# Stochastic Forecast for the South Korean National Pension System

**Hayoung Lee**

A thesis presented for the degree of

Master of Data Science in Quantitative Finance

under the supervision of Len Patrick Dominic M. Garces

School of Mathematical and Physics Sciences

Faculty of Science

University of Technology Sydney

May 26, 2025

# Abstract

This study investigates the long-term financial sustainability of South Korea's National Pension Fund in the face of demographic aging and economic uncertainty. With the working-age population declining and the elderly population growing, concerns over fund depletion and intergenerational fairness have intensified.

To capture the complex dynamics of population change, we apply the Generalized Log-Gamma (GLG) model to forecast fertility rates and the affine mortality model to simulate survival probabilities. Among the mortality models tests, the BS Model with 3 dependent factors demonstrates better in-sample performance and more stable long-term dynamics. These projections are used to stochastically forecast the population and pension fund, providing confidence intervals for risk assessment.

Under the current system, the fund is projected to peak in 2058 and be depleted by 2075 (95% CI: 2072–2077). Sensitivity analysis reveals that changes in economic variables such as wage growth and investment return impact sustainability. Simulations of pension reforms such as raising contribution rates, increasing the pensionable age show the potential to delay depletion. Additionally, a macro slide mechanism based on inflation and demographic shifts offers modest delay effects.

This study emphasizes the importance of quantifying uncertainty and exploring pension reform options to ensure long-term pension sustainability.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation

### 1.1.1 Background and Structural Challenges of the Korean National Pension System

National Pension Service (NPS) in South Korea operates under a mixture system of pay-as-you-go (PAYG) and funding principles. In this system, if the total pension expenditures are less than the total pension contributions, they are covered by the contributions. However, when expenditures exceed contributions, the shortfall is financed through the pension fund. In recent years, South Korea has been facing significant demographic and economic challenges. The birth rate continues to decline, while the population is rapidly aging. In addition, the country is experiencing a slowdown in the economic growth rate and a rise in unemployment. These trends lead to a shrinking working-age population that supports a growing number of pensioners. As a result, there is growing concerns about the potential depletion of the pension fund and whether the current working generation will be able to receive adequate pension benefits later in retirement, although they are paying high contributions into the system.

### 1.1.2 The Need for Accurate Model and Reform-Oriented Simulation

To effectively address longevity risk and uncertainties arising from economic variables, it is essential to simulate the pension system using more sophisticated and accurate models. These models should be capable of capturing demographic dynamics. In addition, sensitivity analyses should be conducted by adjusting contribution rates or replacement rates to ensure the sustainability the fund, providing guidance for pension reform. J. Choi et al. (2016) emphasizes the importance of precise mortality forecasting in projecting future pension liabilities. Using age-period data, he applied mortality rate models such as the Lee-Carter model, the Li and Lee model, and the Cairns, Blake, and Dowd (CBD) model to estimate population dynamics and assess the balance of the NPS. Similarly, J. Choi (2017) highlights that South Korea's population is aging at a faster pace than other countries and stresses the importance of precise and reliable population projections for ensuring the sustainability of the pension system. This study examines the effects of revised population projections on the finance of the NPS, including the dependency ratio, the timing of fiscal deficit, and fund depletion.

### 1.1.3 Limitations of Current Approaches for the Actuarial Valuation

Currently, the actuarial valuation conducted by the NPS is based on a deterministic model, which assumes fixed inputs and yields a single outcome. However, it is not enough to capture the wide range of uncertainties that may occur over the long term. In contrast, stochastic actuarial models allow for random variation in key variables across numerous simulations. Therefore, research into the implementation of stochastic valuation methods is essential to strengthen the long-term forecasting accuracy and resilience of the pension system alongside more refined models for variables such as mortality and economic growth. Moreover, the NPS currently predicts the mortality rate with the expansion model of Lee-Carter model. While this model is widely used due to its simplicity and interpretability, it struggles to reflect the complex patterns of mortality rates. In this context, the affine mortality model has been developed to analyze the randomness in mortality overtime. It represents the short rate of mortality intensity as an exponential affine function and enables closed-form expressions for survival curves. Incorporating this model into population projections needs to be researched to follow more complex patterns of mortality rates.

### 1.1.4 Interpretation of Stochastic Forecasts and Policy Implications

The implementation of stochastic actuarial valuation should be accompanied by appropriate interpretation using key indicators, such as the timing of pension fund depletion. These metrics will be crucial in assessing the long-term sustainability of the pension system under uncertainty. Ultimately, a data-driven and forward-looking approach that integrates stochastic modeling and sensitivity analysis can provide a strong foundation for evidence-based pension reform in South Korea.

## 1.2 Research Problems

This study addresses several research questions concerning the long-term sustainability of the National Pension System under demographic and economic uncertainty. Specifically, it examines whether population can be projected stochastically using fertility and mortality models, how the incorporation of uncertainty alters the interpretation of pension fund forecasts, and to what extent different pension reform strategies may delay fund depletion.

## 1.3 Overview

Before implementing the simulation, Chapter 2 reviews the relevant literature to establish the theoretical and methodological foundations for the analysis. This includes a review of different types of pension systems, approaches to actuarial valuation, and key models used in fertility and mortality forecasting. These references provide justification for the specific modeling choices adopted in this study. Chapter 3 presents the model framework in detail, introducing the Generalised Log-Gamma (GLG) model for fertility and the affine mortality model for survival probabilities. This chapter also describes the key assumptions made for each model, the data sources used, and the R packages and functions applied for estimation. The chapter concludes with the design of the stochastic simulation procedure for projecting population and pension fund dynamics. Chapter 4 presents the simulation results and discusses their implications. This includes baseline projections of population and fund depletion under the current system, a sensitivity analysis to assess the impact of changes in economic variables such as wage growth and investment returns, and an evaluation of various pension reform scenarios—namely, changes in the contribution rate, replacement rate, and pensionable age. The effect of implementing a macro slide system that adjusts benefits in response to inflation and demographic changes is also analyzed. Finally, Chapter 5 summaries the study, discuss the obtained insights, and outlines directions for future research.

# Chapter 2

# Literature Review

## 2.1 Introduction

This chapters review the literature on the actuarial foundations, demographic factors, and economic factors underlying pension forecasting models. The literature review is divided into three sections. The first section 2.2 discusses Actuarial Valuation Models, explaining the structure of the NPS and reviewing both deterministic and stochastic approaches used to evaluate pension finances. It highlights the importance of incorporating uncertainty into long-term projections. The second section 2.3 covers Demographic Models, focusing on how fertility and mortality rates are forecasted. It reviews various modeling approaches including the Generalized Log-Gamma model and Affine mortality model, both of which are used to capture population dynamics in pension projections. The third section 2.4 addresses Economic Variables, discussing wage growth and investment return models. This section reviews prior literature in which these variables have been simulated and forecasted using time-series methods such as autoregressive (AR) and vector autoregressive (VAR) models.

## 2.2 Actuarial valuation model

### 2.2.1 National Pension Service

The National Pension Service is a social security system operated by the government to support citizens in preparing for their retirement. Its main objective is to ensure that individuals can maintain a stable and secure livelihood after retirement when they no longer have a regular income. During their working years, contributors pay premiums, which are later provided as pensions upon retirement. The National Pension Service manages and administers this system. Eligibility for enrollment includes individuals aged between 18 and 59 who have an income. The system offers various types of pensions, including old-age pensions, disability pensions, and survivor pensions to qualified pensioners.

Pension systems are generally divided into three main categories. The first is the Pay-as-you-go (PAYG) system(Samuelson, 1958). In this system, the contributions from the working population are immediately used to finance the benefits for current retirees. The primary characteristics of this method are the transfer of wealth between generations and its dependency on population growth. As the workforce shrinks or the population ages, the system faces increasing financial pressures.

The second category is the funded system(Ando and Modigliani, 2005). Here, individuals contribute to a pension fund throughout their working years, and these contributions are then returned as pension benefits when they retire. Because the system is based on personal savings, it eliminates the generational burden. Moreover, the accumulated funds are invested, ensuring the system's financial stability.

The third approach is the partially funded system(Barr and Diamond, 2009). This system merges the PAYG and funded systems, combining the advantages of them. Some pension payments are covered by the current workers' contributions, while the rest is sourced from the accumulated funds. This dual approach reduces the burden on future generations and enhances long-term financial stability. Additionally, it provides greater flexibility to adapt to demographic shifts, making it a more sustainable model.

The NPS has adapted the partially funded system for two key reasons. First, low birth rates and an aging population make it difficult to maintain pension payments under a PAYG system alone. Incorporating a funded component helps ease the financial burden on future generations. Second, the funded element contributes to the long-term sustainability of the pension system, making it better equipped to handle demographic challenges. This hybrid approach allows the NPS to balance the strengths of both models, ensuring a more secure future for its pension structure.

### 2.2.2 Actuarial Valuation Model

The pension system is represented mathematically in actuarial models(C. Oh (1997), K. Choi (2005)). They introduce actuarial present value concepts to calculate an individual's premium contributions and future pension benefits. On a broader scale, the system needs to account for all contributions and benefits to evaluate the pension fund's status(Plamondon, 2002), which is crucial for several reasons: maintaining financial stability, preparing for demographic and economic changes, ensuring the system's credibility, and assessing whether the fund's investments are appropriately managed.

The NPS operates by requiring employees and their employers to each contribute 4.5% of the employee's monthly salary, totaling 9% for pension premiums. The accumulated funds are invested in a diverse portfolio, including stocks, bonds, and real estate. The pension amount varies depending on the individual's contribution period and income.

The contributions are determined by applying the contribution rate to the average income of contributors, having wage growth and demographic variables. Benefit expenditure is calculated based on the recipient's average wage and contribution period, expressed through the income replacement ratio, indicating what percentage of the average wage is received as a pension. The fund's management is projected using the fund's return rate.

These variables are set deterministically by government institutions using different scenarios for various risk situations. This valuation method is known as deterministic modeling and is widely used in many other countries as well. While it simplifies calculations and is easy to interpret, it does not account for the uncertainty of input variables, which is its limitation.

There are two types of the actuarial valuation model( Cheng et al., 2004); a deterministic model and a stochastic model. Deterministic model is defined a model with specified assumptions for and relationships among variables, which fully determines a single outcome directly reflecting the specifications. In contrast, a stochastic model projects a probability distribution of potential outcomes, allowing for random variation in one or more variables over time. The distributions of them are derived from a large number of simulations.

Both models have advantages and disadvantages. For example, Lee et al. (2003) argues that the deterministic model can be inaccurate due to uncertainties in the assumptions. Tuljapurkar and Steinsaltz (2019) supports the use of stochastic models, which can better capture risks and uncertainty by using probability distributions and historical fluctuations.

### 2.2.3 Stochastic model

Currently, the NPS conducts actuarial valuations every five years to evaluate its financial health with deterministic models. Baek (2016) highlights the need to incorporate stochastic analysis into actuarial evaluations, as deterministic models are hard to address the long-term uncertainties that affect pension sustainability. Kim and Song (2020) adds that sensitivity analyses used to represent uncertainties in the NPS's financial projections are insufficient, suggesting that variables such as investment returns and wage growth rate should be treated as stochastic to provide a clearer understanding of uncertainty in terms of probability. However, there was a lack of detailed explanation regarding the benefit amounts. While a Log-normal distribution was applied to model the fund return rate, there was no clear justification provided for this choice. Additionally, the standard deviation of the distribution was estimated using historical performance data, and then three values chosen by Kim were used for sampling the return rate, but the basis for selecting these values was unclear. Moreover, although wage growth was assumed to follow a normal distribution based on the autocorrelation function and partial autocorrelation function plots, an AR(1) model seems more appropriate.

Furthermore, the literature demonstrates that stochastic models have been applied in other countries. For instance, Lee et al. (2003) utilized the Lee-Carter model for demographic variables and VAR models for economic factors to forecast the U.S. Social Security Trust Fund, treating birth rates, mortality rates, real interest rates, and wage growth as stochastic. However, they did not consider the correlation between these variables. Following this approach, I will also treat these four variables as stochastic in my study.

while Tian and Zhao (2016) applied AR models to the input variables and Tomaš (2020) applied AR models and Lee-Carter models to project public pension systems in China and Croatia, respectively. Both studies lack a clear explanation for why these models were used. In Tian and Zhao (2016), there was no verification of the assumptions necessary for using the AR model, whereas in Thomas's study, the Dickey-Fuller test was applied to check for stationarity. Additionally, neither paper considered the correlation between variables, assuming them to be independent. The U.S. OASDI system also uses AR and VAR models to account for various uncertainties in their financial forecasts.(Cheng et al., 2004)

## 2.3 Population Variables

### 2.3.1 Fertility Rate Model

In forecasting population trends, selecting the appropriate model to estimate the Age-Specific Fertility Rate (ASFR) is critical. The ASFR is the number of births given by women of a specific age divided by the total population of women in that age group. There are several approaches to forecasting ASFR, each with distinct characteristics and applications depending on the data quality and patterns observed. There are three main approaches to forecasting fertility rates: parametric models, non-parametric models, and Bayesian models.

Parameterized models forecast future fertility by applying observed age-specific fertility rates (ASFR) to functions and estimating parameters through time series models. Common examples include the Gamma function (Hoem et al., 1981), Hadwiger function (Hadwiger, 1940), PK1 function (Peristera and Kostaki, 2007), and the Log-Gamma function (Kaneko, 2003). These models are based on the assumption past fertility patterns will continue into the future. making them most effective when the data is complete and of high quality. The main advantage of this approach is that it simplifies the modeling process by using statistical functions to fit the fertility distribution. However, a limitation is that parameter estimates can be unstable.

Another example of a parametric model is the Lee & Carter model (Lee and Carter (1992), Lee (1993), and Lee and Tuljapurkar (1994)). This model utilizes principal component analysis (PCA) as its core methodology, which is particularly useful for modeling fertility patterns over time. Also it separates age and time variables, making it both easy to explain and mathematically simple. It ensures the ease of interpretation and broad applicability in demographic forecasting.

Secondly, Non-parametric methods are typically applied when the ASFR time series is discontinuous or highly volatile. These methods make use of smoothing and weighting techniques to handle the irregularities in the data. They are particularly useful when the data shows significant fluctuations or gaps. For example, Ramsay and Silverman (2002) and Hyndman and Ullah (2007) introduced the Functional Data Model (FDM), which adjusts for irregular patterns caused by factors such as measurement errors, diseases, or wars.

Finally, Bayesian models predict fertility by incorporating past fertility data and updating the predictions as new information becomes available. These models are especially useful in cases where data is scarce or of poor quality. They often reference the three-stage fertility transition observed in developed countries as prior distributions. The major advantage of Bayesian methods is that they incorporate uncertainty, providing probabilistic forecasts rather than a single point estimate, and adapt to changes in patterns over time. Key studies in this area include Alkema et al. (2011) and Raftery et al. (2012).

Each method has its strengths and limitations, and the choice of model depends largely on the data characteristics and the specific forecasting needs.

Statistics Korea (2023) uses the Generalized Log-Gamma (GLG) model(Kaneko, 2003) to estimate the ASFR and the NPS uses this model when forecasting population. Similarly, J. Oh (2020) utilized the GLG model by fitting four parameters using a time-series model. It mentions that the theoretical foundation of the model's probability distribution as a key reason for its use, which enables theoretically grounded statistical approach to parameter estimation. Park et al. (2013) shows that the GLG model has the lowest Mean Absolute Error (MAE) when compared to other models. It also suggests a mixed GLG model using data from France and Germany to account for potential changes in fertility patterns due to pro-natal policies. However, since this model is based on data from other countries, applying it to South Korea introduces the possibility of basis risk.

Given the strong fit of the Gamma model for South Korea's fertility trends, I will employ the GLG model to estimate parameters and predict future ASFR based on a time-series forecast of the model's four parameters. However, I acknowledge a limitation of this model which is that it assumes that past fertility patterns will continue into the future like other parametric models.

### 2.3.2 Migration Model

According to Park et al. (2013), migration rarely impacts on South Korea's population structure due to its relatively low migration rates. As such, it is not expected to significantly influence future population projections. For this reason, I will exclude migration from the population forecast model, focusing instead on fertility and mortality rate.

### 2.3.3 Mortality Rate Model

Mortality models have been developed to analyze the randomness in mortality intensity. These models are generally divided into two categories: discrete-time and continuous-time models. Discrete-time mortality models are commonly used in population projections, where the analysis focuses on age-specific mortality rates over set intervals. On the other hand, continuous-time mortality models treat mortality intensity as a stochastic process, making them useful in pricing life-linked derivatives. While the former is crucial for demographic forecasting, the latter plays a key role in financial applications, particularly for life insurance and pension products.

Discrete-time models, such as the Lee-Carter model (Lee & Carter, 1992), are widely used in population projections. It expresses mortality levels with a single index which requires three parameters: one that captures the shape of mortality, a period effect, and an error term. The model uses the logarithm of central mortality but has limitations such as no smoothness across ages or years and only accounts for age-period effects with one factor. Cairns et al. (2006) builds upon the Lee-Carter model by introducing smoothness between mortality rates at different ages. It assumes that the level affects all ages while the slope impacts higher ages. This model uses two period effects and applies a logit transformation to the mortality rates. Another extension, Renshaw and Haberman (2006) adapted the Lee-Carter model to incorporate cohort effects. It includes not only the main age effects and period effects but also considers the cohort effect, making it a more comprehensive model for mortality projection.

Statistics Korea currently applies the Li-Lee-Gerland extend model (Li et al., 2013) to forecast sex-specific and age-specific mortality rates (Statistics Korea, 2023). This model improves upon the limitations of the original Li-Lee model. The Li-Lee model can be expressed as:

$$\log m_{x,t,i} = a_{x,i} + B_x K_t + b_{x,i} k_{t,i} + \epsilon_{x,t,i}$$

$B_x$ above captures the sensitivity of mortality changes across age groups and reflects the differences in mortality improvement over time. However, in South Korea, recent trends show that mortality improvement has slowed for younger ages while accelerating for older ages. To address this, the extended model replaces the fixed $B_x$ with a time-varying $B_{x,t}$ to capture this rotation in age-specific mortality improvement patterns. Additionally, to correct the rapid decline in infant mortality, the model applies Japan's decline rate for age 0. However, this introduces potential basis risk due to reliance on foreign data.

Continuous-time mortality models focus on capturing mortality intensity over time. One notable example is the Affine mortality model (Milevsky and Promislow, 2001), which employs the affine term structure typically used in interest rate modeling (Duffie and Kan, 1996). This model treats the short rate of mortality intensity as an exponential affine function, offering advantages such as analytical tractability and closed-form expressions for cohort survival curves.

Here are specific research examples that have applied the Affine mortality model. Biffis (2005) modeled both asset prices and mortality dynamics using affine jump diffusions, demonstrating how stochastic modeling of mortality can create a flexible framework for actuarial valuation. Dahl (2004) modeled mortality intensity as a stochastic process, allowing market reserves to incorporate uncertainties in future mortality developments. Schrager (2006), using the univariate Kalman filter(Koopman and Durbin, 2000), simplified the estimation of parameters for the Affine mortality model. Additionally, Huang et al. (2022) and Ungolo et al. (2024) applied the model to historical age-cohort survival data, while Ungolo et al. (2023) developed the R package `AffineMortality` to facilitate comprehensive mortality forecasting and analysis, building on prior studies like Blackburn and Sherris (2013) Huang et al. (2022) and Ungolo et al. (2024).

## 2.4 Economic Variables

### 2.4.1 Wage Growth Rate Model

Kim and Song (2020) assumed that wage growth rates follow a white noise. This conclusion was based on the results of the ACF (autocorrelation function) and PACF (partial autocorrelation function). Similarly, the U.S. OASDI system (Cheng et al., 2004) estimated real wage growth rates using a time-series model, taking into account fluctuations in the Consumer Price Index (CPI).

### 2.4.2 Investment Return Rate

Lee et al. (2003) estimated the return on equity and interest rates by modeling them as vector variables in a vector autoregression (VAR) model. Kim and Song (2020) assumed that investment returns follow a log-normal distribution and used Monte Carlo simulations to estimate the distribution of the fund's trajectory. Tian and Zhao (2016) used a three-year weighted average return rate as a long-term average return rate for projecting pension fund returns.

## 2.5 Conclusion

This study contributes to the literature by addressing a key methodological gap in long-term pension forecasting. While prior research has primarily relied on deterministic projections, this study introduces a stochastic simulation framework that captures uncertainty in both demographic and economic variables. Specifically, it integrates a Generalised Log-Gamma (GLG) fertility model with an affine mortality model to simulate population dynamics under uncertainty. The framework is further extended to the pension fund, incorporating sensitivity analyses on economic assumptions such as wage growth and investment return rates.

# Chapter 3

# Model Framework

## 3.1 Introduction

In this Chapter, we propose a mathematical framework and refer to its assumptions in detail for the forecasts of the National Pension. First, Section 3.2 discusses actuarial valuation model for the long term. Section 3.3 discusses population model.

## 3.2 Actuarial Valuation Model

The Actuarial model of the simplest national pension is composed of contributions, expenditures, and funds.

### 3.2.1 Contributions

Contributions are determined by multiplying contribution rate with the average income and the demography of policyholders. The total in contributions paid by working generations has the model follows

$$C_t = CR_t \times W_{t,a} \times P_{t,a} \times 73.3\%$$

| | |
|---|---|
| $C_t$ | Contributions in year $t$ |
| $CR_t$ | Contribution rate in year $t$ |
| $W_{t,a}$ | Average wages of employed persons in year $t$ at age $a$ |
| $P_{t,a}$ | Population in year $t$ at age $a$ |

Working people from age 18 to age 59 can be insured and should pay the premium with 9% of contribution rate currently. We assume the participation rate in the NPS to be 73.9% throughout the projection period based on the proportion of individuals aged 18 to 59 in 2023(National Pension Service, 2023b).

### 3.2.2 Expenditures

Expenditures are calculated based on average wage multiplying the income replacement ratio, and the number of recipients. The model follows

$$E_t = RR_t \times W_{t,b} \times P_{t,b} \times 51.2\%$$

| | |
|---|---|
| $E_t$ | Expenditures in year $t$ |
| $RR_t$ | Income Replacement rate in year $t$ |
| $P_{t,b}$ | Population in year $t$ at age $b$ |

Retired employees can be get paid over age 65. What percentage of one's average wage that can be paid is called the replacement rate and it is assumed to be 24.2%. The current nominal income replacement rate set by the NPS is 43%, which refers to the proportion of pension benefits relative to average income for an individual who has contributed for 40 years. However, the actual income replacement rate is 24.2% because the average contribution period is only 18.6 years(Jeong, 2022). Hence, it is more reasonable to adopt actual income replacement rate as the assumption. Also, we assume the recipient rate to be 51.2% of individuals aged 65 and older receive pension benefits throughout the projection period, based on the proportion observed in 2023(National Pension Service, 2023b).

### 3.2.3 Pension Fund

Funds are added by the difference between contributions and expenditures and are invested with the return rate of funds. Funds are used when contributions are less than expenditure to compensate the short. The figure 3.1 shows how the fund works. The model of fund follows

$$F_t = (C_t - E_t) + F_{t-1} \times (1 + r_t)$$

| | |
|---|---|
| $F_t$ | Accumulated funding in year $t$ |
| $r_t$ | Rate or return of funding in year $t$ |

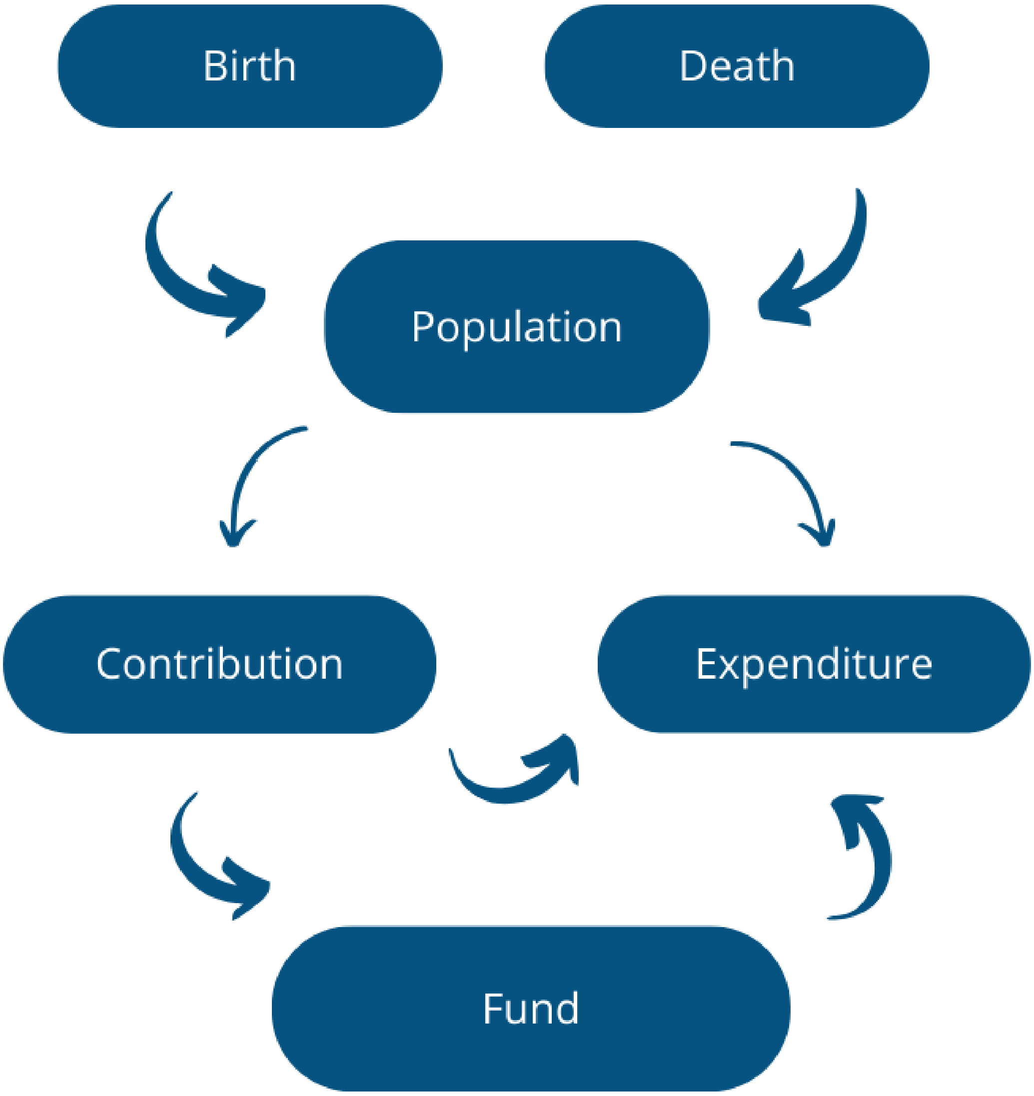


Figure 3.1: Flow of Contributions, Expenditure, and Fund in a Partially Funded Pension System

## 3.3 Population Model

### 3.3.1 Fertility Rate Model - Generalized Log-Gamma Model

The objective is to forecast the number of births in a given year, which is determined by the fertility rates of women aged 15 to 49 in the previous year. Specifically, it is calculated as the sum of Age-Specific Fertility Rate (ASFR) across those ages.

The probability density function (PDF) and the cumulative distribution function (CDF) for the Generalized Log Gamma (GLG) distribution is:

$$g(x) = \frac{|\lambda|}{b\Gamma(\lambda^{-2})} \left(\lambda^{-2}\right)^{\lambda^{-2}} \exp\left[\lambda^{-1}\left(\frac{x-\mu}{b}\right) - \lambda^{-2}\exp\left(\lambda\left(\frac{x-\mu}{b}\right)\right)\right]$$

$$G(x) = 1 - I\left(\lambda^{-2}, \lambda^{-2}\exp\left(\lambda\frac{x-\mu}{b}\right)\right)$$

where $\Gamma$ denotes the gamma function, $\mu, b, \lambda$ are parameters.

Let $F_i(x; C_i, \lambda, \mu, b)$ be a function of age specific cumulative fertility rate of the i-th child at age $x$ with proportion eventually having i-th child $C_i$ and a set of other parameters $\lambda, \mu, b$, then:

$$F_i(x; C_i, \lambda, \mu, b) = C_i G(x; \lambda, \mu, b)$$

In this model, $\mu, b$ represents the mean and standard deviation of age, and $\lambda$ is a shape parameter that defines the distribution. (Kaneko, 2003)

The observed ASFR of the i-th child at age $a$ should be given by:

$$F_i(a+1) - F_i(a)$$

While Kaneko (2003) primarily focuses on cohort-based analyses, Kaneko notes that the performance of the system to predict cohort and period age specific fertility rates seems satisfactory so that it is utilized for country specific precise fertility projection. Based on this and due to the incomplete cohort data, we use the period ASFR from 2000 and 2023 sourced from the Human Fertility Database (HFD). In addition, the reproductive age is typically defined as between 15 and 49 years old, following international standards in demography and public health. (United Nations, 2022) Therefore, we focus on the period ASFR from age 15 to age 49. In each year, the data are classified into five models according to birth order: first, second, third, fourth, and fifth or subsequent births.

For each birth order, four parameters$-C, \lambda, \mu, b-$ are estimated by fitting the GLG model using the nolinear least squares. Since the GLG model is a nonlinear model with multiple parameters, to predict those parameters, I apply the `nlsLM()` function from the `minpack.lm` package (Elzhov et al., 2023), which uses the Levenberg-Marquardt algorithm, combining the stability of gradient descent with the speed of the Gauss-Newton method. It offers robust convergence for complex nonlinear models.

Estimated parameters are then treated as times series. Each parameter's time series is assumed to evolve as random walk following Lee et al. (2003) where the time-varying parameters are modeled as random walks to incorporate long-term uncertainty. Using Monte Carlo simulation based on the random walk models, trajectories of the parameters are generated for the future forecasts.

### 3.3.2 Mortality Model

Mortality model aims to predict the number of people who survive in a given year, based on age-specific mortality patterns. For individuals aged 0–39, mortality rates are assumed to remain fixed at 2023 levels, as the survival probabilities for these age groups are close to one and exhibit no distinct patterns. Therefore, the analysis focuses on individuals aged 40–109, for whom clearer age-specific mortality trends are observed and explicitly modeled. Current demography in South Korea shows a rotation of the previous pattern in which the decreasing speed of mortality of children is getting slower, but the one of elderly is getting quicker. Korean Statistical Information Service (KOSIS) reflect this situation by using the Extended Lee-Carter method to model the rotation of age patterns (LC-ER model) which is an expansion model of Lee-Carter model.(J. Oh, 2020) LC-ER model can be given by:

$$ln(m_{x,t}) = a_x + B_{x,t}K_t + \epsilon_{t,x}$$

where

$$B_{x,t} = \begin{cases} b_x, & e_0^t < 80 \\ (1 - w_s(t))b_x + w_s(t)b_x^u, & 80 \leq e_0^t < e_0^u \\ b_x^u, & e_0^u < e_0^t \end{cases}$$

$B_{x,t}$ is defined as the linear interpolation between $b_x$ from LC model and $b_x^u$. $b_x^u$ is the mortality rate at the highest of the improvement in mortality rate. $e_0^u$ denotes the life expectancy at the end of rotation phenomenon.

$$b_x^u = \begin{cases} \bar{b}_{15-64}, & 0 < x \leq 64 \\ b_x \times \frac{b_{u,60-64}}{b_{65-70}}, & 65 \leq x \end{cases}$$

$$w_s(t) = \left[0.5 \left\{1 + \sin\left(\frac{\pi}{2}(2w(t) - 1)\right)\right\}\right]^p ; w(t) = \frac{e_0^t - 80}{e_o^u - 80}$$

The Affine mortality model can be expressed as the risk-neutral probability that the individual survives up to time T, conditional on being alive at time t as follows.

$$S_a(t,T) := \mathbb{E}^P \left[e^{-\int_t^T \mu_a(s)\,ds} | \mathcal{G}_{\sqcup}\right] = e^{A^P(t,T) + B^P(t,T)^T X(t)}$$

Where $\mu_a(t)$ denotes the instantaneous mortality intensity of an individual aged a at time t, and $A(t,T)$ and $B(t,T)$ are solutions of a system of ordinary differential equations that depend on the coefficients of the model. The latent factor process $\{X(t)\}$ follows Markov process. In particular, we analyze the multifactor Blackburn-Sherris (BS) model(Blackburn and Sherris, 2013), the aribtrage-free Nelson-Siegel (AFNS) model(Christensen et al., 2011), the aribtrage-free generalized Nelson-Siegel (AFGNS) model(Christensen et al., 2009), and the multifactor Cox-Ingersoll-Ross (CIR) model (Brown and Schaefer, 1994,Geyer and Pichler, 1999,Chen and Scott, 2003,Ungolo et al., 2023). First of all, the multifactor Blackburn-Sherris model has an assumption that the latent factors are independent. The AFNS is one of the special cases of AFGNS and it consider the level, slope, and curvature factors to improve the fit. Lastly, the multifactor CIR model is able to capture the heterogeneity of mortality rates at older ages. It describes the short rate of mortality rate in an exponential affine form and mortality rates are driven by level, slope, curvature factors or state variables in the real-world probability space $(\Omega, \mathcal{F}, P)$, where $P$ is the real-world probability measure. The `AffineMortality` R package(Ungolo et al., 2023) is used to fit Mortality rates using age period mortality data for South Korea which is sourced from Human Mortality Database (HMD). It uses univariate Kalman filter approach and it performs via the coordinate ascent algorithm, which maximizes the log-likelihood function iteratively by parameter groups rather than jointly, using the Nelder–Mead simplex method. Four different model are mainly compared:

- BS with independent three factors
- BS with dependent three factors
- AFNS with independent three factors

- AFNS with dependent three factors

The comparison of models is based on the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Root Mean Squared Error (RMSE). Then, future mortality rates are generated via Monte Carlo Simulation by the selected model using the dynamics of latent factor $X(t)$ which follows:

$$dX(t) = -\Delta X(t)dt + \Sigma dW^Q(t)$$

### 3.3.3 Population Projection Model

For the population forecasts stochastically, we use the method called cohort-component method which has a longstanding tradition in demography ( Smith et al., 2013, Alho and Spencer, 1985, Lee and Tuljapurkar, 1994). In addition, the sex ratio at birth is assumed 1.06 male births per female births. It is reasonable because its figure has been stable in recent years by United Nations (2022). The demographic estimation methods refer to Nakazawa (2020).

# Chapter 4

# Results & Discussion

## 4.1 Age Specific Fertility Rate

As a supplementary analysis, Figure 4.1 compares the observed and predicted age-specific fertility rates (ASFR) by birth order for the Korean female cohort born in 1983. Figure 4.2 shows the sum of ASFR by birth order of the cohort. Overall, the model captures the observed trends across all birth orders reasonably well. However, the model shows poorer fit in the age ranges where fertility rates are most concentrated, particularly underestimating ASFR in the early childbearing ages, such as the early twenties. In addition, due to the lack of complete cohort data, it was not possible to compare ASFR beyond age 40 with observed values. It is because South Korea only began collecting cohort-based ASFR data by birth order in the 2000. Further research can be conducted once complete cohort data become available, especially for higher age ranges and later birth orders.

Figure 4.3 presents the fitted and observed period ASFR by birth order and in aggregate for the years 2000 and 2023. It is apparent that there is a clear rightward shift in the fertility distribution, indicating that women tend to postpone childbirth in recent years. In 2000, second-order births were relatively common, whereas in 2023, even second births occurred far less frequently. This is evident from the growing disparity between first and second birth rates in 2023. In terms of model performance, the fitted curves tend to underestimate fertility at very early and very late childbearing ages. This likely contributes to the underestimation of future population size in the simulation. One possible reason is that the log-gamma distribution has a steep decline in its tails, making it less flexible in capturing outlier values. In addition, the

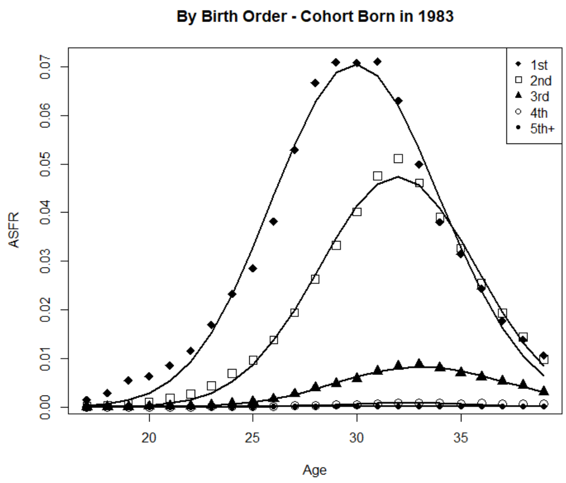


Figure 4.1: Cohort ASFR By birth order

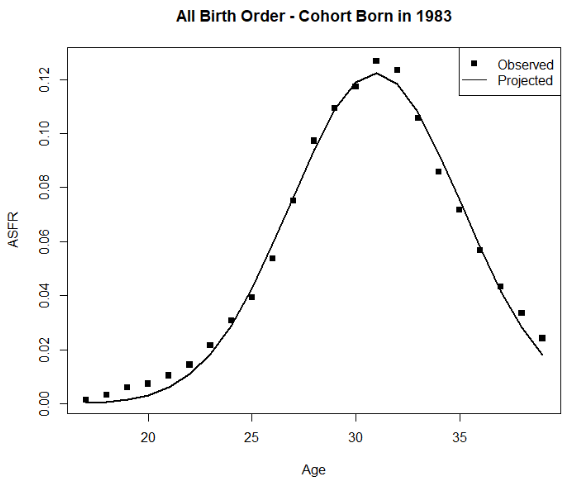


Figure 4.2: Total Cohort ASFR

model failed to accurately capture the peak fertility at modal ages, suggesting that it does not fit well for the most frequent age groups. It is because of limited number of observations, leading to lower accuracy. Future improvement may involve applying additional weights to modal values to better reflect their importance in estimation. To forecast future ASFR, the GLG model was

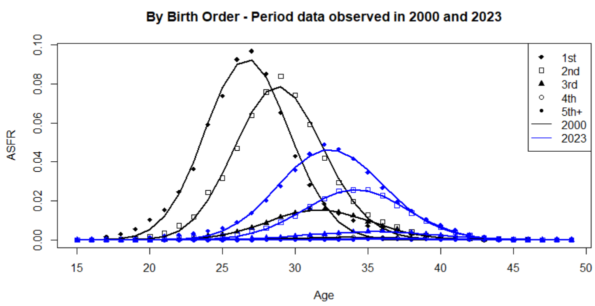


Figure 4.3: ASFR By birth order in 2000 and 2023

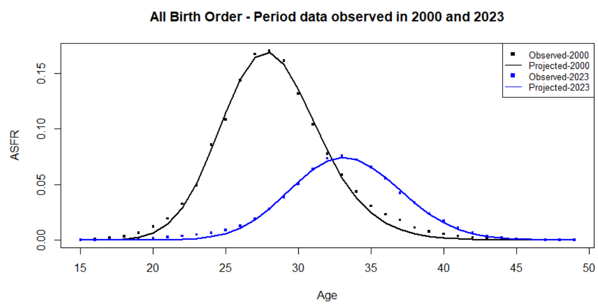


Figure 4.4: Total ASFR in 2000 and 2023

fitted to period ASFR by birth order for each year, and the estimated optimal parameters were treated as time series. From these, parameter values were projected beyond 2024 using time series models. To examine the underlying stochastic process of each parameter, autocorrelation function (ACF) tests were conducted on the parameter series for each birth order at Table 4.1. While some parameters exhibited stationary behavior, for consistency within the stochastic simulation framework, all parameters were assumed to follow a random walk.

| Birthorder | Parameter | Statistic | CriticalValue | Decision |
|---|---|---|---|---|
| ASFR1 | C | -0.51 | -3.00 | Random Walk |
| ASFR1 | u | -0.32 | -3.00 | Random Walk |
| ASFR1 | b | -1.44 | -3.00 | Random Walk |
| ASFR1 | lambda | -1.36 | -3.00 | Random Walk |
| ASFR2 | C | -0.12 | -3.00 | Random Walk |
| ASFR2 | u | -1.71 | -3.00 | Random Walk |
| ASFR2 | b | -0.61 | -3.00 | Random Walk |
| ASFR2 | lambda | -20436.16 | -3.00 | Stationary |
| ASFR3 | C | -0.21 | -3.00 | Random Walk |
| ASFR3 | u | 0.17 | -3.00 | Random Walk |
| ASFR3 | b | 0.31 | -3.00 | Random Walk |
| ASFR3 | lambda | -1.45 | -3.00 | Random Walk |
| ASFR4 | C | -0.99 | -3.00 | Random Walk |
| ASFR4 | u | -0.98 | -3.00 | Random Walk |
| ASFR4 | b | -2.32 | -3.00 | Random Walk |
| ASFR4 | lambda | -2.97 | -3.00 | Random Walk |
| ASFR5 | C | -1.51 | -3.00 | Random Walk |
| ASFR5 | u | -3.50 | -3.00 | Stationary |
| ASFR5 | b | -7.65 | -3.00 | Stationary |
| ASFR5 | lambda | -2.19 | -3.00 | Random Walk |
| ASFR1 | C | -0.51 | -3.00 | Random Walk |
| ASFR1 | u | -0.32 | -3.00 | Random Walk |
| ASFR1 | b | -1.44 | -3.00 | Random Walk |
| ASFR1 | lambda | -1.36 | -3.00 | Random Walk |
| ASFR2 | C | -0.12 | -3.00 | Random Walk |
| ASFR2 | u | -1.71 | -3.00 | Random Walk |
| ASFR2 | b | -0.61 | -3.00 | Random Walk |
| ASFR2 | lambda | -20436.16 | -3.00 | Stationary |
| ASFR3 | C | -0.21 | -3.00 | Random Walk |
| ASFR3 | u | 0.17 | -3.00 | Random Walk |
| ASFR3 | b | 0.31 | -3.00 | Random Walk |
| ASFR3 | lambda | -1.45 | -3.00 | Random Walk |
| ASFR4 | C | -0.99 | -3.00 | Random Walk |
| ASFR4 | u | -0.98 | -3.00 | Random Walk |
| ASFR4 | b | -2.32 | -3.00 | Random Walk |
| ASFR4 | lambda | -2.97 | -3.00 | Random Walk |
| ASFR5 | C | -1.51 | -3.00 | Random Walk |
| ASFR5 | u | -3.50 | -3.00 | Stationary |
| ASFR5 | b | -7.65 | -3.00 | Stationary |
| ASFR5 | lambda | -2.19 | -3.00 | Random Walk |

Table 4.1: Autocorrelation Test Result for parameters of GLG model

| Model | AIC | BIC | RMSE |
|---|---|---|---|
| BS ind.3 | -21557 | -21477 | 0.000324 |
| BS dep. | -22916 | -22805 | 0.000009 |
| AFNS ind | -20270 | -20202 | 0.0001 |
| AFNS dep | -22198 | -22113 | 0.000024 |

Table 4.2: Comparison of in-sample performance for Affine mortality models

## 4.2 Mortality Rate

Model selection was based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Root Mean Squared Error (RMSE). Four affine mortality models were implemented: the BS model with three independent factors, the BS model with three dependent factors, the AFNS model with three independent factors, and the AFNS model with three dependent factors. Among them, the BS model with three dependent latent factors exhibited the best in-sample performance, yielding the lowest values across all numerical criteria as Table 4.2. In addition to in-sample fit, forecast performance was also considered as a criterion for model selection, given the long-term nature of survival probability projections. As illustrated in Figure 4.5, the BS model with three dependent latent factors demonstrates stable long-term forecasts, while the other models become increasingly unstable as the forecast horizon extends where survival probabilities exceed 1, which violates the fundamental definition of a probability. Based on the selected BS model with dependent factors, the survival curve was plotted in Figure4.6. The model exhibits increasing survival probabilities over time, which aligns well with the current demographic trend of rising life expectancy in South Korea. This suggests that the model appropriately reflects the ongoing improvements in longevity.

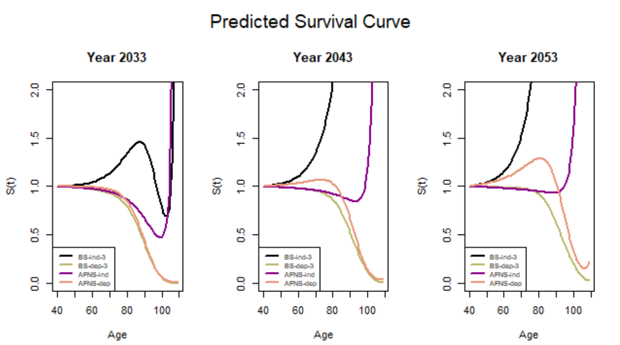


Figure 4.5: Survival Curve Forecast of Affine mortality models

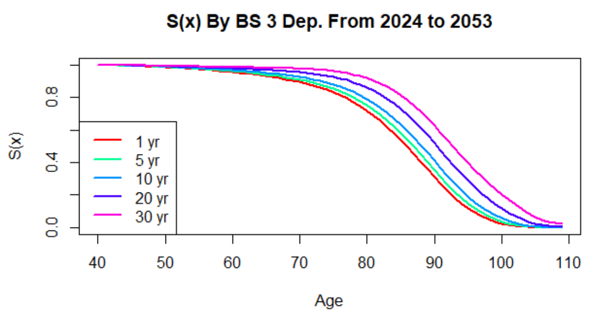


Figure 4.6: Survival Curve Forecast of BS with 3 Dependent factors model

## 4.3 Population

Using the projected fertility and survival rates, the population was stochastically forecasted. A total of 200 iterations were performed to generate a distribution of possible outcomes, from which 95% confidence intervals were derived. This approach allowed for the explicit incorporation of uncertainty in population projections. As shown in the figure 4.7, the shaded blue area represents the 95% confidence interval, while the blue line indicates the median population across simulations. Similar to the population forecast, the old-age dependency ratio was also projected stochastically to capture long-term uncertainty in the figure 4.8. This ratio serves as a key indicator of the burden placed on the working-age population to support the elderly. By quantifying the uncertainty around this metric, the projection provides valuable insights for anticipating future challenges and guiding policy decisions—particularly in the context of pension and tax system reforms. Over the next 30 years, South Korea's population is projected to decline significantly—from approximately 50 million to just 30 million. This dramatic demographic shift is accompanied by a sharp rise in the old-age dependency ratio, which is expected to reach 40. In other words, every 100 working-age individuals will need to support around 40 elderly people, placing increasing financial and social pressure on the economically active population. As the forecast horizon extends, the confidence intervals widen, illustrating the growing uncertainty inherent in long-term demographic projections.

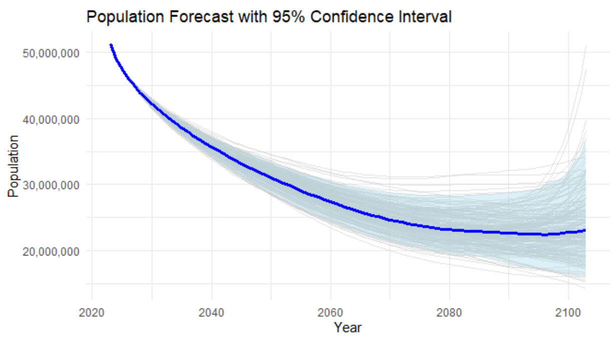


Figure 4.7: Stochastic Population Forecast

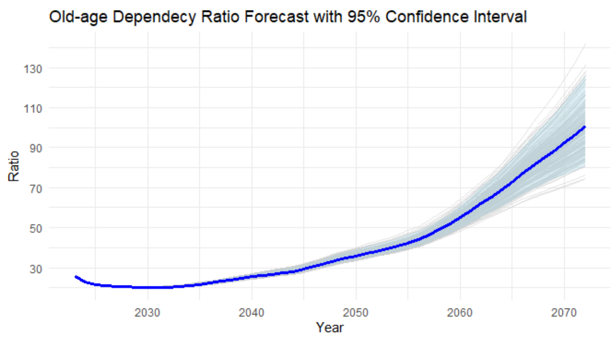


Figure 4.8: Stochastic Old-age Dependency Ratio Forecast

## 4.4 Fund

Using the population forecast along with fixed economic variables, contribution rates, and replaceemnt rates, the pension fund was projected over time. The figure 4.9 illustrates when the

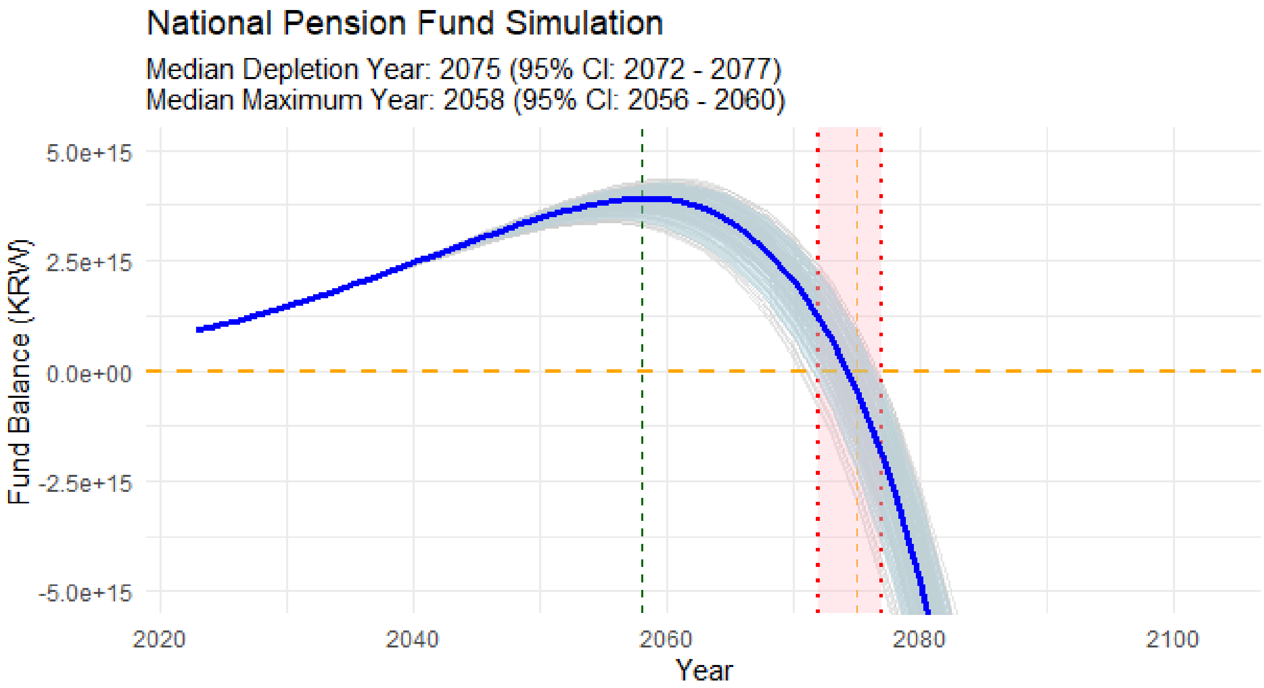


Figure 4.9: Stochastic Fund Simulation

fund is expected to reach its peak and how it gradually declines thereafter until full depletion. Ensuring that the fund does not deplete prematurely is crucial for long-term sustainability. To capture the uncertainty in the timing of depletion, both the median depletion year and the 95% confidence interval were estimated based on stochastic simulations. In 2023, the national pension fund stood at KRW 950 trillion. Under the current system, assuming a wage growth rate of 3.7% and an investment return of 4%, the fund is projected to reach its peak in 2058 and be fully depleted by 2075, with a 95% confidence interval ranging from 2072 to 2077. Following depletion, government subsidies may be required to maintain pension payouts, which could lead to higher tax burdens. Therefore, delaying the depletion through pension reform is essential to ensuring the fund's long-term sustainability.

## 4.5 Comparison with NPS

The table 4.3 compares key projection outcomes from the 2023 National Pension financial report and those estimated in this study. NPS conducts financial projections every five years to assess the long-term sustainability of the pension system(National Pension Service, 2023a). Both projections use the same fertility model, but differ in their mortality assumptions—this study applies an affine mortality model, whereas the NPS report uses an alternative deterministic

| | NPS | This study |
|---|---|---|
| Fertility Model | GLG Model | GLG Model |
| Mortality Model | Li-Lee-Gerland (2013) | Affine Mortality |
| Population in 2070 | 37.7m | 21.4m (95% CI: 21.4m-28.8m) |
| Old-age Dependency Ratio in 2070 | 62.06 | 92.3 (95 % CI: 96.6-113) |
| Depletion Year | 2057 | 2075 (95% CI: 2072-2077) |

Table 4.3: Comparison of Result between NPS and This study

model. In terms of population forecasts, there is a substantial difference after 2070: the NPS projects a population of 37.7 million, while this study estimates a significantly lower 21.4 million. One potential reason for this discrepancy is that migration was not taken into account in this study. With respect to the fund depletion year, the NPS projects depletion in 2057, while this study estimates a later depletion in 2075. This may be attributed to several simplifying assumptions, including fixed economic variables, which may be overly optimistic, and the use of an oversimplified long-term actuarial valuation formula.

## 4.6 Sensitivity Analysis

In the model framework, the wage growth rate and investment return were initially fixed in order to isolate the effects of demographic changes. However, to evaluate the influence of economic factors on the fund's sustainability, a sensitivity analysis was conducted. We implemented a set of scenario combinations in which the wage growth rate ranged from 2% to 5%, and the investment return rate also ranged from 2% to 5%. Figure 4.10 illustrates how the median depletion year varies depending on these two economic parameters. Darker colors represent later depletion years, while lighter colors indicate earlier depletion. Two extreme scenarios help interpret the results. An optimistic scenario with lower wage growth and higher investment return leads to delayed depletion. A pessimistic scenario with higher wage growth and lower return results in earlier depletion. From the gradient pattern in the figure, it can be inferred that a 1 percentage point increase in investment return postpones the depletion year more significantly than a similar change in wage growth. Thus, investment return appears to have a stronger influence on fund sustainability. While economic variables such as wage growth and

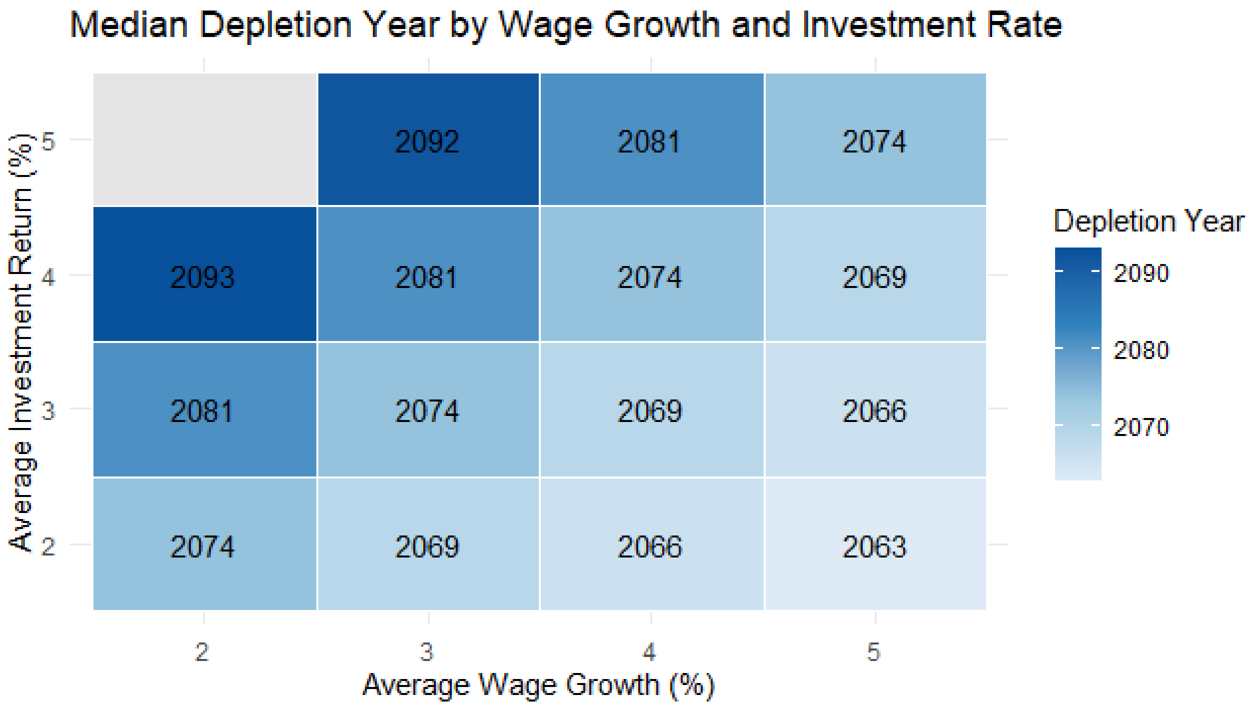


Figure 4.10: Sensitivity Analysis of economic variables

investment return can be influenced by macro-level strategies, they cannot be adjusted directly at will. Efforts to raise wage growth may include expanding labor market participation among older populations and creating more high-quality jobs. Moreover, to increase investment return, pension funds need to strengthen portfolio strategies that consider both long-term profitability and sustainability. Achieving desirable values for these variables requires coordinated and multi-faceted policy efforts. In contrast, parameters that have a direct and immediate impact on pension fund sustainability—such as the contribution rate, replacement rate, and pensionable age—can be adjusted more directly through pension reform.

## 4.7 Pension Reforms

Figure 4.11 illustrates how the depletion year changes under different reform scenarios. The three heatmaps below show how the fund depletion year changes as the pensionable age increases from the current 65 to 66 and 67 from left to right. We also examine the effects of changes in the contribution rate—from the current 9% to 11%, 13%, and 15%—which determines the percentage of income contributed by workers, as well as the replacement rate, which is currently 40% and was adjusted to 30%, 45%, and 50% to explore its impact on fund sustainability. The results indicate that increasing the contribution rate, reducing the replacement rate, and raising

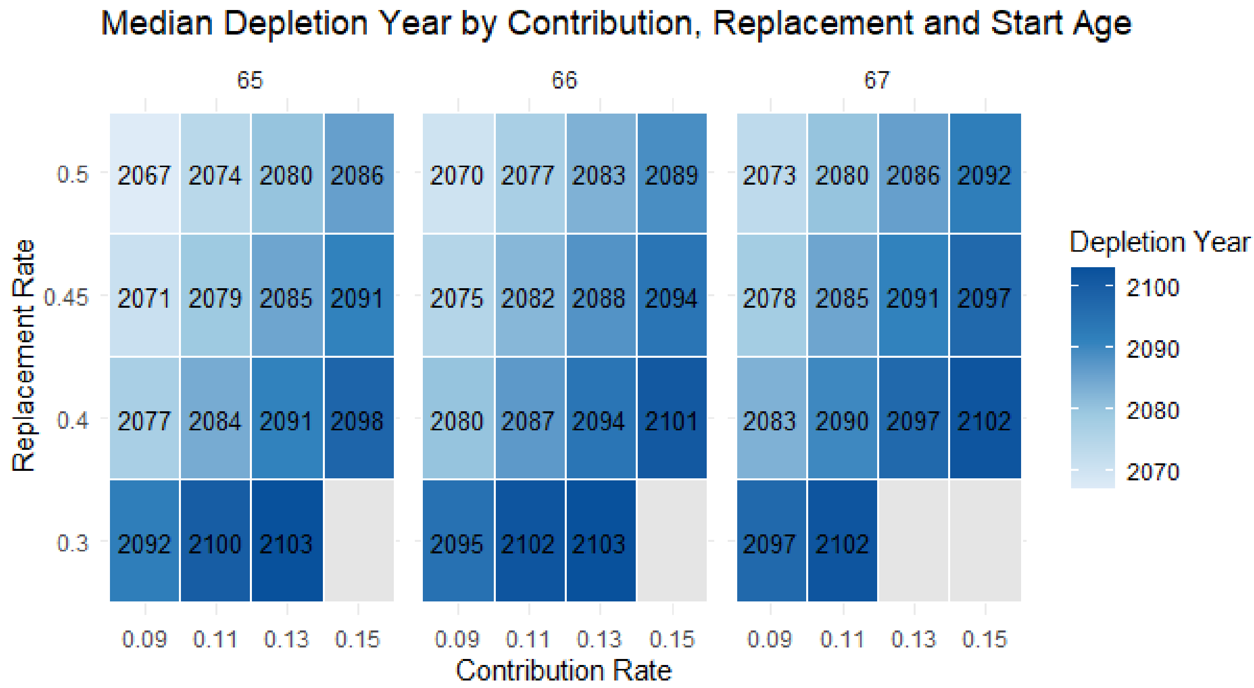


Figure 4.11: Sensitivity Analysis of Pension Reform Scenarios

the pensionable age each contribute to delaying fund depletion. Among the three policy levers, when comparing the maximum effect within the defined range of each variable, we find that raising the contribution rate has the most significant impact on extending the fund.

The National Pension Service is currently discussing the introduction of an automatic adjustment mechanism known as the Macro Slide system. This system aims to improve pension sustainability by automatically reducing pension benefits when inflation (CPI) increases more rapidly than demographic changes. Its primary advantage lies in its ability to gradually adjust benefits downward in response to adverse economic and demographic conditions. To explore how this system might affect the projected depletion year of the pension fund, we incorporated it into the stochastic forecast model. Specifically, past CPI data were fitted using the auto.arima() function in R, which selected an ARMA(1,1) model based on AIC minimisation. The fitted model was then used to generate future CPI values. Demographic changes were captured through two key indicators: the rate of increase in life expectancy and the decline rate of the younger population. When CPI exceeded these demographic indicators, benefits were reduced accordingly. As shown in Figure 4.12, the macro slide is triggered in the early stages of the projection. This is likely because the ARMA(1,1)-based CPI forecasts remain relatively stable between 2.0% and 2.5%, whereas the projected demographic changes fluctuate

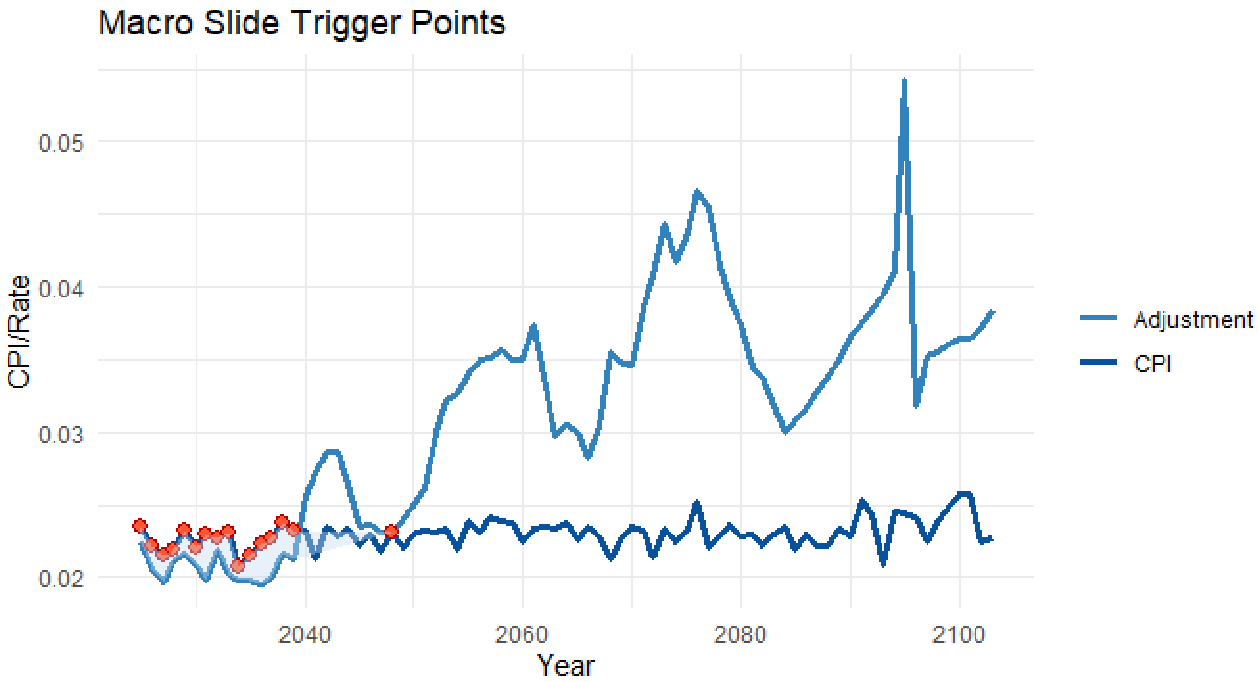


Figure 4.12: Adopt Macro Slide on the Fund

| | Mean | Median | Lower 95 | Upper 95 |
|---|---|---|---|---|
| No Macroslide | 2075.8 | 2076 | 2073 | 2078 |
| With Macroslide | 2076 | 2076 | 2073 | 2079 |

Table 4.4: Macro Slide's Effect on depletion year

more significantly over time. The comparison table in Figure 4.4 shows that the introduction of the macro slide does not significantly change the median depletion year, but it broadens the confidence interval and leads to a slight delay in depletion. This result suggests that while the macro slide may not greatly alter the long-term financial trajectory, it can contribute to absorbing uncertainty and mitigating downside risk.

# Chapter 5

# Conclusions and Future Work

## 5.1 Summary of Outcomes

If the pension fund is depleted, the government will likely need to step in with tax-funded subsidies to cover pension payments. In this study, we applied a Generalised-Log-Gamma fertility model, where parameters are modeled as time series and simulated to forecast age-specific fertility rates (ASFR). We also used an affine mortality model, simulating a Markov process from estimated parameters to project age specific survival probabilities. By combining these models, we stochastically projected population and pension fund dynamics, allowing us to assess long-term financial risk with confidence intervals. Under the current system, the fund is projected to be fully depleted in 2075, with a 95% confidence interval of 2072–2077. However, this projection assumes fixed values for wage growth of 3.7% and investment return of 4%. To account for economic variability, we conducted a sensitivity analysis and found that changes in these economic variables can impact the fund's sustainability and investment return has a greater impact on delaying depletion than wage growth. To explore policy responses, we evaluated various pension reform options — increasing contribution rates, lowering replacement rates, and raising the pensionable age—with contribution rate increases having the most significant effect. Additionally, we also tested a macro slide system that adjusts benefits based on CPI and demographic changes. While the impact was modest, it contributed to a slight delay in depletion and reflected greater uncertainty in projections. In conclusion, quantifying uncertainty is critical for understanding long-term pension sustainability. Stochastic simulation offers a more realistic foundation for evaluating pension risk.

## 5.2 Recommendations & Future Work

As more complete cohort-based fertility and mortality data become available, future work will aim to improve model accuracy by applying a cohort-based approach to both fertility and mortality rates within a unified framework. In addition, rather than assuming fixed economic variables, such as wage growth and investment return, I plan to incorporate stochastic modeling of economic scenarios to better capture the uncertainty inherent in long-term pension forecasting. The current study simplified the pension benefit formula by assuming average wages to focus on demographic long-term simulations. However, the actual South Korean National Pension System involves a complex, nonlinear benefit calculation, which significantly affects individual payouts. Therefore, a key direction for future work is to incorporate the actual pension benefit formula, thereby enhancing both the realism and policy relevance of the simulation results.

# Bibliography


Alho, J. M., & Spencer, B. D. (1985). Uncertain population forecasting. *Journal of the American Statistical Association*, *80*(390), 306–314.

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., & Heilig, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography*, *48*(3), 815–839.

Ando, A., & Modigliani, F. (2005, September). The Life-Cycle Hypothesis of Saving: Aggregate Implications and Tests. In *The Collected Papers of Franco Modigliani, Volume 6*. The MIT Press. https://doi.org/10.7551/mitpress/1923.003.0005

Baek, H. (2016). Enhancing the reliability of the long-term financial projections on the national pension: Introduction to the stochastic models of oasdi and cpp. *Health and Welfare Policy Forum*, *234*, 21–33.

Barr, N., & Diamond, P. (2009). Reforming pensions: Principles, analytical errors and policy directions. *International social security review*, *62*(2), 5–29.

Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: mathematics and economics*, *37*(3), 443–468.

Blackburn, C., & Sherris, M. (2013). Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics*, *53*(1), 64–73.

Brown, R. H., & Schaefer, S. M. (1994). The term structure of real interest rates and the cox, ingersoll, and ross model. *Journal of Financial Economics*, *35*(1), 3–42.

Cairns, A. J., Blake, D., & Dowd, K. (2006). Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin*, *36*(1), 79–120. https://doi.org/10.2143/AST.36.1.2014145

Chen, R.-R., & Scott, L. (2003). Multi-factor cox-ingersoll-ross models of the term structure: Estimates and tests from a kalman filter model. *The Journal of Real Estate Finance and Economics*, *27*, 143–172.

Cheng, A., Miller, M., Morris, M., Schultz, J., Skirvin, J., & of the Chief Actuary, U. S. S. A. O. (2004). *A stochastic model of the long-range financial status of the oasdi program*. Social Security Administration Office of the Chief Actuary. https://books.google.com.au/books?id=75V-0AEACAAJ

Choi, J. (2017). Population projection by adjusting fertility and mortality and its effect on national pension financial prospect. *Journal of The Korean Data Analysis Society*, *19*(1), 371–386.

Choi, J., Kwon, M., & Shin, S. (2016). An analysis of the return rate and financial impact of the national pension based on projected mortality rates and the selection of a mortality projection model. *[NPS] National Pension Research Institute*, *2015*(11).

Choi, K. (2005). Actuarial structure of benefits in the korean national pension system. *Journal of insurance and finance*, *43*, 31–59.

Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2009). An arbitrage-free generalized nelson–siegel term structure model.

Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, *164*(1), 4–20.

Dahl, M. (2004). Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts. *Insurance: mathematics and economics*, *35*(1), 113–136.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, *6*(4), 379–406.

Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. (2023). *Minpack.lm: R interface to the levenberg-marquardt nonlinear least-squares algorithm found in minpack, plus support for bounds* [R package version 1.2-4]. https://CRAN.R-project.org/package=minpack.lm

Geyer, A. L., & Pichler, S. (1999). A state-space approach to estimate and test multifactor cox-ingersoll-ross models of the term structure. *Journal of Financial Research*, *22*(1), 107–130.

Hadwiger, H. (1940). Eine analytische reproduktionsfunktion für biologische gesamtheiten. *Scandinavian Actuarial Journal*, *1940*(3-4), 101–113.

Hoem, J. M., Madien, D., Nielsen, J. L., Ohlsen, E.-M., Hansen, H. O., & Rennermalm, B. (1981). Experiments in modelling recent danish fertility curves. *Demography*, *18*, 231–244.

Huang, Z., Sherris, M., Villegas, A. M., & Ziveyi, J. (2022). Modelling usa age-cohort mortality: A comparison of multi-factor affine mortality models. *Risks*, *10*(9), 183.

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

Jeong, I. (2022). Current status and issues related to the national pension income replacement rate [[In Korean]].

Kaneko, R. (2003). Elaboration of the coale-mcneil nuptiality model as the generalized log gamma distribution: A new identity and empirical enhancements. *Demographic Research*, *9*, 223–262.

Kim, H., & Song, C. (2020). A study on the methodology for long-term financial projections of the national pension considering uncertainty. *[NPS] National Pension Research Institute*.

Koopman, S. J., & Durbin, J. (2000). Fast filtering and smoothing for multivariate state space models. *Journal of time series analysis*, *21*(3), 281–296.

Lee, R. D. (1993). Modeling and forecasting the time series of us fertility: Age distribution, range, and ultimate level. *International Journal of forecasting*, *9*(2), 187–202.

Lee, R. D., Anderson, M. W., & Tuljapurkar, S. (2003). Stochastic forecasts of the social security trust fund. *Michigan Retirement Research Center Research Paper No. WP*, *43*.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting us mortality. *Journal of the American statistical association*, *87*(419), 659–671.

Lee, R. D., & Tuljapurkar, S. (1994). Stochastic population forecasts for the united states: Beyond high, medium, and low. *Journal of the American Statistical Association*, *89*(428), 1175–1189.

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

Milevsky, M. A., & Promislow, S. D. (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, *29*(3), 299–318.

Nakazawa, M. (2020). *Learning demographic analysis using r*. Asakura Publishing Co., Ltd.

National Pension Service. (2023a). National pension fund report 2023 [Accessed: 2025-05-13].

National Pension Service. (2023b). National pension statistics (as of december 2023). https://www.nps.or.kr/pnsinfo/databbs/getOHAF0269M1Detail.do?menuId=MN24000999&pstId=ST20240000000029365&hmpgCd=&hmpgBbsCd=BS20240125&sortSe=FR&pageIndex=2&searchText=&searchGbu=

Oh, C. (1997). The actuarial analysis of benefits and cantributions of the national pension. *Korean Social Security Studies*, *13*(1), 115–153.

Oh, J. (2020). Stochastic population projections on an uncertainty for the future korea. *The Korean Journal of Applied Statistics*, *33*(2), 185–201.

Park, Y., Kim, M., & Kim, S. (2013). Probabilistic fertility models and the future population structure of korea. *The Korean Association for Survey Research*, *14*, 49–78.

Peristera, P., & Kostaki, A. (2007). Modeling fertility in modern populations. *Demographic Research*, *16*, 141–194.

Plamondon, P. (2002). *Actuarial practice in social security*. International Labour Organization.

Raftery, A. E., Li, N., Ševčíková, H., Gerland, P., & Heilig, G. K. (2012). Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences*, *109*(35), 13915–13921.

Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis: Methods and case studies*. Springer.

Renshaw, A., & Haberman, S. (2006). A cohort-based extension to the lee–carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, *38*(3), 556–570. https://doi.org/https://doi.org/10.1016/j.insmatheco.2005.12.001

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of political economy*, *66*(6), 467–482.

Schrager, D. F. (2006). Affine stochastic mortality. *Insurance: mathematics and economics*, *38*(1), 81–97.

Smith, S. K., Tayman, J., & Swanson, D. A. (2013). Overview of the cohort-component method. *A Practitioner's Guide to State and Local Population Projections*, 45–50.

Statistics Korea. (2023). Population projection 2023–2072. *Statistics Korea*.

Tian, Y., & Zhao, X. (2016). Stochastic forecast of the financial sustainability of basic pension in china. *Sustainability*, *8*(1), 46.

Tomaš, I. (2020). A stochastic forecast for the croatian pension system. *Public Sector Economics*, *44*(1), 41–98.

Tuljapurkar, S., & Steinsaltz, D. (2019). Stochastic models for structured populations. In *Handbook of statistics* (pp. 133–155, Vol. 40). Elsevier.

Ungolo, F., Garces, L. P. D. M., Sherris, M., & Zhou, Y. (2023). Affinemortality: An r package for estimation, analysis, and projection of affine mortality models. *Annals of Actuarial Science*, 1–26.

Ungolo, F., Garces, L. P. D. M., Sherris, M., & Zhou, Y. (2024). Estimation, comparison, and projection of multifactor age–cohort affine mortality models. *North American Actuarial Journal*, *28*(3), 570–592.

United Nations. (2022). World population prospects 2022 [[Accessed: 27 April 2025]]. https://population.un.org/wpp/

# Software Documentation

## Code Availability

All scripts and source code used for the simulation and analysis are available here:

https://github.com/HayoungLee-2/UTS_35112