

SFB 649 Discussion Paper 2015-007

Stochastic Population Analysis: A Functional Data Approach

Lei Fang*
Wolfgang K. Härdle**



*Singapore Management University, Singapore
**Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

Stochastic Population Analysis: A Functional Data Approach*

LEI FANG[†]

WOLFGANG K. HÄRDLE[‡]

Humboldt-Universität zu Berlin

Singapore Management University

Abstract

Based on the Lee-Carter (LC) model, the benchmark in population forecasting, a variety of extensions and modifications are proposed in this paper. We investigate one of the extensions, the Hyndman-Ullah (HU) method and apply it to Asian demographic data sets: China, Japan and Taiwan. It combines ideas of functional principal component analysis (fPCA), nonparametric smoothing and time series analysis. Based on this stochastic approach, the demographic characteristics and trends in different Asian regions are calculated and compared. We illustrate that China and Japan exhibited a similar demographic trend in the past decade. We also compared the HU method with the LC model. The HU method can explain more variation of the demographic dynamics when we have data of high quality, however, it also encounters problems and performs similarly as the LC model when we deal with limited and scarce data sets, such as Chinese data sets due to the substandard quality of the data and the population policy.

JEL classification: C14, C32, C38, J11, J13

Keywords: Functional principal component analysis; Nonparametric smoothing; Mortality forecasting; Fertility forecasting; Asian demography; Lee-Carter model, Hyndman-Ullah method

*The authors gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft through the International Research Training Group IRTG 1792 "High Dimensional Non Stationary Time Series" and the Collaborative Research Center CRC 649 "Economic Risk".

[†]Corresponding author. Humboldt-Universität zu Berlin, C.A.S.E. - Center for Applied Statistics and Economics, Unter den Linden 6, 10099 Berlin, Germany. E-mail: fanglei@hu-berlin.de

[‡]Humboldt-Universität zu Berlin, C.A.S.E. - Center for Applied Statistics and Economics, Unter den Linden 6, 10099 Berlin, Germany. Visiting Professor in Sim Kee Boon Institute for Financial Economics, Singapore Management University, 90 Stamford Road, 6th Level, School of Economics, Singapore 178903.

I. INTRODUCTION

IN recent years, population forecasting has received wide-spread attention because we are in a fast-aging society which raises demographic risk in most developed and even some developing countries. Here demographic risk is understood to be an imbalance of the age distribution of a society with the obvious implication of economic growth, social stability, political decisions and resource allocation. The factor, demography, is in particular important for the Asian region since Asia is in a process of continuous economic growth and during which time the age structure of the population could affect the employment, the labour force and even social and economic stability.

China, as a large developing Asian country, is experiencing the transformation to an aging society at an even faster rate, and is therefore a good example with which to study demographic risk. However, due to political reasons and a low-speed national statistical construction system dating from the last century, statisticians always meet the problem of insufficient and unsatisfying Chinese demographic data sets when they apply stochastic demographic models; accordingly it brings about lots of research passion in this field. Japan, China's neighbour, has demonstrated a dramatic demographic change during the last several decades; but fortunately the Japanese government had already set up a complete national statistical system (middle of last century) and thus could provide qualified demographic data sets in longer time horizons to help explore Japan's demographic transition, which can thus be analyzed as a good reference to China as well. Taiwan has also encountered serious demographic problem in last decade, a shrinking teenaged population creating potential risk in Taiwan's education system can be a good example. Declining fertility rates in Taiwan could also not guarantee enough university students, which means that demand for professorships will be adjusted accordingly and higher unemployment rates for academics will result. Similar problems in other industries will also arise.

Since 1980, one of the challenges in demography is to analyze and forecast mortality and fertility in a purely statistical way without involving the subjective opinions of experts. Lee and Carter (1992) firstly proposed the stochastic method with the Singular Value Decomposition technique to explore the unobserved demographic information, which proved insightful and gained a good reputation. In addition, several methods based on stochastic population modelling and forecasting had already been developed, see e.g. Cairns, Blake and Dowd (2008), Tickle (2008) and Booth (2006) for review. Among these methods, they are divided into three approaches to forecast the demographic dynamic movement: extrapolation, expectation and theory-based structural modelling involving exogenous variables. The extrapolation approach is based on the assumption that to some extent the future will repeat the past, and therefore the commonly-used ARIMA modelling is employed. The methods based on expectation make use of the data on the expectations on an individual level and data on the expectations of a group of experts on a whole-population level. A subjective judgement is sometimes involved in this approach. Apart

from the fact that these two approaches that mainly concentrate on the demographic data, the structural models differ in considering the exogenous variables and explaining the demographic rates in terms of the underlying socio-economic variables.

Among all the stochastic models, the most popular one is the Lee and Carter (1992) model, which was used to analyze the U.S. mortality rate from 1933 to 1987. Based on the idea of the LC model, comparisons of different methods and some variants or extensions are developed. Lee and Miller (2001) compared the forecasts of LC model with the U.S. social security system forecasts. Li, Lee and Tuljapurkar (2004) proposed another method when there are few observations at uneven intervals, and applied it to China and South Korea. Hyndman and Ullah (2007) developed a more general method by treating the underlying demographic process as functional data, employing the functional principal components to extract more than one explaining components and providing robust estimation and forecast. The Hyndman-Ullah (HU) technology will be discussed in depth in the main part of this paper.

In light of limited data access combined with fragile quality, less technically refined methods are available for Asian countries compared to for example developed western countries. An exception is the stochastic population approach on Asian data by Li, Lee and Tuljapurkar (2004), who implemented the LC model to sparse data. In their work, they generated the central forecast with just the first and last observations, improved the estimates by additional observations and evaluated its performance with other existing methods. Raftery, Li et al. (2012) proposed a Bayesian method for probabilistic population projections for all countries, where the Bayesian hierarchical models, estimated via Markov chain Monte Carlo, are applied to the United Nations population data. In cases of limited data and similar demographic trends between two populations (regional or national level), the Bayesian stochastic modelling for two populations is proposed by Cairns, Blake et al (2011). This will provide the insights on analyzing Chinese demography via taking Japan into reference.

In this research, we apply the LC and the HU methods to Asian data sets, evaluate the performance of these two methods, compare the similarities of different Asian regions and accordingly propose potential improvements. In the following section, we will discuss the LC and the HU methods in more details. Section 3 focuses on a descriptive analysis of demographic data and the empirical research of Asian data sets, China, Japan and Taiwan for instance. We will further compare the forecasting accuracy between different methods, check the demographic development modes in these Asian regions, and dig out the similarities in an Asian demographic general trend in section 4, which will provide the instructive suggestions to the governments and hints for future research. More discussions on the limitations of these two methods on Asian demographic data and potential directions for improvement will be organised in the last section.

II. METHODS

In this section, we firstly introduce the parameters of interests and then outline the LC and HU methods respectively in details. Furthermore, some extensions based on these methods are covered afterwards.

2.1 Notations and Parameters of Interest

In demography research, three factors are of importance for population forecasting: fertility, mortality and migration. In this research, we concentrate on Asian areas where migration can be assumed not to be a crucial factor, in addition migration data is hardly available; therefore we will only concentrate on the factors of fertility and mortality. The parameters of interests are age-specific fertility rate, age-gender-specific mortality rate and age-gender-specific population counts. We use the symbols f , m and N to denote the fertility rate, mortality rate and population counts. All the parameters are indexed by a one-year age group, denoted by x , and in addition indexed by time, denoted by t . For instance, $m(x, t)$ is the mortality rate for age x in year t while $f(x, t)$ is the fertility rate for age x in year t .

2.2 Lee-Carter (LC) Model

The benchmark LC model employs the Singular Value Decomposition (SVD) to analyse the time series on the log of the age-specific mortality and fertility rates. The method relies on the standard statistical analysis of the time series and thereby ignores the experts' subjective judgment, for example the expectation of next year's fertility rate, which is quite important in traditional demographic forecasting. Nonetheless, the LC model does not fit well in some cases where missing data is common or the horizon of time series is not sufficient, the reason being the assumption of long-term stationarity.

The basic idea is to regress demographic indicators like $m(x, t)$, $f(x, t)$ on non-observable regressors for prediction. The regressors are obtained via SVD of the demographic indicators. It separates the age pattern from the time-dependent components, takes time series analysis on the time-dependent components only and hence forecasts the future trend. Let us take the mortality rate as an example for explanation.

The mortality rate $m(x, t)$ is calibrated via the following model:

$$\log\{m(x, t)\} = a_x + b_x k_t + \varepsilon_{x,t}, \quad (1)$$

or

$$m(x, t) = \exp(a_x + b_x k_t + \varepsilon_{x,t}), \quad (2)$$

where $m(x, t)$ denotes the historical mortality rate at age x in year t , a_x is the derived age pattern averaged across years, b_x stands for the sensitivity of the mortality rates to the change of k_t , reflecting how fast the mortality rate changes over ages, k_t represents the only time-varying index of mortality level, and $\varepsilon_{x,t}$ is the residual term at age x in year t with $E(\varepsilon_{x,t}) = 0$ and $\text{Var}(\varepsilon_{x,t}) = \sigma_\varepsilon^2$.

Three unobserved parameters a_x , b_x and k_t in the single equation (1) means that the LC model is over-parameterized and therefore two normalisation constraints are imposed:

$$\sum k_t = 0, \sum b_x = 1 \quad (3)$$

By SVD, one obtains the parameters k_t and b_x . The parameter k_t can be fitted by the standard ARIMA models, like the Box-Jenkins procedure. After the model identification procedures, Lee and Carter (1992) found out a random walk with drift model usually describes k_t quite well:

$$k_t = k_{t-1} + d + e_t \quad (4)$$

where d is the drift parameter reflecting the average annual change and e_t is an uncorrelated error. Other research taken by Chan, Li and Cheung (2008) found out that the mortality time-varying index k_t can be better fitted with the trend-stationary model for Canada, England and United States when they have a break in mortality rate decline during the 1970s.

Given the h -step ahead forecasting k_{t+h} , we could, in return, forecast the mortality rates in future period $t+h$ via the following formula:

$$m(x, t+h) = \exp(a_x + b_x k_{t+h}). \quad (5)$$

2.3 Hyndman-Ullah (HU) Method

A number of variants and extensions of the LC model exist. One of them is the HU method; it employs functional data analysis techniques on log mortality rates and fertility rates, combined with a nonparametric presmooth. Thereby the HU method with three steps is a generalisation of the LC model and applies widely to lots of time series data. The basic idea is to represent the demographic variables as:

$$y_i(x, t) = s_i(x, t) + \sigma_i(x, t)\varepsilon_{i,t}, \quad (6)$$

$$s_i(x, t) = \mu_i(x) + \sum_{k=1}^K \beta_{i,t,k} \phi_{i,k}(x) + e_i(x, t). \quad (7)$$

Here $y_i(x, t)$ represents the generic variables modelled, like mortality rates or fertility rates at age x in year t , $s_i(x, t)$ is the derived smooth function of x whereas the $\sigma_i(x, t)$ represents the

smooth volatility function, and $\varepsilon_{i,t}$ is the standardised identically independent distributed error term. Similar as a_x in LC model representing the age pattern over years, $\mu_i(x)$ is a measure of the location of $s_i(x, t)$ across years. $\{\phi_{i,k}(x)\}$ is a set of orthonormal basis functions like b_x in LC model reflecting the sensitivity to time-varying index over ages, while $\{\beta_{i,t,k}\}$ are the coefficient functions similar as k_t of the only time-varying index.

For simplicity, we ignore the group index i , like the country or region index and gender index, in the following analysis, and rewrite $y(x, t)$ as $y_t(x)$ and so on.

Step 1: Pre-smoothing

Estimate the smooth functions $s_t(x)$ through the data sets $\{x, y_t(x)\}$ for each t with a nonparametric smoothing method.

$$y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_t$$

For an overview on nonparametric smoothing techniques see e.g. Härdle et al. (2004). The HU approach applied weighted penalised regression spline. For each fixed t ,

$$\hat{s}_t(x) = \underset{\theta_n(x)}{\operatorname{argmin}} \sum_{n=1}^N |y_n(x) - \theta_n(x)| + \lambda \sum_{n=1}^{N-1} |\theta'_{n+1}(x) - \theta'_n(x)|, \quad (8)$$

where n stands for different grid points, while the L_1 loss function and the L_1 – roughness penalty are employed to obtain robust estimates.

The structure of the residual term $\sigma_t(x)\varepsilon_t$ in (6) determines that the weight as the inverse standard deviation $\sigma_t^{-1}(x)$ should be imposed on the loss function.

The mortality rate is approximately binomial distributed:

$$M_t(x) \sim B[N_t(x), m_t(x)], \quad (9)$$

where $M_t(x)$ is the death number of age x in year t .

Thus the variance of the mortality is $\operatorname{Var}[m_t(x)] = N_t^{-1}(x)m_t(x)\{1 - m_t(x)\}$, and the variance of $y_t(x) = \log\{m_t(x)\}$ is obtained by Taylor approximation as $\hat{\sigma}_t^2(x) \approx \{1 - m_t(x)\}N_t^{-1}(x)m_t^{-1}(x)$.

Therefore the standard deviation for mortality is approximated as:

$$\hat{\sigma}_t^2(x) \approx \{1 - m_t(x)\}N_t^{-1}(x)m_t^{-1}(x), \quad (10)$$

where $m_t(x)$ denotes the mortality rate and $N_t(x)$ denotes the total population of age x in year t .

Similarly the standard deviation of fertility is approximated as

$$\hat{\sigma}_t^2(x) \approx \{1000 - f_t(x)\} N_t^{-1}(x) f_t^{-1}(x) \quad (11)$$

where $f_t(x)$ denotes the fertility rate per thousand women of age x in year t .

In addition, a monotonicity constraint is attached to the mortality rate when $x > c$ for some c , like 50 years old. This monotonic increasing constraint helps to reduce the noise from estimation of high ages and meanwhile makes sense that elderly people are of higher risk of death (Wood, 1994). The concavity constraint is adopted to fertility according to the fertility trend graphs which show the properties of the concave function (He and Ng, 1999).

Step 2: Functional principal component analysis

Approximate the bivariate surface $\{s_t(x) - \mu(x)\}$ through the sum of several products of the orthonormal basis functions $\{\phi_k(x)\}$ of age x and coefficients $\{\beta_{t,k}\}$ of time t .

$$s_t(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x)$$

The basis functions can be obtained by principal components, and uncorrelated coefficients is also easily produced via PCA. For a given K , the basis functions $\{\phi_k(x)\}$ will minimise mean integrated squared error (MISE):

$$MISE = n^{-1} \sum_{t=1}^n \int e_t^2(x) dx$$

In the HU model, robust estimation is also proposed so that they use the L_1 -median of the estimated smooth curves $\{\hat{s}_1(x), \dots, \hat{s}_n(x)\}$ for $\mu(x)$:

$$\hat{\mu}(x) = \operatorname{argmin}_{\theta(x)} \sum_{t=1}^n \|\hat{s}_t(x) - \theta(x)\|$$

where $\|g(x)\| = \{\int g^2(x) dx\}^{1/2}$ denotes the L_2 norm of g .

Weighted functional principal component analysis is applied over the $\{\hat{s}_t^*(x)\}$, where $\hat{s}_t^*(x) = \hat{s}_t(x) - \hat{\mu}(x)$ is median-adjusted data. The weight is obtained through PCA by projection pursuit, and it is determined that the outlying years will be assigned with weight 0, otherwise 1.

The weight is calculated by two steps, for details we refer to Hyndman and Ullah (2007).

Step 3: Forecasting

Due to the way the basis functions $\phi_k(x)$ are chosen, the coefficients $\hat{\beta}_{t,k}$ and $\hat{\beta}_{t,l}$ are uncorrelated for $k \neq l$. Therefore we can use univariate time series model, such as ARIMA, to forecast this time-varying scores $\beta_{t,k}$.

Combining the functions (6) and (7):

$$y_t(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x) + \sigma_t(x) \varepsilon_t \quad (12)$$

The h -step ahead forecast of $y_{n+h}(x)$:

$$\hat{y}_{n,h}(x) = E[y_{n+h}(x)] = \hat{\mu}(x) + \sum_{k=1}^K \tilde{\beta}_{n,k,h} \hat{\phi}_k(x) \quad (13)$$

where $\tilde{\beta}_{n,k,h}$ is the h -step ahead forecast of $\beta_{n,k,h}$ via the estimates of $\hat{\beta}_{1,k}, \dots, \hat{\beta}_{n,k}$.

Based on equation (17), the forecast variance is derived as the simple sum of the each component variances due to the approximate orthogonality among these components (Hyndman and Ullah, 2007).

In the presence of outliers, missing data or structural break of time series, the performance of the HU method will be influenced. To model properly and improve the forecasting performance, other extension based on the HU method are proposed, like the Robust HU Method (Hyndman and Ullah, 2007) and Weighted HU Method (Hyndman and Shang, 2009) for instance. For comparisons of these extensions and the LC model, Shang, Booth and Hyndman (2011) analysed ten principal component methods and concluded that the weighted HU method provides the most accurate point and interval forecasts.

III. EMPIRICAL RESULT

In this section we will apply the LC model and the HU method to some Asian regions including China (restricted to the mainland districts), Japan and Taiwan, and define the properties of demographic indicators. After the empirical section, we will continuously compare these two methods and explore the demographic similarities among these regions.

3.1 Data Sets

The demographic data sets are collected from different sources: China data sets are extracted from the China Statistical Year Book, Japan and Taiwan data sets are attained from the Human Mortality Database (HMD) and the Human Fertility Database (HFD). For data sets on China the mortality rates are gender-age specific starting from 0 to 90+ years old, while the mortality rates of Japan and Taiwan are gender-age specific from 0 to 110+ since the Human Mortality Database makes estimates and adjustments on the raw data to extend into wider age group. Similarly the fertility rates in China are age specific from 15- to 49 years old and those from Japan and Taiwan are from 12- to 55+.

As for the sample size, they are all different as well: China mortality data spans from 1994 to 2010 with missing data of years 1996,1997,2001 and 2006, while the fertility data is from 1990 to 2011 with years 1992-1994, 1997 and 2002 missing; Japan mortality data ranges from 1947 to 2012 and the fertility is from 1947 to 2009; for Taiwan, mortality data is available from 1970 to 2010 and fertility is from 1976 to 2010.

To be more clear, the mortality rate is the number of deaths per 1000 living individuals per single calender year, and the fertility rate is the number of births per 1000 women at the same age per single calender year. To make the mortality trend more precise and easy-looking, we take the log scale of all the original mortality data sets.

3.2 Descriptive Analysis

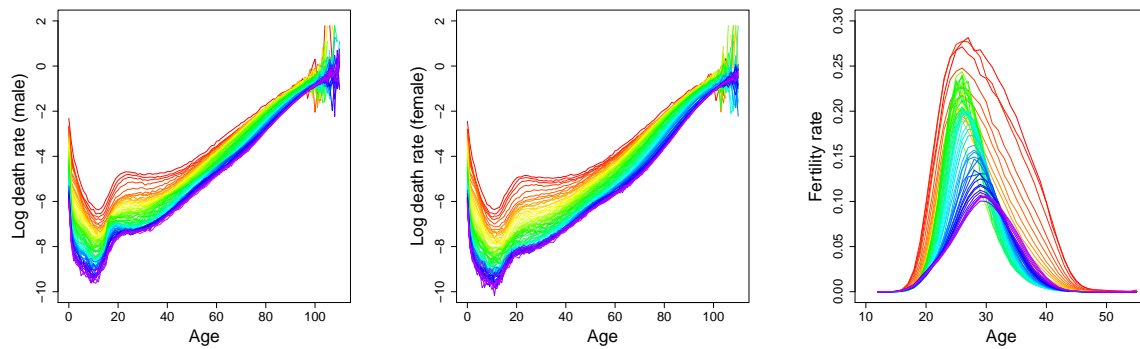


Figure 1: Japan's descriptive demography - male (left) and female (middle) mortality dynamics from 1947 to 2012, fertility (right) dynamics from 1947 to 2009, rates in different years are plotted in rainbow palette order.

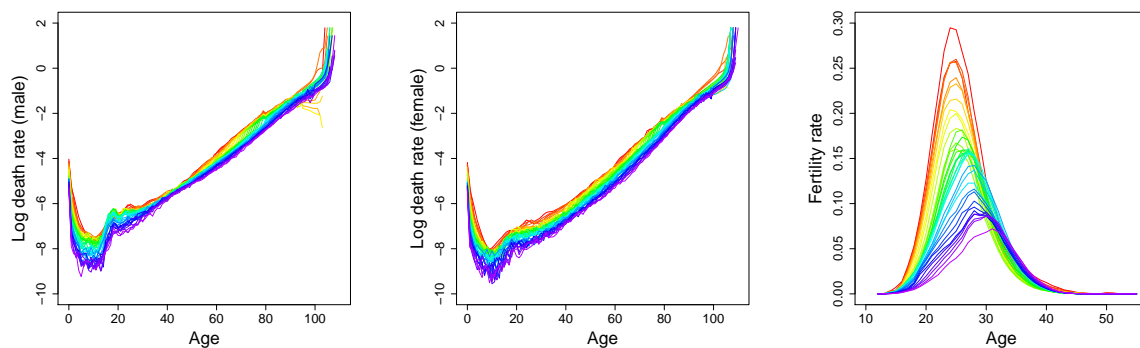


Figure 2: Taiwan's descriptive demography - male (left) and female (middle) mortality dynamics from 1970 to 2010, fertility (right) dynamics from 1976 to 2010, rates in different years are plotted in rainbow palette order.

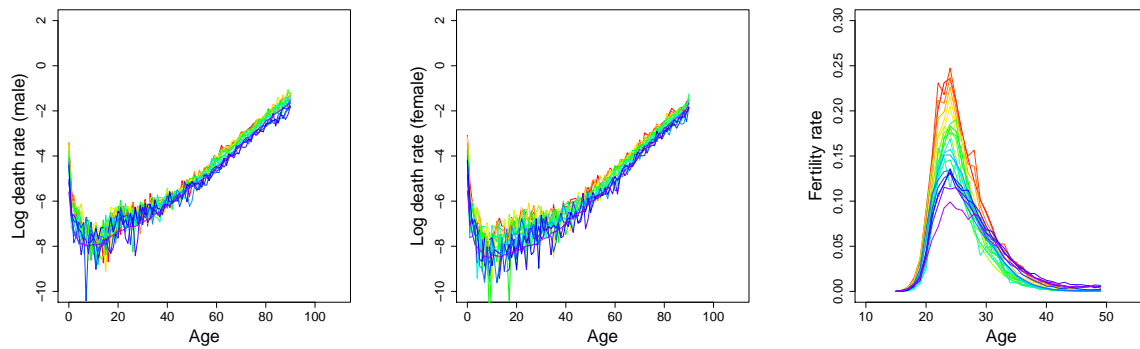


Figure 3: *China's descriptive demography - male (left) and female (middle) mortality dynamics from 1994 to 2010, fertility (right) dynamics from 1990 to 2011, rates in different years are plotted in rainbow palette order.*

We demonstrate the dynamic movement of mortality rates and fertility rates over ages. Each line stands for a fixed year and rates in different years are plotted in a rainbow palette so that the earliest years are red and so on. As displayed in the Fig. 1, Fig. 2 and Fig. 3, Japan and Taiwan both experienced dramatic demographic change, and similar phenomenon from China could be partly revealed by the limited data set as well. Generally, the mortality rates experienced an extreme high value at early ages around 0-1 year old, dropped at a low ebb at 8-10 years old and kept increasing as age continued. The mortality lines meanwhile shifted lower over years, which reflects the improvement of social and medical care conditions. The fertility rates arrived at the peak around 25-30 years old, and the peak points moved to elderly ages in recent years. The fertility rate lines shifted closely to the x-axis dramatically. All these findings confirm our belief that fertility is decreasing and that high-aged baby-bearing is more common.

Specifically, Japan's mortality rates decreased gradually, while the fertility increased for several years after the World War II but declined quite dramatically during the economy boom period. As time moves on, the peak of bearing age shifted to the elderly level, around 30 years old in recent years. Taiwan presents the general decreasing mortality and fertility trends within a shorter time interval, and it also applies to China even with very limited sample size, but more women in China give birth at a relatively younger age compared to Japan and Taiwan.

The obvious difference among them is that the China's data sets contain more wiggly data in a shorter age range, because the data sets of Japan and Taiwan are obtained from Human Mortality Database and Human Fertility Database where the smoothing technique and interpolation method are applied on the original data.

At older ages, the sample size is usually smaller than that of younger ages. Therefore the independently derived number of deaths sometimes becomes larger than the number of exposure-to-risk, eventually resulting in observed log death rates higher than 0. To neglect the influence from estimation of older ages, we only analyse the mortality data under the age of 100 years old

in our research.

3.3 Empirical Findings

In this section, the demographic trend of Japan, Taiwan and China will be analysed respectively. Due to the sample size, only the full-picture demonstration of Japan is displayed and some core parts of Taiwan and China will be illustrated as well. All the other relevant graphs are listed in the Appendix.

3.3.1 Japan

Japan, as a developed country in Asia, has a similar demographic trend as other developed countries: low mortality and fertility rates, increasing baby births of high-aged women and continuously-growing aging risk. To understand the situation better, we take the Japanese female mortality and fertility as an example, apply LC model and HU method and display the plots in Figures 4-10.

Mortality

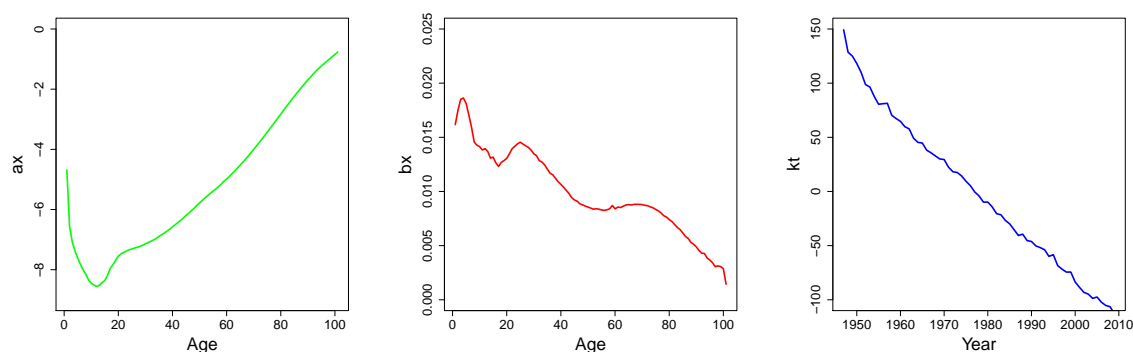


Figure 4: *Japan's female mortality decomposition by LC model*

The general trends in Japan's female mortality rates, decomposed by the LC model and the HU method, are shown in Fig. 4 and Fig. 5 respectively. These figures reveal similar trends explored by the two methods; particularly the first basis function and corresponding coefficients of the HU methods support the findings of the LC model. The variation of Japan's female mortality rates from 1947 to 2009 can be explained 96.1% by the LC model and 99.9% by the HU method with 96.5% from the first component and 3.1% from the second one. From this point of view, the LC model with less parameters is quite convincible compared to HU method. From the general trend of both figures, we can see that the death rates are a little higher at the infant-period of 0-1 age

group and lowest at teenage group of 14-15 years old. After teenage period, the mortality risk monotonically increased as people get older, which of course meets the general expectation. The first component in the graph mainly explains the impact of younger-age groups on the mortality rates, and the corresponding coefficients have declined over past years, which suggests that medical care is gradually improving. Similarly the second component can be viewed as the factor of elderly-age groups, and the coefficients turned into a decreasing trend after 1970s.

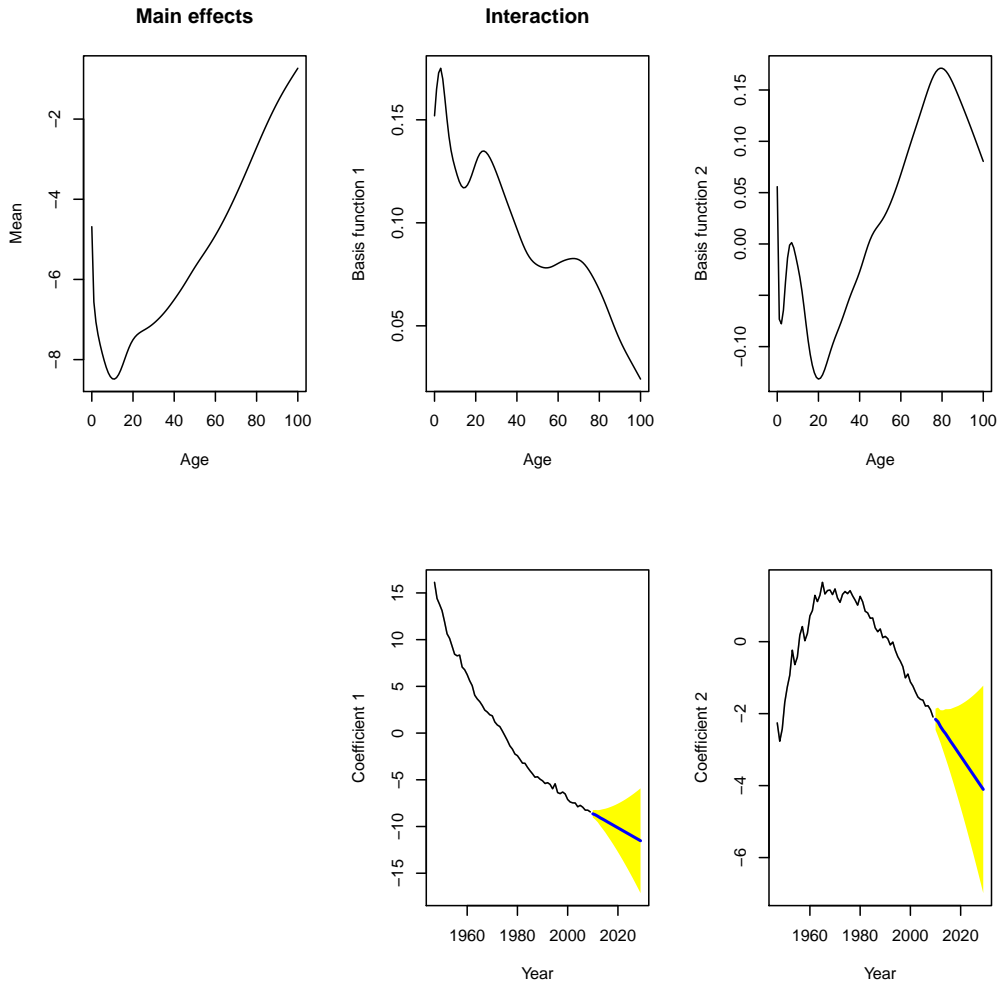


Figure 5: Japan's female mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

The corresponding results for Japan's male mortality rates are given in Fig. 14 and Fig. 15 from the Appendix.

Based on the HU method, we take the out-of-sample test on Japan's female mortality rates from 1947 to 1989, and compare the accuracy of 1-step ahead, 10-step ahead and 20-step ahead forecasts.

The comparisons between forecast rates and actual rates are plotted in Fig. 6. It is clearly shown that the 95% confidence intervals under these three forecast horizons can cover the actual rates rather well, and the forecast rates also exhibits the properties of the actual ones.

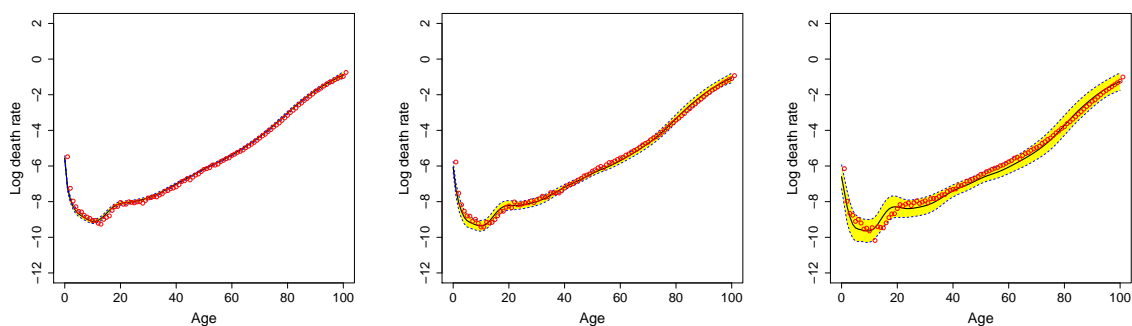


Figure 6: Out-of-sample test on Japan's female mortality (1947-1989): forecast rates (black lines) for 1990 (left), 1999 (middle), 2009 (right), along with 95 % confidence intervals (yellow areas), while actual rates are shown as red circles.

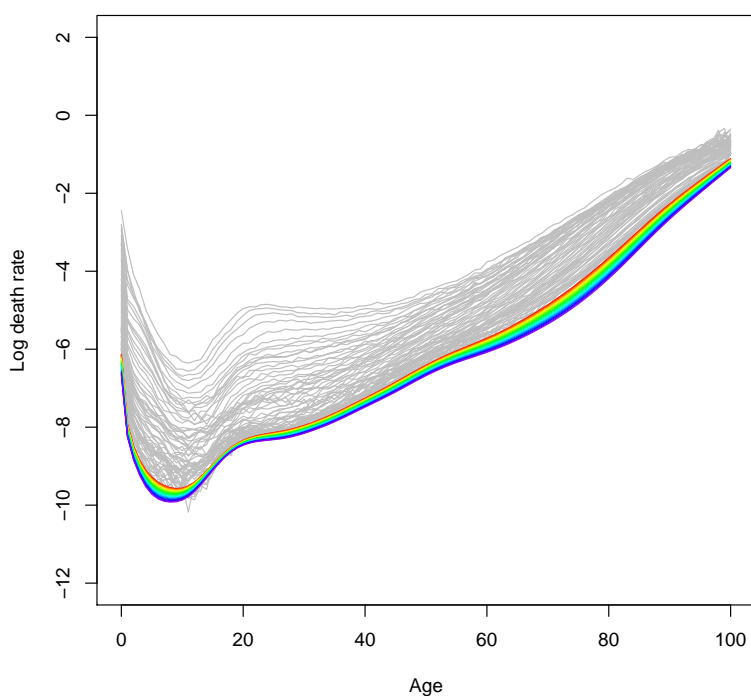


Figure 7: Japan's female mortality forecast from 2010 to 2029 plotted in rainbow palette order, and grey lines represent historical data.

In Fig. 7, the forecasts for the coming 20 years from 2010 to 2029, based on historical data sets, are illustrated. Generally the mortality rates will continue the former trend and keep dropping and

the dynamic declining trend of teenage groups and elderly-age groups will be more obvious than other groups.

Fertility

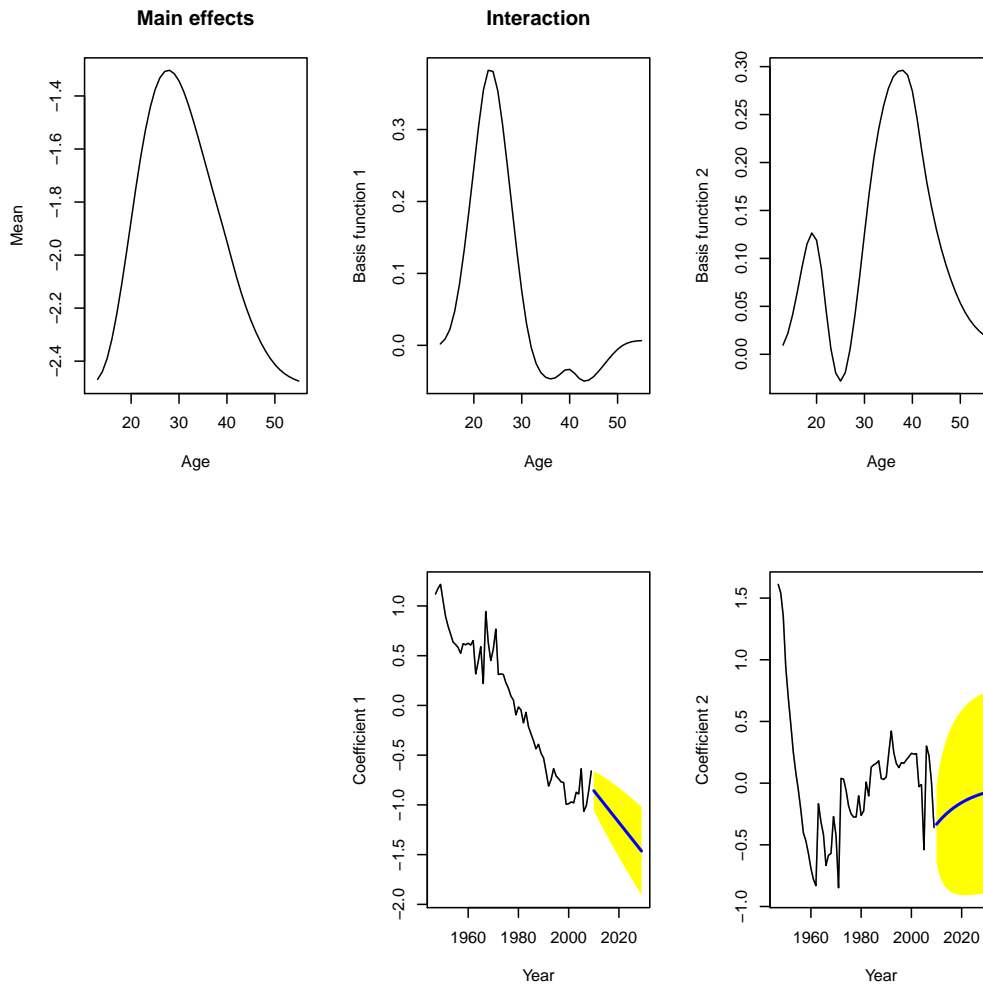


Figure 8: Japan's fertility decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

The analysis on Japan's fertility rates are demonstrated in graphs 8, 9 and 10. The decomposition of Japan's fertility rates are very similar to the decomposition of mortality; the first two components reflect the birth rate of younger-age groups and older-age groups separately, but the coefficients of the first component decline while the coefficients of the second one tend to increase for the future with a higher probability. This result confirms that nowadays more women prefer giving birth at an elder age and that fertility rates are falling due to the decreasing trend of younger-age groups' birth rates (see Fig. 8).

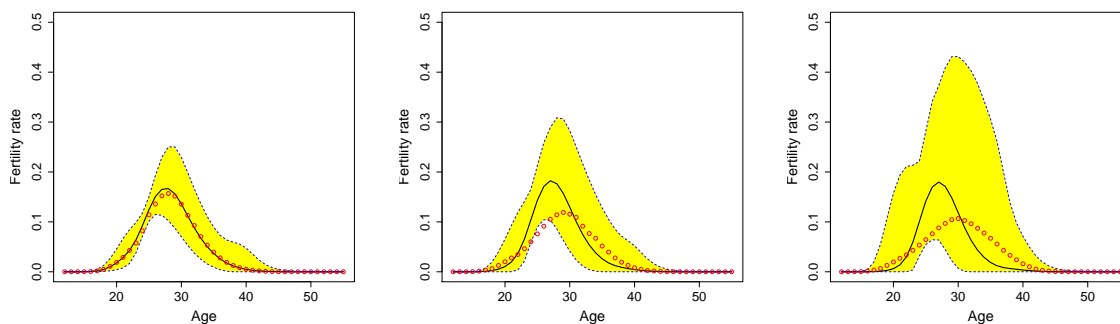


Figure 9: Out-of-sample test on Japan's fertility (1947-1989): forecast rates (black lines) for 1990 (left), 1999 (middle), 2009 (right), along with 95 % confidence intervals (yellow areas), while actual rates are shown as red circles.

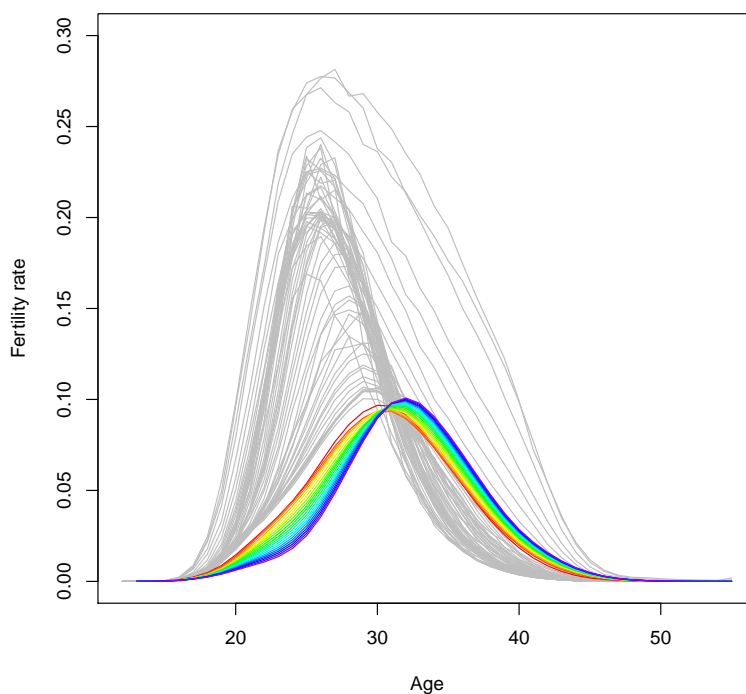


Figure 10: Japan's fertility forecast from 2010 to 2029 plotted in rainbow palette order, and grey lines represent historical data.

The out-of-sample test on Japan's fertility rates based on the historical data from 1947 to 1989 are shown in Fig. 9. Compared to the forecast accuracy of female mortality rates, the forecast performance on fertility rates are not satisfying, in particular when the forecast horizon increases to 20 years. The actual peak of birth rate is around 5 years older than the forecast one and the actual overall birth rates are lower than the expectation as well, when we do the comparison at

20-step ahead forecasting. The main reason leading to this phenomenon lies in that the fertility dynamics are stronger and the rates from the earliest years can not reveal the latest change and will devastate the forecast accuracy. Therefore the sample size should be carefully chosen when coping with more volatile fertility rates.

In Fig. 10, the forecasts from 2010 to 2029 are displayed. It is shown that the peak of birth rates will shift to older ages gradually, around 32 years old. Meanwhile the overall birth rate will change the declining trend and begin to increase a little in following 10 years.

3.3.2 Taiwan

Mortality

It is apparent from Fig. 19 that the mean function basically reflects the mortality trend over different ages, and the basis functions model the female mortality rates in different age ranges: the first basis function models the mortality pattern in children, while the second basis function models the differences between young adults and middle-aged people. The coefficient functions associated with basis functions demonstrate the social effects. In particular, the decrease of the first coefficients function explains the improvement of medical care. But the out-of-sample test is not that good as the test for Japan's mortality data (see Fig. 20 and Fig. 24).

Fertility

In Fig. 26 we explore the mean function, the basis functions and associated coefficient functions. The mean function basically reflects the fertility trend over different ages, and the basis functions model different movements in different age ranges: the first basis function models the early-aged pregnancy, while the second basis function stands for the high-aged pregnancy. Accordingly, the coefficients functions associated with basis functions tell the social effects. We can find that, the decrease of the first coefficients function and the increase of the second function reflect that more women are tending to give birth at older ages, perhaps due to social pressure and medical improvements. These findings tell the same story as Japan. But surprisingly the out-of-sample test is much better than Japan (see Fig. 27); one of the potential reasons is that the more irrelevant data sets from early years are not employed in Taiwan's analysis.

Fig. 28 shows the fertility forecasting from 2011 to 2030. It is out of expectation that total fertility rates will arise in future but the peak concentrates at elderly age, between 30-33 years old. This change will bring great challenges to government health systems.

3.3.3 China

When analysing China, we face a limited data situation that is always a challenge. In light of this problem, we only analyse the general trend explained by these models and do not take any

forecast comparisons.

From Figures 30, 32 and 33, we can conclude that China displays similar trends of mortality and fertility rates as Japan and Taiwan even with the limited date sets. It implies that the general trend of mortality and fertility rates converge to each of the other countries, which will provide a starting point with which to analyse the joint demographic dynamic movement.

3.4 Power of Explanation

Female Mortality (%)		
LC model	HU method	Country
96.1	99.9	Japan
86.3	99.0	Taiwan
41.3	98.9	China

Table 1: *Explained female mortality variance*

Male Mortality (%)		
LC model	HU method	Country
96.2	99.9	Japan
79.9	98.3	Taiwan
39.2	98.5	China

Table 2: *Explained male mortality variance*

To summarize, the power of the LC model and the HU method, the two tables 1 and 2 display the proportion of variance explained by them. From these two tables, it is apparent that the HU method always performs better than the LC model, especially when the sample size is not large enough. When explaining the demographic variation of Japan, the HU method is only slightly better than the LC model but both methods give a satisfying result, and the wider sample range will be one of the main reasons why they both perform well.

		1st	2nd	3rd	4th	5th	6th
Japan	<i>female mortality</i>	96.5	3.1	0.2	0.1	0.0	0.0
	<i>male mortality</i>	97.0	2.0	0.4	0.3	0.1	0.1
	<i>fertility</i>	58.9	31.0	8.5	1.2	0.2	0.1
Taiwan	<i>female mortality</i>	95.1	2.1	0.7	0.5	0.4	0.3
	<i>male mortality</i>	87.6	7.1	2.0	0.8	0.5	0.3
	<i>fertility</i>	90.3	5.5	3.4	0.5	0.1	0.1
China	<i>female mortality</i>	84.8	6.1	2.7	2.7	1.5	1.1
	<i>male mortality</i>	78.5	9.3	5.1	2.7	2	0.9
	<i>fertility</i>	47.3	39.1	9.9	2.5	0.5	0.3

Table 3: *Explained variance from HU method (K = 6)*

To reveal the efficiency of the six factors decomposed by the HU method, Table 3 displays the

variances accounted for by each factor. It is important that all the first explored factors can explain more than half of the variation and even more than 90% in the case of mortality rates; of all the improvement made by the HU method, the explanation power of the first factors for China's mortality is most distinguished: 41.3% (LC model) improved to 84.8% (HU model) for female mortality and 39.2% to 78.5% for male.

IV. COMPARISON ANALYSIS

4.1 Methods Accuracy

To evaluate the forecasting accuracy, we take Japan's mortality data to compare the LC and HU methods. We firstly divide the data set into a fitting period and a forecasting period, and a rolling origin is applied as follows: the initial fitting period is set as 1947-1989, based on which we compute the one-step ahead forecast and calculate the forecast error through comparing the one-step-ahead forecast and the actual out-of-sample data. We then increase the fitting period by one year and repeat the above procedures until it extends to 2008.

From the following two pictures 11 and 12 and the Appendix on Forecast Accuracy, it is obvious that the mean absolute forecasting errors of the HU method is always smaller than those stemmed from the LC model, irrespective of errors averaged by years or by ages. It implies that the HU method provides a more precise forecast.

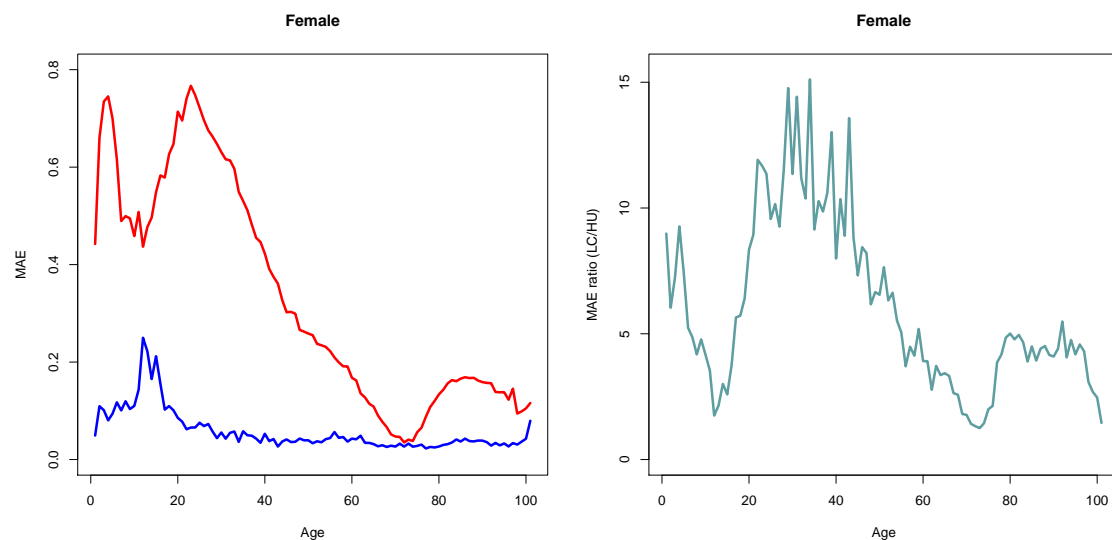


Figure 11: Japan's female mortality Mean Absolute Error for one-step-ahead forecasts averaged over years: LC (red), HU (blue).

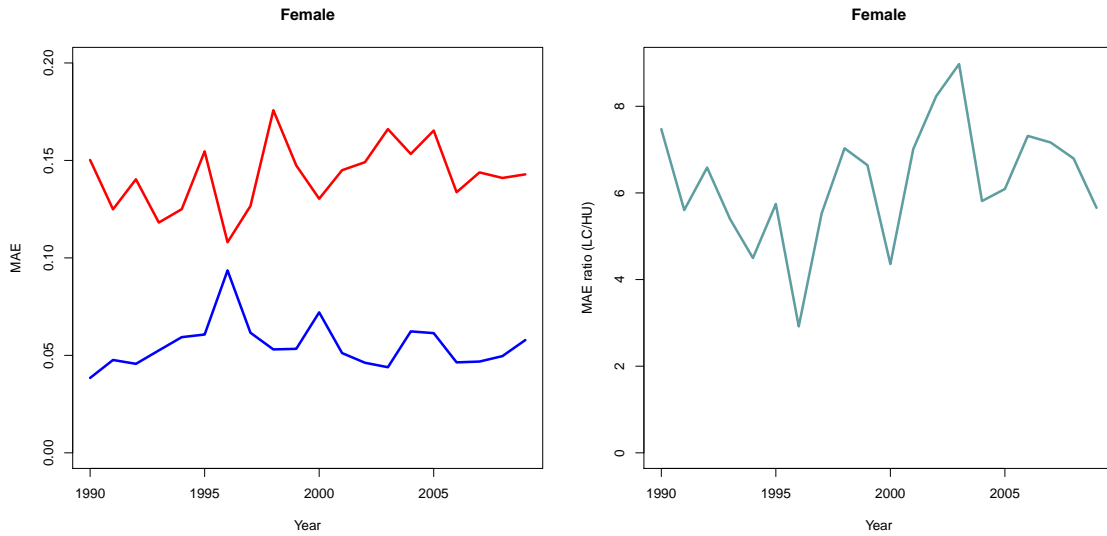


Figure 12: Japan's female mortality Mean Absolute Error for one-step-ahead forecasts averaged over ages: LC (red), HU (blue).

To test for the overall significance of forecasting superiority between the LC model and the HU method in the scale of one year, we apply the Diebold and Mariano (1995) test. Define the loss differential d_t between the one-step ahead mean absolute forecasting errors from both methods as $d_t = d_{1t} - d_{2t}$, where $d_{1t} = |\hat{y}_{LC,t} - y_t|$ and $d_{2t} = |\hat{y}_{HU,t} - y_t|$, $t = 1, 2, \dots, 20$. The two forecasts have equal accuracy if, and only if, the loss differential has an expectation of zero for all t . The null hypothesis is $H_0 : E(d_t) = 0, \forall t$, versus the alternative hypothesis $H_1 : E(d_t) > 0$. The test statistics is

$$DM = \bar{d} / \sqrt{2\pi \hat{f}_d(0) / T} \quad (14)$$

where $\bar{d} = \sum_{t=1}^T d_t$, $\hat{f}_d(0) = \frac{1}{2\pi} \hat{\gamma}_d(0)$, $\hat{\gamma}_d(0) = T^{-1} \sum_{t=1}^T (d_t - \bar{d})^2$ and $T = 20$.

The p -values obtained from the female group and the male group are both smaller than 0.01, so that we can conclude that we have a very strong preassumption against the null hypothesis and believe that the HU method has superiority over the LC model. Similarly we can take the test in the scale of age, and the same conclusion is generated.

4.2 Regional Similarities

Another interesting finding from this research is that China has a demographic trend closer to Japan compared to Taiwan, particularly the mortality trend. On the economic level, China and Japan have been both important economies in last several decades and the development pattern

is also quite similar. K. Hanewald (2011) find that the Lee-Carter mortality index k_t correlates significantly to macroeconomic fluctuations in some periods, which provides a good reference with which to connect the mortality trends between China and Japan.

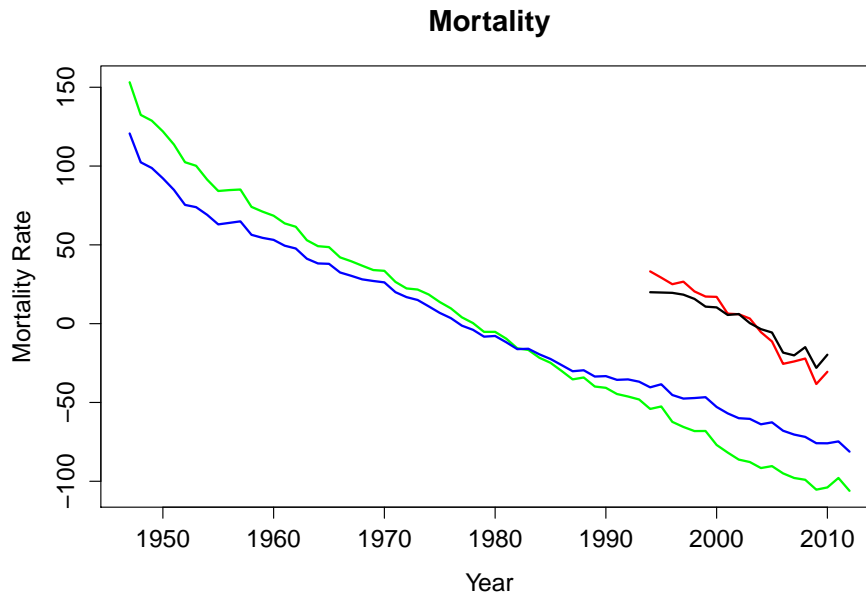


Figure 13: China female mortality (red) vs. Japan female mortality (green), China male mortality (black) vs. Japan male mortality (blue).

The mortality trends are displayed in the above graph. It reads that mortality trends from both gender groups of China correlate with those of Japan respectively, but further research focusing on detailed comparison is needed. The research from W. Härdle and J.S. Marron (1990) on semiparametric comparison of regression curves suggests one potential way.

V. CONCLUSION

In this research we analyze the Lee-Carter (LC) model and the Hyndman-Ullah (HU) method and apply them to Japan, Taiwan and China. In all the three cases the HU method shows better explanation and forecasting power, however, it encounters problems when analyzing China due to the limited data available. However, the empirical findings of a similar trend between China and Japan could open another way to improve the demographic analysis of China, and provide the possibility to forecast the future demographic trend for China jointly with Japan as well.

VI. APPENDIX

6.1 Appendix on Japan

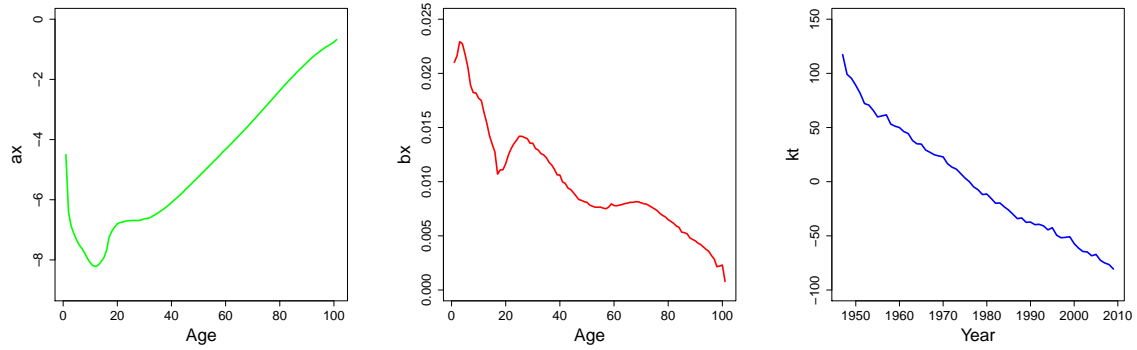


Figure 14: Japan's male mortality decomposition by LC model

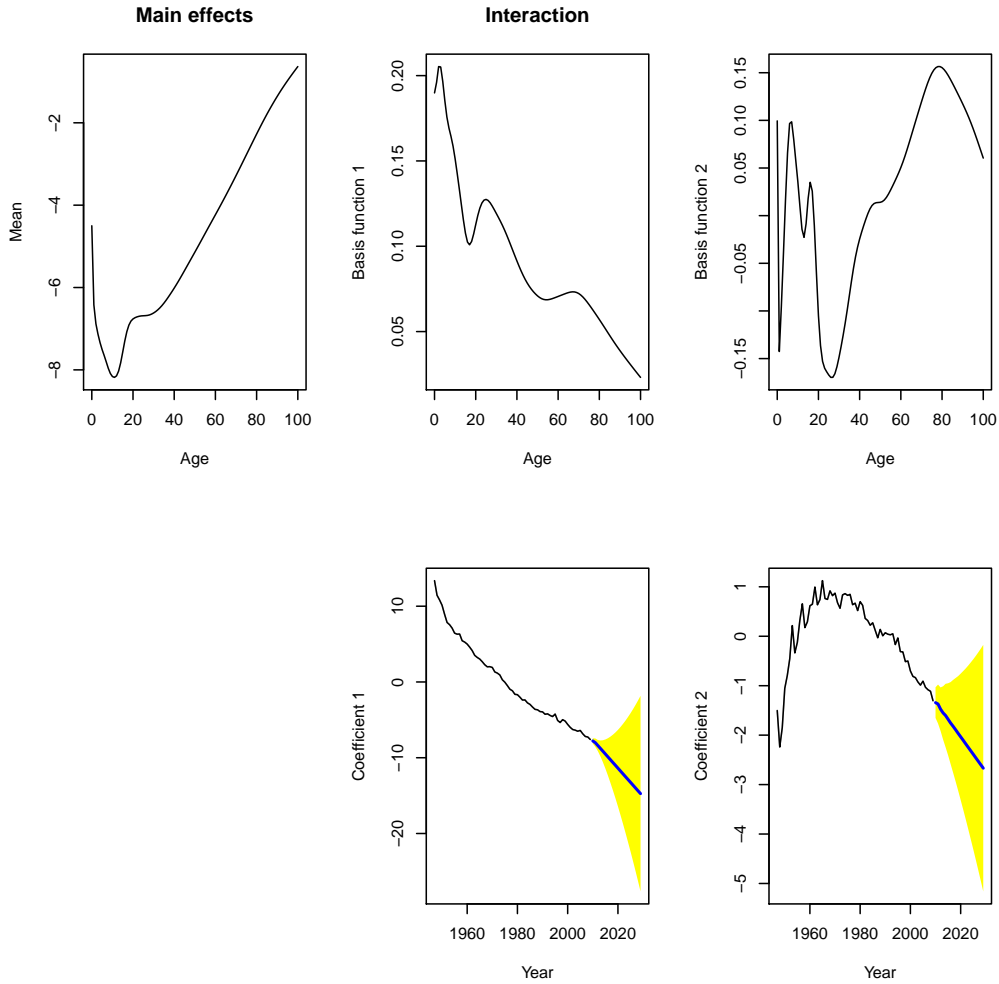


Figure 15: Japan's male mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

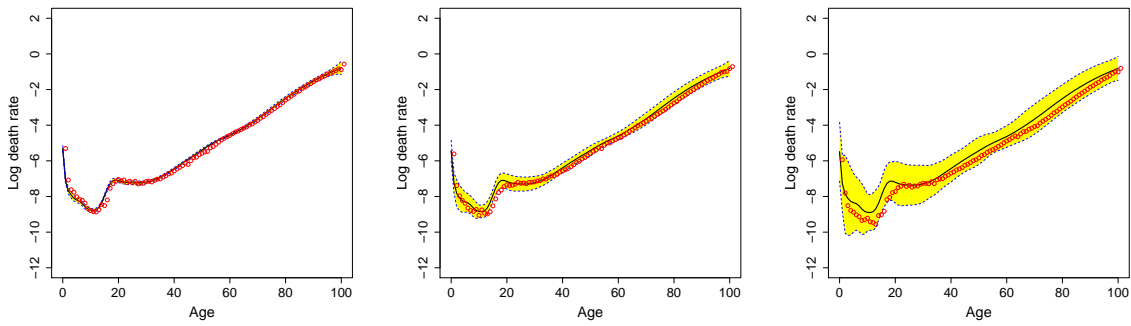


Figure 16: Out-of-sample test on Japan's male mortality (1947-1989): forecast rates (black lines) for 1990 (left), 1999 (middle), 2009 (right), along with 95% confidence intervals (yellow areas), while actual rates are shown as red circles.

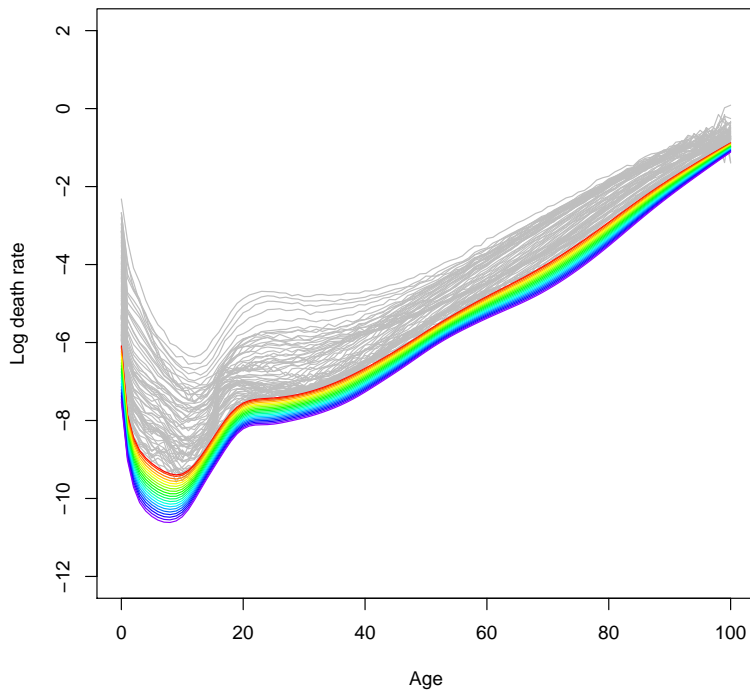


Figure 17: Japan's male mortality forecast from 2010 to 2029 plotted in rainbow palette order, and grey lines represent historical data.

6.2 Appendix on Taiwan

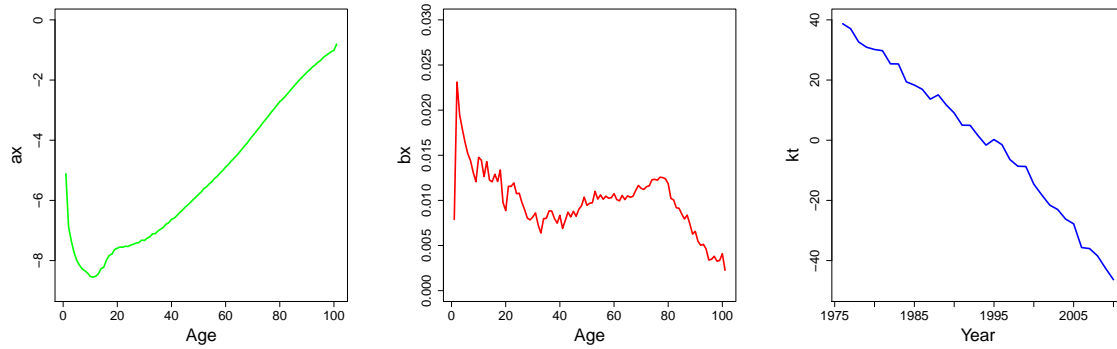


Figure 18: Taiwan's female mortality decomposition by LC model

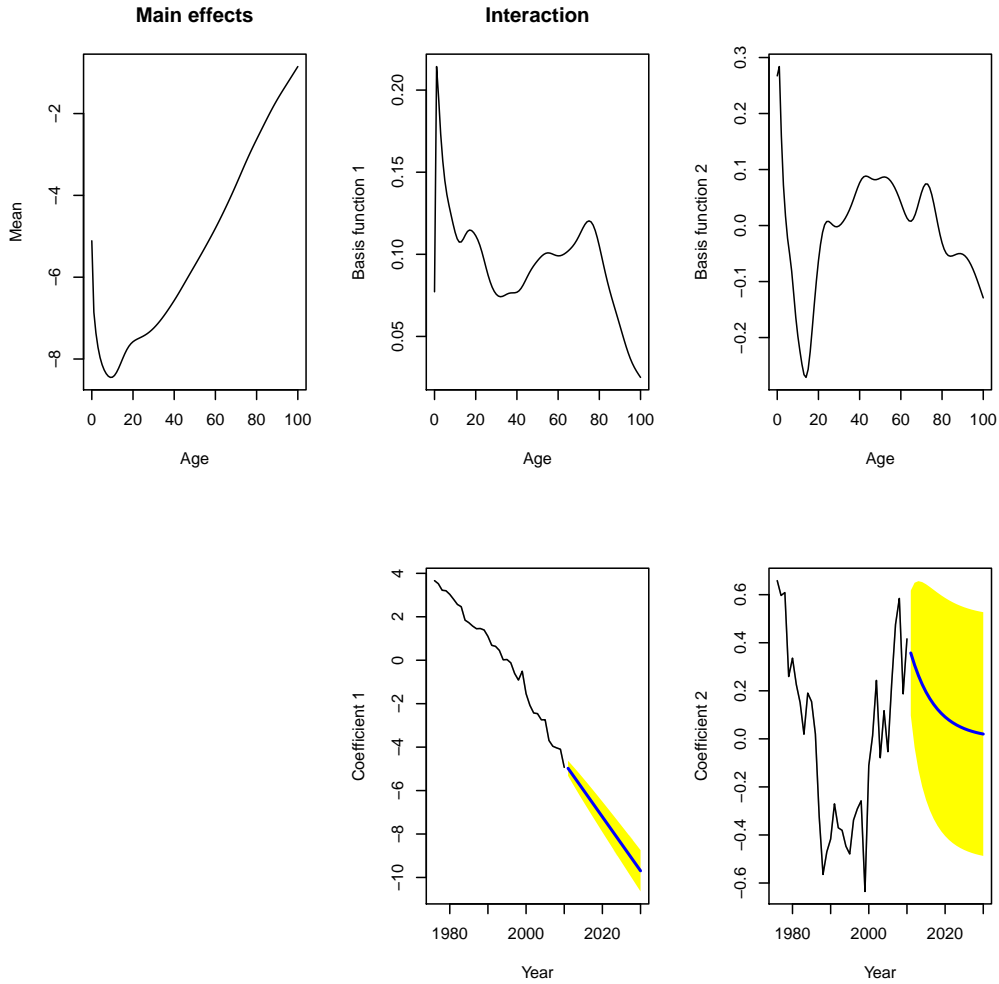


Figure 19: Taiwan's female mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

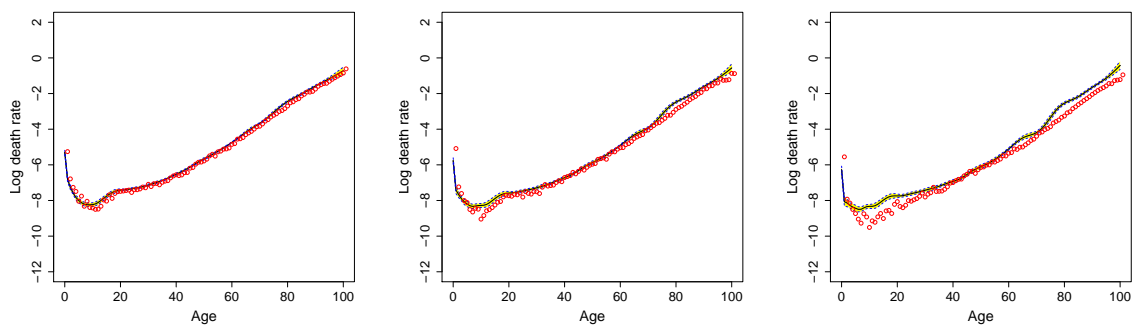


Figure 20: Out-of-sample test on Taiwan's female mortality (1976-1990): forecast rates (black lines) for 1991 (left), 2000 (middle), 2010 (right), along with 95 % confidence intervals (yellow areas), while actual rates are shown as red circles.

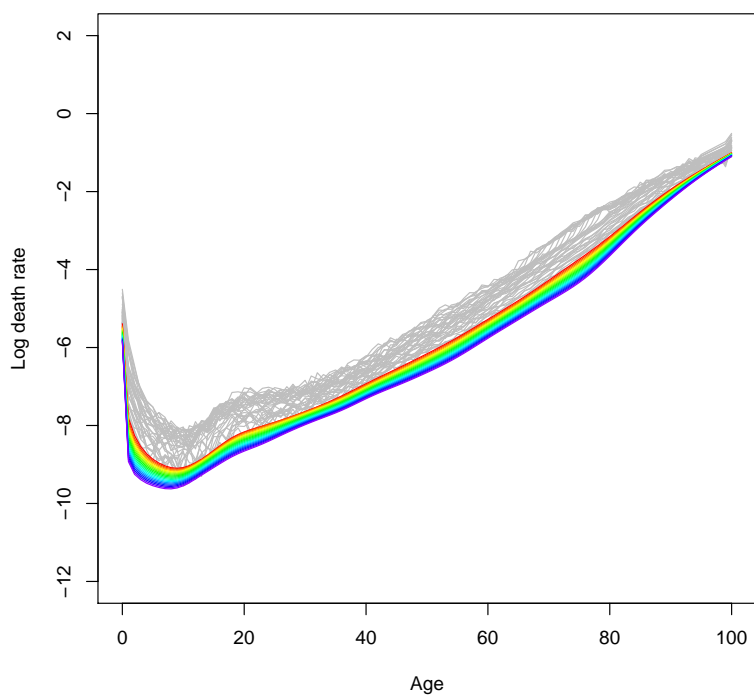


Figure 21: Taiwan's female mortality forecast from 2011 to 2030 plotted in rainbow palette order, and grey lines represent historical data.

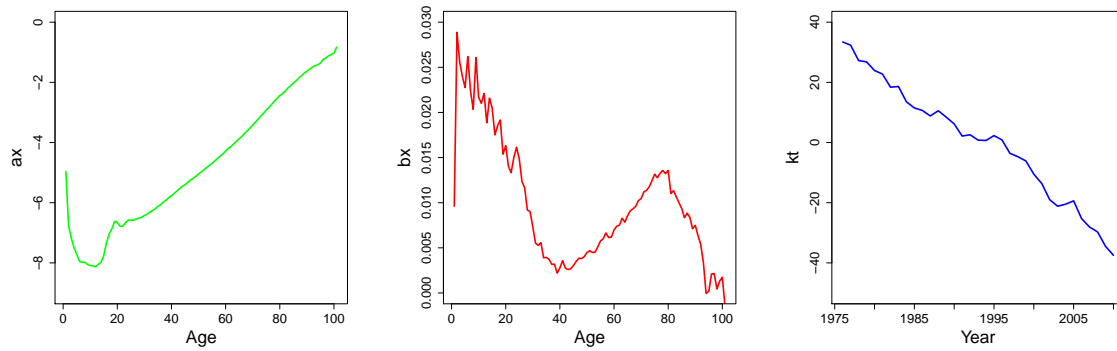


Figure 22: Taiwan's male mortality decomposition by LC model

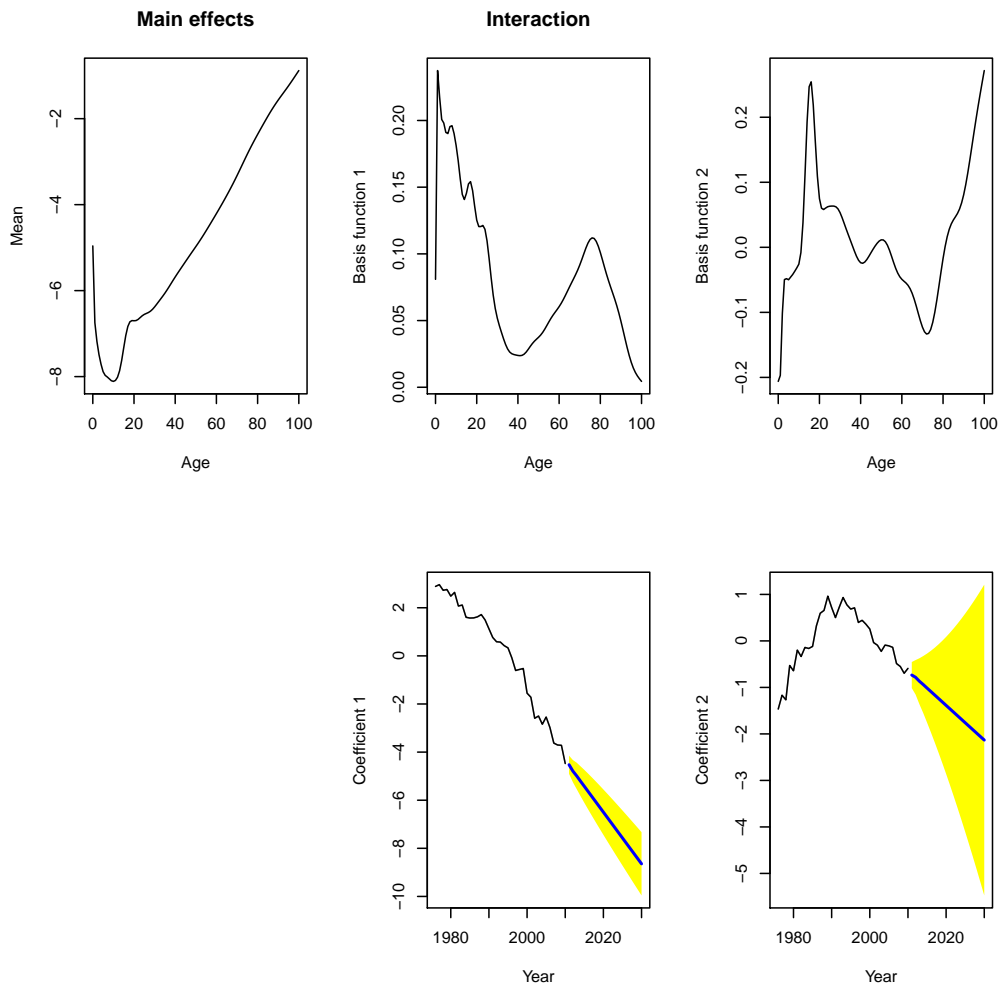


Figure 23: Taiwan's male mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

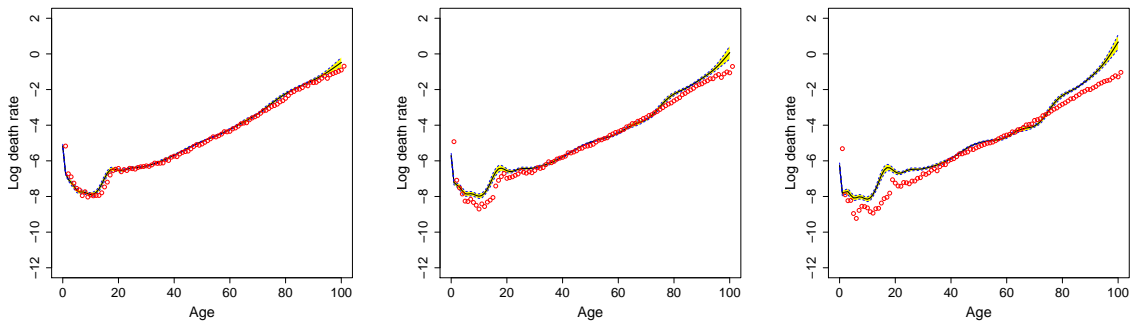


Figure 24: *Out-of-sample test on Taiwan’s male mortality (1976-1990): forecast rates (black lines) for 1991 (left), 2000 (middle), 2010 (right), along with 95 % confidence intervals (yellow areas), while actual rates are shown as red circles.*

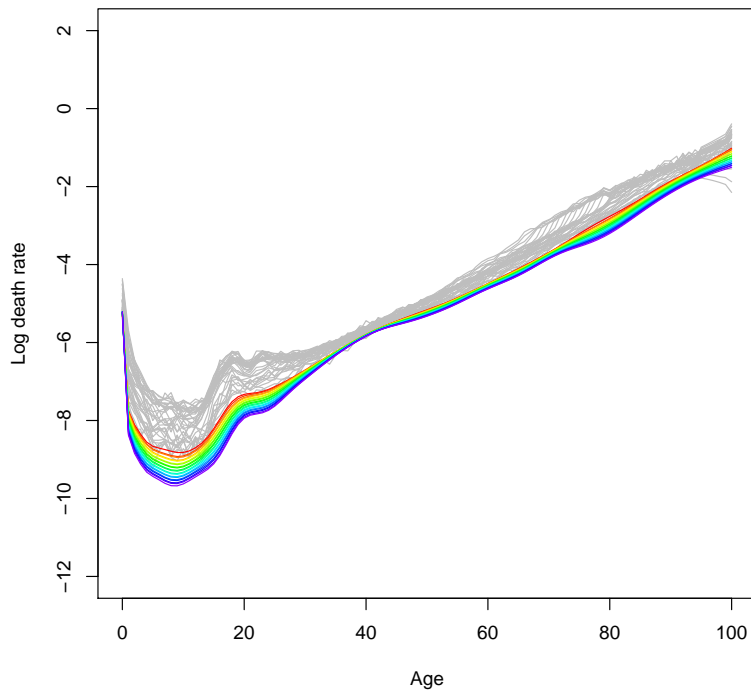


Figure 25: *Taiwan’s male mortality forecast from 2011 to 2030 plotted in rainbow palette order, and grey lines represent historical data.*

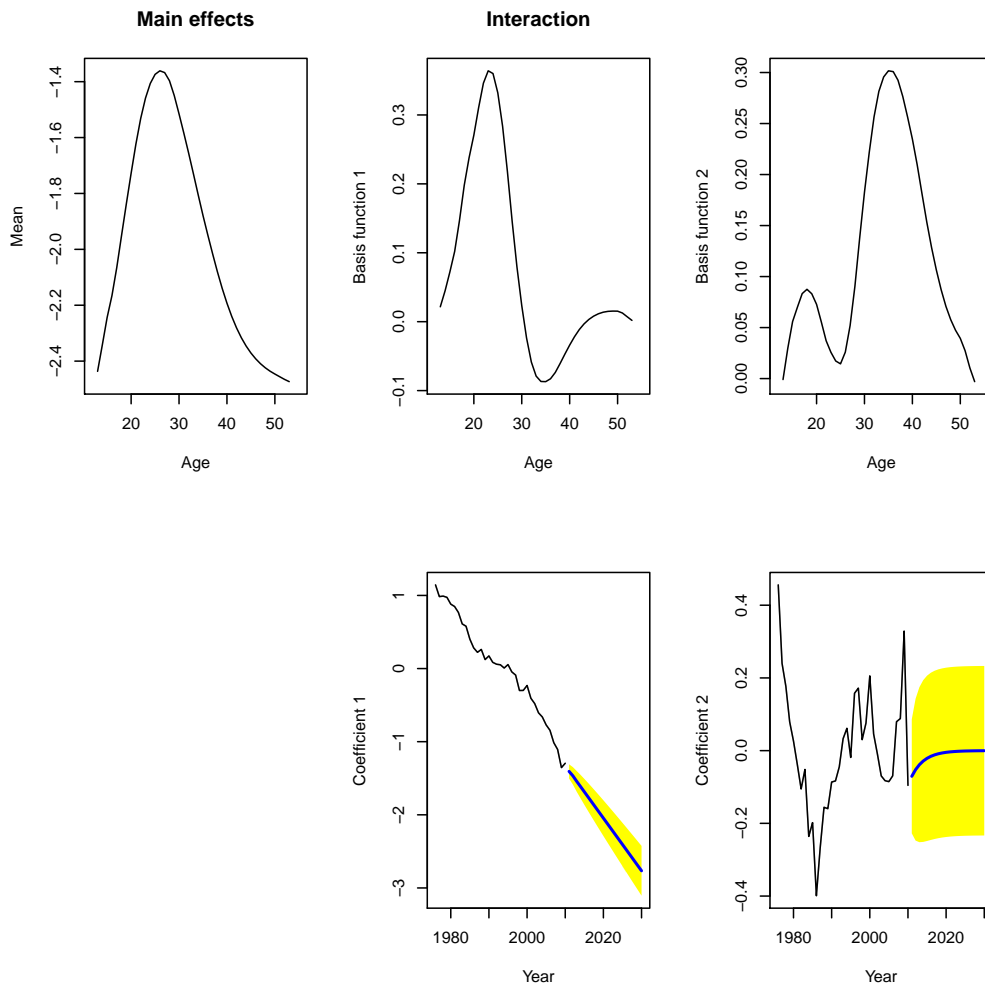


Figure 26: Taiwan's fertility decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

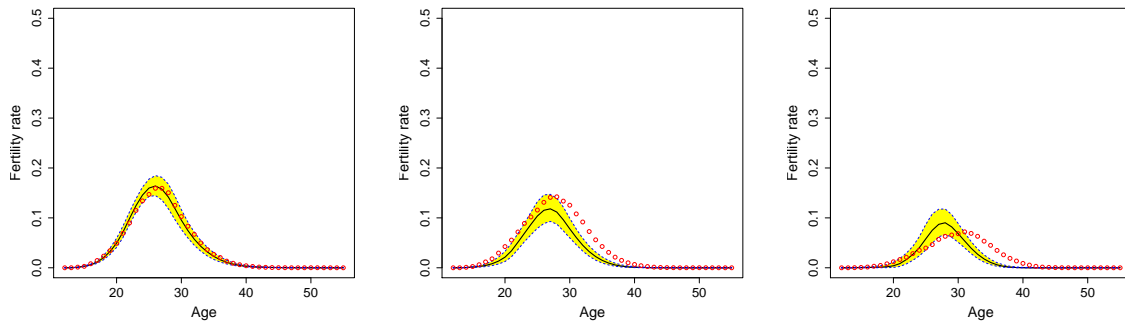


Figure 27: Out-of-sample test on Taiwan's fertility (1976-1990): forecast rates (black lines) for 1991 (left), 2000 (middle), 2010 (right), along with 95 % confidence intervals (yellow areas), while actual rates are shown as red circles.

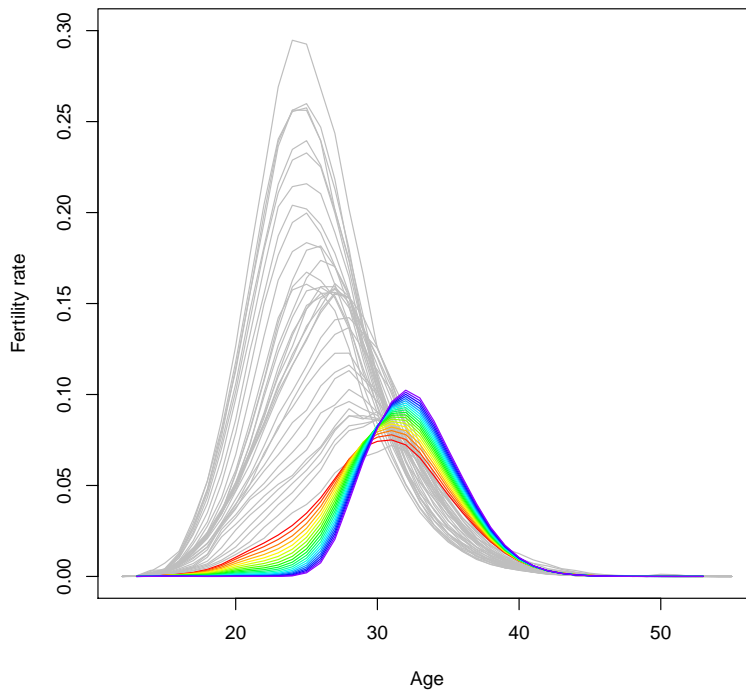


Figure 28: Taiwan's fertility forecast from 2011 to 2030 plotted in rainbow palette order, and grey lines represent historical data.

6.3 Appendix on China

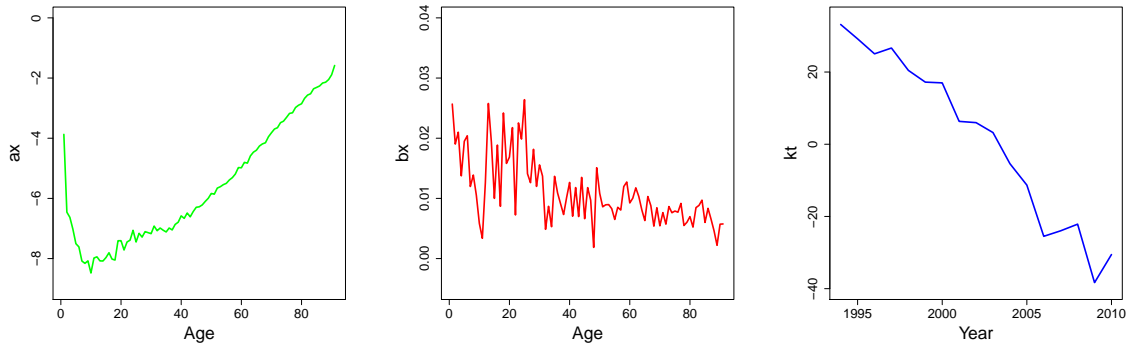


Figure 29: China's female mortality decomposition by LC model

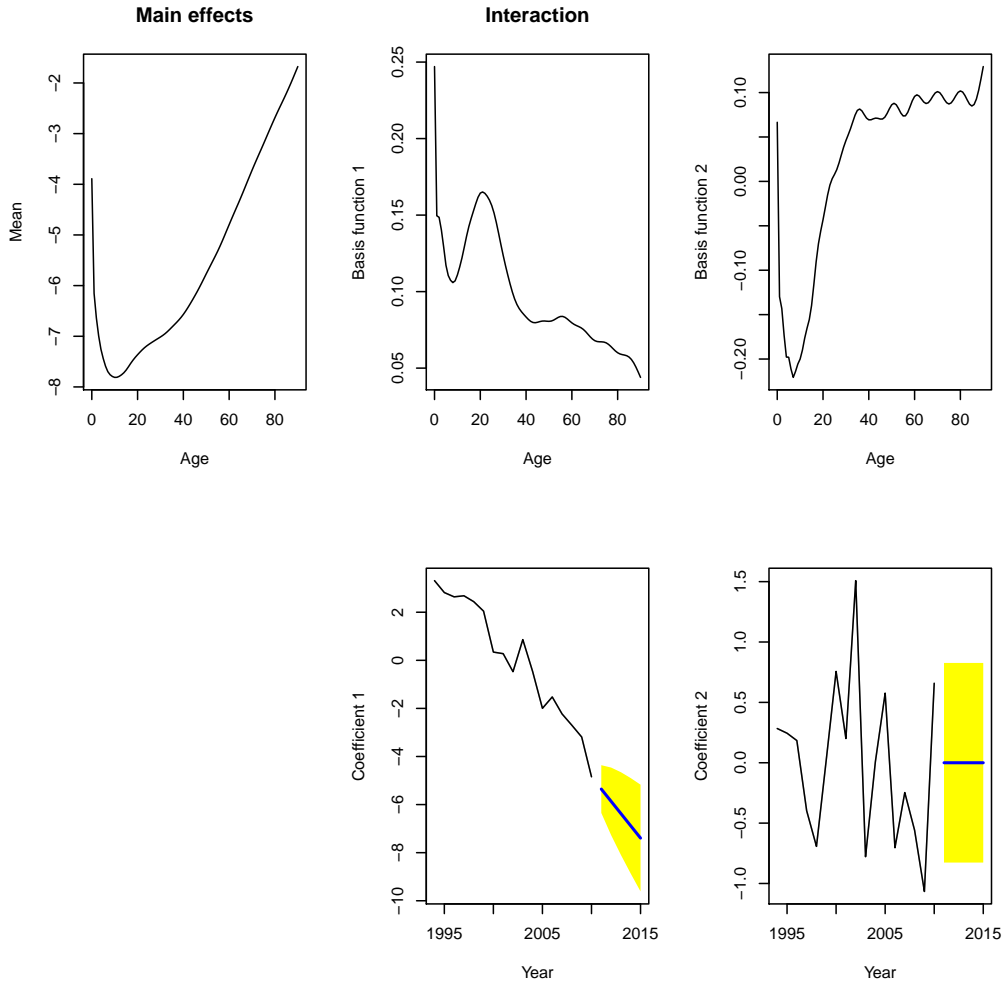


Figure 30: China's female mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

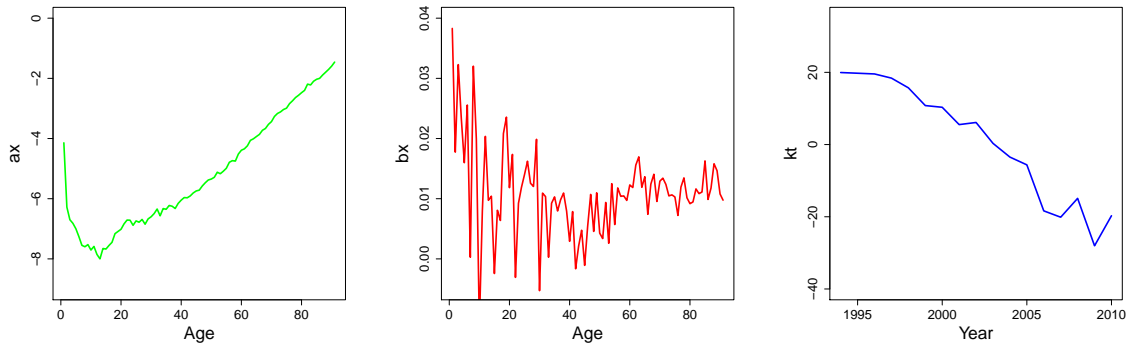


Figure 31: China's male mortality decomposition by LC model

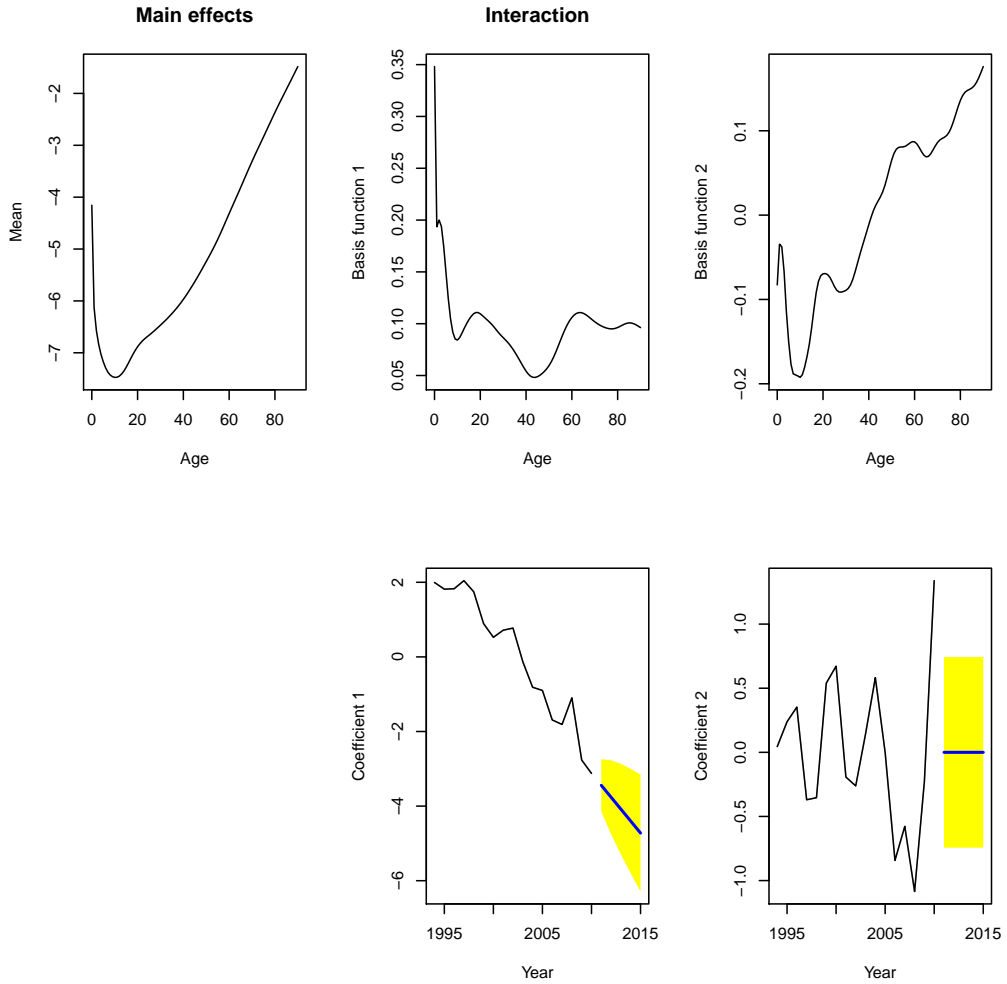


Figure 32: China's male mortality decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

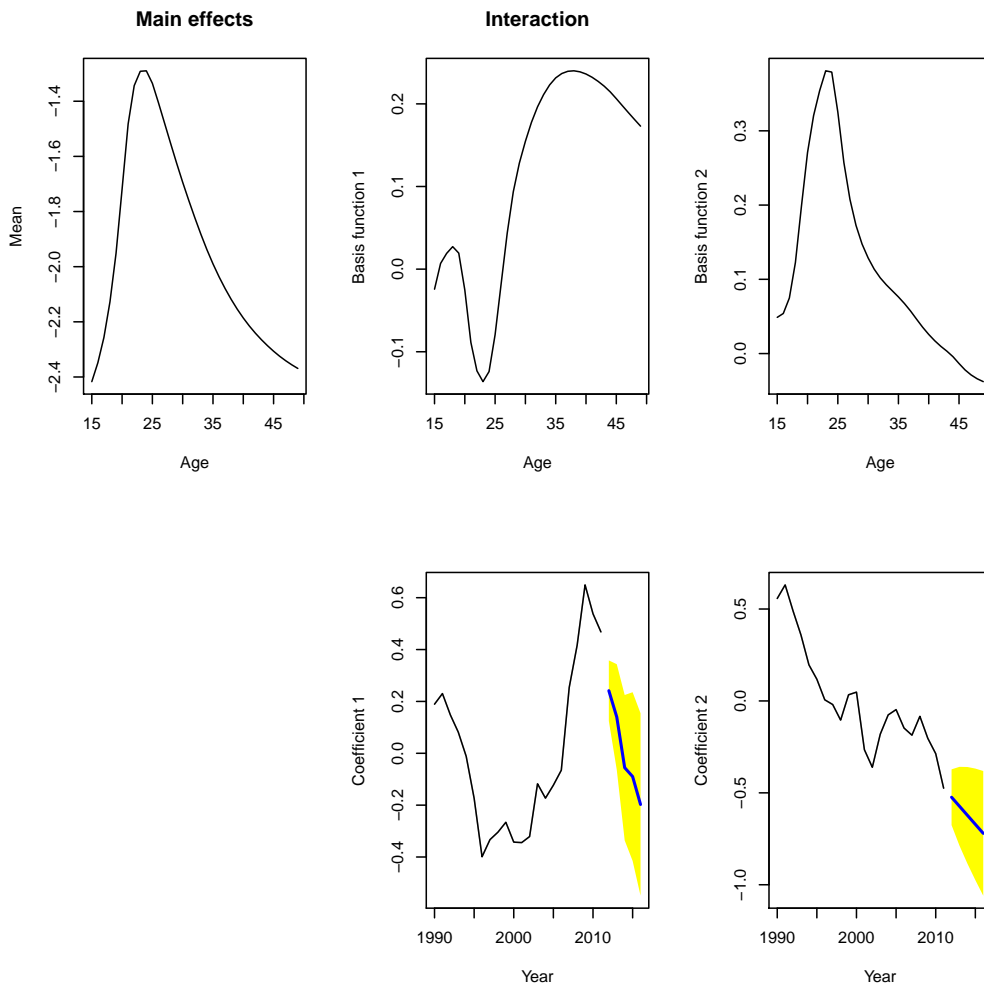


Figure 33: China's fertility decomposition by HU Method: yellow areas represent the 95% confidence intervals for the coefficients forecast.

6.4 Appendix on Forecast Accuracy

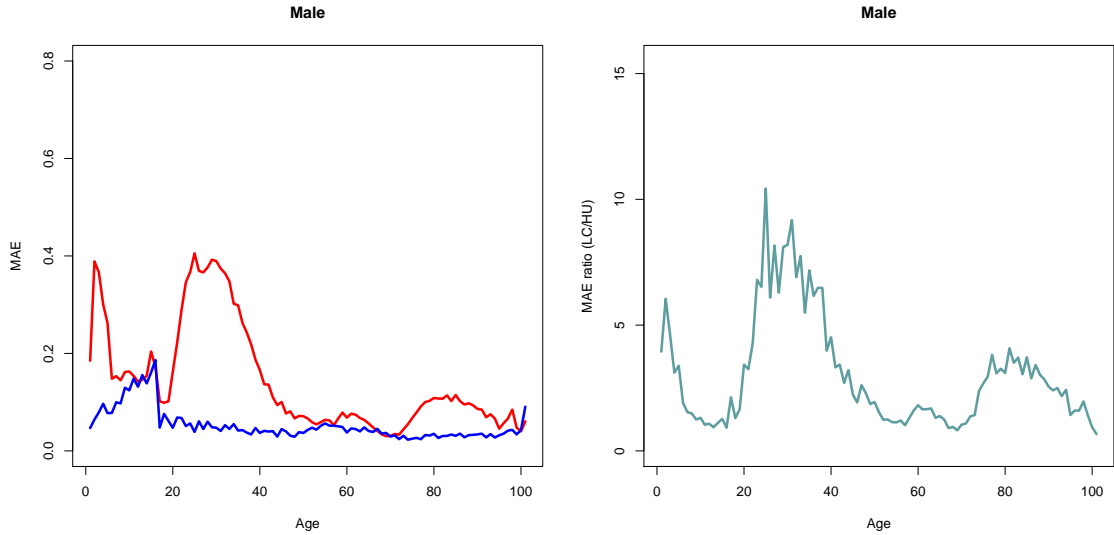


Figure 34: Japan's male mortality Mean Absolute Error for one-step-ahead forecasts averaged over years: LC (red), HU (blue).



Figure 35: Japan's male mortality Mean Absolute Error for one-step-ahead forecasts averaged over ages: LC (red), HU (blue).

REFERENCES

- [1] H. Booth. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22:547-581.
- [2] W.S. Chan, S.H. Li and S.H. Cheung. (2008). Testing deterministic versus stochastic trends in the Lee-Carter mortality indexes and its implications for projecting mortality improvements at advanced ages. *Living to 100*.
- [3] C. Chen and L.M. Liu. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88:284-297.
- [4] F. X. Diebold and R.S. Mariano. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253-263.
- [5] K. Hanewald. (2011). Explaining mortality dynamics: the role of macroeconomic fluctuations and cause of death trends. *North American Actuarial Journal*, 15(2).
- [6] W. Härdle and J.S. Marron. (1990). Semiparametric comparison of regression curves. *Annals of Statistics*, 18:63-89.
- [7] W. Härdle, M. Müller, S. Sperlich and A. Werwatz. (2004). Nonparametric and semiparametric models. *Springer, Berlin*.
- [8] X. He and P. Ng. (1999). COBS: qualitatively constrained smoothing via linear programming. *Computational Statistics*, 14: 315-337.
- [9] R. J. Hyndman and H. Booth. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, 24:323-342.
- [10] R. J. Hyndman and H. L. Shang. (2009). Forecasting functional time series. *Journal of the Korean Statistical Society*, 38:199-211.
- [11] R. J. Hyndman and M. S. Ullah. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, 51:4942-4956.
- [12] R. D. Lee and L. R. Carter. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87:659-671.
- [13] R. D. Lee and T. Miller. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38:537-549.
- [14] G. Li and Z. Chen. (1985). Projection-pursuit approach to robust dispersion matrices and principal components: primary theory and Monte Carlo. *Journal of the American Statistical Association*, 80(391):759-766.

- [15] N. Li, R. Lee and S. Tuljapurkar. (2004). Using the Lee-Carter method to forecast mortality for populations with limited data. *International Statistical Review*, 72:19-36.
- [16] J.O. Ramsay and B.W. Silverman. (2005). Functional data analysis. *Springer, New York*.
- [17] H. L. Shang, H. Booth and R. J. Hyndman. (2011). Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods. *Demographic Research*, 25:173-214.
- [18] S.N. Wood. (1994). Monotonic smoothing splines fitted by cross validation. *SIAM Journal of Scientific Computation*, 15(5):1126-1133.
- [19] S.N. Wood. (2003). Thin plate regression splines. *Journal of Royal Statistical Society*, 65(1):95-114.

SFB 649 Discussion Paper Series 2015

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Pricing Kernel Modeling" by Denis Belomestny, Shujie Ma and Wolfgang Karl Härdle, January 2015.
- 002 "Estimating the Value of Urban Green Space: A hedonic Pricing Analysis of the Housing Market in Cologne, Germany" by Jens Kolbe and Henry Wüstemann, January 2015.
- 003 "Identifying Berlin's land value map using Adaptive Weights Smoothing" by Jens Kolbe, Rainer Schulz, Martin Wersing and Axel Werwatz, January 2015.
- 004 "Efficiency of Wind Power Production and its Determinants" by Simone Pieralli, Matthias Ritter and Martin Odening, January 2015.
- 005 "Distillation of News Flow into Analysis of Stock Reactions" by Junni L. Zhang, Wolfgang K. Härdle, Cathy Y. Chen and Elisabeth Bommers, January 2015.
- 006 "Cognitive Bubbles" by Ciril Bosch-Rosay, Thomas Meissner and Antoni Bosch-Domènech, February 2015.
- 007 "Stochastic Population Analysis: A Functional Data Approach" by Lei Fang and Wolfgang K. Härdle, February 2015.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

