STANDARDIZING THE GIANT: MITIGATING LONGEVITY RISK IN CHINA
THROUGH CAPITAL MARKETS SOLUTIONS

Johnny S.-H. Li¹, Kenneth Q. Zhou¹ and Wai-Sum Chan²

¹ Department of Statistics and Actuarial Science, University of Waterloo, Canada
E-mail: shli@uwaterloo.ca, kenneth.zhou@uwaterloo.ca
² Department of Finance, The Chinese University of Hong Kong, Hong Kong
E-mail: chanws@cuhk.edu.hk

Abstract: As the annuity market in China develops, the Chinese insurance industry is increasingly exposed to longevity risk. A large part of the risk is ‘trend risk’, which cannot be diversified by pooling, but may be transferred to capital markets through derivatives that are written on a certain mortality index. In this paper, we first explore different methods to create a standardized mortality index for China. We then study how Chinese insurers may use such an index to offload a meaningful portion of longevity risk from their annuity books. The performance of the proposed index-based longevity hedge is tested by using a multi-population stochastic mortality model that is estimated to data from different provinces, municipalities and autonomous regions of China. Finally, we investigate the amount of capital relief that can be obtained from an index-based longevity hedge under the China Risk Oriented Solvency System (C-ROSS), which is scheduled to be implemented by 2016.

Keywords: C-ROSS, index-based longevity hedges, longevity risk, multi-population mortality models, securitization.

1 Introduction

Since the launch of the policy of reformation and opening, China has made great strides in reducing mortality. According to the World Bank, the life expectancy at birth for the unisex population of China has increased from 63.61 years in 1970 to 76.52 years in 2012. The reduction in Chinese pension insurers’ mortality is even more remarkable. A comparison between the 1990-1993 and 2000-2003 insurance life tables for Chinese pension insurers reflects an increase in life expectancy of approximately 4.7 years in 10 years of time, which is much faster than the typical rate of increase in life expectancy (3 years per decade) in the developed world. The gain in life expectancy is certainly an important social achievement, but it also poses a threat to the population’s retirement income security. The problem is exacerbated by China’s infamous one-child policy, which leaves the older generations with increased chances of dependency on retirement funds.

The current urban pension system in China is developed from the State Council Document No. 26 in 1997. In line with the recommendations provided by the World Bank (1994) report, it is a three-pillar system comprising of a mandatory publicly-managed pillar, a mandatory privately-managed pillar and a voluntary pillar. The first pillar is a defined-benefit public plan that includes a pay-as-you-go portion financed by employer contributions of 20% of wages plus a funded portion supported by employee contributions of 8% of wages. The second pillar is formed by defined-contribution occupational plans that are known as Enterprise Annuities. According to the Ministry of Labour and Social Security Statement No. 20 in 2004, Enterprise Annuities are funded by employer and employee contributions up to 1/6 of the total salary roll. All contributions go directly to Enterprise Annuity accounts, which are decumulated upon retirement. We refer interested readers to Cai and Cheng (2014) and Dorfman et al. (2012) for further information about pension systems in China.

Although the first pillar of the urban pension system offers individuals some protection against longevity risk through a modest lifetime income, the second pillar provides no protection against longevity risk because the benefit from an Enterprise Annuity must be received as either a lump sum or instalments with a fixed term. To seek additional protection against the risk of outliving financial resources, individuals’ only option is to purchase life annuities from insurance companies. For this reason, the life annuity market in China has a massive growth potential as the coverage of the second pillar broadens. In 2013, the total amount of funds accumulated in Enterprise Annuity accounts is 603.5 billion yuan, which is almost 4 times that in 2007 (Ministry of Human Resources and Social Security, 2014). Adding to the 603.5 billion yuan are the funds from the third (voluntary) pillar. Chen and Zhu (2009) pointed out that the total amount of life annuities purchased with individual savings was 200 billion yuan in 2006. The rising demand of life annuities can already be seen from the total annuity benefit payout of the Chinese insurance industry, which has risen significantly from 14.0 billion yuan in 2010 to 21.5 billion yuan in 2012 (China Insurance Regulatory Commission, 2011, 2013).

A large part of the longevity risk entailed in life annuities is ‘trend risk’, which arises from the uncertainty surrounding the trend in Chinese mortality over time. Because the risk affects all annuitants in the Chinese insurance industry systematically, the more life annuities an insurer sells, the more the insurer is exposed to the risk. Despite
the offsetting exposure in the insurer’s life insurance book may naturally hedge the trend risk acquired from life annuity sales, the effect of such a natural hedge may be limited for reasons such as differences in underwriting, duration and age profile (Zhu and Bauer, 2014), and the (limited) potential of natural hedging may eventually exhaust as the life annuity book grows. Furthermore, the newly introduced China Risk Oriented Solvency System (C-ROSS) specifically requires insurers operating in China to hold longevity risk solvency capital. Coming into effect in 2016, the C-ROSS may possibly compress the ability of the Chinese insurance industry to offer life annuities at affordable prices.

The question arises as to who else can bear the trend risk. One possible candidate is the Chinese government, which may take longevity trend risk exposures from the insurance market explicitly by issuing longevity bonds (Blake et al., 2013) or implicitly by ‘bailing out’ one or more insurance companies in case there is a systemic failure in the insurance industry due to longevity risk (Basel Committee of Banking Supervision, 2013). However, because the Chinese government is already assuming huge longevity trend risk due to its public pension plan, of which the asset amounts to 2826.9 billion yuan at 2013 year-end, it does not seem to have a good capacity to accept further risk of the same kind. A more promising candidate is the capital markets in China. Capital market investors may be interested in taking longevity trend risk exposures, because of the risk premium and diversification benefits they offer. In 2014, the total market capitalization of the equity markets in China is 8.3 trillion USD, while the total notional amount of derivatives traded in Chinese exchanges is 271 trillion USD. These figures suggest that the capital markets in China, in theory, can absorb at least some of the longevity trend risk exposures from the insurance industry.

The OECD (2014) report echoed the potential of capital markets to assist the life insurance industry in continuing to provide longevity protection to individuals. The report further recommended financial institutions to create standardized index-based mortality derivatives, which could resolve the misalignment of incentives between annuity providers, who have an intention to mitigate their longevity trend risk exposures, and capital market investors, who demand liquidity and are likely to be discouraged by the information asymmetry arising from the fact that insurers have better knowledge about the mortality experience of their annuitants. To follow this recommendation, an indispensable prerequisite is the creation of standardized mortality indexes, upon which derivative securities like swaps and forwards can be written. Although there already exist tradable mortality indexes such as the LifeMetrics index provided by the Life and Longevity Markets Association (LLMA), the existing indexes are based on mortality experience in the Western world and may therefore be unsuitable for use in China. We believe that with a population of over 1.35 billion, China deserves its own standardized mortality index.

The first objective of this paper is to investigate how a standardized mortality index for China may be developed. We consider both traditional, non-parametric methods and the recently proposed parametric methods (Chan et al. 2014; Tan et al., 2014). The development of a standardized mortality index for China is made difficult due to the limited availability of historical mortality data. For instance, the age- and gender-specific death counts for the general population of China are available for only a selected number of years: 1986, 1989 and 1994 to 2011. With respect to data limitations, the pros and cons of each index creation method are evaluated.

It is well-known that index-based longevity hedges are imperfect, due primarily to the population basis risk arising from the difference in future mortality improvements between the hedger’s own population of individuals and the population to which the standardized instrument is linked. The population basis risk involved in longevity hedges that are developed form a national standardized mortality index is likely to be significant, because of the substantial socioeconomic differences among different sub-populations in China. Such differences can be clearly seen from Figure 1, in which we show the life expectancies at birth (1990, 2010) and the average urban household incomes (1994, 2011) for various geographical regions in China. It is therefore crucially important to develop a multi-population stochastic mortality model that allows Chinese insurers to assess the potential population basis risk involved in their index-based longevity hedges.

Despite there are several existing multi-population mortality models (Ahmadi and Li, 2014; Cairns et al., 2011; Dowd et al., 2011; Hatzopoulos and Haberman, 2013; Jarner and Kryger, 2011; Li and Hardy, 2011; Li and Lee, 2005; Yang and Wang, 2013; Zhou et al., 2013, 2014), they cannot be applied straightforwardly to China because of, again, data-related issues. In particular, the only available mortality data by geographical regions in China are the values of life expectancy at birth in 1990, 2000 and 2010. The second of objective of this paper is to develop, from the limited available data, a multi-population stochastic mortality model for different provinces, municipalities and autonomous regions of China. The development stems from our parallel study (Li et al, 2015), which attempts to overcome the challenge of inadequate data in part by using information theory (Kullback and Leibler, 1951) and in part by Bayesian methods (Czado et al., 2006; Pedroza, 2006).

To render a standardized mortality index useful, an appropriate hedging strategy is needed. A number of longevity hedging strategies have recently been introduced by researchers including Cairns (2011, 2013), Cairns

---

1Source: Hong Kong Security and Futures Commission, China Security Regulation Commission, World Federation of Exchanges, and the authors’ own calculations.
et al. (2014), Coughlan et al. (2011), Dahl et al. (2008), Li and Luo (2012), Luciano et al. (2012), Tan et al. (2014) and Zhou and Li (2014). The third objective of this paper is to, by adapting the work of Zhou and Li (2014), produce a dynamic hedging strategy that is compatible with the proposed national mortality indexes and the multi-population mortality model. We also demonstrate that these strategies can offload a meaningful portion of longevity risk from insurers’ annuity books.

The final objective of this paper is to study the longevity risk component of C-ROSS, a new solvency system that has drawn considerable attention from both domestic and foreign insurers in recent years (see Zhao, 2014). To this end, we first illustrate how the C-ROSS longevity solvency risk capital is calculated with the prescribed adverse scenario factors. We then demonstrate that the benefit of index-based longevity hedges to insurers by estimating how much C-ROSS solvency capital that such hedges can release.

The rest of this paper is organized as follows. Section 1 describes the mortality data used in this study. Section 2 explains the creation of a standardized mortality index for China. Section 3 details the multi-population stochastic mortality model that is built specifically for assessing population basis risk in China. Section 4 presents the dynamic hedging strategy we consider. Section 5 describes how we estimate the C-ROSS capital relief from an

Figure 1: The life expectancies at birth (1990, 2010) and the average urban household incomes (1994, 2011) for different geographical regions in China. Source: The China Knowledge Resource Integrated Database.
index-based longevity hedge. Finally, Section 6 summarizes the contributions of this study.

2 Data

The accomplishment of the research objectives requires historical mortality data for the entire population and different geographical regions of mainland China. In what follows, we describe the relevant data that are available to us.

The Asia-Pacific Mortality Database managed by the Insurance Risk and Finance Research Centre of Nanyang Technological University provides historical aggregate death rates (i.e., the ratio of total deaths to total population) for the entire population of mainland China from 1960 to 2011.

The World Bank provides historical values of life expectancy at birth for the entire population of mainland China (male, female and unisex) from 1960 to 2012.

The China Knowledge Resource Integrated Database provides age- and gender-specific death and mid-year population counts for the entire population of mainland China for a selected number of years: 1986, 1989 and 1994 to 2011. Data by single years of age generally are available from 0 up to a certain age (99 for years 1986, 1989, 1994, 2001, 2005, 2010; 85 for year 1996; 89 for the other years), beyond which the data are right-censored. A few data values, for example, the death count for females at age 2 in year 2009, are missing. The available nationwide age- and gender-specific mortality data are summarized in the lexis diagrams shown in Figure 2.

Age-specific mortality data by geographical regions in China are unfortunately not available to the authors. At a sub-population level, the only mortality data we have are the values of gender-specific life expectancy at birth in 1990, 2000 and 2010, provided by the China Knowledge Resource Integrated Database. These data cover all 22 provinces, 4 municipalities and 5 autonomous regions of China: Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.³

3 Construction of Standardized Mortality Indexes

We consider three different methods for constructing standardized mortality indexes: non-parametric (aggregate, age-specific) and parametric. With respect to the data-related issues, each method has its pros and cons.

⁴We acknowledge that the Asia-Pacific Mortality Database provides age- and gender-specific death rates and probabilities. Because they are available by 5-year age intervals rather than single years of age, they are not considered in this study.

³The values for Chongqing in 1990 are not available, because this direct-controlled municipality was still not yet established at that time.

Figure 2: Lexis diagrams summarizing the availability of age- and gender-specific data for the entire population of mainland China. The green cells indicate that data are available by single years of age, the blue cells indicate data are available by open age groups, and the red cells indicate data values are missing.
3.1 Non-Parametric Aggregate

Aggregate death rates of the Chinese population (see Figure 3) may be used as a standardized mortality index. An advantage of this quantity is that its historical values are available since 1960, allowing capital market investors to better understand and predict its dynamic over time. However, such an index is subject to three significant limitations. First, it does not distinguish between genders. Second, it does not reflect how the shape of the underlying mortality curve has actually evolved over time. Third, it is perturbed by the information about China’s population structure, which has no relevance to index-based longevity hedges for insurers. For example, an increase in the aggregate death rate may be due entirely to an expansion in the old-age population, whose mortality is higher, rather than an increase in the likelihood of death.

Another option is to use gender-specific life expectancies at birth, just as how Credit Suisse’s mortality index for the US population was created. To transfer longevity risk exposure, one may write an e-forward, a concept that was first proposed by Hunt (2015), on a life expectancy index. Figure 4 depicts the life expectancies at birth for Chinese males and females from 1960 to 2012. Compared to aggregate death rates, life expectancies are advantageous of being purely mortality-related, with no interaction with the population structure. Nevertheless, it still does not indicate any information about the changes in the shape of the underlying mortality curve. Another problem about using the life expectancy at birth as an index is that it is more sensitive to changes in mortality at younger ages than changes in mortality at older ages. Given that the longevity risk exposures associated with life annuities arise predominantly from changes in old-age mortality, an index that is based the life expectancy at birth is less preferred to indexes that are more responsive to changes in old-age mortality.

3.2 Non-Parametric Age-Specific

Similar to the LLMA’s LifeMetrics index, we may also create standardized mortality indexes with age-specific mortality rates. By writing q-forwards on such indexes, longevity risk exposures can be transferred.

Figure 5 displays the mortality rates for Chinese males and females in 1986, 1989 and 2004-2011. A collection of indexes based on death rates at various advanced ages (e.g., 70, 75, etc) can better represent the evolution of the portion of the mortality curve that is relevant to annuity liabilities, but the reliance on multiple indexes may lead to problems in concentrating liquidity. In the work of Li and Luo (2012), it was found that to achieve a over 90% reduction in variance, a static longevity hedge for a single cohort of annuity liability requires 5 non-parametric age-specific indexes. However, this problem is not profound if hedgers adjust their longevity hedges dynamically. Cairns (2011) found that with dynamic adjustments, using 2 non-parametric age-specific indexes can reduce the variance in the values of a single cohort of annuity liability by over 90%.

Note that for Chinese population, non-parametric age-specific indexes are also subject to the limitation that only 20 historical observations are available.
Figure 4: Gender-specific life expectancies at birth for the Chinese population, 1960-2012

Figure 5: Age- and gender-specific death rates for the Chinese population, 1986, 1989 and 1994-2011.
3.3 Parametric

Chan et al. (2014) argued that a parametric (model-based) construction method may improve the information content of mortality indexes. In using a parametric method, mortality indexes are constructed from the time-varying parameters in a stochastic mortality model.

For instance, one may use parameter \( k(t) \) in the original the Lee-Carter model as a mortality index:

\[
\ln(m(x,t)) = a(x) + b(x)k(t) + \epsilon(x,t),
\]

where \( m(x,t) \) denotes the central death rate at age \( x \) and in year \( t \), \( a(x) \) is an age-specific parameter representing the average level of mortality at age \( x \) over time, \( k(t) \) is a time-varying parameter, \( b(x) \) is an age-specific parameter indicating the sensitivity of \( \ln(m(x,t,i)) \) to \( k(t) \), and \( \epsilon(x,t) \) is the error term. We can interpret \( k(t) \) to mean the overall level of mortality in year \( t \). A reduction in \( k(t) \) implies a parallel downward shift of the log-transformed curve of central mortality rates.

As Chan et al. (2014) explained, the model on which the mortality indexes are based must possess the ‘new-data-invariant’ property, which means that when an additional year of mortality data becomes available and the model is updated accordingly to generate a new index value, the index values for the previous years will not be affected. This property is important, because it guarantees the tractability of the resulting mortality indexes. To achieve this property, we may keep the age-specific parameters \( a(x) \) and \( b(x) \) fixed when we update the model with new mortality data.

Figure 6 shows the values of the parametric mortality indexes for Chinese males and females over the period of 1986-2011. Note that the problem of having a smaller number of historical observations still applies to the parametric mortality indexes. For years 1987, 1988, 1990-1993, no mortality data are available and the values of \( k(t) \) are imputed by a Bayesian methodology which will be discussed in Section 4.3.

A conceptual security called K-forward was proposed by Chan et al. (2014) and subsequently implemented by Tan et al. (2014). A K-forward contract is a zero-coupon swap that exchanges on the maturity date a fixed amount for a random amount that is proportional to the value of a parametric mortality index at some future time. Through K-forward contracts, longevity risk exposures can be transferred from one party to another.

4 Developing a Multi-Population Mortality Model for China

As previously mentioned, a standardized longevity hedge that is based on a Chinese national standardized mortality index may be subject to significant population basis risk, because there exist huge socioeconomic differences among different sub-populations in China. To quantify the population basis risk involved, we now build a multi-population stochastic mortality model which captures the co-movement of mortality trends of various sub-populations in China. The materials in this section draw heavily from our parallel study (Li et al., 2015), which
is devoted to investigating how a multi-population stochastic mortality model may be constructed when there is a paucity of data.

In what follows, we first explain how we derive the base mortality tables for various provinces, municipalities and autonomous regions. We then describe the multi-population mortality model and explain how it can be estimated given the limited available data.

### 4.1 Estimating Historical Age-Specific Death Rates for Different Geographical Regions in China

One of the major challenges in this study is that historical age-specific mortality rates for different provinces, municipalities and autonomous regions in China are not available. To build a multi-population mortality model for China, we must first derive age-specific mortality rates for different geographical regions in China from the only demographic quantity (life expectancy at birth) that is available to us.

The method is based on information theory in statistics. The idea behind is to extract as much information as possible from the life expectancy values that are available to us.

In more detail, each life table for the general population provides us with the values of \( q(x) \) for \( x = 0, 1, \ldots \). The value of \( q(x) \) represents the conditional probability that an individual in the general Chinese population dies during the age interval \([x, x+1)\), given that the individual is alive age \( x \). With the values of \( q(x) \), we readily obtain the probability function for the age at death random variable as follows:

\[
\pi(x) = \begin{cases} 
q(x), & x = 0 \\
\prod_{y=0}^{x-1} (1 - q(y)) q(x), & x = 1, \ldots, \omega - 1, \\
0, & \text{otherwise,}
\end{cases}
\]

where \( \omega \) denotes the highest attainable age and \( \pi(x) \) represents the unconditional probability of death during the age interval of \([x, x+1)\). Note that \( \sum_{x=0}^{\omega-1} \pi(x) = 1 \). In our calculations, we set \( \omega \) to 100.

We let \( q^*(x) \) and \( \pi^*(x) \) be the corresponding values of \( q(x) \) and \( \pi(x) \) for a certain province, municipality or autonomous region in China. As previously mentioned, the values of \( q(x) \) (and hence \( \pi(x) \)) are known for a certain number of years, but the values of \( q^*(x) \) and \( \pi^*(x) \) for any year are not.

The only information about the mortality of different provinces, municipalities and autonomous regions is the values of their life expectancy at birth in years 1990, 2000 and 2010. We derive the lifetime distribution for a sub-population in each of these three years by treating the corresponding lifetime distribution for the general population (i.e., \( \pi(x), x = 0, \ldots, \omega - 1 \)) as a prior distribution, which is subsequently updated by incorporating the information contained in the sub-population’s life expectancy.

We let

\[
\sum_{x=0}^{\omega-1} \pi^*(x) \ln \frac{\pi^*(x)}{\pi(x)}
\]

be the Kullback-Leibler information criterion (Kullback and Leibler, 1951) of the age-at-death probability distribution for the sub-population relative to that of the general population. We derive the values of \( \pi^*(x) \) by minimizing expression (2), subject to the constraints

\[
\sum_{x=0}^{\omega-1} \pi^*(x) = 1
\]

and

\[
\sum_{x=0}^{\omega-1} x \pi^*(x) + 0.5 = c^*_\omega,
\]

where \( c^*_\omega \) is the complete period life expectancy at birth for the sub-population. The first constraint ensures that the collection of \( \pi^*(x) \)'s forms a proper probability mass function, while the second constraint ensures that the life expectancy at birth implied by the estimated probability distribution matches that provided by the China Knowledge Resource Integrated Database.\(^4\) It can be shown that the solution to the constrained minimization problem is

\[
\pi^*(x) = \frac{\pi(x) \exp(\lambda_1 x)}{\sum_{x=0}^{\omega-1} \pi(x) \exp(\lambda_1 x)}, \quad x = 0, \ldots, \omega - 1,
\]

where \( \lambda_1 \) is the Lagrange multiplier which can be computed readily by substituting equation (5) into equation (4). Given the estimates of \( \pi^*(x) \) for \( x = 0, 1, \ldots, \omega - 1 \), the values of \( q^*(x) \) for \( x = 0, 1, \ldots, \omega - 1 \) can be calculated recursively by using equation (1) and the fact that \( q(\omega - 1) = 1\).

\(^4\)The first term in equation (4) computes the curtail life expectancy at birth. Assuming uniform distribution of deaths between two consecutive integer ages, adding 0.5 to the first term yields the complete life expectancy at birth.
Figure 7: The estimated age- and gender-specific conditional death probabilities for all 22 provinces, 4 municipalities and 5 autonomous regions in China, 2010.

Figure 7 shows the estimated 2010 age- and gender-specific conditional death probabilities for all 22 provinces, 4 municipalities and 5 autonomous regions in China. As expected, the estimated death probabilities for the most developed geographical regions such as Beijing and Tianjin are consistently lower than the corresponding death probabilities for the general Chinese population. The opposite is true for less developed geographical regions such as Xinjiang.

The life tables estimated in this sub-section will be used in the development of the multi-population stochastic mortality model for various provinces, municipalities and autonomous regions in China.

4.2 The Multi-Population Model

4.2.1 Model Specification

The multi-population model we consider is an adapted version of the augmented common factor model, proposed by Li and Lee (2005).\(^5\) It can be regarded as a multi-population generalization of the classical Lee-Carter

\(^5\)The model we consider is slightly different from that proposed by Li and Lee (2005). In particular, while Li and Lee (2005) used identical parameterization for all populations being modeled, the parameterizations we use for the general population and the sub-populations are not the same.
model (Lee and Carter, 1992).

The model under consideration is specified as follows:

- The general population

\[ \ln m(x,t) = a(x) + b(x)k(t) + \epsilon(x,t), \]  

(6)

- Provinces, municipalities and autonomous regions

\[ \ln m(x,t,i) = a(x,i) + b(x)k(t) + b(x,i)k(t,i) + \epsilon(x,t,i), \]  

(7)

for \( i = 1, 2, \ldots, N \).

In the above, \( m(x,t,i) \) is the central rate of death at age \( x \) and in year \( t \) for sub-population \( i \), \( a(x,i) \) is an age-specific parameter indicating sub-population \( i \)'s average mortality level at age \( x \), \( k(t) \) is a time-varying factor that is specific to sub-population \( i \), \( b(x,i) \) measures the sensitivity of \( \ln m(x,t,i) \) to \( k(t,i) \), \( \epsilon(x,t,i) \) is the error term for population \( i \), and \( N = 30 \) is the total number of sub-populations under consideration.\(^6\) The definitions of \( m(x,t), a(x), b(x) \) and \( k(t) \) remain the same as that in Section 3.3. It is assumed that both \( \epsilon(x,t) \) and \( \epsilon(x,t,i) \) are normally distributed with a zero mean and constant variances of \( \sigma^2_i \) and \( \sigma^2_{i} \), respectively.

The specification implies that the evolution of the general population’s mortality follows the classical Lee-Carter model. The mortality dynamics for sub-population \( i \) follows a normal distribution with a zero mean and a constant variance of \( \sigma^2_i \) and is deemed more biologically reasonable than one that comes with divergent projected trends.

As in the classical Lee-Carter model, the evolution of \( k(t) \) over time is modeled by a random walk with drift:

\[ k(t) = c + k(t - 1) + \zeta(t), \]  

(8)

where \( c \) is a constant and \( \zeta(t) \) follows a normal distribution with a zero mean and a constant variance of \( \sigma^2_i \). For \( i = 1, \ldots, N \), the evolution of \( k(t,i) \) over time is modeled by a first order autoregressive process:

\[ k(t,i) = \phi_0(i) + \phi_1(i)k(t - 1,i) + \zeta(t,i), \]  

(9)

where \( \phi_0(i) \) is a constant, \( \phi_1(i) \) is another constant whose an absolute value is strictly less than 1, and \( \zeta(t,i) \) follows a normal distribution with a zero mean and a constant variance of \( \sigma^2_i \).

The use of an autoregressive process for \( k(t,i) \) implies that \( k(t,i) \) will revert to a long-term equilibrium value of the long-run. In this way, the projected mortality trends for the general population and the sub-populations do not diverge indefinitely. Such a multi-population mortality forecast is considered as ‘coherent’ (Li and Lee, 2005) and is deemed more biologically reasonable that one that comes with divergent projected trends.

Li and Hardy (2011) evaluated the augmented common factor model. Their empirical results indicate that the model fits better and yields more reasonable estimates of population basis risk than its predecessors such as the ‘joint-k’ model introduced by Carter and Lee (1992).

4.2.2 Model Estimation

Because of the missing data values, the model cannot be estimated with simple methods such as singular value decomposition. We overcome the estimation challenge by following the Bayesian method of Pedroza (2006), in which the entire model – equations (6) to (9) – is formulated jointly as a Gaussian state-space model. The time-varying factors \( k(t) \) and \( k(t,i) \) are treated as hidden states, whereas \( a(x), a(x,i), b(x), b(x,i), c, \phi_0(i) \) and \( \phi_1(i), \sigma^2_i, \sigma^2_i(i), \sigma^2_i \) and \( \sigma^2_i(i) \) are considered as model parameters that are assumed to be random themselves.

The iterative estimation procedure consists of the following major components.

Gibbs sampling

It is assumed that \( \ln(m(x,t)) \) and \( \ln(m(x,t,i)) \) for \( i = 1, \ldots, N \) are normally distributed. Under this assumption, the conditional posterior distribution of each parameter can be analytically obtained by using an appropriate conjugate prior of a normal distribution. The conjugate priors we use include normal (for parameters \( a(x), a(x,i), b(x), b(x,i), c, \phi_0(i) \) and \( \phi_1(i) \)) and inverse-gamma (for parameters \( \sigma^2_i, \sigma^2_i(i), \sigma^2_i \) and \( \sigma^2_i(i) \)). From the conditional posterior distributions, we can draw samples of the model parameters readily.

\(^6\)Because Chongqing has a rather short history, we choose not to include this municipality in the multi-population model. The number of sub-populations being modeled are therefore 30, which includes 22 provinces, 3 municipalities and 5 autonomous regions.
Kalman filtering and smoothing

Given a Gaussian state-space formulation, the hidden states \((k(t) \text{ and } k(t, i))\), for \(i = 1, \ldots, N\) and all \(t\) in the calibration window, can be retrieved readily by using a Kalman updating algorithm (to incorporate the information up to and including time \(t\)) and a Kalman smoothing algorithm (to incorporate information beyond time \(t\)).

Imputation of missing data

On the basis of the sample of parameters drawn and the hidden states retrieved in the most recent iteration, simulate the values of ln(m(x, t)) and/or ln(m(x, t, i)) at time points where data are missing. The imputed data and the observed data are combined to form a complete data sample for the Gibbs sampling and Kalman filtering and smoothing in the next iteration.

Enforcement of identifiability constraints

It is well-known that the Lee-Carter model and its variants are subject to the identifiability problem. To stipulate parameter uniqueness, the following constraints are used:

\[
\sum_x b(x) = 1, \quad \sum_x b(x, i) = 1, \quad \sum_t k(t) = 0, \quad \sum_t k(t, i) = 0.
\]

The identifiability constraints are applied at the end of each iteration.

As usual in Bayesian methods, the first batch of 100 samples are regarded as burn-in and therefore discarded. The subsequent samples are used to form the joint empirical posterior distribution of the model parameters. We refer interested readers to Li et al. (2015) for further details about the algorithms for Gibbs sampling and Kalman filtering and smoothing in the estimation procedure.

We estimate the model to the 20 years of data (1986, 1989, 1994-2011) from the general population and the 3 years (1990, 2000, 2010) of estimated age-specific mortality rates from the 30 sub-populations under consideration. We use data for ages 60 and beyond, because the illustrative longevity hedge to be presented in the next section does not depend on the mortality below age 60.

In Figure 8 we show the estimates of \(a(x), b(x)\) and \(k(t)\), parameters that are applicable to both the general population and the sub-populations. The fan chart in each panel shows the central 10% prediction interval for the parameter series with the heaviest shading, surrounded by the 20%, 30%, ..., 90% prediction intervals with progressively lighter shading. The line in the centre of the fan chart represents the best estimate of the parameter series. As expected, the estimate of \(a(x)\) increases with age, reflecting the positive relationship between mortality and age. The downward trend in \(k(t)\) indicates a steady reduction in the overall level of mortality over the past couple of decades.

Figure 9 shows, as an example, the estimates of \(a(x, i), b(x, i)\) and \(k(t, i)\) for males in Guangdong province. Compared to \(a(x), a(x, i)\) is subject to substantially more uncertainty. This outcome arises because there is far less information (only 3 years of data) on which the estimation of \(a(x, i)\) can be based. By construction, \(k(t, i)\) reverts to a long-term equilibrium value, so that the divergence between the projected mortality trends for this province and the general population do not grow indefinitely.

As an illustration, we use the estimated multi-population mortality model to project, for each sub-population, the actuarial present value of a 30-year temporary life annuity immediate of $1 that is issued to a male aged 60 at the end of year 2011. The projection result is displayed in Figure 10. There exist variations in the projected annuity values among different sub-populations, despite the multi-population mortality model we use does not permit an indefinite divergence in expected mortality trends. It is not surprising that the projected annuity values for the more developed geographical regions are generally higher.

5 Hedging Strategies

In this section, we investigate how Chinese insurers can use a national mortality index to offload longevity risk from their balance sheets. We adapt the work of Zhou and Li (2014) to form dynamic delta hedging strategies, whereby the hedge parameters (the deltas) of the insurer’s portfolio and the portfolio of hedging instruments are matched.

We begin this section with a description of the liability being hedged, followed by explanations about how the longevity risk involved in the liability can be mitigated by using instruments written on non-parametric and parametric mortality indexes. We then detail how hedge effectiveness may be measured, and estimate the degrees of hedge effectiveness that longevity hedges for annuity liabilities in different geographical regions of China can

\[\text{The illustrative longevity hedge in Section 5.5 is based on the same annuity liability.}\]
Figure 8: The estimates of $a(x)$, $b(x)$ and $k(t)$ in equation (6), $x = 60, \ldots, 89$ and $t = 1986, \ldots, 2011$, Chinese males.

Figure 9: The estimates of $a(x, i)$, $b(x, i)$ and $k(t, i)$ in equation (7), $x = 60, \ldots, 89$ and $t = 1986, \ldots, 2011$, males in Guangdong province.
achieve. We conclude this section with an analysis of various factors that may affect the performance of a longevity hedge for a specific sub-population. Throughout this section, the multi-population model presented in Section 4 is assumed.

5.1 The Set-up

Let us first define several notation. We let

\[ S_{x,t}^{(i)}(T) = \prod_{s=1}^{T} (1 - q^{(x+s-1,t+s,i)}) \]

be the \textit{ex post} probability that an individual who is from sub-population \( i \) and aged \( x \) at time \( t \) (the end of year \( t \)) would have survived to time \( t + T \), where \( q(x,t,i) \) denotes the probability that an individual from population \( i \) dies between time \( t - 1 \) and \( t \) (during year \( t \)), provided that he/she has survived to age \( x \) at time \( t - 1 \). It is clear from the definitions that \( S_{x,t}^{(i)}(T) \) is not known prior to time \( t + T \), while \( q(x,t,i) \) is not known prior to time \( t \). We also let

\[ p_{x,u}^{(i)}(T,F_t) = \mathbb{E}(S_{x,u}^{(i)}(T)|F_t), \]

where \( u \geq t \) and \( F_t \) represents the information about the evolution of mortality up to and including time \( t \). Because the assumed mortality model is based on central death rates, we need to approximate \( q(x,t,i) \) from \( m(x,t,i) \). We use the relation \( q(x,t,i) = 1 - \exp(-m(x,t,i)) \), which holds exact if the force of mortality between two consecutive integer ages is constant.

Let us suppose that the liability being hedged is a portfolio of life annuities, which are associated with the cohort of individuals who are from sub-population \( i \) and aged \( x_0 \) at time \( t_h \) when the longevity hedge is established. We further assume that the each life annuity pays $1 at the end of each year until death. It follows that the time-\( t \) value of the insurer’s future liabilities (per policyholder at time \( t \)) is

\[ FL_t = \sum_{s=1}^{\infty} (1 + r)^{-s} p_{x_0+t-t_h}^{(i)}(s,F_t), \quad t \geq t_h, \]

Figure 10: The projected actuarial present value of a 30-year temporary life annuity immediate of $1 that is issued to a male aged 60 at the end of year 2011, for each province, municipality and autonomous region of China (except Chongqing).
where \( r \) is the interest rate for discounting purposes.

Suppose that the hedging horizon is \( Y \) years and that the q-forward portfolio is adjusted annually. Due to the dynamic nature of the hedge, the value of \( FL_t \) at the beginning of each of the \( Y \) years has to be computed. As \( FL_t \) takes no analytical form, evaluating the hedge over the hedging horizon requires nested simulations. To reduce computation burden, an approximation formula is used to compute each value of \( FL_t \). The approximation formula is derived by applying a second order Taylor expansion on the probit transformation of \( p_{x,t+s,0} \) about the best estimates of \( k(t) \) and \( k(t,i) \). We refer readers to Cairns (2011) and Zhou and Li (2014) for details concerning the approximation method.

### 5.2 Hedging with a Non-Parametric Age-Specific Mortality Index

We now consider the non-parametric age-specific index described in Section 3.2. As in Section 5.1, we define for the general population

\[
S_x(t) = \prod_{s=1}^{T} (1 - q(x + s - 1, t + s)) \quad \text{and} \quad p_{x,u}(T, \mathcal{F}_t) = \mathbb{E}(S_{x,u}(T)|\mathcal{F}_t),
\]

where \( u \geq t \) and \( q(x,t) \) denotes the probability that an individual from the general population dies between time \( t - 1 \) and \( t \), given that he/she has survived to age \( x \) at time \( t - 1 \).

We suppose here that q-forwards written on the index (age-specific death probabilities for the national population) are used as hedging instruments. A q-forward is a zero-coupon swap with its floating leg proportional to the realized death probability at a certain reference age during the year immediately prior to maturity and its fixed leg proportional to the corresponding forward mortality rate, which is fixed at inception. To hedge the longevity risk involved in the life annuity portfolio, the hedger should participate in the q-forwards as the fixed-rate receiver, so that he/she will receive a net payment from the counterparty when mortality turns out to be lower than expected.

Let us consider a q-forward that is linked to the national population of China and a reference age \( x_f \). Assume that the q-forward is issued at time \( t_0 \) and matures at time \( t_0 + T^* \). By definition, the payoff from the q-forward depends on the realized value of \( q(x_f, t_0 + T^*) \). Let \( q^f(x_f, t_0 + T^*) \) be the corresponding forward mortality rate, which is fixed at \( t = t_0 \) when the q-forward is first launched. At \( t = t_0, \ldots, t_0 + T^* - 1 \), the value of the q-forward (per $1 notional) from the perspective of the hedger (fixed-rate receiver) is given by

\[
Q_t(t_0) = (1 + r)^{-(t_0 + T^* - t)}(q^f(x_f, t_0 + T^*) - \mathbb{E}(q(x_f, t_0 + T^*)|\mathcal{F}_t))
\]

\[
= (1 + r)^{-(t_0 + T^* - t)}(q^f(x_f, t_0 + T^*) - (1 - \mathbb{E}(S_{x_f,t_0+T^*-1}(1)|\mathcal{F}_t)))
\]

\[
= (1 + r)^{-(t_0 + T^* - t)}(q^f(x_f, t_0 + T^*) - (1 - p_{x_f,t_0+T^*-1}(1, \mathcal{F}_t))).
\]

Suppose that at time \( t \) during the hedging horizon, the hedger uses the aforementioned q-forward (with \( t_0 \leq t \)) as the only hedging instrument. The main idea behind the delta hedging strategy is to ensure that the annuity portfolio and the q-forward portfolio have similar sensitivities to changes in \( k(t) \). To achieve this goal, the hedge ratio \( h_t \) (i.e., the notional amount of the q-forward) is chosen in such a way that

\[
\frac{\partial FL_t}{\partial k(t)} = h_t \frac{\partial Q_t(t_0)}{\partial k(t)},
\]

where \( \partial FL_t/\partial k(t) \) and \( h_t \partial Q_t(t_0)/\partial k(t) \) represent the time-t deltas of the annuity portfolio and the (calibrated) q-forward portfolio, respectively.

The hedge portfolio has a value of \( h_t Q_t(t_0) \) at time \( t \) and a value of \( h_t Q_{t+1}(t_0) \) at time \( t + 1 \). At time \( t + 1 \), the q-forward written at time \( t \) is closed out, and another q-forward portfolio is constructed. The process repeats from the beginning to the end of the hedging horizon.

When evaluating such a hedge, we need to compute the value of \( Q_t(t_0) \) for every \( t \) over the hedging horizon, but \( Q_t(t_0) \) cannot be analytically calculated. To avoid the need for nested simulations, an approximation formula is used to calculate \( Q_t(t_0) \). The approximation is based on a first order Taylor’s expansion of the probit transformation of \( p_{x,t+T^* - 1}(1, \mathcal{F}_t) \) about the best estimate of \( k(t) \). We refer readers to Cairns (2011) and Zhou and Li (2014) for further details about the approximation of \( Q_t(t_0) \). The values of \( \partial FL_t/\partial k(t) \) and \( \partial Q_t(t_0)/\partial k(t) \) are calculated on the basis of the approximation formulas for \( FL_t \) and \( Q_t(t_0) \), respectively.

### 5.3 Hedging with a Parametric Mortality Index

We now consider the parametric mortality index introduced in Section 3.3 and suppose that K-forwards written on the index are used as hedging instruments. We define a K-forward with an inception date \( t_0 \) and a maturity
of $T^*$ years as a zero-coupon swap which has a floating leg proportional to the value of $k(t_0 + T^*)$ (implied by the assumed model) and a fixed leg proportional to a constant $k^f(t_0 + T^*)$ that is fixed at inception. The hedger should participate in the contract as the fixed-rate receiver, so that when $k(t_0 + T^*)$ is smaller than expected, which corresponds to lower future mortality and thus more annuity payments, the hedger will receive a net payment from the counterparty of the contract to offset the increase in annuity payments.

Under the assumed stochastic process for $k(t)$, we have $E(k(t_0 + T^*)|F_t) = k(t) + c \times (t_0 + T^* - t)$ for $t = t_0, \ldots, t_0 + T^* - 1$. Hence, for $t = t_0, \ldots, t_0 + T^* - 1$, the value of the K-forward (per $1$ notional) from the hedger’s perspective can be expressed as

$$K_t(t_0) = (1 + r)^{-(t_0 + T^* - t)}(k^f(t_0 + T^*) - E(k(t_0 + T^*)|F_t))$$

$$= (1 + r)^{-(t_0 + T^* - t)}(k^f(t_0 + T^*) - k(t) + c \times (t_0 + T^* - t)).$$

Suppose that at time $t$, the hedger uses the aforementioned K-forward (with $t_0 \leq t$) as the only hedging instrument. The hedge ratio $h_t$ (i.e., the notional amount of the K-forward) is chosen in such a way that

$$\frac{\partial FL_t}{\partial k(t)} = h_t \frac{\partial K_t(t_0)}{\partial k(t)},$$

where $\partial FL_t/\partial k(t)$ and $h_t \partial K_t(t_0)/\partial k(t)$ are regarded as the time-$t$ delus of the annuity portfolio and the (calibrated) K-forward portfolio, respectively. The value of the K-forward is $h_t K_t(t_0)$ at time $t$ and becomes $h_t K_{t+1}(t_0)$ at time $t + 1$. At time $t + 1$, the K-forward written at time $t$ is closed out, and another K-forward is written. The process repeats until the end of the hedging horizon is reached.

Technically speaking, it is easier to evaluate a K-forward hedge than a q-forward hedge. This is because the time-$t$ value of a K-forward is simply a linear function of $k(t)$, thereby sparing us from the need of nested simulations or approximations. For the same reason, the partial derivative of $K_t(t_0)$ can be calculated straightforwardly as follows:

$$\frac{\partial K_t(t_0)}{\partial k(t)} = (1 + r)^{-(t_0 + T^* - t)}.$$

### 5.4 Measuring Hedge Effectiveness

We can evaluate the effectiveness of a dynamic longevity hedge by simulating a large number of mortality scenarios from the assumed multi-population mortality model.

We let $PL_{t_h} = FL_{t_h}$ and

$$PL_t = \sum_{s=1}^{t-t_h} (1 + r)^{-s} S_{x_0,t_h}^{(i)}(s) + (1 + r)^{-(t-t_h)} S_{x_0,t_h}^{(j)}(t) FL_t, \quad t = t_h + 1, \ldots, t_h + Y.$$

We can interpret $PL_t$ to mean the value of all annuity payments at time $t_h$ when the hedge is established, given the information up to and including time $t$. For $t > t_h$, the value of $PL_t|F_{t_h}$ is random in part because the value of $S_{x_0,t_h}^{(i)}(s)$ depend on the realizations of $k(t_h + 1), \ldots, k(t_h + s)$ and $k(t_h + 1, i), \ldots, k(t_h + s, i)$, and in part because the value of $FL_t$ depends on the realizations of $k(t)$ and $k(t, i)$.

Define by $PA_t$ the time-$t_h$ value of the assets backing the pension plan at time $t$, where $t \geq t_h$. We assume that the asset value equals the liability value when the hedge is established; i.e., $PA_{t_h} = PL_{t_h}$.

Let us consider a q-forward dynamic hedge. To simplify exposition, we assume that all q-forwards used have the same maturity $T^*$ and reference age $x_f$. We also assume that at every time point $t$ when the hedge portfolio is adjusted, a freshly launched q-forward is written (i.e., $t_0 = t$ for $t = t_h, \ldots, t_h + Y - 1$). Under these assumptions, we have

$$PA_t = PA_{t-1} + (1 + r)^{-(t-t_h)} h_{t-1} Q_t(t - 1)$$

for $t = t_h + 1, \ldots, t_h + Y$. The asset process for a K-forward hedge can be obtained by replacing $Q_t(t - 1)$ with $K_t(t - 1)$. The potential deviation between $PA_t$ and $PL_t$ is the residual risk that is not eliminated by the longevity hedge. Hence, we may measure hedge effectiveness by the following metric:

$$HE_u = 1 - \frac{\mathrm{Var}(PA_{t_h+u} - PL_{t_h+u}|F_{t_h})}{\mathrm{Var}(PL_{t_h+u}|F_{t_h})}, \quad u = 1, \ldots, Y,$$

which is close to 1 if the hedge is effective and 0 if not.
5.5 An Illustration

In this sub-section, we illustrate the use of a standardized national mortality index to hedge the longevity risk associated with annuity portfolios that are located in different provinces, municipalities and autonomous regions in China. The following assumptions are made in the illustration.

1. The liability being hedged is a portfolio of life annuities that are sold to males who are aged 60 at the end of year 2011. Each annuity pays $1 at the end of each year until the annuitant dies or reaches age 90, whichever is the earliest.

2. The mortality experience of the annuitants is the same as that of the males in the province, municipality or autonomous regions to which they belong.

3. The hedge begins at the end of year 2011 and the hedging horizon is 30 years. The hedge portfolio is adjusted annually.

4. The hedging instruments used are q-forwards that are linked to the national population of China. They all have a time-to-maturity of 10 years and a reference age of 75.

5. All q-forwards used have a zero risk premium, which means $q_f(x, t_0 + T^*) = E(q(x, t_0 + T^*))$. This working assumption has no impact on the resulting hedge effectiveness.

6. The market for q-forwards is liquid and no transaction cost is required.

7. The interest rate for all durations is $r = 4\%$ per annum and remains constant over time. The hedger can invest or borrow at this rate.

8. The evaluation of hedge effectiveness is based on 1,000 mortality scenarios that are generated from the multi-population mortality model presented in Section 4.2.2.

9. There is no small sample risk.

The hedging results are presented in Figure 11. In each panel, the grey (larger) fan chart shows the distributions of $PL_{t_h + u | F_{t_h}}$ for $u = 1, \ldots, Y$, whereas the green (smaller) fan chart depicts the distributions of $PA_{t_h + u} - PL_{t_h + u | F_{t_h}}$ for $u = 1, \ldots, Y$. The difference between the widths of the two fan charts reflects the amount of longevity risk that is removed from the dynamic q-forward hedge. The corresponding value of $HE_{30}$ is displayed on the top of each panel.

Interestingly, the longevity hedges in most developed geographical regions are not necessarily the most effective. As of 2011, Shanghai and Tianjin were ranked first and fourth in terms of the per capita income of urban households. However, our simulation results indicate that the longevity hedge in Shanghai is only moderately effective (with $HE_{30} = 0.2659$), while the longevity hedge in Tianjin performs even worse (with $HE_{30} = 0.1540$). We believe that rather than having a positive relationship with the sub-population’s level of economic development, the effectiveness of a longevity hedge in a sub-population depends on the economic and demographic proximities between the sub-population and the general population. This conjecture can be verified by considering the information contained in Figures 1 and 11 jointly. For the three geographical regions (Guangdong, Guangxi and Anhui) in which the longevity hedges are the most effective, the incomes and/or life expectancies were either close to or moving closer to the national averages. By contrast, the opposite is true for the three geographical regions (Jilin, Tianjin and Zhejiang) in which the longevity hedges are the least effective.

For the 30 sub-populations of China under consideration, the values of $HE_{30}$ range from 14% to 40%. These values are relatively low in comparison those of Zhou and Li (2014), who applied a similar longevity hedge to pension liabilities in 25 Western countries. The relatively low level of hedge effectiveness may be attributed to the extents of population basis risk and parameter risk (including that arising from the missing data) that we are confronting. To assess the extent of population basis risk, in Figure 12 we present the hedging results under the hypothetical situation that the annuitants’ mortality experience is the same as that of the general male population of China (the q-forwards’ reference population). By comparing Figures 11 and 12, we can infer the degrees of population basis risk that the longevity hedges in different sub-populations are subject to. In terms of $HE_{30}$, population basis risk erodes hedge effectiveness by 50 to 74 percentage points, depending on which sub-population the annuity liability is associated with. The extent of population basis risk is admittedly not small, despite a mortality index that is tailor-made for the general population of China is used. The issue of parameter risk will be studied in greater depth in the next sub-section.
5.6 A Decomposition of Risks

The full model used for generating the results in the previous sub-section incorporates various sources of risk, including trend risk (the uncertainty arising from \( \zeta(t) \) and \( \zeta(t, i) \)), model-error risk (the uncertainty arising from \( \epsilon(x, t) \) and \( \epsilon(x, t, i) \)) and parameter risk (the uncertainty in estimating the parameters in the Lee-Carter structure and the time-series processes). To better understand how they contribute to the erosion in hedge effectiveness, we now re-evaluate the longevity hedge by using restricted models in which some of the stochastic components are switched off. In particular, the following four scenarios with different levels of conservatism are considered.

Scenario 1: Trend risk only

This scenario in the most optimistic. In this scenario, it is assumed that the Lee-Carter structure captures mortality patterns perfectly and that all parameters are accurately and precisely estimated. The simulation procedure incorporates only the randomness arising from \( \zeta(t) \) and \( \zeta(t, i) \).

Scenario 2: Trend risk and trend-related parameter risk only

Compared to Scenario 1, this scenario is less optimistic. Other than the randomness arising from \( \zeta(t) \) and \( \zeta(t, i) \), we also consider in this scenario the uncertainty about the parameters that are associated with the time-series processes (i.e., parameters \( c, \phi_0(i), \phi_1(i), \sigma^2_\zeta \) and \( \sigma^2_\epsilon(i) \)).
Scenario 3: All but model-error risk

It is assumed in this scenario that the Lee-Carter structure – equations (6) and (7) – describes mortality patterns perfectly. The hedging results under this scenario are obtained by assuming $\sigma^2 = \sigma^2(i) = 0$ in the simulation procedure. All other sources of risk are retained.

Scenario 4: All sources of uncertainty

This scenario is the most conservative. The full simulation model that incorporates all sources of uncertainty is used. The results under this scenario are identical to those obtained in the previous sub-section.

As an example, we display the hedging results for Guangzhou under the four scenarios in Figure 14. Let us begin with the most optimistic view. If parameter and model-error risks are assumed to be non-existent, then there would be less uncertainty associated with both the future mortality rates of Guangzhou and the future mortality differentials between Guangzhou and the general population, leading to narrower fan charts for both the hedged and unhedged liabilities. Overall, the hedge effectiveness becomes significantly higher. The value of $HE_{30}$ under this scenario is close to 85%, which is more than 2 times that when all sources of uncertainty are taken into account. The results generated under this scenario are the most comparable with those of Cairns (2011) and Zhou and Li (2014) in which parameter and model-error risks were not considered.

We then move on to Scenarios 2 and 3. If we take a slightly less optimistic view by incorporating the portion of parameter risk that is related to the time-series processes for $k(t)$ and $k(t, i)$, then the value of $HE_{30}$ would reduce to approximately 68%, which is still substantially higher than that in the most conservative scenario. If we incorporate also the portion of parameter risk that is related to the rest of the model, then the value of $HE_{30}$ would drop to about 53%. We can interpret the difference (32 percentage points) between the values of $HE_{30}$ in Scenarios 1 and 3 to mean the impact of parameter risk on the effectiveness of the longevity hedge.

The only difference between Scenarios 3 and 4 is the incorporation of model-error risk, so the difference (13 percentage points) between the values of $HE_{30}$ in Scenarios 3 and 4 can be regarded as the erosion of hedge effectiveness due to model-error risk. The impact of model-error risk is significant, but is less significant than that of parameter risk.

6 Longevity Risk Solvency Capital under C-ROSS

The C-ROSS was first introduced by the China Insurance Regulatory Commission (CIRC) in 2012 to supersede the former Insurance Company Solvency Regulations (ICSR) established in 2008. The C-ROSS can be seen as the Chinese version of Europe’s Solvency II, in which regulations and capital requirements are emphasized on a risk-oriented system rather than on a factor-based system. Last year, some domestic and foreign insurance companies in...
China tried implementing C-ROSS when preparing their reserve calculations. In February 2015, the CIRC released the official version of C-ROSS.

Similar to Solvency II, C-ROSS adopted a regulatory framework with three pillars: Quantitative Capital Requirements, Qualitative Supervisory and Market Discipline Mechanism. In the first pillar, the calculation of the minimum capital requirement (MCR) for insurers’ quantifiable risks, namely insurance risk, market risk and credit risk, is explicitly specified using actual versus minimum capital assessment standards. In the second pillar, the Solvency Aligned Risk Management Requirements and Assessment (SARMRA) is introduced to evaluate insurers’ overall solvency level through an integrated risk rating system on qualitative risks including operational risk, strategic risk, reputational risk and liquidity risk. The third pillar imposes supervision of insurance companies from rating agencies, financial reports, media and the general public by enforcing risk disclosure, risk transparency and market disciplines.

In terms of longevity risk management, the C-ROSS classifies mortality and longevity risks as part of insurance risk, and explicitly specifies the calculation of the MCR for these risks. To ease exposition, in what follows we ignore insurance risks other than mortality and longevity risks. We use $MCR^M$ and $MCR^L$ to represent the C-ROSS MCRs for mortality and longevity risks, respectively.

For an unhedged insurance/annuity liability, we have

$$MCR^{(i)} = \max(V((1 + SF^{(i)}) m)) - V(m), \quad i = M, L,$$

where $V(\cdot)$ is the present value of all cash flows from the insurance/annuity liability evaluated at a certain mortality curve, $m$ is the best-estimate mortality curve for the duration of the liability, $SF^{(i)}$ is the adverse scenario factor. In C-ROSS, $SF^{(M)}$ is a parallel shock to the mortality curve reflecting the Value-at-Risk at a certain conservative confidence level:

$$SF^{(M)} = \begin{cases} \frac{10}{100}, & N > 200, \\ \frac{15}{100}, & 100 < N \leq 200, \\ \frac{20}{100}, & N \leq 100, \end{cases}$$
where $N$ denotes the number of contracts; and $SF^{(L)}$ is specified as follows:

\[
SF^{(L)} = \begin{cases} 
(1 - 3\%)^t - 1, & 0 < t \leq 5, \\
(1 - 3\%)^5(1 - 2\%)^{t-5} - 1, & 5 < t \leq 10, \\
(1 - 3\%)^5(1 - 2\%)^5(1 - 1\%)^{t-10} - 1, & 10 < t \leq 20, \\
(1 - 3\%)^5(1 - 2\%)^5(1 - 1\%)^{10} - 1, & t > 20,
\end{cases}
\]

where $t$ is the number of years after the assessment date. Finally, to calculate the MCR for both mortality and longevity risk, the following formula is used:

\[
MCR = \sqrt{\Sigma MM'},
\]

where

\[
M = (MCR^{(M)}, MCR^{(L)})
\]

and

\[
\Sigma = \begin{pmatrix} 1 & -0.25 \\ -0.25 & 1 \end{pmatrix},
\]

which is predetermined by the CIRC.

The portion labeled ‘the unhedged liability’ in Table 1 displays the time-0 values of $MCR^{(M)}$ and $MCR^{(L)}$ for the annuity liability defined in Section 5.5. The value of $m$ is taken as the best-estimate mortality curve implied by the assumed mortality model. Given how the liability is structured, the calculated value of $MCR^{(M)}$ is always 0 no matter what value of $N$ is assumed. Because $MCR^{(M)} = 0$, $MCR = MCR^{(L)}$ and the correlation matrix $\Sigma$ is not involved in the calculations. In general, when the insurer’s portfolio contains both insurance and annuity liabilities, the value of $MCR^{(M)}$ is not necessarily 0.

Similarly, the values of $MCR^{(M)}$ and $MCR^{(L)}$ for a portfolio of hedging instruments is given by

\[
MC^{(i)} = \max(H((1 + SF^{(i)}(m)) - H(m), 0), \quad i = M, L,
\]

where $H$ is the present value of all cash flows from the portfolio of hedging instruments, evaluated at the best-estimate mortality curve $m$. Finally, for a hedged portfolio, we have

\[
MC^{(i)} = \max(V((1 + SF^{(i)}(m)) - V(m) - H((1 + SF^{(i)}(m)) + H(m), 0), \quad i = M, L.
\]

The specifications of $SF^{(M)}$, $SF^{(L)}$ and $\Sigma$ also apply to the calculations of MCRs for a portfolio of hedging instruments and a hedged portfolio.

The calculated time-0 values of $MCR^{(M)}$ and $MCR^{(L)}$ for the calibrated q-forward portfolio (with $N = 1$ contract) and the hedged annuity liability are shown in Table 1. Note that the q-forward portfolio would incur a loss only if future mortality turns out to be higher than expected, so its value of $MCR^{(L)}$ is always 0. Overall, the longevity hedge removes all $MCR^{(L)}$ from the annuity portfolio, but introduces some $MCR^{(M)}$ to it.

In Figure 14 we show the percentage of MCR reduced by the longevity hedge. Depending on the geographical location of the annuity liability, the percentage reduction in MCR ranges from 69% to 74%. Our results indicate that a standardized longevity hedge can significantly reduce a Chinese insurer’s required capital. The results in this section make a strong case for launching a standardized tradable mortality index in China, despite such an index does not lead to an impressive reduction in portfolio variance when all sources of risk are taken into account.

7 Concluding Remarks

China’s rapidly shifting demographics have created a huge potential demand for life annuities, which offer individuals a protection against longevity risk. However, due to the systematic nature of longevity risk, there is a limit to which insurers can accept the risk. To maintain insurers’ ability to sell life annuities at affordable prices and to reduce the risk of a systemic failure of the insurance industry due to an excessive exposure to longevity risk, it has been suggested to develop markets for standardized mortality-linked securities, through which longevity risk can be transferred from insurers to capital market investors. In this paper, we study the possibility of developing a market for standardized mortality-linked securities in China, with a focus on four issues that largely determine the success or failure of such a market.

The first issue is the creation of a Chinese national mortality index, upon which standardized mortality-linked securities can be written. We qualitatively evaluate several possible methods for creating mortality indexes, including non-parametric aggregate, non-parametric age-specific and parametric. We argue that for a number of reasons, non-parametric aggregate indexes such as life expectancies at birth and aggregate death rates are not the most suitable for hedging purposes, despite they have a long history for market participants to track. Non-parametric
The unhedged liability

<table>
<thead>
<tr>
<th>Region</th>
<th>2023 MCR</th>
<th>2024 MCR</th>
<th>2021 MCR</th>
<th>2025 MCR</th>
<th>2026 MCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td>0.6538</td>
<td>0.6999</td>
<td>0</td>
<td>0.1757</td>
<td>0</td>
</tr>
<tr>
<td>Tianjin</td>
<td>0.6711</td>
<td>0.7277</td>
<td>0.1837</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hebei</td>
<td>0.7196</td>
<td>0.8051</td>
<td>0.2104</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shanxi</td>
<td>0.7212</td>
<td>0.8063</td>
<td>0.2106</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>0.7325</td>
<td>0.8252</td>
<td>0.2180</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Liaoning</td>
<td>0.7083</td>
<td>0.7842</td>
<td>0.2029</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jilin</td>
<td>0.7051</td>
<td>0.7777</td>
<td>0.1995</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>0.7136</td>
<td>0.7908</td>
<td>0.2040</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.6402</td>
<td>0.6794</td>
<td>0.1681</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>0.7043</td>
<td>0.7765</td>
<td>0.2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>0.6912</td>
<td>0.7567</td>
<td>0.1930</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anhui</td>
<td>0.7217</td>
<td>0.8079</td>
<td>0.2112</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fujian</td>
<td>0.7188</td>
<td>0.8054</td>
<td>0.2124</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>0.7324</td>
<td>0.8242</td>
<td>0.2168</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shandong</td>
<td>0.7087</td>
<td>0.7853</td>
<td>0.2033</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Henan</td>
<td>0.7280</td>
<td>0.8186</td>
<td>0.2149</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hubei</td>
<td>0.7261</td>
<td>0.8157</td>
<td>0.2155</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hunan</td>
<td>0.7289</td>
<td>0.8187</td>
<td>0.2150</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Guangdong</td>
<td>0.7096</td>
<td>0.7872</td>
<td>0.2042</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Guangxi</td>
<td>0.7297</td>
<td>0.8217</td>
<td>0.2163</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hainan</td>
<td>0.7193</td>
<td>0.8033</td>
<td>0.2101</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sichuan</td>
<td>0.7292</td>
<td>0.8211</td>
<td>0.2170</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Guizhou</td>
<td>0.7559</td>
<td>0.8675</td>
<td>0.2326</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yunnan</td>
<td>0.7635</td>
<td>0.8813</td>
<td>0.2371</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tibet</td>
<td>0.7689</td>
<td>0.8910</td>
<td>0.2407</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>0.7251</td>
<td>0.8112</td>
<td>0.2120</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gansu</td>
<td>0.7410</td>
<td>0.8408</td>
<td>0.2231</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Qinghai</td>
<td>0.7589</td>
<td>0.8731</td>
<td>0.2345</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ningxia</td>
<td>0.7334</td>
<td>0.8284</td>
<td>0.2185</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>0.7463</td>
<td>0.8487</td>
<td>0.2255</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: The values of $MCR^{(M)}$ and $MCR^{(L)}$ for the unhedged liability, the q-forward portfolio and the hedged liability, on the basis of projected mortality rates for different provinces, municipalities and autonomous regions of China.

The second issue is the problem of population basis risk, which concerns hedgers because the mortality experience of the national population may be different from that of the individuals associated with their own portfolios. To address this problem, we consider a multi-population stochastic mortality model for all provinces, municipalities and autonomous regions of China, except Chongqing which has a rather short history. The model captures the mortality dynamics of all sub-populations jointly, thereby allowing one to quantify the amount of population basis risk that a longevity hedge based on a national mortality index is subject to.

The third issue is the achievable level of hedge effectiveness, which is related to the previous two issues as it depends on how well the standardized index can capture the evolution of mortality and the extent of population basis risk. By adapting the work of Zhou and Li (2014), we formulate a dynamic hedging strategy and examine how much of a stylized life annuity portfolio’s variance can be eliminated when the hedging strategy is implemented. The resulting hedge effectiveness (measured by the proportion of variance eliminated) varies considerably among different sub-populations of China, ranging from 14% to 40%. The effectiveness of a longevity hedge in a sub-population appears to depend on the economic and demographic proximities between the sub-population and the general population.

The fourth issue is the amount of capital relief that can potentially be obtained from an index-based longevity hedge under C-ROSS. To address this issue, we compare the C-ROSS MCRs for insurance risks when a longevity hedge is absent and when the longevity hedging strategy we formulate is implemented. It is found that the percentage reduction in MCR ranges from 69% to 74%, depending on the geographical location of the annuity liability. Despite being subject to population basis risk, longevity hedges created from a Chinese national mortality index still yield significant reductions in C-ROSS MCR, suggesting that the development of a market for standardized

age-specific and parametric indexes may better capture the evolution of mortality at pensionable ages and hence more suitable for hedging the longevity risk associated with annuity liabilities, but they are subject to problem of having only a small number of historical observations.

The second issue is the problem of population basis risk, which concerns hedgers because the mortality experience of the national population may be different from that of the individuals associated with their own portfolios. To address this problem, we consider a multi-population stochastic mortality model for all provinces, municipalities and autonomous regions of China, except Chongqing which has a rather short history. The model captures the mortality dynamics of all sub-populations jointly, thereby allowing one to quantify the amount of population basis risk that a longevity hedge based on a national mortality index is subject to.

The third issue is the achievable level of hedge effectiveness, which is related to the previous two issues as it depends on how well the standardized index can capture the evolution of mortality and the extent of population basis risk. By adapting the work of Zhou and Li (2014), we formulate a dynamic hedging strategy and examine how much of a stylized life annuity portfolio’s variance can be eliminated when the hedging strategy is implemented. The resulting hedge effectiveness (measured by the proportion of variance eliminated) varies considerably among different sub-populations of China, ranging from 14% to 40%. The effectiveness of a longevity hedge in a sub-population appears to depend on the economic and demographic proximities between the sub-population and the general population.

The fourth issue is the amount of capital relief that can potentially be obtained from an index-based longevity hedge under C-ROSS. To address this issue, we compare the C-ROSS MCRs for insurance risks when a longevity hedge is absent and when the longevity hedging strategy we formulate is implemented. It is found that the percentage reduction in MCR ranges from 69% to 74%, depending on the geographical location of the annuity liability. Despite being subject to population basis risk, longevity hedges created from a Chinese national mortality index still yield significant reductions in C-ROSS MCR, suggesting that the development of a market for standardized
longevity risk transfers does worth Chinese insurers’ attention.

To our knowledge, this study represents the first attempt to investigate the possibility of developing a market for standardized mortality-linked securities in China. We believe that many areas in this topic deserve further research. In terms of indexation, it would be interesting to study the benefit of creating separate mortality indexes for certain groups of sub-populations. For instance, given that cities such as Beijing, Shanghai and Tianjin have incomes and life expectancies that are far away from the national averages, it may be worthwhile to create a separate mortality index for these cities. Along this line, it would also be interesting to study how population basis risk may average out should an insurers’ portfolio contain individuals located in different parts of China.

In terms of hedging strategies, future research warrants a further extension of the work of Cairns (2011) and Zhou and Li (2014) to incorporate not only life annuities but also life insurances. Such an extension is important, because at present Chinese insurers generally run both life insurance and life annuity lines. With this extension, one can gauge the effect of natural hedging and hence determine the need for transferring the residual longevity risk to capital markets. Another direction of future research is to consider more sophisticated hedging strategies, including ‘delta-gamma’ hedging which incorporates not only the first but also the second order partial derivatives with respect to $k(t)$.

Finally, it terms of modeling, it is warranted to study the factors affecting the improvement of Chinese mortality and the reasons behind the observed convergence of some sub-populations’ life expectancies to the national average. The outcome of such a study may shed light on how a multi-population stochastic mortality model for different sub-populations of China may be better structured. Also, with a better understanding about the drivers behind a Chinese national mortality index, capital market investors may feel more comfortable to trade securities written on the index.

Acknowledgements

The authors are grateful to Yiran Cao for her research assistance. The authors would also like to thank the Asia-Pacific Mortality Database in Nanyang Technological University’s Insurance Risk and Finance Research Centre for their provision of part of the mortality data used in this study. This work is financially supported by the Global Risk Institute, the Society of Actuaries Center of Actuarial Excellence Programme and the Natural Sciences and
References


