, Chile, England, France, German, Japan, Sweden and selected developing countries with high quality long term data

(Alho, 2008; Booth et al., 2006; Härdle and Myšičková, 2009; Lee, 2000; Lee and Miller, 2001; Tuljapurkar et al., 2000). Nevertheless, due to the incapacity of the Lee-Carter model to deal with the lack of high quality and long term data, there are very limited studies on developing countries including China. On the other hand, the failure of the standard Lee-Carter model to capture those external factors which substantially influence human's mortality behaviour has aroused our interest as well.

Under the pressure of rapid population ageing and insufficient personal savings, China's Central government is currently undertaking extensive pension reforms. This requires accurate pension liability estimation, and hence accurate life expectancy forecasts at both national and regional levels. However, there have been few studies on life expectancy and mortality improvements in China due to a lack of high quality, long term data. Furthermore, given the vast regional differences in socioeconomic development in China and demonstrated strong correlations between socioeconomic vaiables and mortality behaviour, it raises questions about how severe the regional differences in life expectancy are in mainland China and the impact of socioeconomic factors on China's regional differences. This research addresses these identified gaps by extending the Lee-Carter model to incorporate of socioeconomic variables and to fit the limited provincial mortality data for China. The methodology is explained in more detail in next section.

CHAPTER 3

DATA & METHODOLOGY

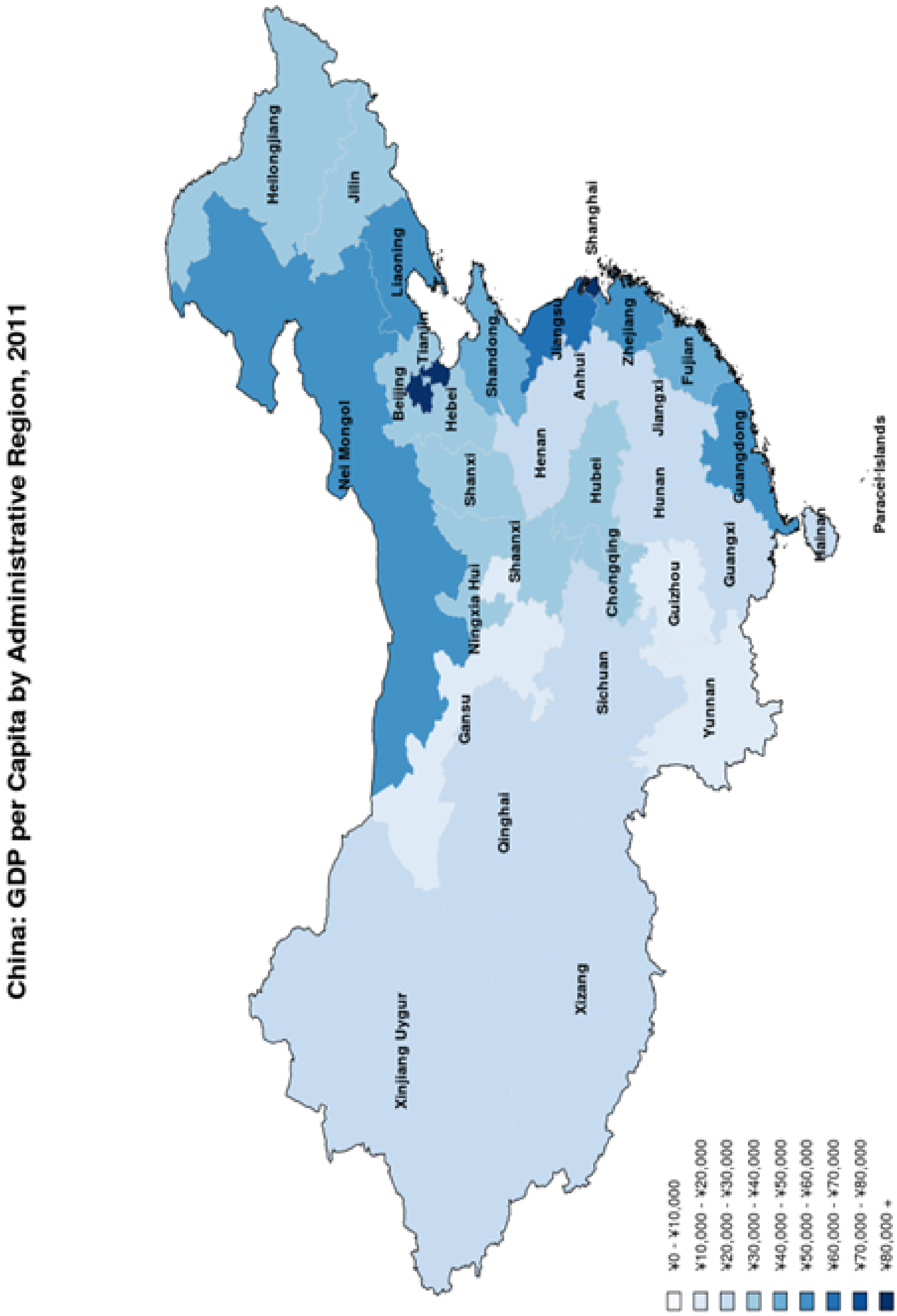
This chapter discusses the sources of data and the methodology applied. First, the study describes construction of a unique set of life tables for China on a regional basis for both urban and rural populations. A barrier to the completion of this project was the absence of life tables at the regional level (by province and for urban and rural populations) as there was no formal data source on a regional basis. The relevant data only was available from different sources as PDF documents, so all data had to be manually input into excel spreadsheets. From this initial data, the regional life tables for each province and for the urban and rural populations were constructed following Cai (2005). Next, the methodology to be applied is discussed. A variant of the extended Lee-Carter model developed by Li et al. (2004) is applied to investigate regional differences in mortality behaviour. Finally, the relationship between the index of the intensity of mortality, k_t, and changes in socioeconomic fluctuations, such as GDP, unemployment rate and number of doctors per 10,000 people, is incorporated following Hanewald (2011).

3.1 Data

As there is no existing life table available in China at the regional level, this research applies the methodology presented in Cai (2005) to construct regional life tables based on Chinese Census data from 1987, 1990, 1995, 2000, 2005, and 2010. These data was collected from four sources, including the National Bureau of Statistics of China, the National Population and Reproductive Health Science Data Center, China's National Knowledge Infrastructure and China Data Online. Both population and death by age, sex and region are available at the regional level (by province and for urban and rural populations) for infants and standard five-year age groups up to age 100, and for the group age 100 and over. There are 31 provincial level administrative regions in Mainland China, including 22 provinces, 4 municipalities (Beijing, Tianjin, Shanghai, Chongqing) and 5 autonomous regions (Guangxi, Inner Mongolia, Tibet, Ningxia, Xinjiang). Data from ten provinces are classified into three major groups according to their economic development status – that is, 'developed' (Beijing, Shanghai and Guangdong), 'developing' (Liaoning, Henna and Anhui) and 'least developed' (Gansu, Tibet, Yunnan and Sichuan) (See Figure 3.1).[1] Meanwhile, national level mortality data from 1994 to 2010 was collected from China's National Knowledge Infrastructure. National level socioeconomic data from 1994 to 2010 was collected from the National Bureau of Statistics of China. These data allowed us to investigate the relationship between socioeconomic variables and mortality improvements (discussed further in Section 3.2.3). Since the majority of data was provided in PDF format, the data had to be manually collected and imported into a consolidated database. Extensive cleaning and data quality check was conducted to guarantee the high quality of mortality and socioeconomic data. The data was then cleaned and reconstructed into the usual format. A unique dataset was constructed by compiling the regional level panel data. It represents the first dataset containing national and regional mortality data for both urban and rural populations for China.

[1]Economic status was compared using regional GDP per capita.

Figure 3.1: Regional GDP per Capita in China (2011).



Source: National Bureau of Statistics of China

3.2 Methodology

The aim of this research is to extend and apply the Lee-Carter model in order to estimate mortality at both the national and regional level for China (and to draw implications for age-related policies, such as pensions). This involved three key steps:

1. **Construction of Regional Life Tables**: This involved building regional life tables (by province and for urban and rural populations) using raw data published in six Chinese Censuses (1987, 1990, 1995, 2000, 2005, and 2010). The overall effect of the stochastic variation of life table functions was then tested using the standard errors of life expectancies.

2. **Prediction Based on Limited Data**: The aim of this step was to fit the Lee-Carter model to the regional life tables constructed in the first step. The extended time series methods developed by Li et al. (2004) were then used to model the Lee-Carter mortality index k_t. Next, the forecast of the future mortality rate and life expectancy were undertaken and reported on a disaggregated basis. This allowed an assessment of the future trend of the change in China's regional differences in life expectancy.

3. **Incorporation of Socioeconomic Variables**: In this step, the relationship between the index of the intensity of mortality, k_t, and changes in socioeconomic fluctuations was analyzed for males and females at national level. The significant variables were incorporated into the time series model as developed by Li et al. (2004). A sensitivity analysis was implemented to the prediction of future regional difference in mortality improvements.

Finally, the projected life expectancies will be used to assess the sustainability of the PAYG financed and fragmented pension system.

3.2.1 Construction of Regional Life Tables

Using this data, the regional life tables (by province and for urban and rural populations) were constructed for males, females and the total population for the years, 1987, 1990, 1995, 2000, 2005, and 2010, with columns displaying the following:

1. The age interval, $(x, x+n)$
2. The central death rate in the interval, ${}_n\mathrm{m}_x$
3. The death rate in the interval, ${}_nq_x$
4. The number alive at exact age x, l_x
5. The number of deaths in the interval, ${}_n\mathrm{d}_x$
6. Total person-years lived in the interval, ${}_n\mathrm{L}_x$
7. Total person-years lived beyond age x, T_x
8. Life expectancy at age x, e_x

To construct the Life Tables, we first calculated the central age-specific death rate, ${}_n\mathrm{m}_x$, using the formula proposed in Cai (2005).

$$ {}_n\mathrm{m}_x = \frac{{}_n\mathrm{D}_x}{{}_n\mathrm{P}_x^m} \tag{3.1} $$

where ${}_n\mathrm{D}_x$ is observed deaths between age x and $x+n$ in the Census. ${}_n\mathrm{P}_x^m$ is the person-years lived for those between x and $x+n$. This can be approximated as follows:

$$ {}_n\mathrm{P}_x^m = \frac{1}{2}({}_n\mathrm{P}_x +_n \mathrm{P}_{x+1}) + \frac{{}_n\mathrm{a}_x}{n} \times_n \mathrm{D}_x \tag{3.2} $$

where ${}_n\mathrm{P}_x$ is the population between age x and $x+n$ and ${}_n\mathrm{P}_{x+1}$ represents the population between age $x+1$ and $x+1+n$. In the study, ${}_n\mathrm{P}_x$ was used to approximate $\frac{1}{2}({}_n\mathrm{P}_x +_n \mathrm{P}_{x+1})$ as ${}_n\mathrm{P}_{x+1}$ was not available in collected data. $\frac{{}_n\mathrm{a}_x}{n}$ equals 'the average number of years lived in the year prior to the Census for those who died in the year prior to the census' (Cai, 2005). The reason to add $\frac{{}_n\mathrm{a}_x}{n}$ is to solve the asynchronous reporting issue in which the live population is reported at the time of the Census but the death population is collected in the year prior to the Census. In addition, we assumed the distribution of birth day over years is uniformed.

According to Cai (2005), ${}_n\mathrm{a}_x$ is chosen to equal $\frac{n}{2}$, except for age 0 and age 1 - 4. For the two youngest age groups, we apply the adapted Coale and Demeny formula recommended in Cai (2005).

Table 3.1: Coale and Demeny Formula

		Males	Females
$_1\mathrm{a}_0$	if $_1\mathrm{m}_0 \geq 0.107$,	0.330	0.350
	otherwise	0.045+2.684 $_1\mathrm{m}_0$	0.053+2.800 $_1\mathrm{m}_0$
$_4\mathrm{a}_1$	if $_1\mathrm{m}_0 \geq 0.107$,	1.352	1.361
	otherwise	1.651 - 2.816 $_1\mathrm{m}_0$	1.522 - 1.518 $_1\mathrm{m}_0$

By combining equation 3.2 and Table 3.1, $_1\mathrm{m}_0$ can be obtained through solving the following equation:

$$_1\mathrm{m}_0 = \frac{_1\mathrm{D}_0}{\frac{1}{2}\left(_1\mathrm{P}_0 +_1 \mathrm{P}_1\right) + \left(\alpha + \beta \times \ _1\mathrm{m}_0\right) {_1\mathrm{D}_0}} \tag{3.3}$$

where α and β are the coefficients in the Coale and Demeny formula. For the situation of two solutions of $_1\mathrm{m}_0$ from equation 3.3, we choose the one falling in the restriction range between 0 and the ceiling value (0.330 for male and 0.350 for female). In addition, $_1\mathrm{P}_0$ was used to approximate $\frac{1}{2}(_1\mathrm{P}_0 +_1 \mathrm{P}_1)$ as $_1\mathrm{P}_1$ was not available in collected data. Assuming the death rate is uniformly distributed in the interval, $_nq_x$ can be calculated using the following equation:

$$_nq_x = \frac{n \times_n \mathrm{m}_x}{1 + (n -_n \mathrm{a}_x)\ _n\mathrm{m}_x} \tag{3.4}$$

In the next step of constructing the Life Tables, we set the radix of the table as 100,000 and calculate the other columns in the table as follows:

$$_n\mathrm{d}_x = l_x \times_n q_x \tag{3.5}$$

$$l_{x+n} = l_x -_n \mathrm{d}_x \tag{3.6}$$

$$_n\mathrm{L}_x = n \times (_n\mathrm{d}_x \times \frac{_n\mathrm{a}_x}{n} + l_{x+n}) \tag{3.7}$$

$$_\infty\mathrm{L}_{100} = \frac{_\infty\mathrm{d}_{100}}{_\infty\mathrm{m}_{100}} \tag{3.8}$$

$$T_x = \sum_{i=x} \infty_n \mathrm{L}_i \tag{3.9}$$

$$e_x = \frac{T_x}{l_x} \tag{3.10}$$

There were a number of specific modifications taken to adapt to collected data as follows:

1. For high age groups with no death observed, ${}_n\mathrm{L}_x$ was set equivalent to l_x.

2. For some regions in some years for age greater than 80, original data set were collected as compiling all number of deaths into one group for age greater than 80, whereas survival population were collected separately as 80, 85, 90, 95, 100. As a result, the survival population for age greater than 80 needed to be compressed into one age group.

3. For ${}_n\mathrm{q}_x > 1$ at high age groups, such as age 90 and 95, we combined the age groups into one group and set ${}_n\mathrm{q}_x$ to 1.

Ultimately, we gauge the overall effect of stochastic variation of Life Table functions using the standard errors of life expectancies according to Cai (2005).

$$S^2_{{}_nq_x} = \frac{n \times_n \mathrm{m}_x \times (1 -_n \mathrm{a}_x \times_n \mathrm{m}_x)}{{}_n\mathrm{P}^m_x \times [1 + (n -_n \mathrm{a}_x) \times_n \mathrm{m}_x]^3} \tag{3.11}$$

$$S^2_{\widehat{e_x}} = \frac{\sum_{x=y}^{95} l_x^2 \times [\widehat{e_x} + (n -_n \mathrm{a}_x))]^2 \times S^2_{{}_nq_x}}{l_x^2} \qquad y = 0, 1, 5, 10, ..., 95 \tag{3.12}$$

A complete set of life tables was derived. [2]

3.2.2 Prediction Based on Limited Data

Next, we turn to model developing and fitting. In this stage, the mortality behaviour of each region was fitted through the standard Lee-Carter model of mortality as discussed in

[2] The derived life tables can be obtained by emailing jichen.li@unsw.edu.au.

Section 2.1:

$$ln(m_{x,t}) = a_x + b_x \times k_t + \varepsilon_{x,t} \tag{3.13}$$

where $m_{x,t}$, which is the central death rate for age x in year t, is calculated as in the previous section; a_x and b_x are age-specific constants; k_t, as a time-varying index, is the index of the intensity of mortality; $\varepsilon_{x,t}$ is the error term with mean zero and variance σ_{ε^2}. The unique least square solution for the model was obtained under two specific constraints:

$$\sum_{x=1}^{w} b_x = 1 \qquad and \qquad \sum_{t=1}^{n} k_t = 0 \tag{3.14}$$

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

Next, the extended time series method developed by Li et al. (2004) was used to deal with the issue of limited data for China. The extended model of the mortality index, k_t, was modelled as follows:

$$k_{t,\phi,\theta} = k_T + c(t-T) + \sigma \sum_{s=T+1}^{t} e(s) \tag{3.15}$$

where k_T represents the most recent index of the 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. c and σ can be obtained as follows:

$$\widehat{c} = \frac{k(u(T)) - k(u(0))}{u(T) - u(0)} \tag{3.16}$$

$$\widehat{\sigma} = \frac{\sum_{t=1}^{T}\{k(u(t)) - k(u(t-1)) - \widehat{c}[u(t) - u(t-1)]\}^2}{u(T) - u(0) - \frac{\sum_{t=1}^{T}[u(t) - u(t-1)]^2}{u(T) - u(0)}} \tag{3.17}$$

where $u(T)$ represents the year of the last data point and $u(0)$ represents the year of the first data point.

The standard error in estimating c and σ were obtained according to Li et al. (2004), as follows:

$$var(\widehat{c}) = \frac{\widehat{\sigma}^2}{u(T) - u(0)} \tag{3.18}$$

$$var(\widetilde{\sigma}) = \frac{\widehat{\sigma}}{2 \times \{u(T) - u(0)\frac{\sum_{t=1}^{T}[u(t) - u(t-1)]^2}{u(T) - u(0)}\}} \tag{3.19}$$

Then,

$$re(\widehat{\sigma}) = \frac{\sqrt{var(\widetilde{\sigma})}}{\widetilde{\sigma}} \tag{3.20}$$

$$c = \widehat{c} - \widehat{\sigma}\frac{1 - re(\widehat{\sigma}) \times \theta}{\sqrt{u(T) - u(0)}}\phi \tag{3.21}$$

$$\sigma = \widehat{\sigma}[1 - re(\widehat{\sigma}) \times \theta] \tag{3.22}$$

where ϕ and θ are two independent standard normal variables which account for parameter risk. θ reflects estimated errors in using historical data. According to Li et al. (2004), the forecast 95% probability intervals of k_t were obtained by taking θ=0. The wide and narrow bounds of the forecast 95% probability intervals were obtained by using θ=1.96 and θ =-1.96 respectively.

Then the prediction of mortality behaviour was calculated using:

$$ln(m_{x,t}) = ln(m_{x,T}) + b_x[k_t - k_T] \tag{3.23}$$

where k_t is forecast from equation 3.15. $m_{x,T}$ reflects the most recent observed age-specific central death. Furthermore, the 95% probability intervals of $m_{x,t}$ were derived from the uncertainty in the forecast of k_t (Li et al., 2004).

The fitted Lee-Carter model was then applied to investigate the mortality behaviour in each region. furthermore, investigation was implemented into developed, developing and least developed areas after aggregating populations according to their economic status. In addition, the different mortality performances of urban and rural populations were assessed as well.

3.2.3 Incorporation of Socioeconomic Variables

Hanewald (2011) found that socioeconomic variables had strong correlations with mortality improvements. In this context, China has experienced dramatic economic growth in past twenty years which significantly improved residents' living standards. Therefore, an investigation of the relationships between socioeconomic variables, such as GDP, the unemployment rate and the number of doctors per 10,000 people, and mortality improvements in China is motivated. Due to limited mortality data available at the regional level, the analysis here focused on national level first. This was done in two steps.

First, the Phillips-Perron test, a generalization of the augmented Dickey-Fuller test, was applied to analyze the trend behavior of the mortality index k_t, which is obtained from equation 3.13 and 3.14 in the previous section, and socioeconomic indicators for each region (Hanewald, 2011). The socioeconomic variables chosen to be investigated were GDP, GDP per capita, the unemployment rate, the number of doctors per capita, the number of healthcare assistants per capita, public healthcare expenditure and the number of residents who have completed tertiary education. This stationary test is not only applied to untransformed indicators but also to the indicators after taking differences or logarithms. Second, the pairwise Pearson correlation coefficients were evaluated between the stationary form of the mortality index and the socioeconomic factors (Hanewald, 2011). The hypothesis test on whether estimated correlations are equal to zero is implemented using a t-test with a significant level of 5% (Hanewald, 2011). The t-statistic is expressed as:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2} \tag{3.24}$$

where r is the estimated correlation and n represents the number of observations. A number of socioeconomic variables which have significant impacts on the index of the intensity of mortality are then selected to be incorporated into modelling of k_t. This is expected to overcome the shortcoming of the ordinary Lee-Carter model which does not take account of the external factors influencing mortality rates, such as macroeconomic growth and advances in medicine. Here we implement the idea presented in Hanewald

(2011), using the linear regression modelling of k_t with the incorporation of important socioeconomic indicators. For instance, if the first difference of k_t and the logarithm of some crucial socioeconomic variables are found to be stationary, then the regression model can be constructed as follows:

$$\Delta(k_t) = \theta + \sum_{i=1}^{n} \beta_i \Delta(ln(X_{t,i})) + \varepsilon_{x,t} \tag{3.25}$$

where $X_{t,i}$ represents important socioeconomic indicators 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 previously were compared based on the goodness-of-fit measure, R^2, (Cameron and Windmeijer, 1997).

$$R^2 = 1 - \frac{\sum_{t=0}^{T}[\Delta(k_t) - \theta - \sum_{i=1}^{n}\beta_i\Delta(ln(X_{t,i}))]^2}{\sum_{t=0}^{T}[\Delta(k_t) - \Delta(\bar{k_t})]^2} \tag{3.26}$$

where w is the maximum age and T is the maximum year. This measure tells us how much variance in the data is explained by the fitted models.

Furthermore, the tractability and predictability of the models were taken into account in the comparison. Next, by assuming all socioeconomic factors have the same impacts on local mortality improvements in each region, the selected models were used to forecast the future mortality rate and life expectancy in each region based on their fitted mortality indexes, k_t, and estimated future socioeconomic indicators. In addition, a sensitivity analysis was implemented for the prediction of future regional difference in mortality improvements based on various changes of socioeconomic factors. Consequently, a better understanding of the overall future trend of the change in China's regional differences in life expectancy can be obtained.

3.3 Conclusion

Overall, this chapter has presented a discussion of how the collected data was used to construct national and regional life tables (by province and for urban and rural populations). Then a technique to estimate the regional differences in mortality improvements using the extended Lee-Carter model developed by Li et al. (2004) was discussed. Next, this chapter described the method to investigate the relationships between selected socioeconomic variables and the index of the intensity of mortality, k_t. Finally, it explained how to conduct a sensitivity analysis of the prediction of future regional difference in mortality movement. Results and discussion of these results are presented in next two chapters.

CHAPTER 4

RESULTS AND DISCUSSION PART I — CONSTRUCTION OF LIFE TABLES

We constructed 800 national and regional life tables (for each province and for urban and rural populations) using the method discussed in Section 3.2.1. In this chapter, we report and discuss the historical trends in life expectancy by gender and by region for both urban and rural populations. Historical trends in life expectancy at birth and at age 65 are mainly investigated. We then discuss the potential implications of these trends for China's pension arrangements.

4.1 Results and Analysis

Sex specific life tables were constructed at national and regional (by province and for urban and rural populations) levels for the years - 1987, 1990, 1995, 2000, 2005 and 2010, following the steps outlined in Section 3.2.1. Due to missing mortality data in some provinces in some specific years, a total of 800 sex specific life tables were constructed at national and regional levels [1]. There are 19 administrative regions with more than five data panels. These account for 72.29% of total GDP and 63.90% of the total population (around 889 million people) in China as illustrated in Figure 4.1). These provinces cover most of the Chinese population and therefore provide a reliable database for the study and analysis. Examples of the life tables built are shown in the Appendix in Tables A.1, A.2 and A.3. These tables show survival probability, death rate and life expectancy for each age group and gender at national and regional level (by province and for urban and rural populations). On average, our calculations of life expectancy at birth are about two years higher than the result published by World Health Organization (WHO) (Lopez, 1999). This is due to the different methodology applied. The age specific death rate, ${}_nq_x$, is assumed to be uniformly distributed in the interval in this paper. However, Lopez (1999) used the Brass relational method to calculate mortality rates for WHO. This poses questions about the possible underestimation of future pension liabilities by Chinese government due to their reference to the lower predictions of life expectancy by WHO. Another thing worth noting is the volatility of mortality rate at the old age group, which might be because of the small sampling error for the higher age groups which have small number of survival and death population collected at regional level (Cai, 2005). One possible improvement could be conducted by combining data for those high age groups (age greater than 85) with few observed deaths and a smaller survival population.

[1] There are 304, 248 and 248 sex specific national and regional life tables constructed for the overall, the urban and the rural populations respectively. In total, we get 304 + 248 + 248 = 800 sex specific national and regional life tables for 1987, 1990, 1995, 2000, 2005 and 2010.

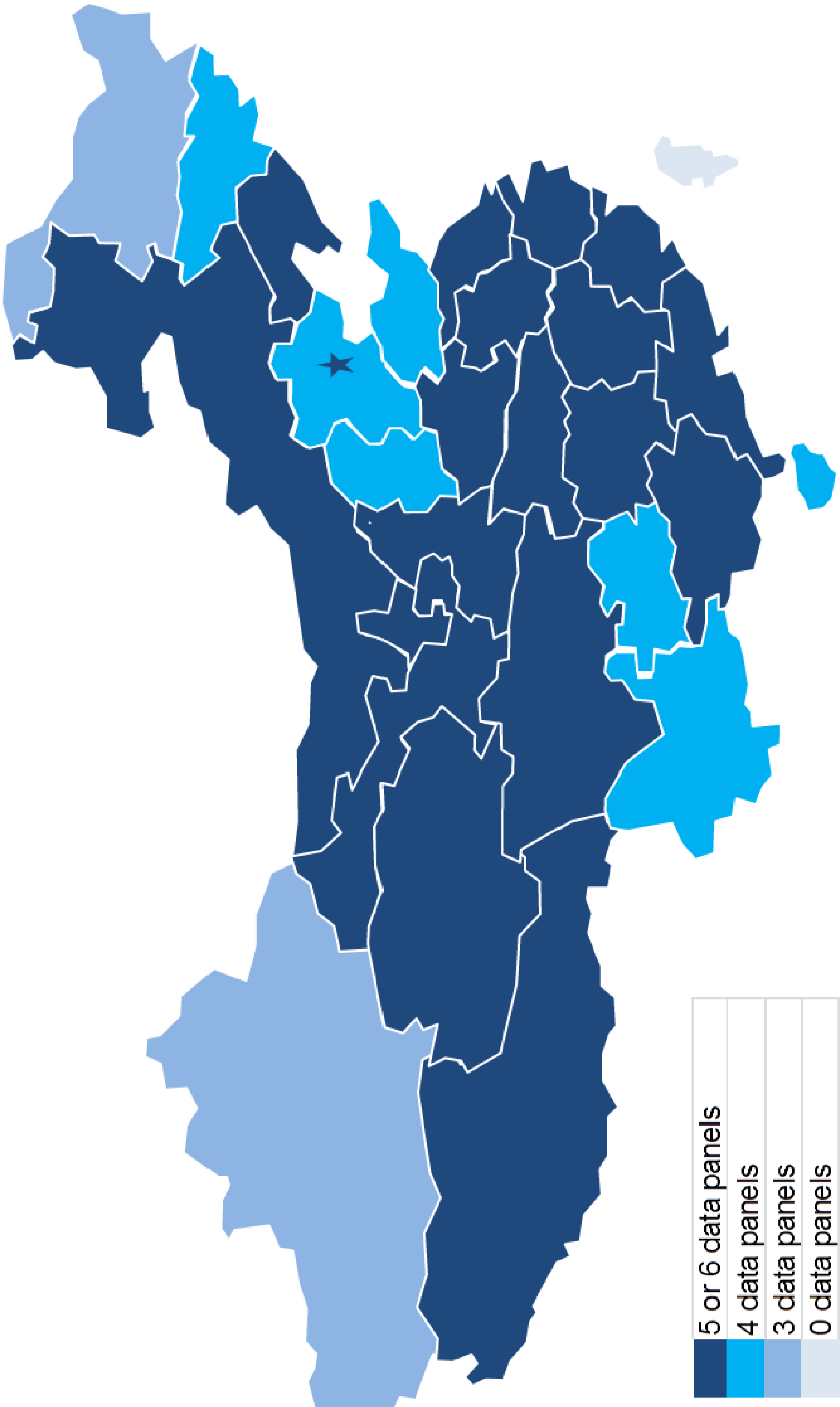


Figure 4.1: Number of data panels available for provincial level administrative regions

Source: Conducted by authors.

Some interesting findings regarding the projected life expectancy by region and by gender are discussed next.

Figure 4.2 comprises six charts of life expectancy at each age stage, by gender for each of the years, 1987, 1990, 1995, 2000, 2005 and 2010. [2] It shows that life expectancy has converged for males and females at older ages (age over 85). For instance, the difference in sex specific life expectancy was about 5 years for infants in 1987. The difference then reduced to about 3 years at age 65. For age over 85, life expectancy for male and female became the same. This is consistent with demographic experiences in Australia and United Kingdom. This implies that there are no significant sex specific differences in mortality rates for the current cohort at older ages, especially for those aged 85 and over. As a result, there is no immediate need for the standard pension policy to focus on gender difference at older ages. However, as only period life tables are available (rather than cohort life tables), it is unclear if this trend will continue to be the case for the future generations or is driven by cohort effects due to the mortality experience of the old age groups. These age groups experienced particular types of social events that are unlikely to be experienced by future generations, such as a series of wars from the 1920s to the 1980s, the Cultural Revolution and China's three years of natural disasters. However, more evidence is required to support this argument.

Figure 4.3 comprises six charts showing national and regional life expectancy at birth and at age 65, by gender for each of the years, 1987, 1990, 1995, 2000, 2005 and 2010. [3] This reports that life expectancy at each age group for both genders decreased from 1987 to 1990, then increased for the remainder of the period under consideration. For example, life expectancy at birth decreased from about 75 to 73 between 1987 and 1990. Then it increased to about 78 by 2010. In particular, for the age groups 0, 1-4, 20-24 and 60 and over, the death rate increased from 1987 to 1990 (see Table 4.1). After 1990, the death rate for all age groups started to drop, except for age 0-4, for which the death rate still increased from 1990 to 1995. This unusual demographic change can be partially explained by the impact of a number of natural disasters which occurred between 1988 and 1990 across the country, noting here that the death population was collected one year before the census. In 1988, Hunan and Hainan Provinces experienced floods which caused around 2.3 million people to be short of water (Wei et al., 2012). In the same year, drought occurred

[2] These illustrative examples represent a subset of the data.

[3] These illustrative examples represent a subset of the data.

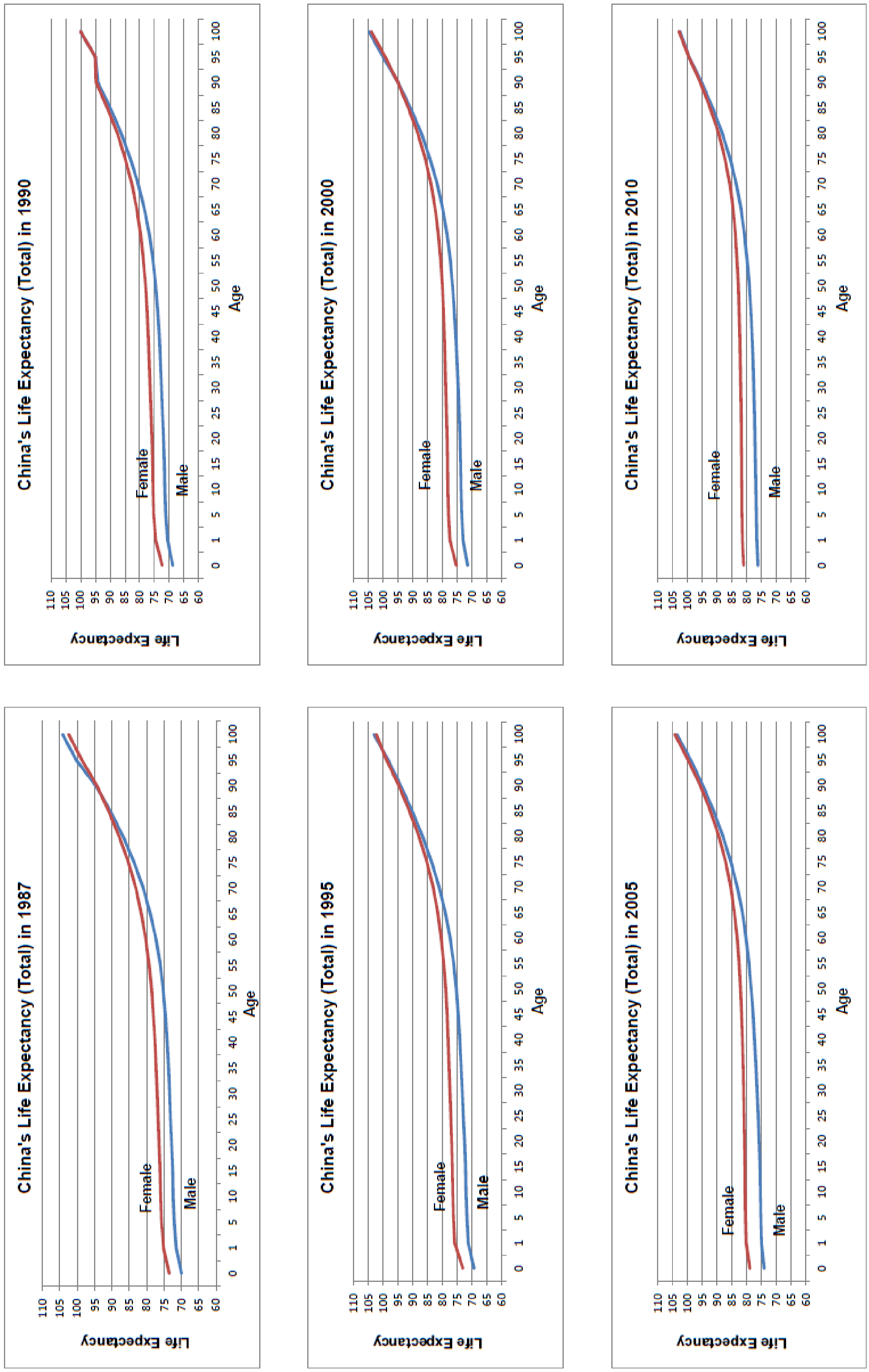
in Shandong, Inner Mongolia, and the area south of the Yellow River, which left about 90 million people short of food (Hualou and Jian, 2010; Wei et al., 2012). In 1989, there was also severe drought in Liaoning, Jilin, Heilongjiang, Inner Mongolia, Hebei, Hunan, Shanxi and Guangxi, which left 9.6 million people short of essential food (Davis, 2008).

Table 4.1: Growth rate of age specific death rate of China (Male, total population) from 1987 to 1990.

Age Group	Growth Rate of Age Specific death rate of China (Male) from 1987 to 1990
0	36.50%
1	10.85%
20	13.53%
60	4.31%
65	7.04%
70	7.96%
75	6.66%
80	9.30%
85	4.84%
90	7.62%

Source: Calculated using the method described in Chapter 3

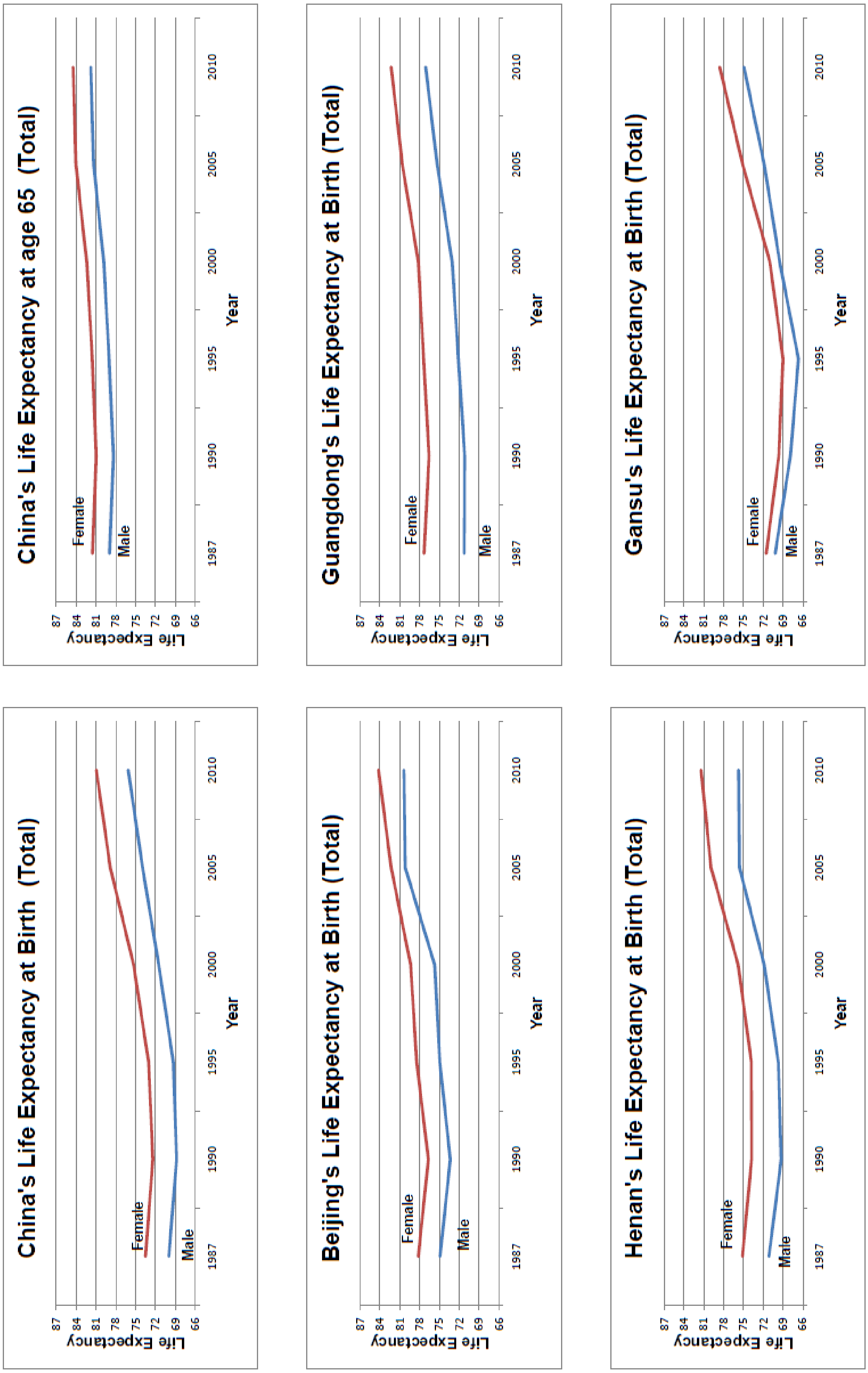
Figure 4.2: China's sex specific life expectancy at each age stage (total population) in 1987, 1990, 1995, 2000, 2005 and 2010.



Source: Calculated using the method described in Chapter 3

Notes: Demonstrated examples present a subset of the data.

Figure 4.3: National and regional life expectancy at birth and at age 65 (total population) in 1987, 1990, 1995, 2000, 2005 and 2010.



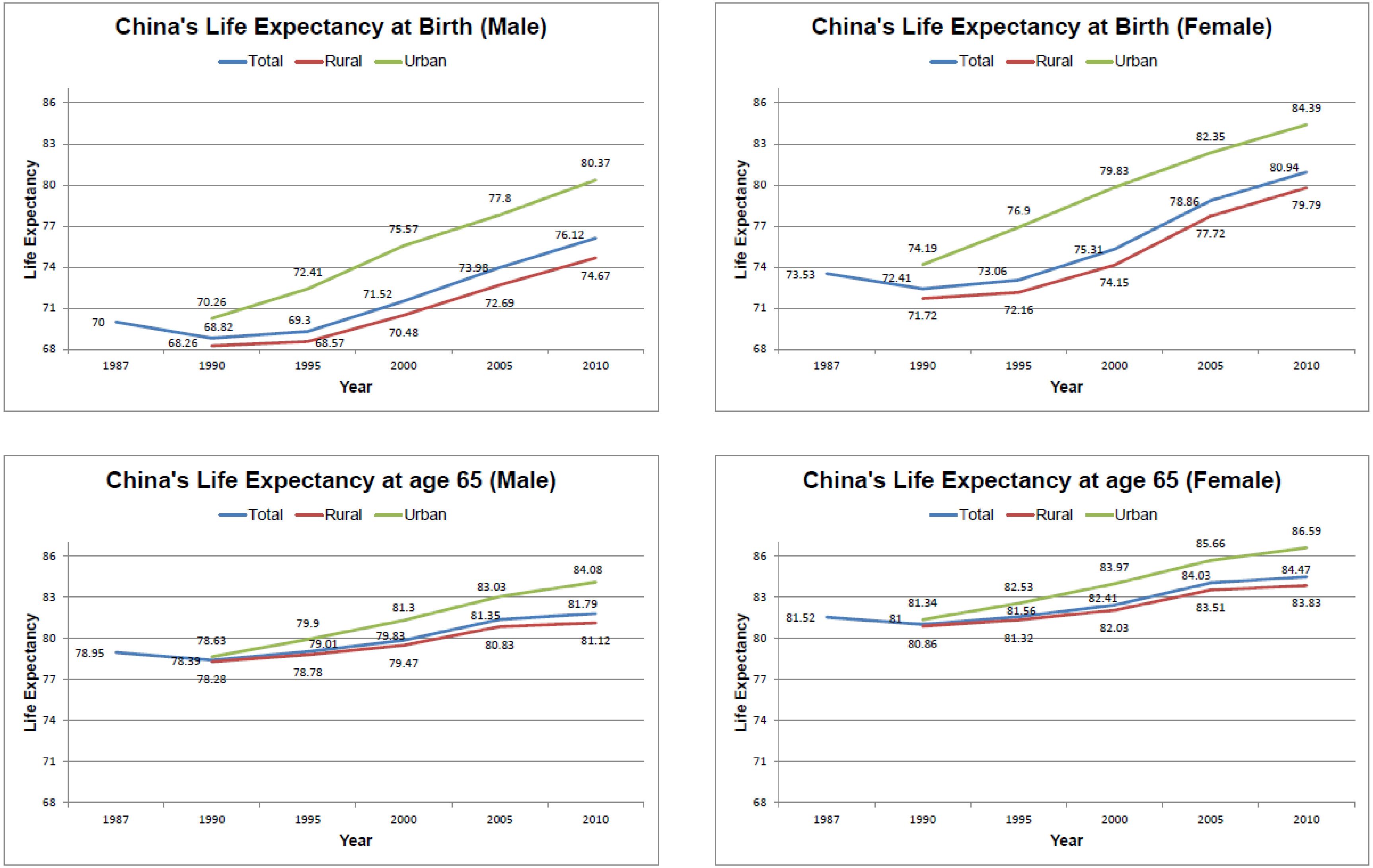
Source: Calculated using the method described in Chapter 3

Notes: Demonstrated examples present a subset of the data.

The life tables specially constructed for this project can be represented in numerous formats. An illustrative selection are shown in Figures 4.4, 4.5, 4.6 and 4.7. These figures illustrate the change in life expectancy at birth from 1987 to 2010 for males and females in urban and rural populations for all of mainland China (Figure 4.4), an example of a developed region (Beijing, Figure 4.5), a developing region (Henan, Figure 4.6) and a least developed region (Gansu, 4.7). It shows an increased difference in life expectancy at birth between the urban and rural population from 1990 to 2010. For males, at the national level, residents of urban areas have tended to experience higher life expectancy at birth (by around 4 years) than the population as a whole. People living in rural areas have experienced relatively lower life expectancy at birth than the overall population (ie, about 1 year lower). The difference in life expectancy for males at birth between rural and urban areas has continued to increase over the past 20 years. For females, at the national level, residents of urban areas have also experienced higher life expectancy at birth (of about 4 years) than the overall population. Similarly, females living in rural areas have experienced comparably lower life expectancy at birth (by about 1 year) than the overall population. The difference in life expectancy at birth between rural and urban areas increased between 1990 and 2000, then fell back to a stable level (of around 4 and a half years) until 2010. Similar differences are evident at the regional level. Overall the rural and urban differences in life expectancy at birth have increased between 1990 and 2010. This finding is consistent with studies in France, England and United States by Woods (2003), Hartley (2004) and Singh and Siahpush (2006).

Another age group of interest are the 65 year olds, as this is the most likely future official retirement age in China (Oksanen, 2010). Overall the difference in life expectancy at age 65 between rural and urban areas increased between 1990 to 2010 as seen in Figure 4.4 and 4.5. This finding is coincident with Woods (2003), Hartley (2004) and Singh and Siahpush (2006). However, this difference is smaller than for life expectancy at birth. For both males and females, at the national level, residents of urban areas have generally experienced higher life expectancy at age 65 than the overall population. People living in rural areas experience lower life expectancy at age 65 than the overall population. The difference in life expectancy at age 65 between rural and urban area has continued to increase over the past 20 years. Similar findings are evident at the regional level as well.

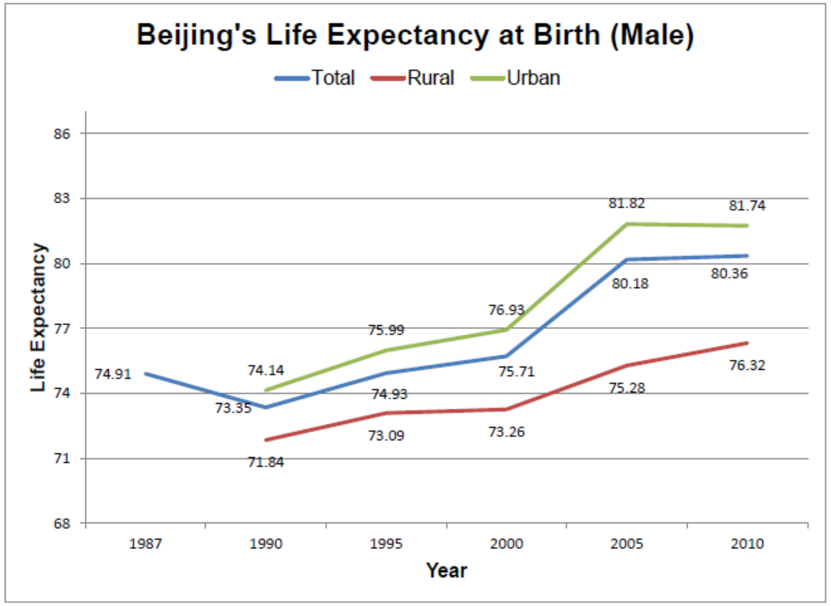
Figure 4.4: National life expectancy at birth and age 65 for total, urban and rural population from 1987 to 2010.

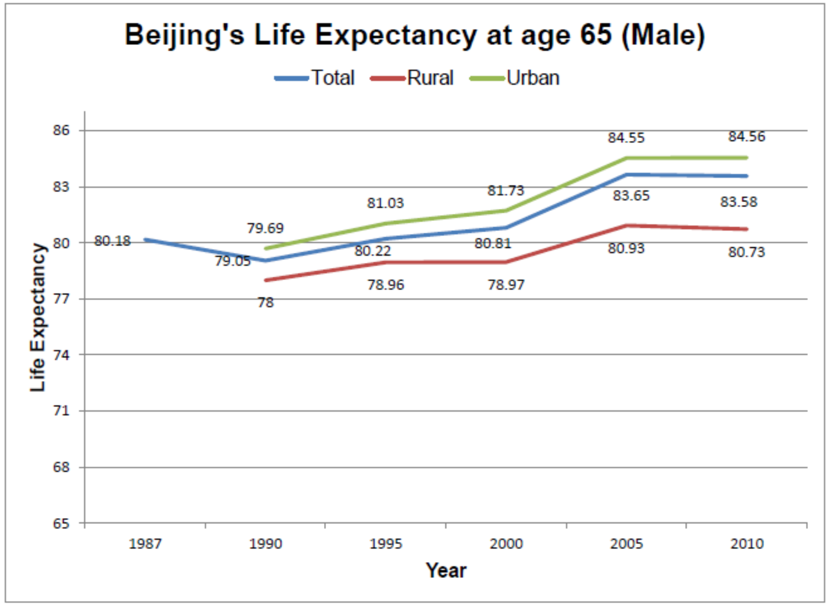


Source: Calculated using the method described in Chapter 3

Notes: Mortality data at urban and rural level missed in 1987

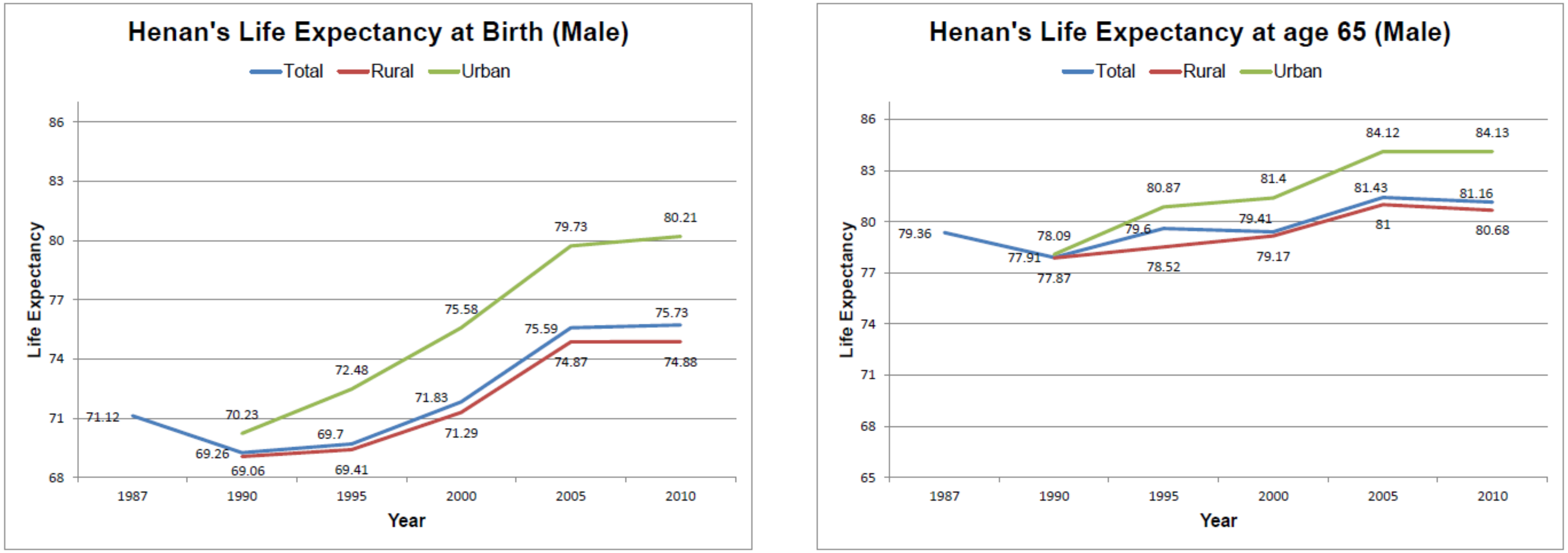
Figure 4.5: Life expectancy at birth and age 65 for total, urban and rural population from 1987 to 2010 — Beijing (an example of a developed region).





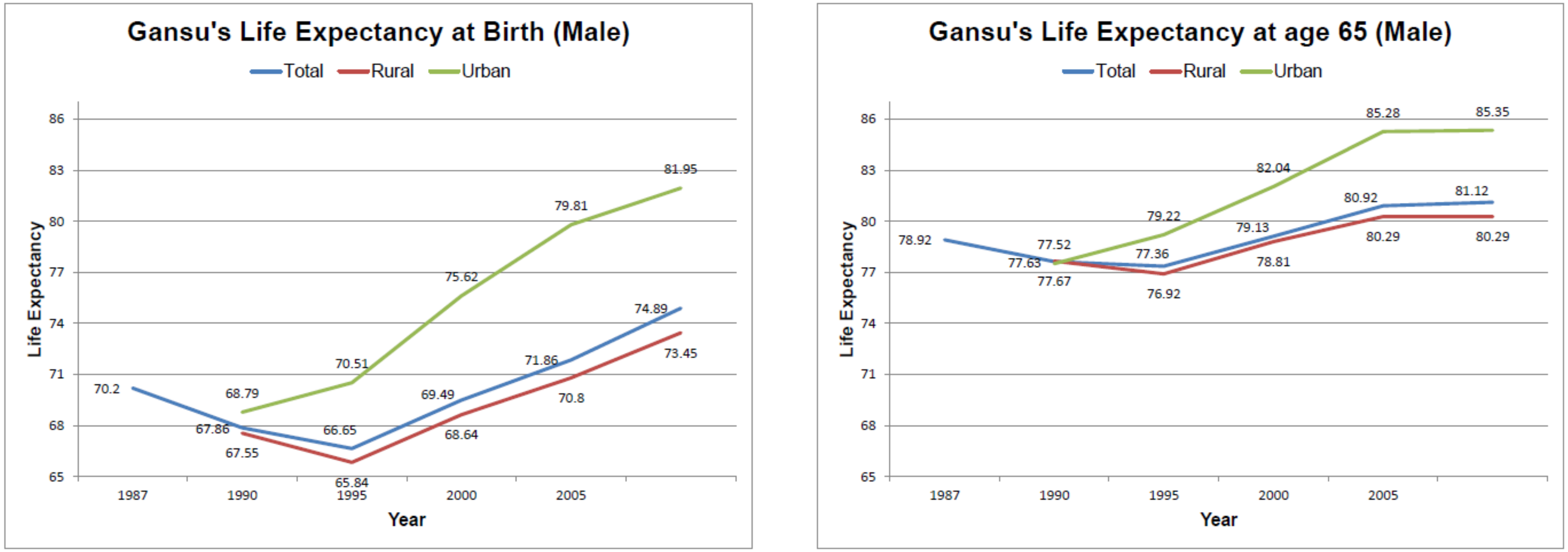
Source: Calculated using the method described in Chapter 3
Notes: Mortality data at urban and rural level missed in 1987

Figure 4.6: Life expectancy at birth and age 65 for total, urban and rural population from 1987 to 2010 — Henan (an example of a developing region).



Source: Calculated using the method described in Chapter 3
Notes: Mortality data at urban and rural level missed in 1987

Figure 4.7: Life expectancy at birth and age 65 for total, urban and rural population from 1987 to 2010 — Gansu (an example of a least developed region).

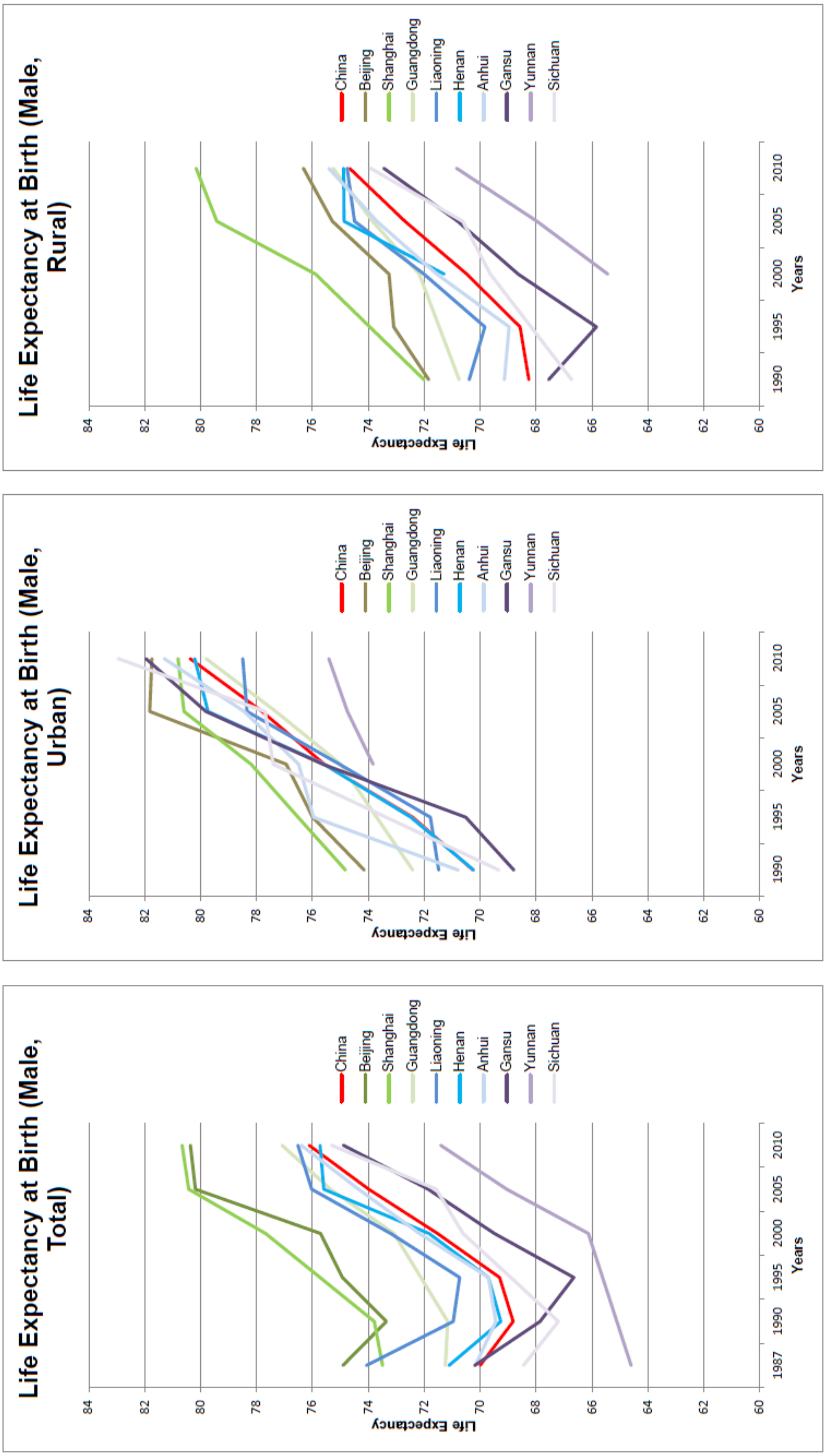


Source: Calculated using the method described in Chapter 3
Notes: Mortality data at urban and rural level missed in 1987

We then used the national level of life expectancy as the baseline to assess comparable regional differences in life expectancy. Regions were classified according to their economic status. That is, 'developed' (Beijing, Shanghai and Guangdong), 'developing' (Liaoning, Henna and Anhui) and 'least developed' (Gansu, Tibet, Yunnan and Sichuan). Figures 4.8 and 4.9 present life expectancy at birth at national and regional level (by province and for urban and rural populations). These figures illustrate that there was no significant difference in the urban populations of the various administrative regions, whereas the differences in the rural populations was quite marked. Therefore, the differences in life expectancy at birth and at age 65 for populations from different administrative regions appear to be mainly driven by the difference in life expectancy of the rural population in each region. [4] This indicates that the rural population should be the target of any initiatives to reduce the significant regional differences in life expectancy. Moreover, this finding fills the current literature gap as to the best of our knowledge, there is no study yet which explains disparities of which groups of population resulting in the regional differences in life expectancy. In addition, this raises questions about the impact of China's urbanization on the trend of regional differences in mortality improvements over coming decades. The vast regional differences in the development of urbanization can be seen in Figures 4.10 and 4.11. Furthermore, as current pension benefits are predetermined by a formula of which a benefit multiple is determined actuarially based on expected future mortality rate, pension policy reform should consider the calculation of benefits according to locally projected mortality rate. In addition, the regional differences in life expectancy at age 65 were much smaller than the differences in life expectancy at birth.

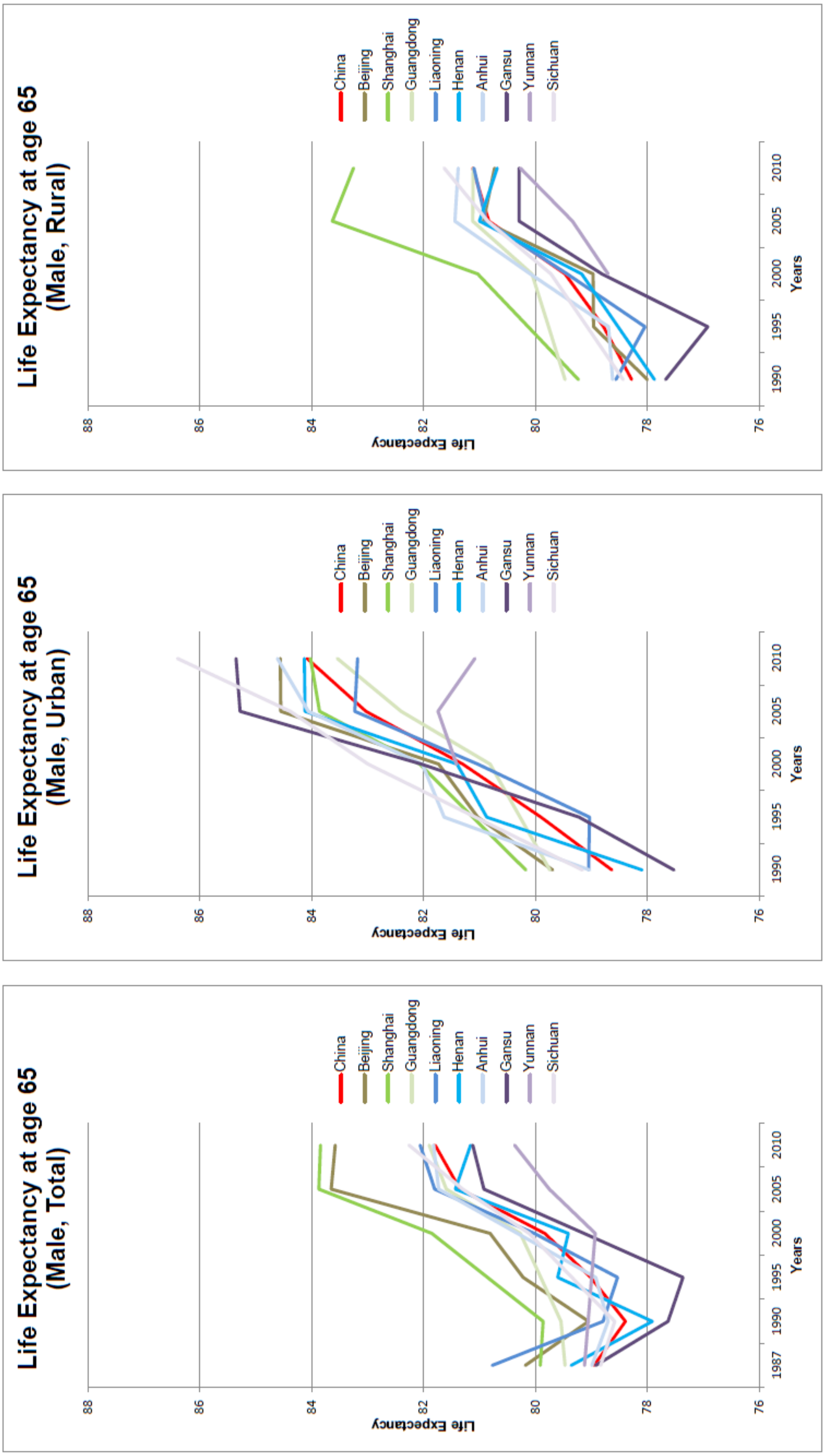
[4] For years, in which mortality data is missing, we apply the linear approximation to the life expectancy. For only one year of missing data, such as 1995, $e_{x,1995} = \frac{1}{2}(e_{x,1990}+e_{x,2000})$. If there are two years missing, such as 1995, 2000, $e_{x,1995} = \frac{1}{3}(e_{x,2005}-e_{x,1990}) + e_{x,1990}$ and $e_{x,2000} = \frac{2}{3}(e_{x,2005}-e_{x,1990}) + e_{x,1990}$.

Figure 4.8: Comparison of regional life expectancy at birth (Male) for total, urban and rural population from 1990 to 2010.



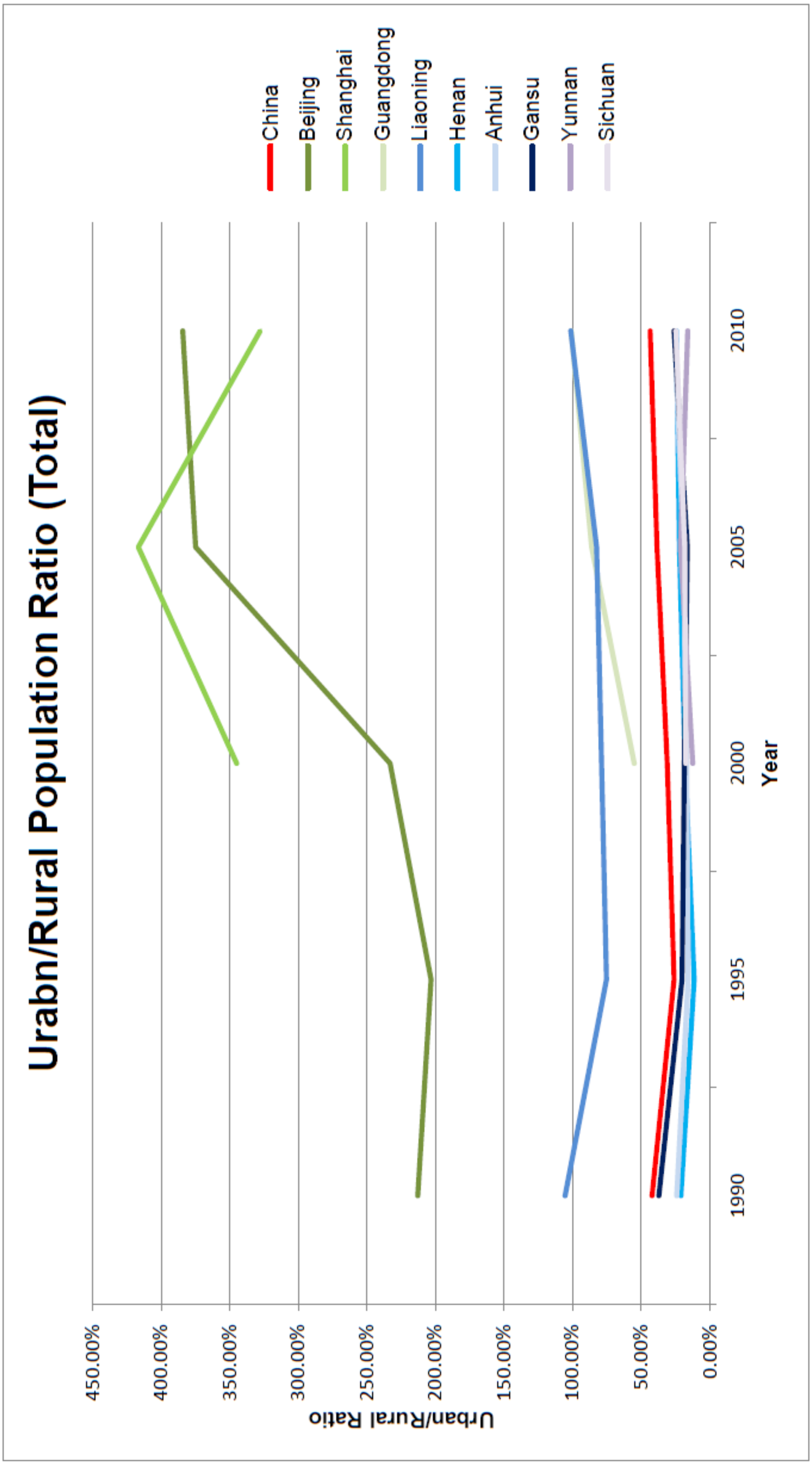
Source: Calculated using the method described in Chapter 3

Figure 4.9: Comparison of regional life expectancy at age 65 (Male) for total, urban and rural population from 1990 to 2010.



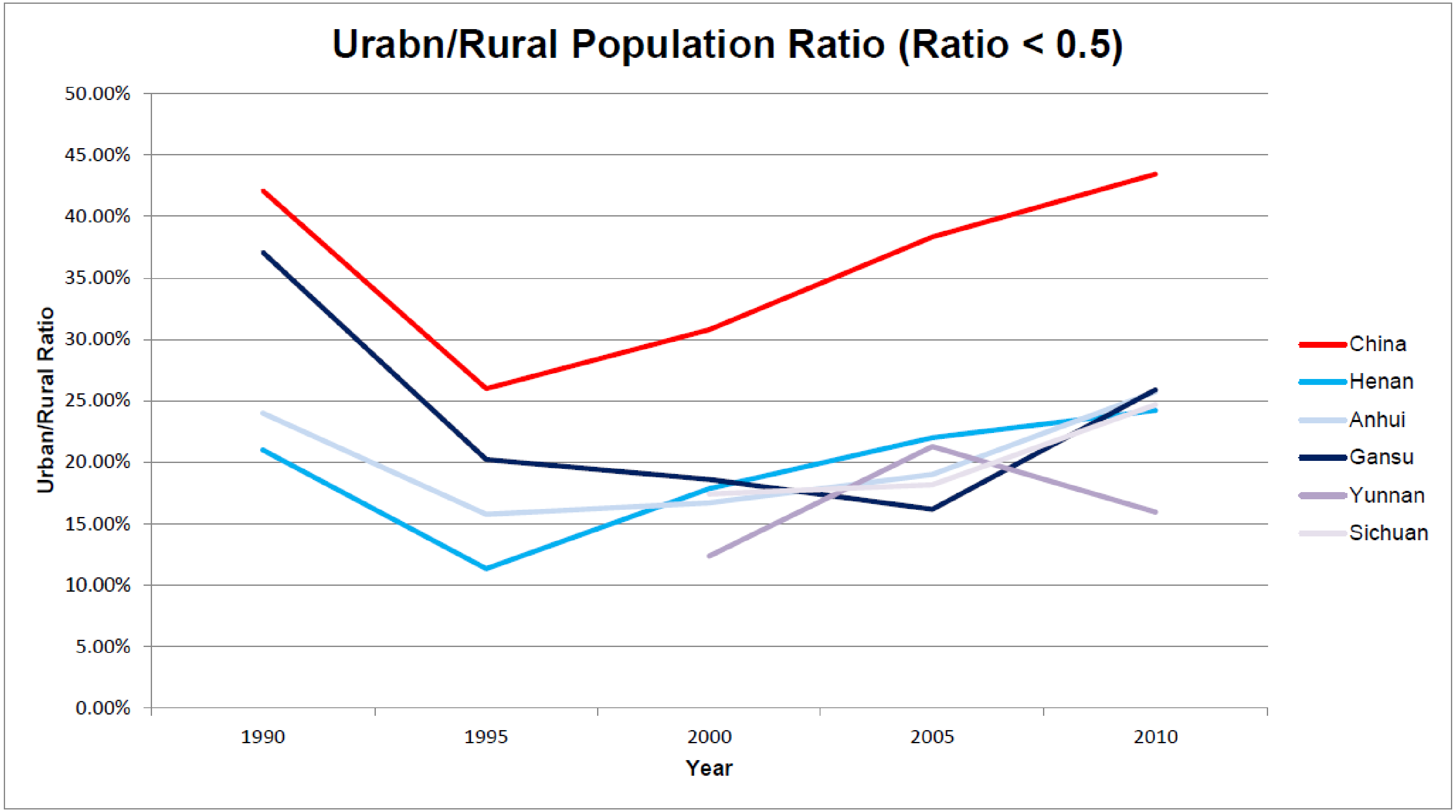
Source: Calculated using the method described in Chapter 3

Figure 4.10: Urban/Rural population ratio from 1990 to 2010 — Total.



Source: Calculated using the method described in Chapter 3

Figure 4.11: Urban/Rural population ratio from 1990 to 2010 — Ratio < 0.5.



Source: Calculated using the method described in Chapter 3

It is also found that differences in mortality rates are concentrated in the young and middle age groups. This helps explain the smaller difference in life expectancy at older age groups. To investigate this further, we first compared age specific death rates for the urban and rural populations (see Table 4.2). This table displays that younger age groups in rural areas experience much higher death rates than those people in urban area. For instance, the death rate for rural males at age 20-24 was 307% higher than the urban males in 2010. Next, the rural age specific mortality rates were further compared at the regional level as seen in Table 4.3. These shows that higher death rates are experienced at the younger age groups of rural population in the less developed regions. For example, the death rate for rural males at age 20-24 in Sichuan was 51% higher than that in Beijing in 2010. Through those comparisons, it is clear that the young and middle age groups had much larger dispersion in mortality than the older age groups.

4.2 Discussion and Conclusion

In this chapter, 800 national and regional life tables (by province and for urban and rural populations) were constructed using the unique dataset described in Chapter 3. The historical trends of national and regional mortality changes were then investigated. The study found that sex specific life expectancy converged at older ages (age over 85) which implied there were no significant sex specific differences in mortality rates at older ages. As a result, there may be no need for the standard pension policy to focus on gender difference at older ages. However, this issue might need cohort life tables for further investigations. Next, this study illustrated that the differences in life expectancy at birth and at age 65 for populations from different regions were mainly driven by the differences in life expectancy of the rural population in each region. In addition, the overall rural and urban differences in life expectancy at birth increased between 1990 and 2010. Nevertheless, the regional differences in life expectancy at age 65 were much smaller than the differences in life expectancy at birth. These raised questions about the impact of China's urbanization on the trend of regional difference in mortality improvement in next decades. Furthermore, as current pension benefits are predetermined by a formula of which a benefit multiple is determined actuarially based on expected future mortality rates, pension policy reform should consider calculating benefits according to locally projected mortality rates (rather than using national rates). Finally, the study also found that differences in mortality rates were concentrated in the young and middle age groups, which helps to explain the smaller differences in life expectancy at the older age groups.

Table 4.2: Comparison of age specific death rates at urban and rural level.

Percentage difference in age specific death rate (China, Male)

	1990	1995	2000	2005	2010
0 Year	41%	115%	183%	102%	58%
1-4 Years	52%	66%	160%	72%	123%
5-9 Years	23%	68%	83%	64%	159%
10-14 Years	18%	26%	51%	131%	152%
15-19 Years	20%	46%	143%	105%	246%
20-24 Years	23%	30%	128%	199%	307%
25-29 Years	24%	32%	96%	195%	263%
30-34 Years	29%	40%	78%	138%	232%
35-39 Years	25%	58%	58%	105%	168%
40-44 Years	23%	42%	50%	68%	126%
45-49 Years	16%	37%	41%	38%	82%
50-54 Years	13%	6%	40%	42%	68%
55-59 Years	13%	32%	36%	47%	61%
60-64 Years	10%	24%	37%	35%	61%
65-69 Years	7%	20%	29%	31%	52%
70-74 Years	6%	13%	27%	36%	50%
75-79 Years	3%	12%	21%	35%	39%
80-84 Years	2%	14%	18%	26%	33%
85-89 Years	0%	14%	13%	11%	21%
90-94 Years	0%	-11%	8%	13%	15%
95-99 Years		-9%	2%	11%	17%
100 Years and Over		0%	0%	0%	0%

Percentage difference in age specific death rate (China, Female)

	1990	1995	2000	2005	2010
0 Year	42%	141%	224%	174%	75%
1-4 Years	66%	89%	199%	164%	133%
5-9 Years	42%	62%	105%	95%	139%
10-14 Years	28%	33%	75%	42%	136%
15-19 Years	31%	75%	184%	152%	208%
20-24 Years	37%	119%	185%	244%	253%
25-29 Years	45%	55%	151%	205%	216%
30-34 Years	47%	82%	111%	190%	202%
35-39 Years	37%	56%	84%	102%	156%
40-44 Years	30%	63%	71%	72%	122%
45-49 Years	22%	31%	62%	73%	88%
50-54 Years	18%	48%	53%	42%	85%
55-59 Years	15%	47%	48%	43%	83%
60-64 Years	14%	23%	43%	40%	78%
65-69 Years	9%	17%	36%	42%	67%
70-74 Years	9%	24%	30%	47%	60%
75-79 Years	6%	10%	26%	35%	43%
80-84 Years	4%	13%	19%	30%	32%
85-89 Years	1%	5%	13%	14%	17%
90-94 Years	0%	4%	9%	5%	10%
95-99 Years		28%	8%	-12%	9%
100 Years and Over		0%	0%	0%	0%

$$\text{Formula: } = \frac{\textit{rural age specific death rate} - \textit{urban age specific death rate}}{\textit{urban age specific death rate}}$$

(Area in red for difference rate greater 10%)

Source: Calculated by authors

Table 4.3: Comparison of rural age specific death rates at regional level.

Percentage difference in rural age specific death rate (Henan vs Beijing, A vs B respectively, Male)

	1990	1995	2000	2005	2010
0 Year	85%	51%	409%	-8%	-49%
1-4 Years	123%	75%	244%		-28%
5-9 Years	19%	108%	30%		94%
10-14 Years	5%	140%	-5%		-15%
15-19 Years	79%	485%	40%	-51%	25%
20-24 Years	64%	99%	46%	61%	65%
25-29 Years	39%	178%	65%	-26%	127%
30-34 Years	46%	131%	29%	72%	116%
35-39 Years	50%	93%	15%	269%	64%
40-44 Years	58%	27%	19%	125%	55%
45-49 Years	51%	-4%	24%	-24%	37%
50-54 Years	31%	-7%	18%	-28%	39%
55-59 Years	40%	9%	12%	28%	23%
60-64 Years	34%	35%	9%	-15%	24%
65-69 Years	15%	23%	9%	17%	16%
70-74 Years	7%	-9%	5%	-10%	13%
75-79 Years	-6%		-5%	0%	2%
80-84 Years	-6%		-7%	-14%	0%
85-89 Years	-12%		-15%	-25%	-18%
90-94 Years	0%		-11%		-23%
95-99 Years			-14%		-30%
100 Years and Over			0%		0%

Percentage difference in rural age specific death rate (Sichuan vs Beijing, Male)

	1990	1995	2000	2005	2010
0 Year	31%		150%		78%
1-4 Years	171%	-100%	41%		41%
5-9 Years	88%	105%	42%		35%
10-14 Years	35%		12%		18%
15-19 Years	91%	229%	17%	-100%	-1%
20-24 Years	73%	71%	5%	445%	51%
25-29 Years	53%	135%	-8%	269%	94%
30-34 Years	90%	-10%	-15%	38%	119%
35-39 Years	88%	213%	3%	432%	103%
40-44 Years	57%	76%	20%	33%	59%
45-49 Years	62%	13%	11%	84%	36%
50-54 Years	55%	94%	15%	21%	25%
55-59 Years	40%	7%	18%	39%	41%
60-64 Years	42%	18%	24%	-26%	38%
65-69 Years	34%	-2%	14%	0%	35%
70-74 Years	27%	-2%	15%	-10%	25%
75-79 Years	15%		8%	19%	12%
80-84 Years	10%		3%	47%	2%
85-89 Years	2%		-3%	-9%	-1%
90-94 Years	0%		-4%		-22%
95-99 Years			-32%		-30%
100 Years and Over			0%		0%

$$\text{Formula:} = \frac{\textit{age specific death rate of Province A} - \textit{age specific death rate of Province B}}{\textit{age specific death rate of Province B}}$$

(Area in red for difference rate greater 10%)

Source: Calculated by authors

CHAPTER 5

RESULTS AND DISCUSSION PART II — EXTENDED LEE-CARTER MODEL

This chapter presents results of two extensions to the standard Lee-Carter model (proposed in Chapter 3) to fit the regional mortality data for China discussed in the previous chapter. First, the extended Lee-Carter model developed by Li et al. (2004) was used to address the limited availability of mortality data in China. Second, we incorporated selected socioeconomic factors into the standard Lee-Carter model following the methodology proposed by Hanewald (2011). Using these extensions, we first estimated the future trend of the change in China's life expectancy by region and for urban and rural populations. Next, we investigated the relationship between the index of the intensity of mortality, k_t and changes in socioeconomic fluctuations for males and females at the national level. Selected variables were then incorporated into the standard Lee-Carter model. The analysis concluded with a sensitivity analysis of the prediction of future regional differences in mortality improvements.

5.1 Extended Lee-Carter Model

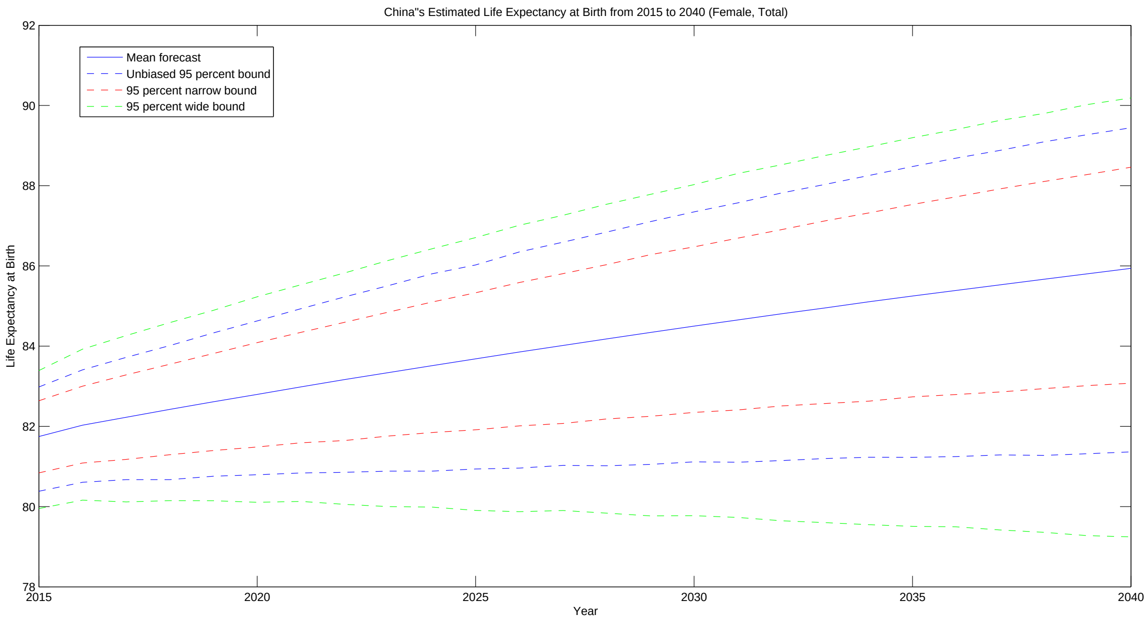
This section applied the extended Lee-Carter model to investigate differences in mortality improvements by region and for total, urban and rural populations. Three groups of regions were classified according to their economic status as discussed in previous chapter. We compiled provincial level administrative regions with similar socioeconomic backgrounds together aiming to avoid small sample errors.[1] On the other hand, in order to eliminate the unusual shock in 1990 and the impact of health care system reform during 1990s (Li et al., 2004), we applied the extended Lee-Carter model to fit mortality data from 2000 to 2010, and to forecast future mortality improvements from 2015 to 2040.

Figure 5.1 shows a prediction of future mortality improvements for the overall female population in China. Setting $\theta = 0$, the most likely forecast of future life expectancy at birth was obtained as the solid line. The other dash lines were its corresponding unbiased 95 percent confidence bound, 95 percent narrow bound (by setting $\theta = 1.96$) and 95 percent wide bound (by setting $\theta =$ -1.96). The relative estimating error, $re(\widehat{\sigma})$, was about 0.1655, which was quite high but consistent with the findings in Li et al. (2004). This relatively high estimating error was mainly because of the limited planes of available data. With more mortality data collected in the future, the forecast will be more reliable and accurate.

Next, investigation was undertaken into the trend of urban and rural differences in life expectancy over next twenty-five years. This is illustrated in Figures 5.2 and 5.3 which show a stable difference between the urban and rural population. A potential factor behind this phenomenon was the increasing urbanization across China. The Chinese government has upgraded most of its towns and villages into cities which are expected to significantly lift living standards in rural areas. This has caused significant mortality improvements for the rural population. On the other hand, for rural residents who enjoyed the fastest development, their hometowns might be upgraded and reclassified as cities. As a result, they were recounted as urban population which could be reflected as a dramatic rise of urban population ratio, from 26% of total population in 1990 to 52.6% in 2010 (Zheng, 2012). This resulted in a stable difference in life expectancies between urban and rural populations.

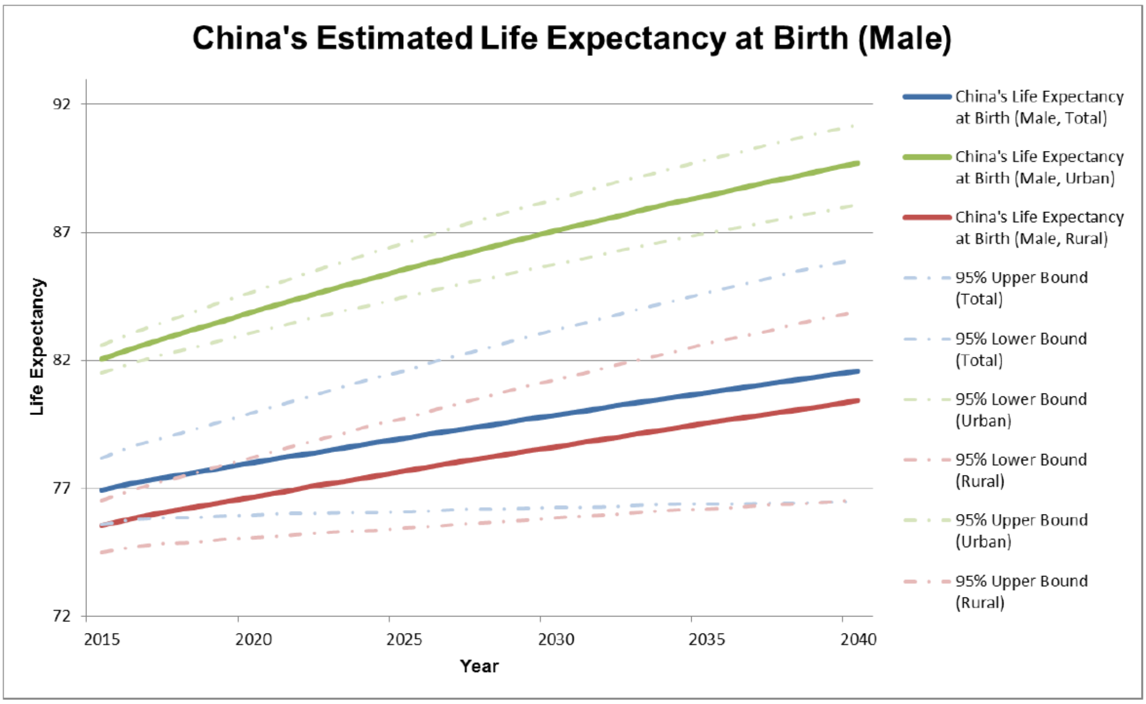
[1] Shanghai and Beijing were combined together to form the developed region. Henan and Anhui were combined together to form the developing region. Gansu and Yunnan were combined together to form the least developed region.

Figure 5.1: China's estimated life expectancy at birth from 2015 to 2040 (Female, Total).



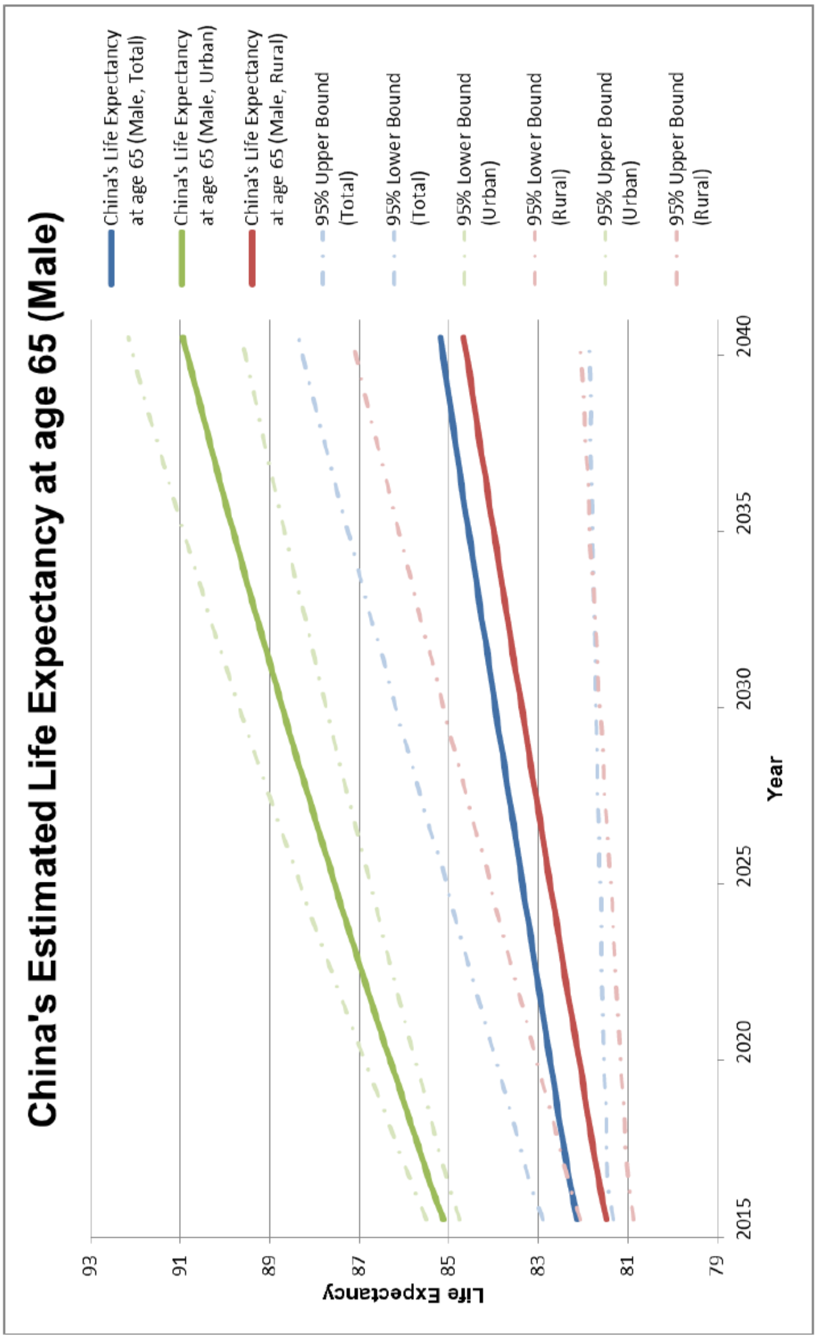
Source: Calculated using the method described in Chapter 3

Figure 5.2: China's estimated life expectancy at birth from 2015 to 2040 (Male).



Source: Calculated using the method described in Chapter 3

Figure 5.3: China's estimated life expectancy at age 65 from 2015 to 2040 (Male).



Source: Calculated using the method described in Chapter 3

Regional differences in life expectancy at birth and at age 65 were assessed according to the prediction using the extended Lee-Carter model as well. Figures 5.4, 5.5 and 5.6 illustrate estimated life expectancy improvements in developed, developing and least developed regions for total, urban and rural populations over next twenty-five years. From Figures 5.4, 5.5 and 5.6, a relative constant regional difference in life expectancy for the total population is observed, whereas the regional difference for the urban population is expected to converge. In comparison, the regional difference in rural areas is found to slightly diverge over next twenty-five years. These results suggest that a different treatment of urban and rural pension policies will be necessary. As the benefits provided by the PAYG pension system in China are predetermined locally by a formula based on individuals' earnings history, tenure of service, age and a benefit multiple designed to reflect the life expectancy of the residents in the local area (Dickson et al., 2013). This emphasizes the necessity of an accurate estimate of local mortality improvement to be able to monitor the sustainability of the public pension system. Therefore, future rural pension benefits should be aligned with local life expectancies to enhance sustainability. Nevertheless, further investigations are required to provide more reliable explanations.

In addition, as an extension to the commonly used model, the Li and Lee (2005) method which took account of the regional differences in life expectancy could be applied. Instead of modelling the mortality behaviour of each province separately, one common factor for each region could be incorporated into the modelling.

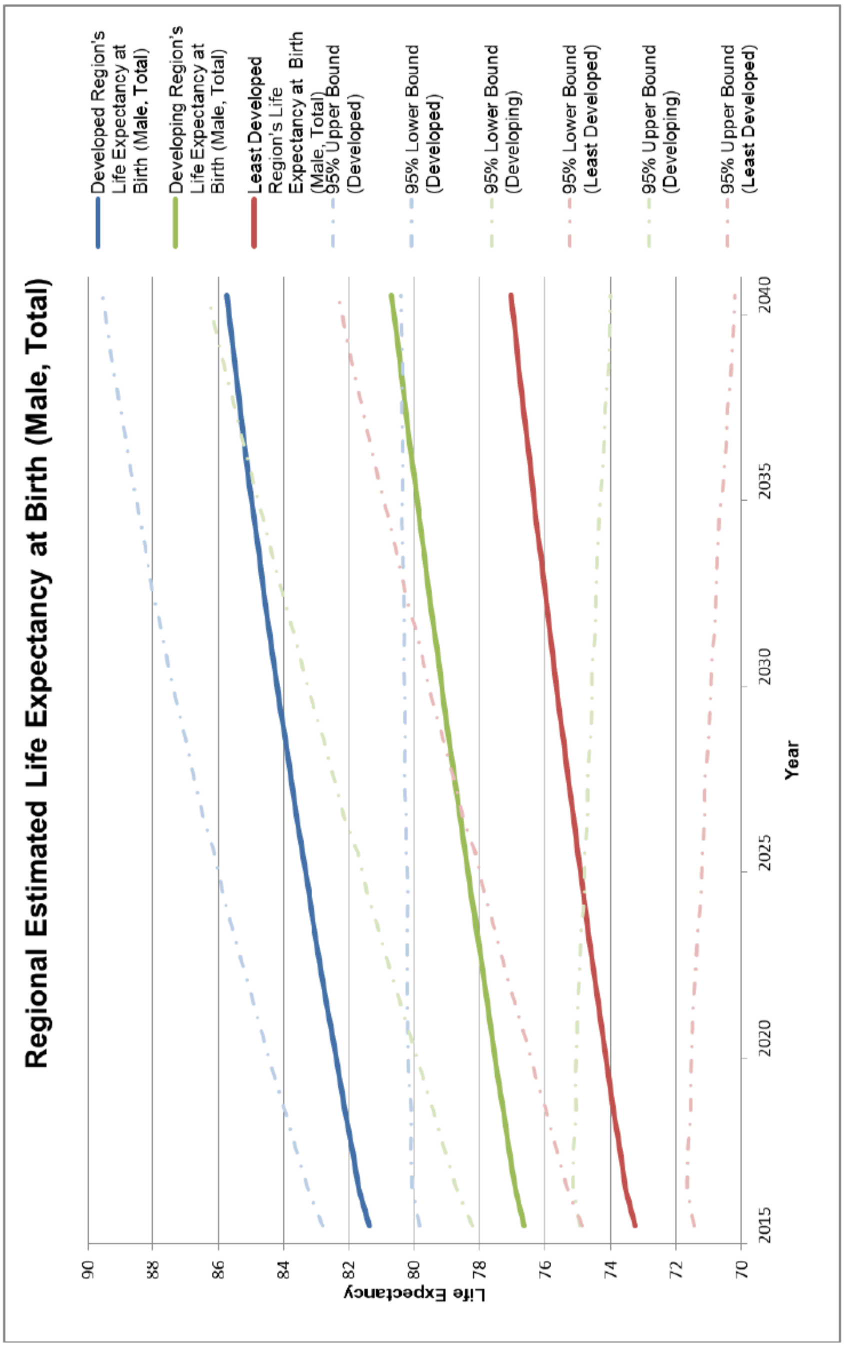
$$ln(m_{x,t,i}) = a_{x,i} + B_x \times K_t + b_{x,i} \times k_{t,i} + \varepsilon_{x,t,i} \tag{5.1}$$

where B_x and K_t, which are the common factors for all provinces, can be calculated though the application of the standard Lee-Carter model to the mortality behaviour of the entire nation. $a_{x,i}$, $b_{x,i}$ and $k_{t,i}$, reflect the specific factors for each province. $a_{x,i}$ can be obtained as follows:

$$a_{x,i} = \frac{\sum_{t=0}^{T} ln(m_{x,t,i})}{T+1} \tag{5.2}$$

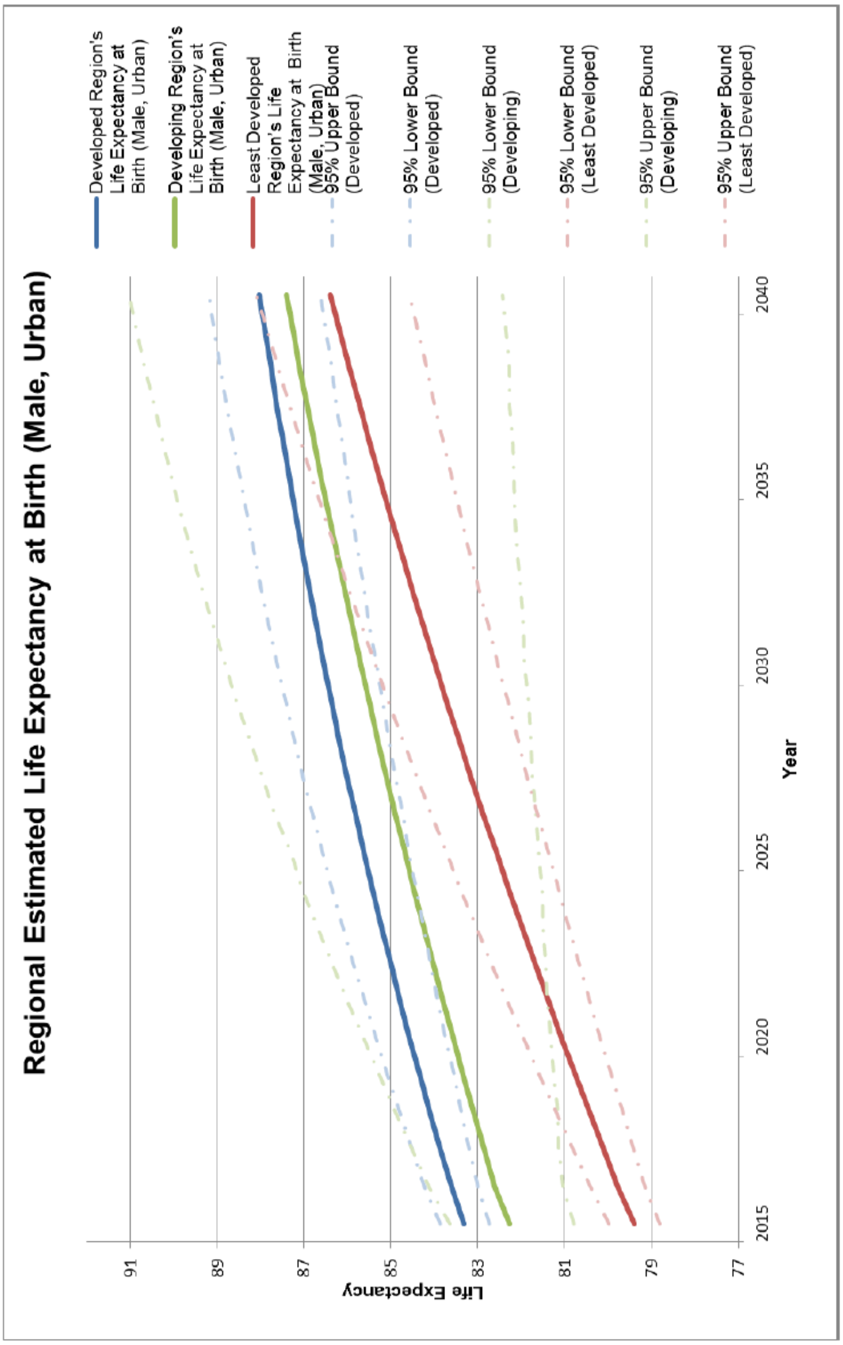
Next $b_{x,i}$ and $k_{t,i}$ can be estimated using the SVD method. Then the prediction of K_t and $k_{t,i}$ follows the procedures mentioned before based on Li et al. (2004).

Figure 5.4: Regional estimated life expectancy at birth from 2015 to 2040 (Male, Total).



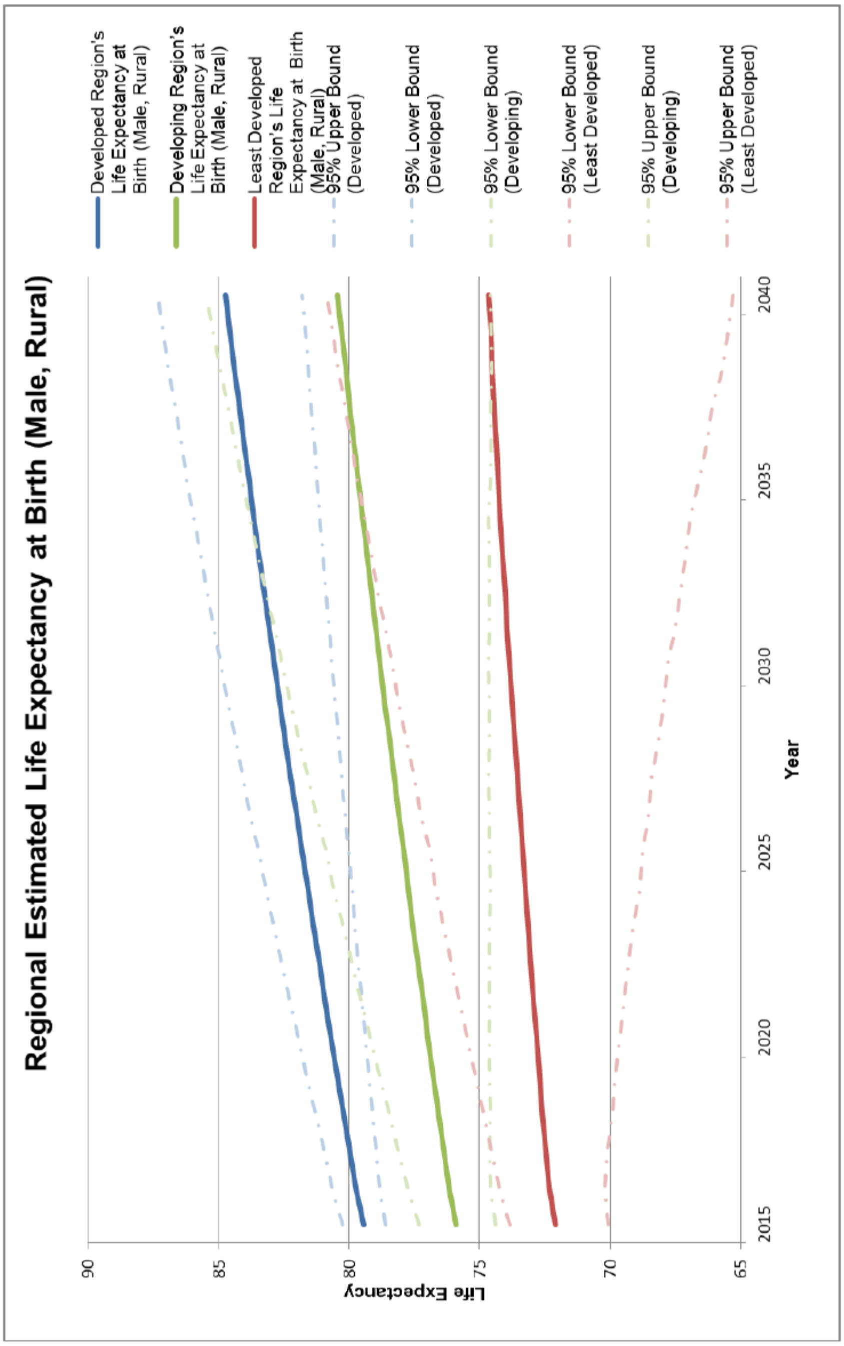
Source: Calculated using the method described in Chapter 3

Figure 5.5: Regional estimated life expectancy at birth from 2015 to 2040 (Male, Urban).



Source: Calculated using the method described in Chapter 3

Figure 5.6: Regional estimated life expectancy at birth from 2015 to 2040 (Male, Rural).



Source: Calculated using the method described in Chapter 3

This extension might provide a more detailed explanation of the regional differences in life expectancy and their causes. Nevertheless, this model provides accurate results only under enough planes of data. Therefore, this model was not considered here due to the very limited availability of relevant Chinese data.

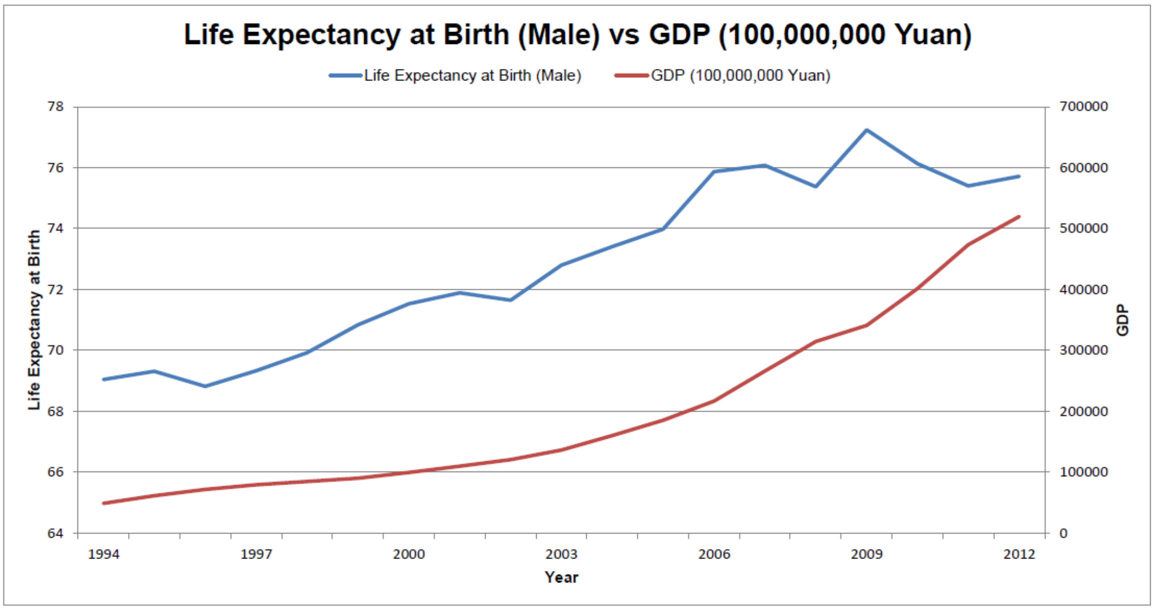
Overall, the extended Lee-Carter model developed by Li et al. (2004) successfully addressed the limited availability of mortality data in China. The prediction by using the extended Lee-Carter model shows a stable difference between the urban and rural population over next twenty-five years. In addition, the study displays a relative constant future regional difference in life expectancy at birth and age 65 for total population, through the simulation using the extended Lee-Carter model. However, the regional difference for urban population is found to narrow in the future. By comparison the regional difference in rural area is expected to enlarge in next twenty-five years. These implies a different treatment to urban and rural pension policies might be necessary as future rural pension benefits should be aligned with local life expectancies to enhance sustainability. Nevertheless, further investigations are required to provide more reliable explanations.

5.2 Socioeconomic Variables

One limitation of the standard Lee-Carter model is its failure to take account of important socioeconomic variables which may have significant impacts on mortality improvements, such as GDP per capita and public healthcare expenditure. For example, Figure 5.7 shows a positive correlation between China's GDP and life expectancy at birth for total male population. Hanewald (2011) also finds that real GDP growth rate and the unemployment rate have a substantial influence on mortality improvements. In this section, the relationship between the mortality index k_t and changes in socioeconomic variables, such as GDP per capita, the unemployment rate and a proxy for medical care (the number of doctors per 10,000 people), for males and females in different age groups will be investigated.

As there is very limited mortality data available at the regional level, the study here has focused on national mortality data. Mortality data by gender and age group at the national level from 1994 to 2012 were collected from China's National Knowledge Infrastructure. Annual mortality data is available at national level.

Figure 5.7: Life expectancy at birth for total male population in China versus China's GDP (100,000,000 Yuan) from 1994 to 2012.



Source: Calculated by authors

The methodology used is as follows. First, the mortality index k_t, was obtained from equations 3.13 and 3.14 as presented in Section 3.2.2 (see Figures 5.8 and 5.9). Next, the stationary test (Phillips-Perron test) was applied to the fitted mortality index k_t and a set of selected socioeconomic variables, including GDP, GDP per capita, income per capita, the unemployment rate, the number of healthcare assistants per 10,000 people, the number of doctors per 10,000 people, public healthcare expenditure and the number of residents who completed tertiary education per 10,000 people. These variables were chosen because GDP, GDP per capita, income per capita and the unemployment rate are most commonly used economic variables, and the number of healthcare assistants per 10,000 people, the number of doctors per 10,000 people, public healthcare expenditure and the number of residents are good social indicators. The data was sourced from the National Bureau of Statistics of China.

This stationary test was not only applied to untransformed indicators but also to the indicators after taking differences or logarithms (see Table 5.1). Then the pairwise Pearson correlation coefficients and a t-test with a significant level of 5% were evaluated between the stationary form of the mortality index and the socioeconomic factors as discussed in Section 3.2.3 (see Table 5.2). According to Table 5.2, all of the selected socioeconomic indicators failed the statistical test implying the null hypothesis of zero estimated correlation was not able to be rejected. On the other hand, the growth rate of GDP, GDP per capita and income per capita and the change of number of doctor per 10,000 people were found to have a relatively strong positive correlation with the change of $k_{t,male}$ and $k_{t,female}$. This implies a negative impact of economic and social development on mortality improvements, which is consistent with findings in Edwards (2008), Gerdtham and Ruhm (2006), Tapia Granados (2005) and Tapia Granados (2008), in their studies of other countries, such as the United States and Japan.

Regression analysis was then applied to the stationary form of the mortality index and those socioeconomic factors with relatively strong correlation with it. The preferred models were selected according to fitting errors, R^2, tractability and predictability of the models. The selected models for the gender specific mortality index are as follows:

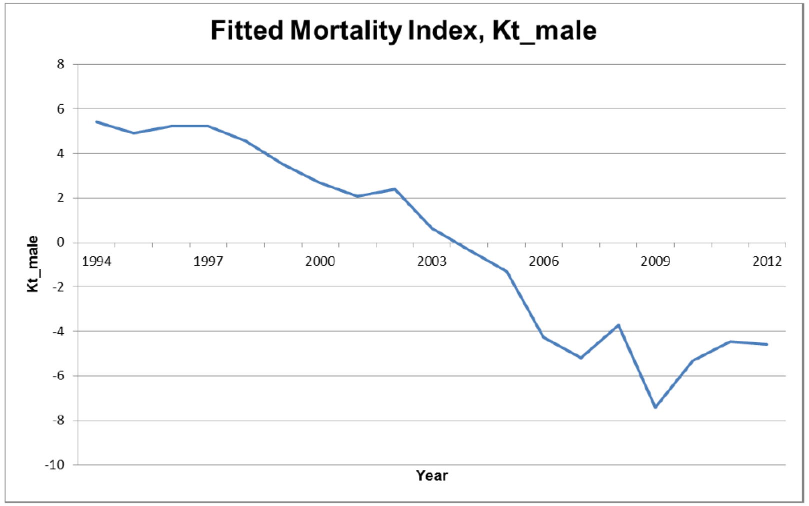
$$\Delta k_{t,male} = \underset{(0.98601)}{-1.50265} + \underset{(7.04117)}{7.16987} \times \Delta ln(GDP_t) + \varepsilon_t \qquad R^2 - 0.0609 \tag{5.3}$$

$$\Delta k_{t,female} = \underset{(0.87286)}{-1.47319} + \underset{(6.23313)}{5.12425} \times \Delta ln(GDP_t) + \varepsilon_t \qquad R^2 - 0.0405 \tag{5.4}$$

where Δ represents first differences. ε_t is the error term following the standard normal distribution. These two models were selected because of their small fitting errors and better predictability compared to the other models. the growth rate of GDP was incorporated because of its significant influence on the change of $k_{t,male}$ and $k_{t,female}$. In addition, there are better and more accurate projections of the growth rate of GDP available. Hence, it has a better predictability compared to the other variables.

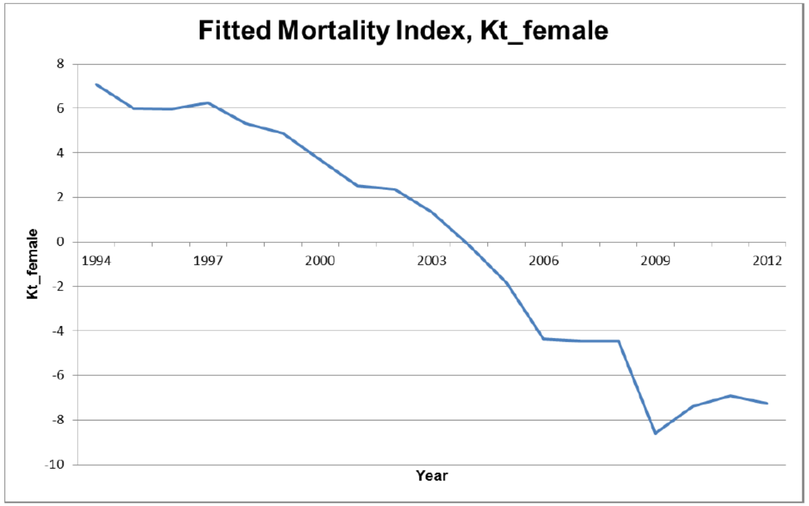
Under the assumption that the growth rate of GDP has the same impacts on local mortality improvements in each region, the selected models were used to forecast the future mortality rate and life expectancy in classified regions based on fitted mortality indexes, k_t, and estimated future socioeconomic indicators. 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 as reported in Table 5.3. Then future mortality improvements were estimated and sensitivity analysis was conducted.

Figure 5.8: Fitted mortality index $k_{t,male}$.



Source: Calculated using the method described in Chapter 3

Figure 5.9: Fitted mortality index $k_{t,female}$.



Source: Calculated using the method described in Chapter 3

Table 5.1: Stationary test applied to fitted mortality index k_t and selected socioeconomic variables.

Variable	Tau	Pr < Tau
$\Delta k_{t,male}$	-4.26	0.0002
$\Delta k_{t,female}$	-3.4	0.0018
$\Delta Log(GDP)_t$	-2.2	0.0296
$\Delta Log(GDP\ per\ capita)_t$	-2.19	0.0303
$\Delta Log(Income\ per\ capita)_t$	-2.22	0.0284
$\Delta(Unemployment)_t$	-3.05	0.0042
$\Delta Log(No.of\ HC\ Assit\ per\ 10000\ ppl)_t$	-1.88	0.0575
$\Delta(No.of\ Doctor\ per\ 10000\ ppl)_t$	-4.9	<.0001
$\Delta\Delta(Government\ Health\ Expenditure)_t$	-2.37	0.0207
$\Delta(No.of\ Residents\ Completed\ Teritary\ Education\ per\ 10000\ ppl)_t$	-3.33	0.0022

Table 5.2: Pearson correlation coefficients between the stationary form of the mortality index and the socioeconomic factors and corresponding p-values of the t-test.

Pearson Correlation Coefficients & t-test								
	D_Log_GDP	D_Long_GDP per Capita	D_Log_Income per Capita	D_Unemployment rate (%)	D_Log_Number of healthcare assis	D_Number of doctor per 10000 people	DD_Log_Government Health Expenditure	D_Log_Number of people who completed tertiary education per 10000 people
D_k_male	0.2467	0.24365	0.1881	-0.00304	-0.07141	0.20778	-0.08058	-0.15171
	(0.3237)	(0.3299)	(0.4548)	(0.9905)	(0.7783)	(0.408)	(0.7506)	(0.5479)
D_k_female	0.20132	0.19536	0.12476	-0.06496	-0.11092	0.16461	0.07863	-0.11867
	(0.4231)	(0.4372)	(0.6218)	(0.7979)	(0.6613)	(0.5139)	(0.7565)	(0.6391)

1. D represents for taking difference once. DD represents for taking difference twice.
2. The rows corresponding to D_k_male and D_k_female show the Pearson correlation coefficents.
3. The numbers in the bracket show the p-value of the t-test.

Source: Calculated using the method described in Chapter 3

Table 5.3: Estimated GDP Growth Rate for Different Regions from 2015 to 2040.

	Year	Optimistic Scenario	Base Scenario	Pessimistic Scenario
Developed	2015-2025	8.8%	7.3%	5.8%
	2025-2035	6.8%	5.8%	3.8%
	2035-2040	4.8%	3.8%	2.3%
Developing	2015-2025	10.5%	9.8%	8.3%
	2025-2035	8.7%	7.9%	6.8%
	2035-2040	7.5%	6.2%	5.2%
Least Developed	2015-2025	13.7%	12.2%	11.3%
	2025-2035	11.7%	10.7%	9.8%
	2035-2040	9.7%	8.7%	7.8%

Source: Calculated using the method described in Chapter 3

This analysis finds that predicted life expectancies are higher than using the extended Lee-Carter model as described in the previous section (see Figures 5.10, 5.11 and 5.12). The simulations also displays that the impacts of different scenarios on future mortality improvements are relatively small in each region. However, life expectancy at birth in the least developed regions is found to catch up with life expectancy at birth in developing region by 2040 (see Figure 5.13). This shows that the least developed regions experience a strong influence on mortality improvements from economic development. This finding confirms with the original expectation of the study which predicted the life expectancies in different regions would converge because the marginal impact of socioeconomic factors was expected to decrease as the level of economic development increased.

However, further investigations should be taken in the future. One possible extension is to take account of lagged effects of socioeconomic variables because current and future mortality improvements in each region might be accumulated effects of past development in those areas. The other possible extension is to incorporate the impact of internal labour migration. The local mortality data is currently collected based on the population living in the region over one year, whereas China has a very high level of labour mobility, which might influence mortality improvements elsewhere in China.

Figure 5.10: Estimated Life Expectancy at Birth from 2015 to 2040 (Developed Region, Male, Total).

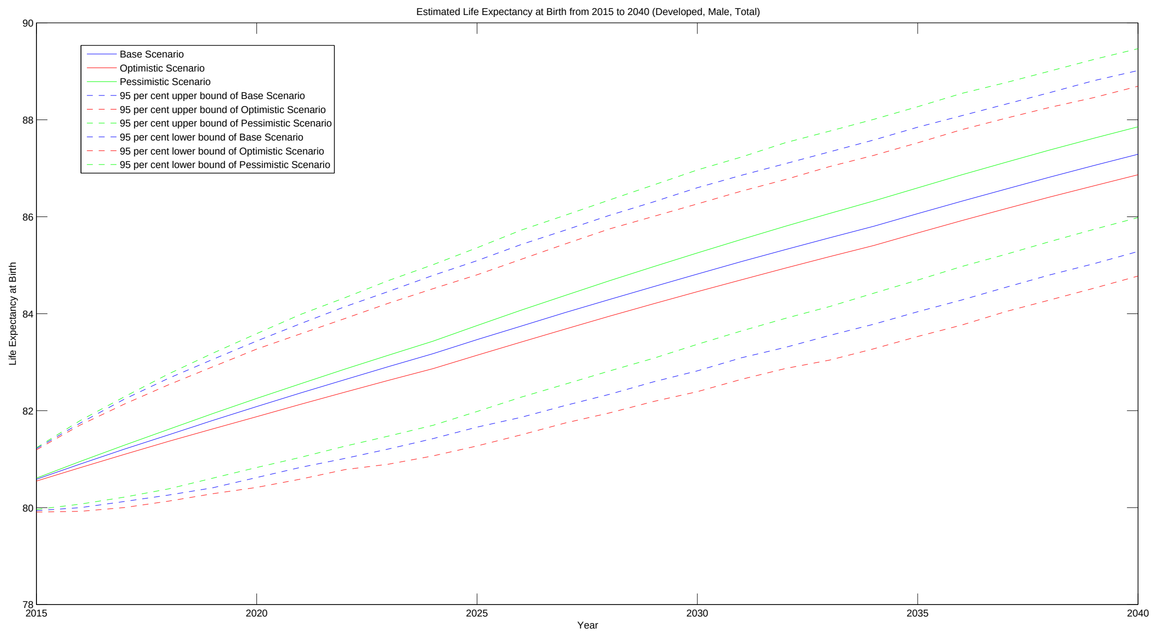
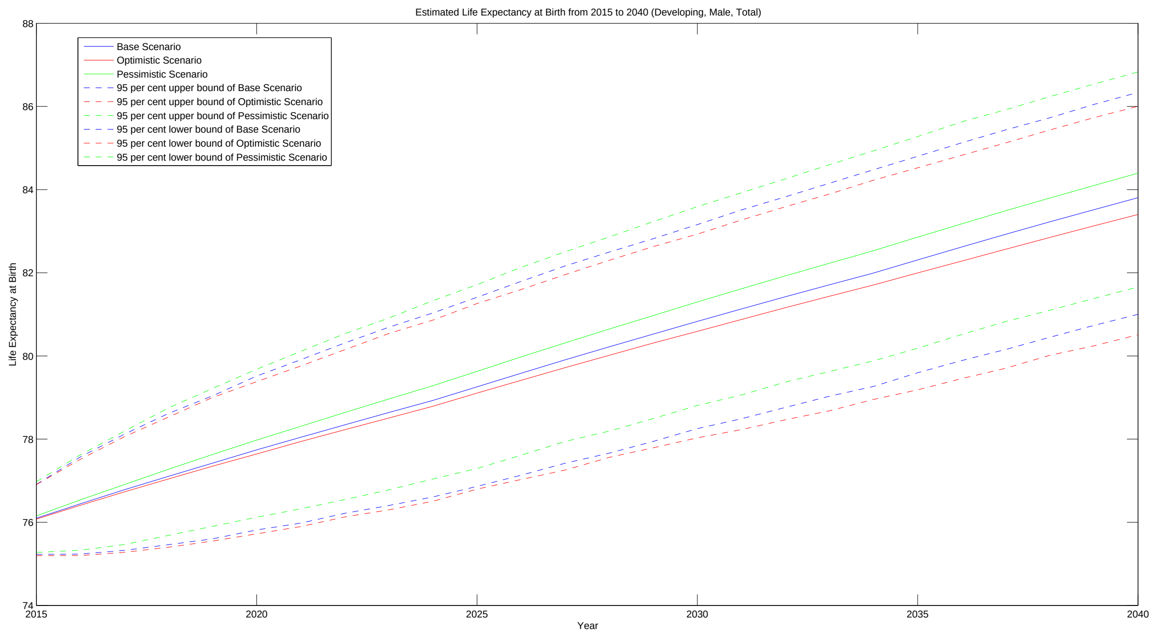


Figure 5.11: Estimated Life Expectancy at Birth from 2015 to 2040 (Developing Region, Male, Total).



Source: Calculated using the method described in Chapter 3

Figure 5.12: Estimated Life Expectancy at Birth from 2015 to 2040 (Least Developed Region, Male, Total).

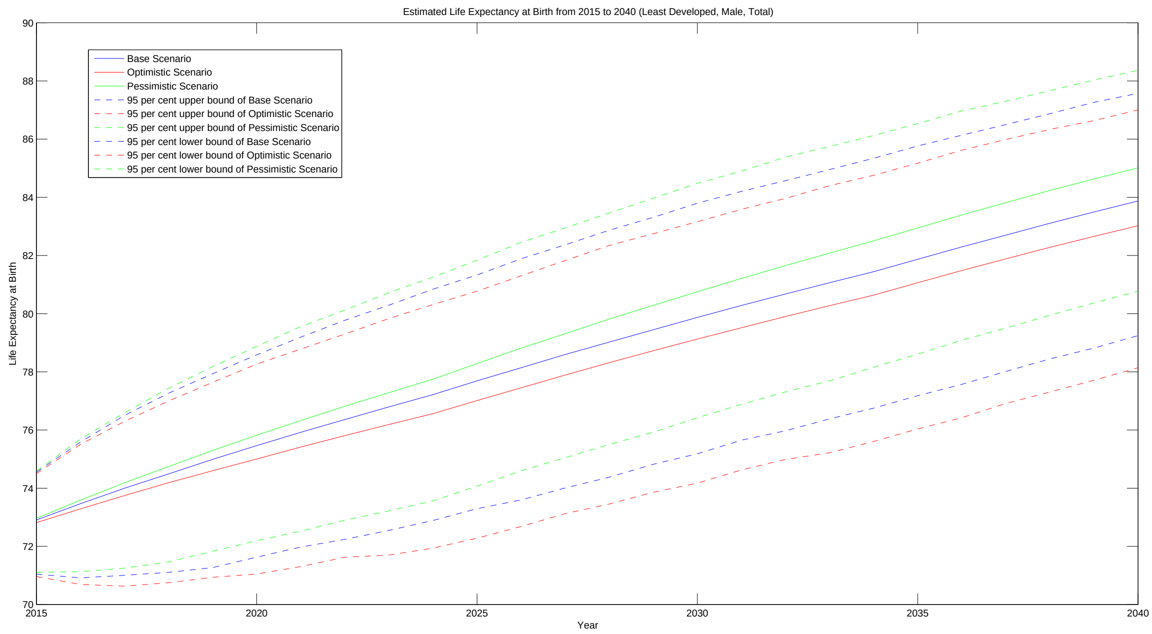
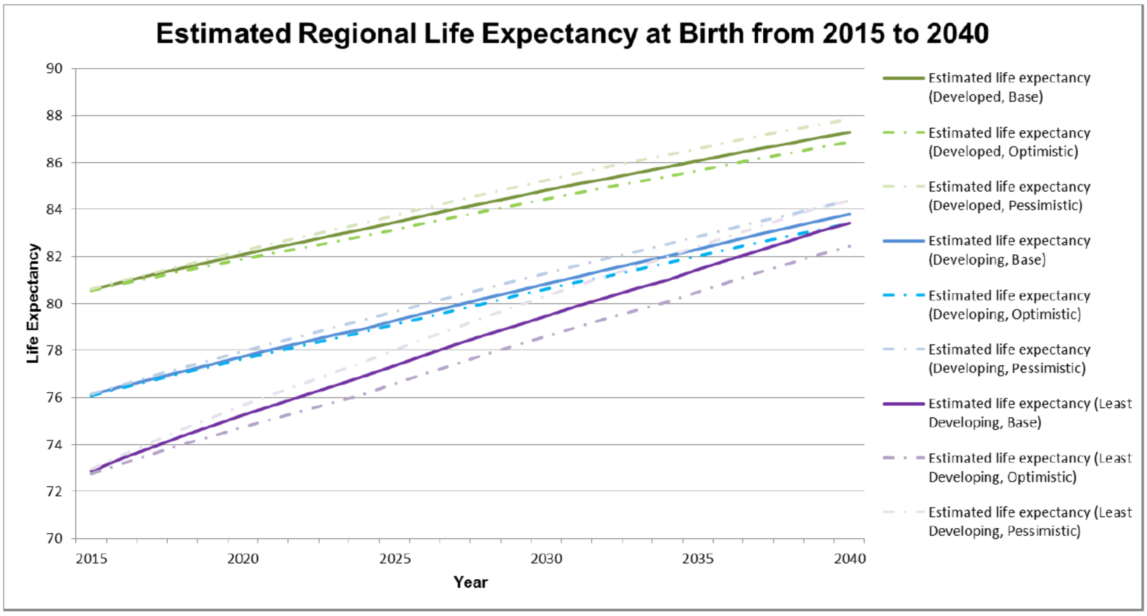


Figure 5.13: Estimated regional life expectancy at Birth from 2015 to 2040.



Source: Calculated using the method described in Chapter 3

5.3 Discussion and Conclusion

In this chapter, the prediction by using the extended Lee-Carter model shows a stable difference between the urban and rural population. A potential factor behind this phenomenon was the increasing urbanization across China. The Chinese government has upgraded most of its towns and villages into cities which are expected to significantly lift living standards in rural areas. This has caused significant mortality improvements for the rural population. On the other hand, for rural residents who enjoyed the fastest development, their hometowns might be upgraded and reclassified as cities. As a result, they were recounted as urban population which resulted in a stable difference in life expectancies between urban and rural populations.

In addition, the study displays a relative constant future regional difference in life expectancy at birth and age 65 for total population, through the simulation using the extended Lee-Carter model. However, the regional difference for urban population is found to narrow in the future. By comparison the regional difference in rural area is expected to enlarge in next twenty-five years. These imply a different treatment of urban and rural pension policies will be necessary as future rural pension benefits should be aligned with local life expectancies to enhance sustainability. Nevertheless, further investigations are required to provide more reliable explanations.

Finally, the study shows that the growth rate of GDP, GDP per capita and income per capita and change of number of doctor per 10,000 people had relatively strong correlations with the change of $k_{t,male}$ and $k_{t,female}$. After the incorporation of the growth rate of GDP, it is found that different scenarios have relatively small impacts on the differences in future mortality improvements of each region. Nevertheless, the study also reveals that the least developed regions experience a strong influence on mortality improvements from economic development. This finding confirms with the original expectation of the study which predicted the life expectancies in different regions would converge because the marginal impact of socioeconomic factors was expected to decrease as the level of economic development increased. Further investigations should be taken to consider lagged effects of socioeconomic variables and the impact of internal labour migration.

CHAPTER 6

CONCLUSION

Due to the pressure of the demographic challenge and the shortcomings of the current pension system, the Chinese government is currently undertaking a range of reforms to its pension system in order to address the equity, adequacy & sustainability issues. This requires accurate pension liability estimation, and hence, accurate life expectancy forecasts at both national and regional levels. Nevertheless, due to the incapacity of the Lee-Carter model to deal with limited data, there have not been many studies on China's future mortality trends because of the lack of high quality, long term data. Furthermore, there are vast regional differences in life expectancy in China which raises questions about the impact of socioeconomic factors on mortality improvements. Unfortunately, the standard Lee-Carter model fails to capture those external factors which substantially influence human's mortality behaviour as well. Hence, the lack of prior research in this area has become a significant obstacle to policy formulation.

This research addresses these literature gaps by extending the Lee-Carter model in two ways. First, the extended time series methods developed by Li et al. (2004) was used to overcome limited availability of mortality data in China. Second, socioeconomic variables were incorporated in the original Lee-Carter model to explore external factors which may have vital impacts on mortality (as undertaken in Hanewald (2011)). The developed models were then used to estimate future mortality improvements at both the national and regional levels. In the study, 800 national and regional life tables (by province and for urban and rural populations) were constructed to enable the analysis, based on the unique dataset collected manually.

Using the unique constructed, the trends of national and regional mortality improvements were investigated. Interesting findings include.

- The study shows that sex specific life expectancy converged at older ages. This is consistent with demographic experience in Australia and United Kingdom. This implied that there are no significant sex specific differences in mortality rates for the current cohort at older ages. As a result, there is no immediate need for the standard pension policy to focus on gender difference at older ages. However, as only period life tables are available (rather than cohort life tables) in the study, it is unclear if this trend will continue to be the case for the future generations or is only driven by cohort effects due to the mortality experience of the old age groups. These age groups experienced particular types of social events that are unlikely to be experienced by future generations, such as a series of wars from the 1920s to the 1980s, the Cultural Revolution and China's three years of natural disasters. However, more evidence is required to support this argument.

- The study finds that the rural and urban differences in life expectancy at birth and at age 65 have increased between 1990 and 2010. This finding is consistent with studies in France, England and United States by Woods (2003), Hartley (2004) and Singh and Siahpush (2006). For males, at the national level, residents of urban areas have tended to experience higher life expectancy at birth (by around 4 years) than the population as a whole. People living in rural areas have experienced relatively lower life expectancy at birth than the overall population (ie, about 1 year lower). However, the differences in life expectancy at age 65 is smaller than for life expectancy at birth. The difference in life expectancy for males at birth between rural and urban areas has continued to increase over the past 20 years. Similar differences are evident for females and at the regional level.

- The study illustrates that there was no significant difference in the urban populations of the various administrative regions, whereas the differences in the rural populations was quite marked. Therefore, the differences in life expectancy at birth and at age 65 for populations from different administrative regions appear to be mainly driven by the difference in life expectancy of the rural population in each region. This indicates that the rural population should be the target of any initiatives to reduce the significant regional differences in life expectancy. Moreover, this finding fills the current literature gap as to the best of our knowledge, there is no study yet which explains disparities of which groups of population resulting in the regional differences in life expectancy. In addition, this raises questions about the impact of China's urbanization on the trend of regional differences in mortality improvements over coming decades. Furthermore, as current pension benefits are predetermined by a formula of which a benefit multiple is determined actuarially based on expected future mortality rate, pension policy reform should consider the calculation of benefits according to locally projected mortality rate.

- The study also found that differences in mortality rates are concentrated in the young and middle age groups, which helps explain the smaller differences in life expectancy at older age groups.

Using the constructed life tables, the Lee-Carter model was extended to address the limited availability of relevant data and used to estimate future mortality improvements. Simulation using the extended Lee-Carter model shows a stable difference between the urban and rural population. A potential factor behind this phenomenon is the increasing urbanization across China. The Chinese government has upgraded most of its towns and villages into cities which are expected to significantly lift living standards in rural areas. This has caused significant mortality improvements for the rural population. On the other hand, for rural residents who enjoyed the fastest development, their hometowns might be upgraded and reclassified as cities. As a result, they were recounted as urban population which resulted in a stable difference in life expectancies between urban and rural populations. In addition, the study displays a relative constant future regional difference in life expectancy at birth and age 65 for total population, through the simulation using the extended Lee-Carter model. However, the regional difference for urban population is found to narrow in the future. By comparison the regional difference in rural area is expected to enlarge in next twenty-five years. These suggest a different treatment of urban and rural pension policies will be necessary as future rural pension benefits should be aligned with local life expectancies to enhance sustainability. Nevertheless, further investigations are required to provide more reliable explanations.

The analysis was then extended further to investigate the impact of socioeconomic factors on the regional differences in life expectancy. The growth rate of GDP, GDP per capita, income per capita and the change in the number of doctors per 10,000 people were chosen as proxies of economic and social development. These variables all show relatively strong correlations with the change of the index of $k_{t,male}$ and $k_{t,female}$. Sensitivity analysis conducted through the incorporation of the growth rate of GDP into the standard Lee-Carter model shows that life expectancy at birth in the least developed regions catches up with life expectancy at birth in the developing regions. This indicates that the least developed regions experience a strong influence on mortality improvements from economic development. This finding is in line with the original expectation of the study which predicted the life expectancies in different regions would converge because the marginal impact of socioeconomic factors was expected to decrease as the level of economic development increased. Further investigations should be taken to consider lagged effects of socioeconomic variables and the impact of internal labour migration.

The key contributions of this thesis are as follows.

- 800 national and regional life tables (by province and for urban and rural populations) were constructed based on a unique hand collected mortality dataset. To the best of our knowledge, this is the first dataset at both national and regional level, by province and for urban and rural populations. This extends investigations of life expectancy beyond just the national level.
- An extended Lee-Carte model has succeeded to address the limited availability of data in China. Its predictions of future regional mortality improvements have brought a better understanding of China's regional differences in life expectancy, which is crucial to better estimate the future sustainability of a pension system built around regional pools.
- Moreover, the investigation and incorporation of socioeconomic variables into the standard Lee-Carter model clearly shows how socioeconomic factors (proxied by the growth rate of GDP) influences the regional differences in life expectancy for total,

urban and rural populations.

Overall, this study provides a more accurate forecast of future mortality improvements by regions, which also enables more accurate estimates of the fiscal sustainability of China's pension system and provides a framework for future reform proposals regarding pensions and the implications of an ageing population more generally.

One limitation of the study is the limited availability of mortality data for China. As mentioned before, the Lee-Carter model needs at least the past thirty years' mortality data to provide a good fit. While Li et al. (2004) is applied to addresses this limited data issue, more data collected in the future will help provide more reliable results. However, this study represents a substantial contribution to existing research and will provide a platform for future work as more disaggregated data for China becomes available.

APPENDIX A

APPENDIX

This Appendix provides demonstrated examples of constructed national and regional life tables (by province and for urban and rural populations). A complete set of life tables was derived. [1]

[1] The derived life tables can be obtained by emailing jichen.li@unsw.edu.au.

Table A.1: Life table for male, Beijing, 2010 — Total.

Year	Regions	Age Group	nMx	nqx	lx	ndx	nLx	Tx	ex
2010	Beijing	0 Year	0.00134	0.00134	100000	134	99873	8035703	80.36
2010	Beijing	1-4 Years	0.00027	0.00107	99866	106	399215	7935830	79.46
2010	Beijing	5-9 Years	0.00011	0.00053	99760	52	498669	7536615	75.55
2010	Beijing	10-14 Years	0.00019	0.00097	99708	97	498297	7037945	70.59
2010	Beijing	15-19 Years	0.00022	0.00111	99611	110	497780	6539648	65.65
2010	Beijing	20-24 Years	0.00022	0.00109	99501	109	497233	6041868	60.72
2010	Beijing	25-29 Years	0.00022	0.0011	99392	109	496689	5544635	55.79
2010	Beijing	30-34 Years	3.00E-04	0.00148	99283	147	496047	5047946	50.84
2010	Beijing	35-39 Years	6.00E-04	0.00297	99136	295	494942	4551899	45.92
2010	Beijing	40-44 Years	0.00111	0.00554	98841	548	492836	4056957	41.05
2010	Beijing	45-49 Years	0.00201	0.01002	98293	985	489006	3564121	36.26
2010	Beijing	50-54 Years	0.00354	0.01755	97309	1708	482274	3075115	31.6
2010	Beijing	55-59 Years	0.00495	0.02445	95601	2337	472161	2592841	27.12
2010	Beijing	60-64 Years	0.00817	0.04006	93264	3736	456979	2120680	22.74
2010	Beijing	65-69 Years	0.01425	0.06878	89528	6158	432245	1663702	18.58
2010	Beijing	70-74 Years	0.0256	0.12031	83370	10030	391775	1231457	14.77
2010	Beijing	75-79 Years	0.04366	0.19682	73340	14435	330614	839682	11.45
2010	Beijing	80-84 Years	0.0725	0.30688	58905	18077	249335	509068	8.64
2010	Beijing	85-89 Years	0.11832	0.45655	40829	18640	157542	259733	6.36
2010	Beijing	90-94 Years	0.19413	0.65349	22188	14500	74692	102191	4.61
2010	Beijing	95-99 Years	0.24745	0.76439	7688	5877	23750	27500	3.58
2010	Beijing	100 Years and Over	0.48312	1	1811	1811	3750	3750	2.07

Source: Calculated using the method described in Chapter 3

Table A.2: Life table for male, Beijing, 2010 — Urban.

Year	Regions	Age Group	nMx	nqx	lx	ndx	nLx	Tx	ex
2010	Beijing	0 Year	0.00116	0.00115	100000	115	99890	8173617	81.74
2010	Beijing	1-4 Years	2.00E-04	0.00081	99885	80	399349	8073726	80.83
2010	Beijing	5-9 Years	1.00E-04	0.00048	99804	48	498902	7674377	76.89
2010	Beijing	10-14 Years	0.00013	0.00065	99756	65	498619	7175475	71.93
2010	Beijing	15-19 Years	0.00017	0.00085	99691	84	498245	6676856	66.98
2010	Beijing	20-24 Years	0.00016	8.00E-04	99607	80	497835	6178611	62.03
2010	Beijing	25-29 Years	0.00017	0.00085	99527	85	497424	5680776	57.08
2010	Beijing	30-34 Years	2.20E-04	0.00112	99442	112	496932	5183352	52.12
2010	Beijing	35-39 Years	4.40E-04	0.00221	99331	219	496105	4686420	47.18
2010	Beijing	40-44 Years	9.00E-04	0.0045	99111	446	494443	4190315	42.28
2010	Beijing	45-49 Years	0.00176	0.00877	98666	865	491167	3695872	37.46
2010	Beijing	50-54 Years	0.0032	0.01587	97801	1552	485125	3204705	32.77
2010	Beijing	55-59 Years	0.00423	0.02092	96249	2014	476211	2719580	28.26
2010	Beijing	60-64 Years	0.00692	0.03401	94235	3205	463163	2243368	23.81
2010	Beijing	65-69 Years	0.01199	0.05822	91030	5300	441901	1780205	19.56
2010	Beijing	70-74 Years	0.02184	0.10355	85730	8877	406458	1338305	15.61
2010	Beijing	75-79 Years	0.03824	0.17452	76853	13412	350734	931847	12.13
2010	Beijing	80-84 Years	0.06574	0.28231	63441	17910	272428	581113	9.16
2010	Beijing	85-89 Years	0.10468	0.41484	45531	18888	180433	308685	6.78
2010	Beijing	90-94 Years	0.18081	0.62262	26643	16588	91742	128252	4.81
2010	Beijing	95-99 Years	0.23978	0.74958	10054	7537	31431	36510	3.63
2010	Beijing	100 Years and Over	0.49573	1	2518	2518	5079	5079	2.02

Source: Calculated using the method described in Chapter 3

Table A.3: Life table for male, Beijing, 2010 — Rural.

Year	Regions	Age Group	nMx	nqx	lx	ndx	nLx	Tx	ex
2010	Beijing	0 Year	0.00201	0.002	100000	200	99810	7632243	76.32
2010	Beijing	1-4 Years	0.00052	0.00207	99800	206	398713	7532433	75.48
2010	Beijing	5-9 Years	0.00014	0.00069	99593	69	497794	7133720	71.63
2010	Beijing	10-14 Years	0.00039	0.00195	99524	194	497137	6635926	66.68
2010	Beijing	15-19 Years	0.00041	0.00204	99331	203	496146	6138789	61.8
2010	Beijing	20-24 Years	0.00046	0.00228	99128	226	495073	5642643	56.92
2010	Beijing	25-29 Years	0.00046	0.00228	98901	226	493943	5147569	52.05
2010	Beijing	30-34 Years	0.00065	0.00323	98676	319	492581	4653626	47.16
2010	Beijing	35-39 Years	0.0012	0.00598	98357	589	490312	4161045	42.31
2010	Beijing	40-44 Years	0.00179	0.00891	97768	871	486664	3670733	37.55
2010	Beijing	45-49 Years	0.00284	0.0141	96897	1366	481072	3184069	32.86
2010	Beijing	50-54 Years	0.00465	0.02301	95531	2198	472162	2702997	28.29
2010	Beijing	55-59 Years	0.0072	0.03536	93334	3300	458417	2230835	23.9
2010	Beijing	60-64 Years	0.01182	0.05741	90033	5169	437246	1772417	19.69
2010	Beijing	65-69 Years	0.0208	0.09885	84865	8389	403353	1335172	15.73
2010	Beijing	70-74 Years	0.03895	0.17745	76476	13571	348454	931819	12.18
2010	Beijing	75-79 Years	0.06521	0.28034	62906	17635	270440	583365	9.27
2010	Beijing	80-84 Years	0.10125	0.40399	45270	18289	180630	312925	6.91
2010	Beijing	85-89 Years	0.17751	0.61475	26982	16587	93441	132294	4.9
2010	Beijing	90-94 Years	0.25935	0.78669	10395	8178	31530	38853	3.74
2010	Beijing	95-99 Years	0.29536	0.84951	2217	1884	6377	7323	3.3
2010	Beijing	100 Years and Over	0.35294	1	334	334	945	945	2.83

Source: Calculated using the method described in Chapter 3

BIBLIOGRAPHY

Abeysinghe, T. and Rajaguru, G. (2004), 'Quarterly real GDP estimates for China and ASEAN with a forecast evaluation', *Journal of Forecasting* **23**(6), 431–447.

Alexander, M., Harding, M. and Lamarche, C. (2011), 'Quantifying the impact of economic crises on infant mortality in advanced economies', *Applied Economics* **43**(24), 3313–3323.

Alho, J. (2008), 'Aggregation across countries in stochastic population forecasts', *International Journal of Forecasting* **24**(3), 343–353.

Alho, J. M. (1990), 'Stochastic methods in population forecasting', *International Journal of Forecasting* **6**(4), 521–530.

Alonso, J., Caballero, M. A., Hui, L., Llanes, M. C., Tuesta, D., Hu, Y.-W. and Cao, Y. (2011), 'Potential outcomes of private pension developments in China', *BBVA Bank, Economic Research Department* .

Banister, J. and Hill, K. (2004), 'Mortality in China 1964–2000', *Population Studies* **58**(1), 55–75.

Bateman, H. and Liu, K. (2013), 'Pension reform in China: Racing against the demographic clock', *UNSW Australian School of Business Research Paper* (2013ACTL22).

Booth, H. (2006), 'Demographic forecasting: 1980 to 2005 in review', *International Journal of Forecasting* **22**(3), 547–581.

Booth, H., Hyndman, R. J., Tickle, L., De Jong, P. et al. (2006), Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions, Technical report, Monash University, Department of Econometrics and Business Statistics.

Booth, H., Maindonald, J. and Smith, L. (2002), 'Applying Lee-Carter under conditions of variable mortality decline', *Population Studies* **56**(3), 325–336.

Booth, H. and Tickle, L. (2008), 'Mortality modelling and forecasting: A review of methods', *Annals of Actuarial Science* **3**(1-2), 3–43.

Brandt, L. and Rawski, T. G. (2008), *China's great economic transformation*, Cambridge University Press.

Brouhns, Natacha D, M. V. J. K. et al. (2002*a*), 'Measuring the longevity risk in mortality projections', *Bulletin of the Swiss Association of Actuaries* **2**, 105–130.

Brouhns, Natacha, D. M. V. J. K. et al. (2002*b*), 'A poisson log-bilinear regression approach to the construction of projected lifetables', *Insurance: Mathematics and Economics* **31**(3), 373–393.

Cai, Y. (2005), 'National, provincial, prefectural and county life tables for China based on the 2000 Census', *CSDE Working Paper 05-03* .

Cairns, A. J., Blake, D. and Dowd, K. (2008), 'Modelling and management of mortality risk: a review', *Scandinavian Actuarial Journal* **2008**(2-3), 79–113.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D. and Khalaf-Allah, M. (2011), 'Bayesian stochastic mortality modelling for two populations', *Astin Bulletin* **41**(01), 29–59.

Cameron, C. A. and Windmeijer, F. A. (1997), 'An R-squared measure of goodness of fit for some common nonlinear regression models', *Journal of Econometrics* **77**(2), 329–342.

Davis, L. A. (2008), *Natural disasters*, Infobase Publishing.

De Jong, P. and Tickle, L. (2006), 'Extending Lee-Carter mortality forecasting', *Mathematical Population Studies* **13**(1), 1–18.

Demurger, S. (2001), 'Infrastructure development and economic growth: an explanation for regional disparities in china?', *Journal of Comparative economics* **29**(1), 95–117.

Dickson, D. C., Hardy, M., Hardy, M. R. and Waters, H. R. (2013), *Actuarial mathematics for life contingent risks*, Cambridge University Press.

Dorfman, M. C., Holzmann, R., O'Keefe, P., Wang, D., Sin, Y. and Hinz, R. (2013), *China's pension system: A vision*, World Bank Publications.

Edwards, R. (2008), 'Who is hurt by procyclical mortality?', *Social Science & Medicine* **67**(12), 2051–2058.

Gerdtham, U.-G. and Ruhm, C. J. (2006), 'Deaths rise in good economic times: evidence from the OECD', *Economics & Human Biology* **4**(3), 298–316.

Granados, J. A. T. (2005), 'Increasing mortality during the expansions of the US economy, 1900–1996', *International Journal of Epidemiology* **34**(6), 1194–1202.

Hainaut, D. (2012), 'Multidimensional Lee-Carter model with switching mortality processes', *Insurance: Mathematics and Economics* **50**(2), 236–246.

Hanewald, K. (2011), 'Explaining mortality dynamics: The role of macroeconomic fluctuations and cause of death trends', *North American Actuarial Journal* **15**(2), 290–314.

Hanewald, K., Post, T. and Gründl, H. (2011), 'Stochastic mortality, macroeconomic risks and life insurer solvency', *The Geneva Papers on Risk and Insurance-Issues and Practice* **36**(3), 458–475.

Härdle, W. K. and Myšičková, A. (2009), 'Stochastic population forecast for germany and its consequence for the german pension system', *SFB 649 Discussion Paper* .

Hartley, D. (2004), 'Rural health disparities, population health, and rural culture', *American Journal of Public Health* **94**(10), 1675.

Hualou, L. and Jian, Z. (2010), 'Farmland destroyed by natural hazards in china: Spatio-temporal pattern and integrated prevention and treatment system', *Disaster Advances* **3**(4), 592–597.

Hyndman, R. J., Ullah, S. et al. (2007), 'Robust forecasting of mortality and fertility rates: a functional data approach', *Computational Statistics & Data Analysis* **51**(10), 4942–4956.

Jackson, R., Nakashima, K., Howe, N. and Zhou, J. (2009), 'China's long march to retirement reform', *Hallym International Journal of Aging* **2**(2), 135–147.

Jiang, Q., Song, W. and Sánchez-Barricarte, J. J. (2013), 'Forecasting China's mortality', *Population Review* **52**(2).

Keilman, N. (2001), 'Demography: uncertain population forecasts', *Nature* **412**(6846), 490–491.

Keyfitz, N. (1944), 'Experiments in the projection of mortality', *UALibraries Site Administrator Test Journal 2* **18**(2), 1–17.

Lee, R. (2000), 'The lee-carter method for forecasting mortality, with various extensions and applications', *North American Actuarial Journal* **4**(1), 80–91.

Lee, R. D. (1998), 'Probabilistic approaches to population forecasting', *Population and Development Review* pp. 156–190.

Lee, R. D. and Carter, L. R. (1992), 'Modeling and forecasting US mortality', *Journal of the American Statistical Association* **87**(419), 659–671.

Lee, R. D. and Nault, F. (1993), Modeling and forecasting provincial mortality in canada, *in* 'World Congress of the International Union for Scientific Study of Population, Montreal'.

Lee, R. and Miller, T. (2001), 'Evaluating the performance of the Lee-Carter method for forecasting mortality', *Demography* **38**(4), 537–549.

Li, J. S.-H., Chan, W.-S. and Cheung, S.-H. (2011), 'Structural changes in the Lee-Carter mortality indexes: detection and implications', *North American Actuarial Journal* **15**(1), 13–31.

Li, N. and Lee, R. (2005), 'Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method', *Demography* **42**(3), 575–594.

Li, N., Lee, R. and Gerland, P. (2013), 'Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections', *Demography* **50**(6), 2037–2051.

Li, N., Lee, R. and Tuljapurkar, S. (2004), 'Using the Lee-Carter method to forecast mortality for populations with limited data', *International Statistical Review* **72**(1), 19–36.

Li, Q., Reuser, M., Kraus, C. and Alho, J. (2009), 'Ageing of a giant: a stochastic population forecast for China, 2006–2060', *Journal of Population Research* **26**(1), 21–50.

Lopez, A. D. (1999), *Life tables for 191 countries: data, methods and results*, World Health Organization.

Murphy, M. (1995), 'The prospect of mortality: England and Wales and the United States of America, 1962–1989', *British Actuarial Journal* **1**(02), 331–350.

Niu, G. and Melenberg, B. (2012), 'Trends in mortality decrease and economic growth'.

OECD (2012), 'Medium and long-term scenarios for global growth and imbalances', *OECD Economic Outlook* **2012**, 1.

Oksanen, H. (2010), The chinese pension system-first results on assessing the reform options, Technical report, Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.

Pozen, R. C. (2013), *Tackling the Chinese Pension System*, Paulson Institute.

Renshaw, A. E. and Haberman, S. (2003), 'Lee-Carter mortality forecasting with age-specific enhancement', *Insurance: Mathematics and Economics* **33**(2), 255–272.

Renshaw, A. E. and Haberman, S. (2006), 'A cohort-based extension to the lee–carter model for mortality reduction factors', *Insurance: Mathematics and Economics* **38**(3), 556–570.

Ruhm, C. J. (2007), 'A healthy economy can break your heart', *Demography* **44**(4), 829–848.

Russolillo, M., Giordano, G. and Haberman, S. (2011), 'Extending the Lee-Carter model: a three-way decomposition', *Scandinavian Actuarial Journal* **2011**(2), 96–117.

Salditt, F., Whiteford, P. and Adema, W. (2008), 'Pension reform in China', *International Social Security Review* **61**(3), 47–71.

Singh, G. K. and Siahpush, M. (2006), 'Widening socioeconomic inequalities in us life expectancy, 1980–2000', *International Journal of Epidemiology* **35**(4), 969–979.

Tapia Granados, J. A. (2005), 'Increasing mortality during the expansions of the US economy, 1900–1996', *International Journal of Epidemiology* **34**(6), 1194–1202.

Tapia Granados, J. A. (2008), 'Macroeconomic fluctuations and mortality in postwar Japan', *Demography* **45**, 323–343.

Tapia Granados, J. A. and Ionides, E. L. (2008), 'The reversal of the relation between economic growth and health progress: Sweden in the 19th and 20th centuries', *Journal of Health Economics* **27**(3), 544–563.

Tuljapurkar, S., Li, N. and Boe, C. (2000), 'A universal pattern of mortality decline in the G7 countries', *Nature* **405**(6788), 789–792.

United Nations (2012), 'World urbanization prospects, the 2011 revision', *Final Report with Annex Tables. New York, NY: United Nations Department of Economic and Social Affairs* .

United Nations (2013), 'World population prospects, the 2012 revision', *New York: Department for Economic and Social Affairs* .

Wei, Y.-M., Jin, J.-L. and Wang, Q. (2012), 'Impacts of natural disasters and disasters risk management in China: The case of China's experience in wenchuan earthquake', *ERIA Research Project Report,* pp. 641–675.

Wilmoth, J. R. (1993), Computational methods for fitting and extrapolating the lee-carter model of mortality change, Technical report, Technical report, Department of Demography, University of California, Berkeley.

Woods, R. (2003), 'Urban-rural mortality differentials: An unresolved debate', *Population and Development Review* **29**(1), 29–46.

Yuan, J., Zhao, C., Yu, S. and Hu, Z. (2007), 'Electricity consumption and economic growth in China: cointegration and co-feature analysis', *Energy Economics* **29**(6), 1179–1191.

Zhang, H., Zhang, H. and Zhang, J. (2013), 'Demographic transition and economic growth: Evidence from Chinese provinces'.

Zhao, B. B., Liang, X., Zhao, W. and Hou, D. (2013), 'Modeling of group-specific mortality in China using a modified Lee-Carter model', *Scandinavian Actuarial Journal* **2013**(5), 383–402.

Zheng, W. Y. F. K. (2012), 'Mingmei; the development trend of China urbanization—China international forum on urbanization in 2012', *Urban Studies* **6**.