

Article

Mortality Heterogeneity and Pension Redistribution Across Spatial Scales: Evidence from Japan

Ning Zhang, Chenlu Deng and Lingyu He * 

College of Finance and Statistics, Hunan University, Changsha 410006, China; ningzhang@hnu.edu.cn (N.Z.); dengchenlu@hnu.edu.cn (C.D.)

* Correspondence: helingyu@hnu.edu.cn

Abstract

Under a unified pension system, subnational differences in longevity can translate into implicit pension redistribution through differences in the expected duration of pension receipt. Using Japan as a case study, this paper examines the implicit pension redistribution induced by subnational longevity heterogeneity and analyzes how its measurement varies across spatial scales. Accurate measurement of future redistribution requires forecasts of subnational mortality rates, from which future remaining life expectancy can be derived. Because mortality series at different spatial scales are linked by hierarchical aggregation relationships, independent forecasts may violate aggregation coherence and affect the comparability of redistribution estimates. To address this issue, we introduce and compare several forecast reconciliation methods and select the empirical minimum trace (EMinT) method based on its forecasting performance. Using the reconciled mortality forecasts, we estimate future post-retirement remaining life expectancy across subnational areas and subsequently measure implicit pension redistribution. The results show that subnational longevity heterogeneity will persist throughout the forecast period, generating implicit pension redistribution with clear spatial clustering. Redistribution is substantially greater at the prefectural level than at the regional level, and the difference is projected to widen over time. These findings indicate that analyses conducted at more aggregated spatial scales may underestimate the true extent of pension redistribution across areas. This study provides quantitative evidence for the spatial evaluation of pension systems and public policymaking.

Keywords: pension redistribution; mortality heterogeneity; forecast reconciliation; spatial scales



Academic Editor: Montserrat Guillén

Received: 24 May 2026

Revised: 4 July 2026

Accepted: 6 July 2026

Published: 11 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

As population aging accelerates, the sustainability and fairness of pension systems have become central issues in global public policy. Under uniform contribution and benefit rules, systematic differences in life expectancy may give rise to implicit pension redistribution. Specifically, among individuals with identical contribution histories, longer-lived individuals receive pensions for a longer period and consequently obtain greater lifetime pension benefits, whereas shorter-lived individuals receive lower lifetime benefits, resulting in an implicit transfer from the short-lived to the long-lived (Ayuso et al. 2016; Bravo et al. 2021).

Existing studies have primarily examined longevity disparities associated with individual characteristics such as income, occupation, and sex, as well as their resulting pension

redistribution effects. Specifically, [Sanchez-Romero et al. \(2020\)](#) found that longevity gaps between high- and low-income groups generated implicit transfers from low-income to high-income individuals, thereby affecting the actuarial fairness of pension systems. [Ayuso et al. \(2016\)](#) systematically reviewed longevity heterogeneity across dimensions such as income, wealth, and education, showing that less-educated and low-income groups bore implicit taxes, whereas highly educated and high-income groups received implicit subsidies. [Ayuso et al. \(2017\)](#) showed that when a uniform pension benefit formula did not account for gender differences in life expectancy, women received implicit subsidies because of their longer life expectancy, whereas men bore implicit taxes. Using Swedish data, [Shi and Kolk \(2022\)](#) further found that mortality inequality had a more pronounced effect on lifetime pension inequality among men, primarily because life expectancy disparities across educational and income groups were larger among men than among women. However, compared with the focus on individual characteristics discussed above, the geographical dimension of pension redistribution has received insufficient attention. Existing studies of subnational longevity have primarily examined spatial disparities in mortality and life expectancy and their determinants. [Chetty et al. \(2016\)](#) found substantial geographical differences in life expectancy among low-income individuals in the United States and showed that these differences were associated with health behaviors and local socioeconomic conditions. [Dwyer-Lindgren et al. \(2017\)](#) further documented persistent disparities in life expectancy across US counties and linked them to socioeconomic, behavioral, and health-related factors. Using Australian regional data, [Atalay et al. \(2023\)](#) found that spatial mortality inequalities were partly associated with urban-rural longevity disparities and unequal access to healthcare. Although these studies identified subnational disparities in mortality and life expectancy and their determinants, they have yet to examine the redistributive effects of such spatial heterogeneity under a unified pension system. Under a unified pension system, subnational longevity disparities may further translate into implicit spatial redistribution of pension resources. Moreover, longevity heterogeneity may manifest differently across spatial scales, making spatial scale an important consideration when measuring pension redistribution. Such implicit redistribution directly concerns the equitable distribution of pension benefits among residents of different areas. Accurately measuring its intensity, spatial pattern, and future evolution can facilitate a comprehensive assessment of how subnational longevity disparities affect the distribution and fairness of pension benefits, while providing quantitative evidence for the spatial evaluation of pension systems and the design of complementary fiscal support measures.

Accurate measurement of pension redistribution depends on reliable estimates of life expectancy. Because life expectancy is calculated from age-specific mortality rates using life tables, assessing future pension redistribution first requires forecasts of subnational age-specific mortality rates. However, subnational mortality exhibits a clear hierarchical aggregation structure ([Shang and Hyndman 2017](#)). Forecasting each area independently may produce results that fail to satisfy aggregation coherence across spatial levels, thereby undermining the comparability and reliability of pension redistribution measures ([Li and Hyndman 2021](#); [Shang and Yang 2021](#)). To address this issue, we introduce forecast reconciliation methods to preserve subnational mortality heterogeneity while ensuring coherent forecasts across spatial levels, thereby providing a reliable forecasting basis for accurately measuring future spatial pension redistribution.

Based on this background, this study primarily addresses two research questions: First, to what extent do subnational longevity disparities give rise to implicit pension redistribution, and what are their spatial patterns and future evolutions? Second, how do the measurements of pension redistribution differ across spatial scales? To address these questions, we conduct an empirical analysis using Japan as a case study. Japan

provides an ideal research setting due to its high overall life expectancy, distinct longevity disparities across prefectures, and a unified national pension system, which allow for a comprehensive examination of implicit pension redistribution driven by subnational mortality heterogeneity (Nomura et al. 2025). We construct a multilevel mortality dataset covering three spatial scales, including national, regional, and prefectural levels, and evaluate the forecasting performance of various forecast reconciliation methods to select the most suitable approach for Japan's future mortality projections. The results show that the empirical minimum trace (EMinT) method generally achieves better out-of-sample mortality forecasting performance than the other compared approaches across most spatial levels and forecasting horizons. Based on its forecasts, we find that subnational longevity heterogeneity in Japan will persist throughout the forecast period and generate implicit pension redistribution with a clear pattern of spatial clustering under the unified pension system. Longer-lived areas generally receive implicit subsidies, whereas shorter-lived areas bear implicit tax burdens. Further comparisons reveal that the extent of redistribution is more pronounced at the prefectural level than at the regional level, with the disparity widening over time. This suggests that higher levels of spatial aggregation may lead to an underestimation of the true magnitude of pension redistribution across subnational units.

The contributions of this study are threefold: First, while existing studies on pension redistribution mainly focus on individual socioeconomic characteristics such as income, education, and occupation, this study examines implicit redistribution under a unified pension system from the perspective of subnational longevity heterogeneity, thereby extending the literature to the spatial dimension. Second, this study integrates multilevel mortality forecast reconciliation with pension redistribution measurement, constructing a comprehensive analytical framework that links coherent mortality forecasting, estimation of post-retirement remaining life expectancy, and measurement of future pension redistribution. By ensuring that pension redistribution estimates at different spatial levels are derived from coherent mortality forecasts, this framework improves the comparability of redistribution measures across spatial scales. Third, this study compares pension redistribution at the regional and prefectural levels and reveals the effect of spatial aggregation on redistribution measurement, thereby providing quantitative evidence for the spatial evaluation of pension systems and related policy analysis.

The remainder of this paper is organized as follows: Section 2 introduces the data used in this study, constructs a multilevel spatial mortality data system, and presents the basic characteristics of subnational mortality differences. Section 3 introduces the pension redistribution measurement framework, including the construction of pension redistribution indicators, mortality forecasting methods, forecast reconciliation methods for ensuring coherence, and life expectancy estimation based on standard life table methods. Section 4 presents the empirical results, including the comparison and selection of different forecast reconciliation methods, the projected results of future subnational life expectancy, and the pension redistribution induced by longevity heterogeneity, together with further comparisons of pension redistribution outcomes across different spatial scales. Section 5 summarizes the main findings and further discusses the related policy implications and limitations of the study.

2. Data and Descriptive Analysis

This section introduces the data used in this study and provides a preliminary description of subnational mortality disparities in Japan. First, we outline the data sources and the spatial structure. Subsequently, we conduct a descriptive analysis of mortality heterogeneity across areas, establishing the empirical basis for the subsequent measurement of pension redistribution.

2.1. Data Source and Spatial Structure

Japan combines a nationally unified National Pension system with relatively complete subnational mortality data, providing a suitable empirical setting for examining the implicit pension redistribution induced by subnational longevity heterogeneity. Specifically, Japan's public pension system consists of two tiers. The first tier is the National Pension, which provides basic pension coverage to all residents aged 20–59. The second tier is the Employees' Pension Insurance, which covers employees and provides earnings-related pension benefits in addition to the basic pension. This study focuses on the basic pension layer represented by the National Pension. Its contribution requirements, eligibility conditions, and benefit standards are established nationally, without prefecture-specific contribution rates or benefit formulas (Ministry of Health, Labour and Welfare 2025). The National Pension operates within Japan's broader public pension financing framework, which is financed primarily on a pay-as-you-go basis, with pension reserves maintained to support long-term financial stability. According to the Government Pension Investment Fund (GPIF), over the approximately 100-year financial projection period, about 90% of the resources for public pension benefits are expected to come from current pension premiums and government contributions, while about 10% is expected to be financed by pension reserves (Government Pension Investment Fund 2025). In this setting, subnational differences in age-specific mortality translate into differences in remaining life expectancy after retirement and thus in the expected present value of pension benefits, generating implicit pension redistribution across areas.

To measure the above implicit redistribution and its evolution over time, this study uses Japanese mortality data obtained from the Japanese Mortality Database (Japanese Mortality Database 2024). The database provides age-specific death counts, exposure-to-risk data, and mortality rates at both the national and prefectural levels from 1975 to 2022. This study uses prefectural-level data as the basis for constructing a multilevel spatial analytical framework. Long-term mortality and population exposure data are available at the prefectural level, with population sizes sufficient to support the estimation of age-specific mortality rates. Moreover, prefectures are fundamental administrative and statistical units in Japan, giving prefectural-level results substantial policy relevance. Based on these data, we aggregate the 47 prefectures into 8 regions according to Japan's standard geographical classification, thereby constructing a three-level spatially nested structure comprising the national, regional, and prefectural levels. Considering data completeness and comparability, we analyze single-year mortality rates for ages 0–99, while ages 100 and above are combined into an open-ended age group.

We simultaneously consider the national, regional, and prefectural levels, and further distinguish total, female, and male populations in order to characterize mortality differences across areas and genders. This produces a set of age-specific mortality series covering multiple spatial levels and population groups. The detailed structure of the data is summarized in Table 1 and forms the basis for the subsequent mortality forecasting, remaining life expectancy estimation, and pension redistribution analysis.

Table 1. Structure of the Japanese mortality data.

Series Category	Composition	Number of Series
National population	Total, female, male	3
Regional population	8 regions $\times$ 3 groups	24
Prefectural population	47 prefectures $\times$ 3 groups	141
Total	—	168

2.2. Mortality Heterogeneity Across Areas

Subnational mortality heterogeneity is an important source of differences in remaining life expectancy after retirement and also forms the basis for measuring implicit pension redistribution. To illustrate such differences more intuitively, Figure 1 presents the ranking of mortality rates at age 65 across prefectures in 2022 for the total, female, and male populations. Since age 65 generally corresponds to the eligibility age for receiving the National Pension, mortality differences at this age provide intuitive evidence of post-retirement longevity differences and support the subsequent analysis of pension redistribution.

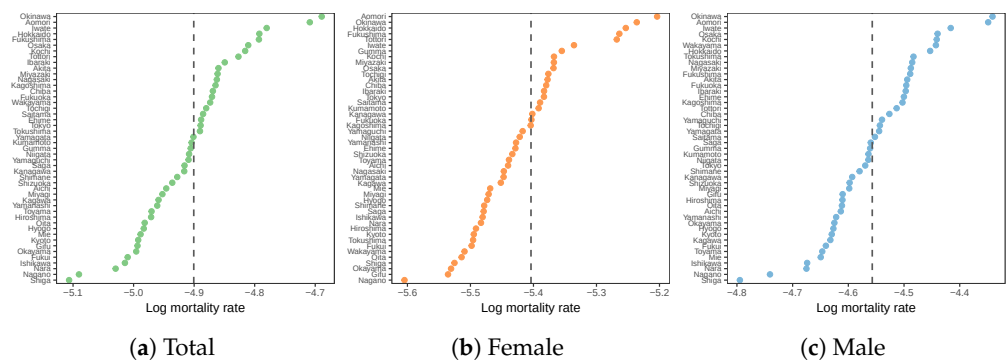


Figure 1. Ranking of log mortality rates at age 65 across Japanese prefectures in 2022.

Overall, substantial mortality differences can still be observed across Japanese prefectures, with mortality levels showing a certain degree of variation rather than clustering around similar values. In particular, prefectures such as Okinawa and Aomori exhibit relatively high mortality levels, whereas Shiga and Nagano show relatively low mortality levels. This indicates that residents in different areas face different mortality risks around retirement age.

From the gender perspective, the overall ranking patterns for males and females are broadly similar, although some prefectures still exhibit differences in their relative positions across gender groups. In addition, mortality levels for males are consistently higher than those for females, which is consistent with the generally higher mortality risk observed among male populations. At the same time, the gap between the highest- and lowest-mortality prefectures is relatively larger for males, indicating more pronounced subnational differences in male longevity. In comparison, female mortality levels are relatively closer across prefectures. The mortality level for the total population generally lies between those of the male and female populations.

In summary, there is clear mortality heterogeneity at the subnational level in Japan. Since remaining life expectancy after retirement is jointly determined by age-specific mortality rates after retirement age, the above mortality differences will be further reflected in differences in post-retirement remaining life expectancy across areas. These differences may then lead to unequal expected durations of pension benefit receipt across areas, thereby generating implicit pension redistribution under the unified National Pension system.

3. Pension Redistribution Measurement Framework

In this section, we present a framework for measuring the implicit pension redistribution under a unified pension system. First, we introduce the indicators used to measure such redistribution. Second, we describe the mortality forecasting model used to project future age-specific mortality rates. Third, to ensure coherence in forecast results across different spatial levels, forecast reconciliation methods are employed to adjust mortality forecasts across different levels. Finally, we use life table methods to translate the reconciled

mortality forecasts into estimates of remaining life expectancy after retirement, thereby providing the foundational actuarial basis for measuring pension redistribution.

3.1. Pension Redistribution Measure

Under a unified pension system, residents across different areas generally follow the same contribution and benefit rules. However, due to differences in longevity across areas, the duration of pension benefit receipt after retirement may vary across populations, thereby generating implicit pension redistribution across areas. From a life-cycle actuarial perspective, an individual's expected net pension benefit depends on the difference between the present value of expected lifetime pension benefits after retirement and the present value of contributions made during working life. When contribution and benefit rules are uniform nationwide, systematic differences in remaining life expectancy across areas directly affect the present value of post-retirement benefits. Specifically, residents in areas with higher longevity are expected to receive pensions for a longer period, so their lifetime expected present value of benefits is relatively higher, making them more likely to be net beneficiaries within the system. In contrast, residents in areas with lower longevity have a shorter expected benefit duration and a relatively lower lifetime expected present value of benefits, and may therefore receive lower expected net pension benefits. Consequently, in the process of unified pension pooling and benefit payment, part of the contributions made by residents in shorter-lived areas may, in an actuarial sense, support the additional benefits received by residents in longer-lived areas due to their longer benefit duration. Under the same contribution and benefit rules, residents in shorter-lived areas therefore systematically obtain lower expected net benefits and, in effect, bear a relatively higher implicit burden.

To characterize the pension redistribution mechanism induced by longevity differences, we follow the framework of Ayuso et al. (2016) by translating subnational longevity differences into differences in pension liabilities, and further measure the pension redistribution arising from these differences. At the per capita level, pension liabilities mainly depend on the pension benefit level, the growth rate of pension benefits, the discount rate, and post-retirement survival time. For simplicity, the pension benefit per period is normalized to 1, the pension benefit growth rate is denoted by g , the discount rate by r , and the remaining life expectancy at retirement age x by e_x . The per capita pension liability can then be expressed as

$$PL_x = 1 + \frac{1+g}{1+r} + \left(\frac{1+g}{1+r}\right)^2 + \cdots + \left(\frac{1+g}{1+r}\right)^{e_x-1} = \sum_{t=1}^{e_x} \left(\frac{1+g}{1+r}\right)^{t-1}.$$

The above expression implies that, under a unified pension system, subnational differences in pension liabilities are primarily driven by differences in post-retirement survival duration. Areas with longer post-retirement life expectancy correspond to higher per capita pension liabilities, whereas areas with shorter survival duration are associated with relatively lower pension liabilities.

Furthermore, when the benefit growth rate equals the discount rate, i.e., $g = r$, the above expression can be simplified to $PL_x = e_x$. We adopt this benchmark assumption to eliminate the influence of benefit growth and discounting on comparisons across areas, thereby isolating the effect of differences in post-retirement remaining life expectancy. In this case, per capita pension liabilities are linearly related to post-retirement life expectancy. Therefore, under a unified pension system, subnational longevity differences are further translated into differences in pension liabilities, thereby generating implicit pension redistribution across areas.

To further evaluate the relative gains or burdens across different areas within the pension system, we take the per capita pension liability under the national average longevity level as the benchmark and define the pension redistribution effect for group s in area i at retirement age x as

$$\tau_{x,s,i} = \frac{PL_{x,s,i} - PL_{x,s,J}}{PL_{x,s,J}} = \frac{e_{x,s,i}}{e_{x,s,J}} - 1,$$

where $PL_{x,s,i}$ denotes the per capita pension liability for group s in area i at age x , calculated based on the local longevity level, while $PL_{x,s,J}$ denotes the corresponding per capita pension liability under the national average longevity level. $e_{x,s,i}$ denotes the remaining life expectancy at age x for group s in area i , and $e_{x,s,J}$ denotes the corresponding remaining life expectancy under the national average level. The index s represents the population group, including the total, male, and female populations.

When $\tau_{x,s,i} > 0$, the corresponding group in area i has a higher life expectancy than the national average, implying a higher per capita pension liability and therefore receiving an implicit pension subsidy under the unified pension system. In contrast, when $\tau_{x,s,i} < 0$, the corresponding group has a lower pension liability than the national average and therefore bears an implicit pension tax under the unified pension system. The magnitude of $|\tau_{x,s,i}|$ reflects the intensity of pension redistribution arising from subnational longevity differences.

3.2. Mortality Forecasting

Based on the above indicator definitions, evaluating pension redistribution requires projections of the future remaining life expectancy after retirement for each population group across areas. However, remaining life expectancy is obtained through a nonlinear transformation of age-specific mortality rates using life table methods. Therefore, to obtain reliable estimates of future remaining life expectancy, the primary requirement is the accurate projection of age-specific mortality rates. To this end, the classical Lee-Carter model is employed to model mortality dynamics (Lee and Carter 1992). The model has been widely applied in mortality forecasting research because of its interpretability and ability to capture long-term mortality trends across age and time. Let $m_t(x)$ denote the central death rate at age x in year t , where $x = 0, 1, \dots, P$ and $t = 1, 2, \dots, T$. The Lee-Carter model is specified as follows:

$$\log m_t(x) = a(x) + b(x)k_t + \epsilon_t(x),$$

where $a(x)$ is age-specific invariant and is estimated by the average of $\log m_t(x)$ over time, k_t is a time index capturing the overall trend in mortality, $b(x)$ measures the sensitivity of the log mortality rate at age x to k_t , and $\epsilon_t(x)$ is an error term representing the component that cannot be explained.

An identification issue arises in the Lee-Carter model. Following Lee and Carter (1992), the following constraints are imposed to ensure identifiability:

$$\sum_{x=0}^P b(x) = 1, \sum_{t=1}^T k_t = 0.$$

For parameter estimation, singular value decomposition (SVD) is employed to estimate the parameters of the Lee-Carter model. Following Lee and Carter (1992), the estimated mortality index k_t is further modeled as a random walk with drift:

$$k_t = d + k_{t-1} + e_t,$$

where d is the drift parameter, and e_t is a white noise error term. Under this specification, the future value of the mortality index at horizon h can be written as

$$k_{T+h} = k_T + hd + \sum_{j=1}^h e_{T+j}.$$

When forecasting future mortality, the conditional expectation of this random walk process is commonly used as the baseline forecast. Since the error term has mean zero, the forecast of k_{T+h} is given by

$$\tilde{k}_{T+h} = E(k_{T+h} | \mathcal{F}_T) = \hat{k}_T + h\hat{d},$$

where $\mathcal{F}_T$ denotes the information available up to year T . The corresponding forecast of the log central death rate is then given by

$$\log \tilde{m}_{T+h}(x) = \hat{a}(x) + \hat{b}(x)\tilde{k}_{T+h}.$$

The forecasted central death rate $\tilde{m}_{T+h}(x)$ is obtained by exponentiating this value.

3.3. Forecast Coherence Across Spatial Levels

Considering that this study simultaneously involves the national, regional, and prefectural levels, while further distinguishing the total, male, and female populations, mortality forecasting needs to be conducted separately across different spatial levels and population groups in order to obtain the base forecasts for each level. Since aggregation relationships exist across both spatial levels and population groups, the mortality forecasts should remain coherent in terms of both spatial and population structures. For mortality rates at age x in year t , the female and male mortality rates should separately satisfy the following aggregation relationships across different spatial levels:

$$m_{A,t}^F(x) = \sum_{j \in A} w_{A,j,t}^F(x) m_{j,t}^F(x).$$

$$m_{A,t}^M(x) = \sum_{j \in A} w_{A,j,t}^M(x) m_{j,t}^M(x).$$

Furthermore, the total mortality rate for an upper-level spatial unit can be obtained as the weighted aggregation of female and male mortality rates across its lower-level areas, that is,

$$m_{A,t}^T(x) = \sum_{j \in A} w_{A,j,t}^F(x) m_{j,t}^F(x) + \sum_{j \in A} w_{A,j,t}^M(x) m_{j,t}^M(x),$$

where A denotes an upper-level spatial unit, and $j \in A$ represents the lower-level areas contained within A . T , F , and M denote the total, female, and male populations, respectively. $m_{j,t}^F(x)$, $m_{j,t}^M(x)$, and $m_{j,t}^T(x)$ denote the mortality rates for the female, male, and total populations in area j at age x and year t , respectively. $w_{A,j,t}^F(x)$ and $w_{A,j,t}^M(x)$ denote the corresponding exposure-based population weights for the female and male populations in area j within the upper-level spatial unit A , respectively.

However, since forecasts at different levels are generated independently, the resulting forecasts may not automatically satisfy the above aggregation relationships, thereby leading to incoherence across spatial levels and population groups. To ensure forecast coherence across different spatial levels, we further employ forecast reconciliation methods to adjust the base forecasts (Athanasopoulos et al. 2024; Hyndman et al. 2011).

Let $\hat{\mathbf{y}}_{T+h}(x)$ denote the base forecasts across all levels, and let $\tilde{\mathbf{y}}_{T+h}(x)$ denote the corresponding coherent forecasts satisfying the aggregation constraints. Following Wickramasuriya et al. (2019), a general linear reconciliation framework can be expressed as

$$\tilde{\mathbf{y}}_{T+h}(x) = \mathbf{S}_{T+h}(x) \mathbf{G}_h \hat{\mathbf{y}}_{T+h}(x),$$

where $\mathbf{S}_{T+h}(x)$ denotes the summing matrix that describes the aggregation structure across different spatial levels and population groups. $\mathbf{G}_h$ denotes the reconciliation matrix, which maps the base forecasts into adjusted bottom-level forecasts. These adjusted bottom-level forecasts are then aggregated by $\mathbf{S}_{T+h}(x)$ to obtain coherent forecasts at all levels. Since the construction of $\mathbf{G}_h$ is not unique, different specifications of the reconciliation matrix correspond to different forecast reconciliation methods.

In this study, we consider several widely used forecast reconciliation methods, including bottom-up (BU), ordinary least squares (OLS), weighted least squares (WLS), minimum trace (MinT), and empirical minimum trace (EMinT) (Hyndman et al. 2016, 2011; Wickramasuriya 2021; Wickramasuriya et al. 2019). Among them, the BU method only utilizes information from the bottom-level series and obtains higher-level forecasts through upward aggregation. The OLS method linearly adjusts the base forecasts under the unbiasedness constraint while assuming equal forecast error variances across different series. The WLS method further accounts for heterogeneity in forecast error variances across series by constructing reconciliation weights through variance scaling. The MinT method further exploits the forecast error covariance structure and obtains reconciled forecasts by minimizing the trace of the covariance matrix of reconciled forecast errors. Since the number of series is relatively large compared with the number of available time observations, the sample covariance matrix of forecast errors may be unstable. We therefore use the shrinkage covariance version of MinT, referred to as MinTs, to improve the stability of covariance estimation. The EMinT method further relaxes the unbiasedness constraint within the MinT framework and constructs an empirical reconciliation solution by minimizing the trace of the covariance matrix of reconciled forecast errors. The specifications of the reconciliation matrices corresponding to different forecast reconciliation methods are provided in Appendix A.

To obtain more accurate mortality forecasts while ensuring forecast coherence, an expanding-window approach is adopted for out-of-sample forecast evaluation. Different forecast reconciliation methods are compared based on the root mean squared forecast error (RMSFE). Based on the RMSFE comparison, the forecast reconciliation method with the lowest RMSFE is selected and applied to the subsequent pension redistribution analysis.

3.4. Life Expectancy Estimation

After obtaining coherent mortality forecasts for each area using the methods described above, this section further applies standard life table methods to translate these forecasts into remaining life expectancy after retirement, thereby providing the necessary actuarial input for the measurement of implicit pension redistribution.

Specifically, the life-table calculation uses the forecasted central death rate $m_t(x)$ at age x in year t as the mortality input. The construction of the life table begins by converting $m_t(x)$ into the one-year probability of death $q_t(x)$. Under the assumption that deaths are uniformly distributed within each age interval, the relationship between the two can be expressed as

$$q_t(x) = \begin{cases} \frac{m_t(x)}{1 + 0.5m_t(x)}, & x < \omega, \\ 1, & x = \omega. \end{cases}$$

where ω denotes the highest open-ended age group in the life table. Since all individuals surviving to this age group are assumed to eventually die within the open-ended interval, the probability of death for this group is strictly set equal to 1.

Setting the life table radix to $l_t(x_0) = 100,000$, for ordinary ages with $x < \omega$, the number of survivors reaching age $x + 1$ in year t can be recursively calculated as

$$l_t(x + 1) = l_t(x)[1 - q_t(x)], \quad x < \omega.$$

Furthermore, the number of person-years lived between ages x and $x + 1$ in year t , denoted by $L_t(x)$, is given by

$$L_t(x) = \begin{cases} \frac{l_t(x) + l_t(x + 1)}{2}, & x < \omega, \\ \frac{l_t(\omega)}{m_t(\omega)}, & x = \omega. \end{cases}$$

Based on the person-years lived at each age, the total number of remaining person-years for individuals aged x in year t , denoted by $T_t(x)$, is obtained by summing $L_t(j)$ from age x to the open-ended age group

$$T_t(x) = \sum_{j=x}^{\omega} L_t(j).$$

Finally, the remaining life expectancy after retirement at age x in year t , denoted by $e_t(x)$, can be calculated as

$$e_t(x) = \frac{T_t(x)}{l_t(x)}.$$

Through the above life table conversion, the reconciled age-specific mortality forecasts are further translated into remaining life expectancy after retirement. In the empirical analysis, we set the retirement age to 65, and the resulting remaining life expectancy at age 65 serves as the direct actuarial input for the subsequent quantification of subnational implicit pension redistribution.

4. Empirical Results

In this section, we present the empirical results based on Japanese mortality data. First, different forecast reconciliation methods are compared based on their out-of-sample forecasting performance in order to identify the method most suitable for the characteristics of Japanese subnational mortality data. Second, based on the selected method, future mortality rates are projected and remaining life expectancy is estimated, and the spatial distribution of remaining life expectancy at age 65 across prefectures is presented. On this basis, the pension redistribution induced by longevity heterogeneity across prefectures is further quantified. Finally, pension redistribution results across different spatial scales are compared in order to examine the scale effects in the measurement of pension redistribution.

4.1. Model Comparison and Selection

To identify the forecast reconciliation method most suitable for the characteristics of Japanese mortality data, this subsection compares the out-of-sample forecasting performance of different reconciliation methods. Following the methodology introduced in the previous section, an expanding-window forecasting scheme is adopted for evaluation. Specifically, mortality data from 1975 to 2007 are initially used for model estimation, and forecasts are generated for 2008–2022. The estimation window is then recursively expanded year by year, with data from 1975 to 2008 used to forecast 2009–2022, and this procedure

continues until data from 1975 to 2021 are used to forecast 2022. Forecast accuracy is evaluated using the root mean squared forecast error (RMSFE), defined as follows:

$$\text{RMSFE}_k(h) = \frac{1}{m_k} \sum_{i=1}^{m_k} \sqrt{\frac{1}{(P+1) \times (H+1-h)} \sum_{\zeta=h}^H \sum_{x=0}^P \left[m_{T+\zeta}^i(x) - \tilde{m}_{T+\zeta}^i(x) \right]^2},$$

where m_k denotes the number of series at the k th level for $k = 1, \dots, K$; T denotes the endpoint of the initial estimation window and is set to 2007; h denotes the forecast horizon; P is the maximum age; and H denotes the longest forecast horizon and is set to 15. $m_{T+\zeta}^i(x)$ and $\tilde{m}_{T+\zeta}^i(x)$ denote the observed mortality rate and its h -step-ahead forecast, respectively, for age x in year $T + \zeta$ for the i th series at the k th level. This measure enables a comprehensive evaluation of forecast errors at different forecast horizons and spatial levels, thereby allowing a systematic comparison of the out-of-sample forecasting performance of different forecast reconciliation methods.

Table 2 summarizes the mean and median RMSFE for each method across different levels, calculated from 1- to 15-step-ahead forecasts. The results show that forecast reconciliation methods generally reduce the errors of the independent forecasts while maintaining forecast coherence. In particular, the EMinT method has lower mean and median RMSFE values at all levels, indicating better overall forecasting performance. Meanwhile, the forecast errors of all methods are relatively small at the national level but increase at the regional and prefectural levels, suggesting that mortality forecasting uncertainty increases as the spatial hierarchy becomes more disaggregated.

Table 2. Mean and median out-of-sample RMSFEs ($\times 100$) for base and reconciliation methods across levels.

Level	Base	BU	OLS	WLS	MinTs	EMinT
<i>Mean</i>						
Total	1.33	1.07	1.15	1.17	1.07	0.81
Sex	1.30	1.27	1.35	1.37	1.26	0.99
Region	1.45	1.17	1.26	1.28	1.16	0.85
Region + Sex	1.47	1.42	1.49	1.52	1.40	1.11
Prefecture	1.51	1.28	1.37	1.38	1.28	1.01
Prefecture + Sex	1.70	1.70	1.77	1.79	1.70	1.48
<i>Median</i>						
Total	1.25	0.99	1.07	1.09	1.02	0.72
Sex	1.21	1.18	1.25	1.28	1.18	0.89
Region	1.41	1.13	1.21	1.23	1.13	0.77
Region + Sex	1.42	1.35	1.42	1.45	1.33	1.00
Prefecture	1.45	1.24	1.32	1.34	1.25	0.93
Prefecture + Sex	1.64	1.64	1.71	1.73	1.64	1.37

Building on this, Figure 2 further illustrates the variation in RMSFE across 1- to 15-step-ahead forecasts. Unlike the aggregate metrics in Table 2, Figure 2 captures the dynamic characteristics of how forecast errors evolve with the forecast horizon. It can be observed that forecast errors for all methods generally increase with the forecast horizon, reflecting the greater uncertainty inherent in medium- to long-term mortality forecasting. Throughout the evaluation period, the RMSFE curve for the EMinT method consistently lies at the lowest position across most spatial levels and forecast steps. This indicates that the performance of EMinT is not driven by an isolated result, but that it generally performs better than other methods across different spatial levels and forecast horizons.

Based on the above comparison of forecasting performance, we use the EMinT method for subsequent mortality forecasting, and then use the forecast results to estimate remaining life expectancy and measure pension redistribution. Since remaining life expectancy is derived from the complete age-specific mortality schedule through a unified life-table procedure, more accurate and coherent mortality forecasts provide a more reliable demographic basis for the subsequent pension redistribution analysis.

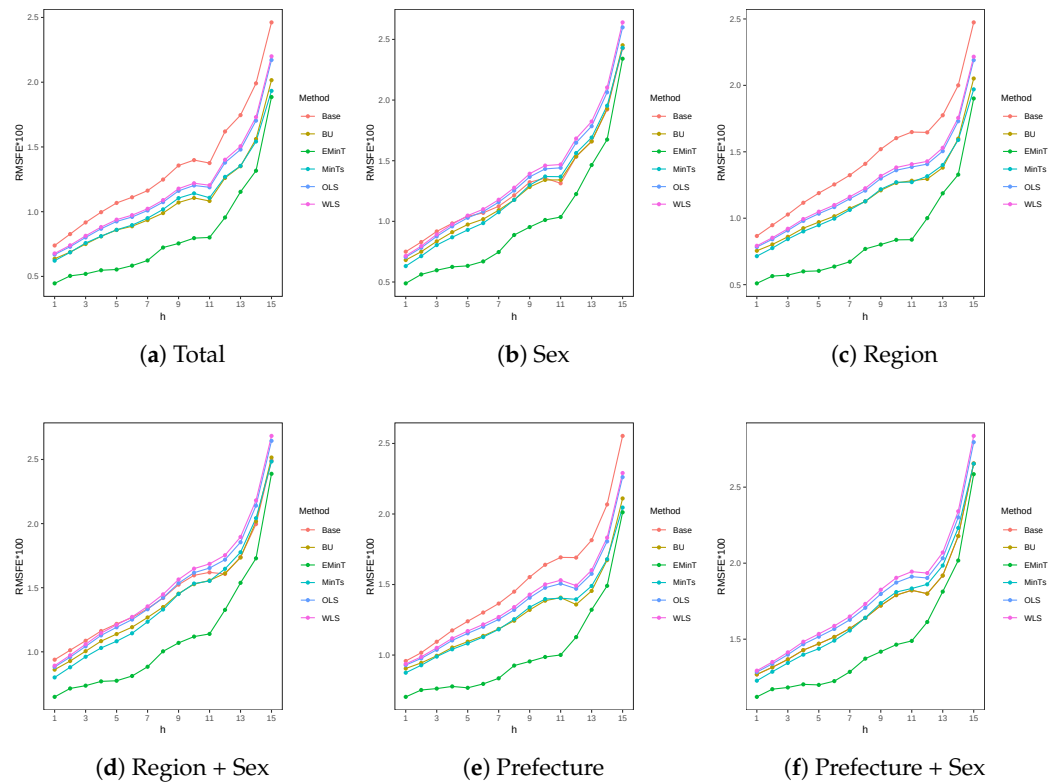


Figure 2. Out-of-sample RMSFEs ($\times 100$) by forecast horizon and method across levels.

4.2. Future Longevity Heterogeneity

Building upon the EMinT-based reconciled mortality forecasts, we apply the life table method from Section 3.4 to derive future remaining life expectancy from the forecasted mortality rates, thereby examining the temporal evolution and spatial distribution of longevity heterogeneity. Figures 3 and 4 illustrate the characteristics of future longevity heterogeneity in Japan from the perspectives of temporal evolution and spatial distribution, respectively. Specifically, Figure 3 presents the trends in remaining life expectancy at age 65 for the total population over the period 2022–2032 for several representative prefectures, including relatively high-longevity prefectures such as Shiga and Oita and relatively low-longevity prefectures such as Okinawa and Aomori, along with the national average for comparison. Figure 4 further illustrates the spatial distribution of remaining life expectancy at age 65 for the total, male, and female populations across prefectures in Japan in 2032.

From the perspective of temporal evolution, the remaining life expectancy at age 65 across Japan is projected to exhibit a continuous upward trend in the future, indicating that mortality conditions are expected to continue improving. However, longevity differences do not show clear signs of convergence as overall longevity increases. The gap between high-longevity and low-longevity prefectures persists throughout the projection period and even shows a slight widening trend. Prefectures with relatively high longevity levels consistently maintain higher expected remaining life expectancy over the entire projection period, whereas prefectures with relatively low longevity levels continue to exhibit

lower longevity levels. The national average generally lies between the high-longevity and low-longevity prefectures. These findings suggest that longevity heterogeneity across prefectures in Japan is likely to persist in the foreseeable future.

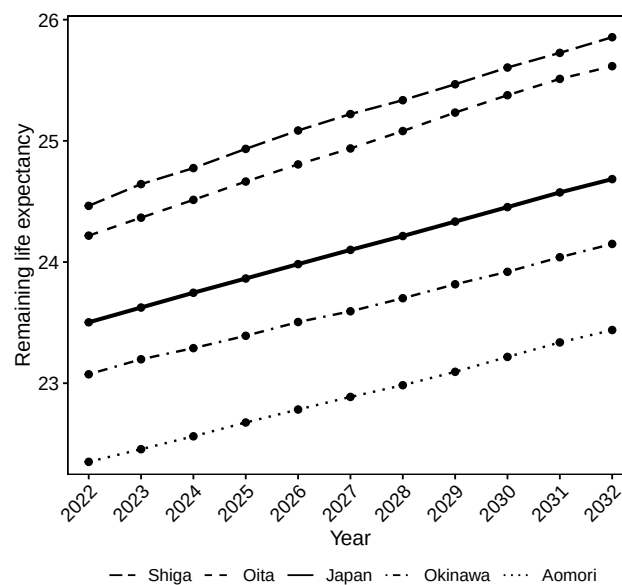


Figure 3. Projected remaining life expectancy at age 65 for the total population in selected prefectures, 2022–2032.

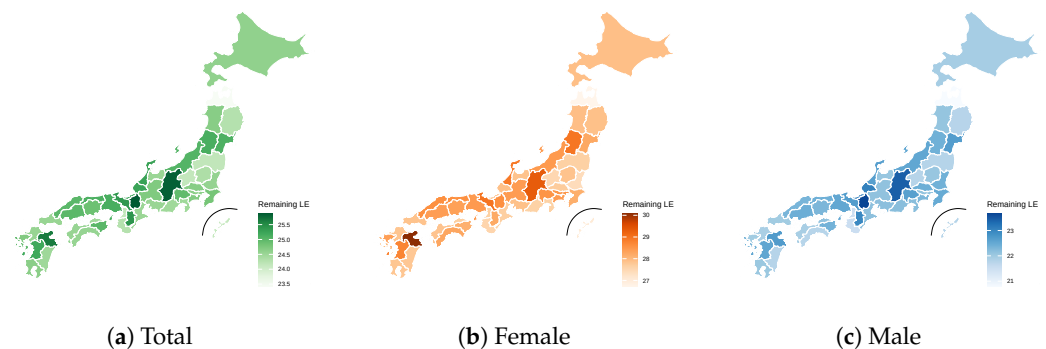


Figure 4. Projected remaining life expectancy at age 65 across prefectures in Japan, 2032.

Figure 4 further illustrates the spatial distribution of longevity heterogeneity in 2032. Overall, substantial longevity differences can still be observed across prefectures in Japan, exhibiting a relatively stable spatial distribution pattern. For the total population, prefectures such as Shiga, Nagano, and Oita correspond to relatively high remaining life expectancy at age 65, whereas prefectures such as Aomori and Okinawa exhibit relatively lower longevity levels. From the spatial perspective, high-longevity and low-longevity prefectures display clear geographic clustering patterns. In particular, high-longevity prefectures are mainly concentrated in the Chubu and Kinki regions, whereas low-longevity prefectures are more concentrated in the Tohoku region. These findings suggest that future longevity heterogeneity is not randomly distributed, but instead exhibits a clear degree of spatial association.

From the gender perspective, females generally exhibit higher remaining life expectancy at age 65 than males across most prefectures, while the longevity level of the total population lies broadly between those of the male and female populations. At the same time, substantial longevity differences can still be observed within both the male and female populations, suggesting that longevity heterogeneity is not merely driven by differences in gender composition, but also exists within each gender group.

Due to substantial differences in remaining life expectancy across subnational areas, the duration of pension benefit receipt after retirement also varies significantly among residents in different areas. Residents in areas with higher longevity are generally expected to receive pension benefits for a longer period after retirement, whereas those in areas with lower longevity tend to have relatively shorter benefit durations. Under a unified pension system, such differences further lead to systematic disparities in pension liabilities across areas, thereby generating corresponding pension redistribution.

4.3. Spatial Patterns of Implicit Pension Redistribution

Based on the projected remaining life expectancy obtained in the previous sections, we further quantify the implicit pension redistribution arising from longevity heterogeneity. Figure 5 illustrates the spatial patterns of implicit tax rates and implicit subsidy rates at the prefectural level for the total, male, and female populations in 2032.

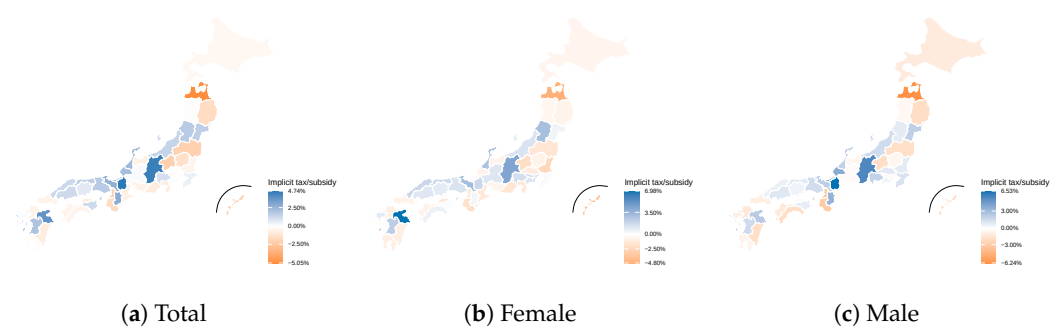


Figure 5. Prefecture-level implicit pension redistribution at age 65 in Japan, 2032.

Overall, substantial pension redistribution differences can be observed across prefectures in Japan. Prefectures with relatively high longevity levels generally correspond to implicit subsidies, whereas prefectures with relatively low longevity levels are more likely to exhibit implicit tax burdens. This indicates that longevity differences across prefectures have already been translated into resource redistribution within the pension system. In terms of redistribution intensity, the gap between implicit tax rates and implicit subsidy rates across prefectures is also considerable. For the total population, the highest implicit subsidy rate reaches 4.74% in 2032, whereas the highest implicit tax burden reaches -5.05% , implying a gap of nearly 10 percentage points. This suggests that longevity heterogeneity across prefectures has generated substantial implicit resource transfers under the unified pension system.

Further from the spatial perspective, prefectures such as Shiga and Nagano correspond to relatively high implicit subsidy rates, whereas prefectures such as Aomori and Okinawa exhibit relatively high implicit tax burdens. From the geographic distribution pattern, implicit subsidy prefectures are mainly concentrated in the Chubu and Kinki regions, whereas the Tohoku region is more characterized by implicit tax burdens. Overall, the pension redistribution effect exhibits a clear spatial clustering pattern, in the sense that neighboring prefectures tend to display similar redistribution characteristics.

From the gender perspective, the spatial patterns for the male, female, and total populations are broadly similar. However, compared with the female population, the differences between implicit tax rates and implicit subsidy rates are generally larger for the male population. The gap between the highest implicit subsidy rate and the highest implicit tax burden reaches 12.77% for males, whereas the corresponding gap for females is 11.78%. This corresponds to the more pronounced longevity differences observed across areas within the male population. In contrast, although females generally exhibit higher

longevity levels, the longevity differences across areas are relatively smaller, and therefore the corresponding implicit pension redistribution differences are also relatively weaker.

In summary, the long-standing longevity heterogeneity across prefectures in Japan has been further translated into implicit redistribution within the pension system. Prefectures with relatively high longevity levels generally correspond to implicit subsidies, whereas prefectures with relatively low longevity levels are more likely to exhibit implicit tax burdens, with clear spatial clustering patterns. These findings suggest that, under a unified pension system, longevity differences across prefectures systematically influence the spatial redistribution of pension resources.

4.4. Scale Effects in Measuring Pension Redistribution

Previous results indicate substantial pension redistribution differences at the prefectural level in Japan. However, longevity heterogeneity may be represented differently across spatial scales, which may further lead to differences in the measurement of pension redistribution. Therefore, spatial scale itself may constitute an important factor in the measurement of pension redistribution. To further examine the scale effects in the measurement of pension redistribution, we compare pension redistribution results at the prefectural and regional levels.

Figure 6 illustrates the distribution of pension redistribution across different spatial scales in 2032. From a cross-sectional perspective, pension redistribution at the prefectural level exhibits substantially greater dispersion, with the ranges of implicit pension subsidies and implicit tax burdens being considerably wider than those observed at the regional level. In contrast, redistribution results at the regional level are more concentrated, and both extreme subsidies and extreme tax burdens become noticeably weaker. This suggests that spatial scale itself can influence the measured pension redistribution. One possible explanation is that regional-level aggregation smooths redistribution differences across prefectures with heterogeneous longevity levels. Since Japan's 8 major regions still contain prefectures with substantially different longevity and redistribution characteristics, aggregation at the regional level may partially offset redistribution differences across lower-level areas. Consequently, pension redistribution measured at the regional level is substantially weaker than that measured at the prefectural level. These findings imply that analyses conducted solely at higher spatial scales may underestimate the true extent of pension redistribution across lower-level areas.

Building upon the cross-sectional comparison, Figure 7 further illustrates the temporal evolution of pension redistribution intensity across different spatial scales. Here, overall redistribution intensity is measured by the gap between the highest implicit subsidy rate and the highest implicit tax burden, and the corresponding trends at the prefectural and regional levels during 2022–2032 are examined. The results show that redistribution intensity at the prefectural level remains consistently higher than that at the regional level throughout the entire projection period, indicating that pension redistribution measured at higher spatial scales is smoothed to some extent. Meanwhile, redistribution intensity at the prefectural level exhibits a gradual upward trend over time, suggesting that pension redistribution induced by longevity differences across areas may become further intensified in the future. In contrast, redistribution intensity at the regional level displays a certain degree of convergence. Consequently, the gap in redistribution intensity between the prefectural and regional levels may continue to widen over time. This further suggests that spatial aggregation can not only underestimate measured pension redistribution at a given point in time, but also obscure the increasing divergence in redistribution intensity over time. As a result, redistribution patterns measured at higher spatial scales may deviate further from the underlying redistribution differences that exist across lower-level areas.

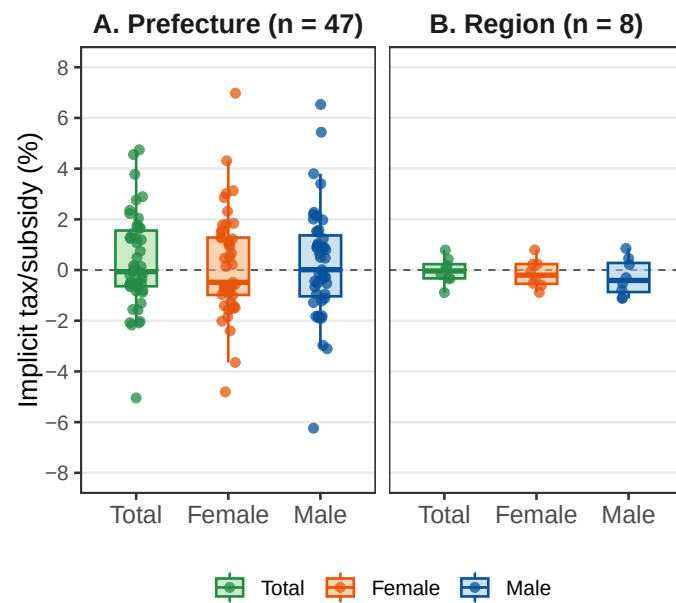


Figure 6. Distribution of implicit pension redistribution across spatial scales in Japan, 2032.

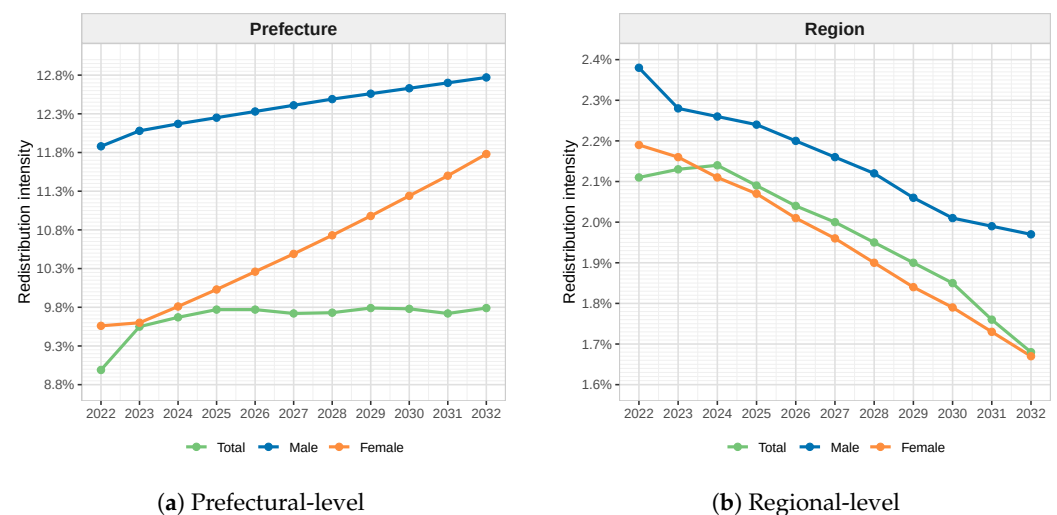


Figure 7. Evolution of pension redistribution intensity across spatial scales in Japan, 2022–2032.

From the gender perspective, the total, male, and female populations exhibit similar patterns across different spatial scales. In particular, the redistribution intensity associated with the male population remains consistently higher than that of the female population across spatial scales, which is consistent with the more pronounced longevity heterogeneity observed within the male population in the previous analysis.

Overall, pension redistribution results differ substantially across spatial scales. Compared with the regional level, redistribution measured at the prefectural level exhibits greater dispersion and stronger redistribution intensity, suggesting that finer spatial units are able to capture more pronounced longevity heterogeneity and the associated pension redistribution patterns across areas. These findings indicate that analyses based solely on more aggregated spatial levels may underestimate the true extent of pension redistribution across lower-level areas.

5. Discussion and Conclusions

Using subnational mortality data from Japan, this study examines longevity disparities across subnational areas and the implicit pension redistribution induced by these dispar-

ities under a unified pension system. Since the analysis involves multiple spatial levels, including the national, regional, and prefectural levels, mortality forecasts need to satisfy spatial aggregation coherence. To this end, we compare several forecast reconciliation methods and select the EMinT method for subsequent mortality forecasting based on its out-of-sample forecasting performance. On this basis, we further estimate post-retirement remaining life expectancy across areas. The results show that longevity heterogeneity across Japanese areas will persist throughout the forecast period, with no clear convergence between longer-lived and shorter-lived areas, and that these disparities exhibit a certain degree of spatial clustering.

Based on the projected post-retirement remaining life expectancy, we further measure implicit pension redistribution under the unified pension system. The results indicate that areas with higher longevity generally receive implicit pension subsidies, whereas areas with lower longevity bear implicit pension taxes, suggesting that subnational longevity disparities are translated into systematic redistribution differences within the unified pension system. The redistribution results also display clear spatial clustering, indicating that implicit subsidies and taxes are not randomly distributed across areas. From the perspective of spatial scale, prefectural-level redistribution outcomes are more dispersed and exhibit stronger redistribution intensity than those at the regional level, and this intensity increases over the forecast period. These findings suggest that analyses based solely on national average life expectancy or more aggregated spatial levels may underestimate the true extent of pension redistribution across areas.

These findings demonstrate that subnational longevity heterogeneity and its spatial distribution are crucial for evaluating the redistributive effects of a unified pension system. Our estimates provide a quantitative basis for the spatial evaluation of pension policy. First, they can serve as supplementary information for assessing distributional fairness within the unified pension system. More specifically, these estimates can be used in periodic actuarial reviews of the National Pension system to assess whether subnational longevity differences lead to persistent spatial inequalities in expected lifetime pension benefits under uniform contribution and benefit rules. Second, the estimates can inform the design of complementary fiscal support measures. For example, policymakers could consider tax-financed central subsidies or contribution subsidies targeted at areas that systematically occupy disadvantaged redistributive positions. These measures would help mitigate the implicit burden arising from shorter post-retirement survival while preserving the nationwide uniformity of contribution and benefit rules. In this sense, the proposed measurement framework provides practical evidence for pension policy evaluation and for the design of complementary fiscal support mechanisms under a unified pension system.

Several limitations should be acknowledged. First, this study focuses on pension redistribution induced by subnational mortality differences and does not explicitly model the socioeconomic determinants of longevity heterogeneity. Although the supplementary analysis in Appendix B does not show a significant positive association between prefectural income and remaining life expectancy at age 65, future research could further incorporate other factors that could better explain in a more comprehensive way the differences observed by prefecture, such as income, education, sex or climate conditions. Second, migration is not explicitly modeled. Although the use of annual resident population exposures and death counts partly reflects changes in population composition, future migration flows may still affect prefectural mortality profiles and pension redistribution patterns. Third, the pension liability measure is based on the benchmark assumption that the benefit growth rate equals the discount rate. Future studies may conduct sensitivity analyses under alternative assumptions regarding discount rates and pension indexation. Finally, potential future pension reforms are not incorporated into the analysis. Future

research may further integrate demographic, economic, and institutional factors to provide a more comprehensive evaluation of the long-term redistribution of pension systems under population aging.

Author Contributions: Conceptualization, N.Z., C.D. and L.H.; methodology, C.D.; validation, L.H.; formal analysis, C.D. and L.H.; writing—original draft preparation, C.D.; writing—review and editing, N.Z. and L.H.; supervision, N.Z. and L.H.; project administration, N.Z.; funding acquisition, L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under grant number 72301097 and the Key Project of the National Social Science Fund of China under grant number 24AJL005.

Data Availability Statement: The data presented in this study are openly available in the Japanese Mortality Database at <https://www.ipss.go.jp/p-toukei/JMD/index-en.asp> (accessed on 10 October 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Reconciliation Matrices for Forecast Reconciliation Methods

This appendix presents the reconciliation matrix forms corresponding to different forecast reconciliation methods.

Let $\hat{y}_{T+h}(x)$ denote the independent h -step-ahead forecasts at all levels, and let $\tilde{y}_{T+h}(x)$ be the corresponding reconciled h -step-ahead forecasts which satisfy all aggregation constraints. The general form of forecast reconciliation can be written as

$$\tilde{y}_{T+h}(x) = S_{T+h}(x)G_h\hat{y}_{T+h}(x),$$

where $S_{T+h}(x)$ denotes the summing matrix and G_h denotes the reconciliation matrix. It should be noted that the $S_{T+h}(x)$ matrix contains exposure-to-risk ratios, whose future values are projected following the approach of Shang and Hyndman (2017) to construct future aggregation matrices. Different reconciliation methods correspond to different choices for G_h .

The bottom-up (BU) reconciliation method only utilizes forecasts from the bottom-level series and obtains higher-level forecasts through aggregation (Dangerfield and Morris 1992; Zellner and Tobias 2000). The corresponding reconciliation matrix is given by

$$G_h = (\mathbf{0}_{m \times (n-m)}, \mathbf{I}_m),$$

where m denotes the number of bottom-level series, n denotes the total number of series in the hierarchy, and $\mathbf{0}_{m \times (n-m)}$ is a zero matrix of dimension $m \times (n - m)$.

The ordinary least squares (OLS) reconciliation method jointly uses base forecasts from all hierarchical levels and reconciles them through ordinary least squares to obtain coherent forecasts, while assuming forecast errors across series have equal importance (Hyndman et al. 2011). The reconciliation matrix is given by

$$\hat{G}_h = (S_{T+h}^\top(x)S_{T+h}(x))^{-1}S_{T+h}^\top(x).$$

The weighted least squares (WLS) reconciliation method extends the OLS method by further accounting for heterogeneity in forecast error variances across series (Hyndman et al. 2016). Specifically, different weights are assigned according to the magnitude of forecast error variances across series, and the base forecasts are adjusted accordingly to obtain coherent forecasts. The reconciliation matrix can be expressed as

$$\hat{G}_h = \left(S_{T+h}^\top(x)\Lambda_h^{-1}S_{T+h}(x) \right)^{-1}S_{T+h}^\top(x)\Lambda_h^{-1},$$

where W_h denotes the covariance matrix of the h -step-ahead base forecast errors, and $\Lambda_h = \text{diag}(W_h)$ denotes the diagonal matrix containing their variances.

The minimum trace (MinT) reconciliation method further incorporates the covariance structure among forecast errors across different series and obtains coherent forecasts by minimizing the trace of the reconciled forecast error covariance matrix (Wickramasuriya et al. 2019). The corresponding reconciliation matrix is given by

$$\hat{G}_h = (S_{T+h}^\top(x) W_h^{-1} S_{T+h}(x))^{-1} S_{T+h}^\top(x) W_h^{-1},$$

where W_h denotes the covariance matrix of the h -step-ahead base forecast errors. In practice, W_h can be estimated in different ways. Under general settings, the in-sample forecast error covariance matrix is typically used directly. However, in high-dimensional settings, the in-sample covariance matrix may become unstable or non-invertible. In such cases, a shrinkage estimator of the covariance matrix can be adopted, corresponding to the MinTs method.

The preceding forecast reconciliation methods ensure that the reconciled forecasts are unbiased when the base forecasts are unbiased. However, base forecasts may be biased in practical applications. To relax this condition, Wickramasuriya (2021) proposed the empirical minimum trace (EMinT) reconciliation method, which estimates the reconciliation matrix by minimizing the trace of the reconciled forecast error covariance matrix without imposing the unbiasedness constraint. Assuming that the series are jointly weakly stationary, the reconciliation matrix can be written as

$$\hat{G}_h = B_h^\top(x) \hat{Y}_h(x) (\hat{Y}_h^\top(x) \hat{Y}_h(x))^{-1},$$

where $B_h(x) = [b_{h,1}, \dots, b_{h,T}]^\top$ denotes the matrix of bottom-level observations, and $\hat{Y}_h(x) = [\hat{y}_{h|0}, \dots, \hat{y}_{h|T-h}]^\top$ denotes the matrix of corresponding base forecasts.

Appendix B. Correlation Between Prefectural Income and Remaining Life Expectancy

As a supplement to the analysis of subnational longevity disparities, this appendix examines the relationship between remaining life expectancy at age 65 and income levels using data for all 47 Japanese prefectures. Income levels are measured by per capita prefectural income for fiscal year 2022, obtained from the Prefectural Accounts published by the Economic and Social Research Institute, Cabinet Office, Government of Japan (Cabinet Office 2026). Period remaining life expectancy at age 65 is calculated from mortality rates for the total population in 2022.

The Pearson correlation coefficient and Spearman rank correlation coefficient are used to measure the linear and monotonic relationships between the two variables, respectively. Figure A1 presents the scatter plot of per capita prefectural income and period remaining life expectancy at age 65.

Table A1 reports the two correlation coefficients and their corresponding p -values.

Table A1. Correlation between per capita prefectural income and period remaining life expectancy at age 65.

Correlation Measure	Coefficient	p -Value
Pearson correlation	−0.027	0.859
Spearman rank correlation	−0.106	0.479

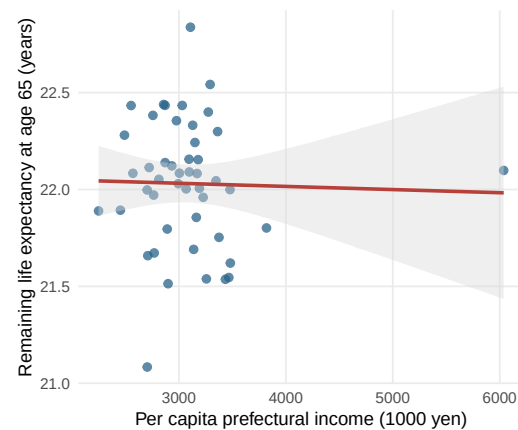


Figure A1. Relationship between per capita prefectural income and period remaining life expectancy at age 65.

Neither correlation coefficient is statistically significant, and the scatter plot shows no clear positive association between the two variables. Overall, remaining life expectancy at age 65 across prefectures does not exhibit a clear income gradient.

References

- Atalay, Kadir, Rebecca Edwards, and Fiona Georgiakakis. 2023. Mortality inequality, spatial differences and health care access. *Health Economics* 32: 2632–54. [\[CrossRef\]](#) [\[PubMed\]](#)
- Athanasopoulos, George, Rob J. Hyndman, Nikolaos Kourentzes, and Anastasios Panagiotelis. 2024. Forecast reconciliation: A review. *International Journal of Forecasting* 40: 430–56. [\[CrossRef\]](#)
- Ayuso, Mercedes, Jorge Miguel Bravo, and Robert Holzmann. 2016. *On the Heterogeneity in Longevity Among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes*. Technical report, IZA Discussion Paper. Bonn: Institute for the Study of Labor (IZA).
- Ayuso, Mercedes, Jorge Miguel Bravo, and Robert Holzmann. 2017. Addressing longevity heterogeneity in pension scheme design. *Journal of Finance and Economics* 6: 1–21. [\[CrossRef\]](#)
- Bravo, Jorge M., Mercedes Ayuso, Robert Holzmann, and Edward Palmer. 2021. Addressing the life expectancy gap in pension policy. *Insurance: Mathematics and Economics* 99: 200–21. [\[CrossRef\]](#)
- Cabinet Office. 2026. *Prefectural Accounts: Per Capita Prefectural Income*. Tokyo: Economic and Social Research Institute, Government of Japan.
- Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. The association between income and life expectancy in the United States, 2001–2014. *JAMA* 315: 1750–66. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dangerfield, Byron J., and John S. Morris. 1992. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. *International Journal of Forecasting* 8: 233–41. [\[CrossRef\]](#)
- Dwyer-Lindgren, Laura, Amelia Bertozzi-Villa, Rebecca W. Stubbs, Chloe Morozoff, Johan P. Mackenbach, Frank J. Van Lenthe, Ali H. Mokdad, and Christopher J. L. Murray. 2017. Inequalities in life expectancy among US counties, 1980 to 2014: Temporal trends and key drivers. *JAMA Internal Medicine* 177: 1003–11. [\[CrossRef\]](#) [\[PubMed\]](#)
- Government Pension Investment Fund. 2025. *Annual Report Fiscal Year 2024*. Technical report. Tokyo: Government Pension Investment Fund.
- Hyndman, Rob J., Alan J. Lee, and Earo Wang. 2016. Fast computation of reconciled forecasts for hierarchical and grouped time series. *Computational Statistics & Data Analysis* 97: 16–32. [\[CrossRef\]](#)
- Hyndman, Rob J., Roman A. Ahmed, George Athanasopoulos, and Han Lin Shang. 2011. Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis* 55: 2579–89. [\[CrossRef\]](#)
- Japanese Mortality Database. 2024. *The Japanese Mortality Database*. Tokyo: National Institute of Population and Social Security Research.
- Lee, Ronald D., and Lawrence R. Carter. 1992. Modeling and forecasting US mortality. *Journal of the American Statistical Association* 87: 659–71. [\[CrossRef\]](#)
- Li, Han, and Rob J. Hyndman. 2021. Assessing mortality inequality in the US: What can be said about the future? *Insurance: Mathematics and Economics* 99: 152–62. [\[CrossRef\]](#)

- Ministry of Health, Labour and Welfare. 2025. *Annual Health, Labour and Welfare Report 2025: Pension Security*. Tokyo: Ministry of Health, Labour and Welfare, Government of Japan.
- Nomura, Shuhei, Michio Murakami, Santosh Kumar Rauniyar, Naoki Kondo, Takahiro Tabuchi, Haruka Sakamoto, Yasuharu Tokuda, Nishali Patel, Jose Navarro De Pablo, Joseph L. Dieleman, and et al. 2025. Three decades of population health changes in Japan, 1990–2021: A subnational analysis for the global burden of disease study 2021. *The Lancet Public Health* 10: e321–e332. [\[CrossRef\]](#)
- Sanchez-Romero, Miguel, Ronald D. Lee, and Alexia Prskawetz. 2020. Redistributive effects of different pension systems when longevity varies by socioeconomic status. *The Journal of the Economics of Ageing* 17: 100259. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shang, Han Lin, and Rob J. Hyndman. 2017. Grouped functional time series forecasting: An application to age-specific mortality rates. *Journal of Computational and Graphical Statistics* 26: 330–43. [\[CrossRef\]](#)
- Shang, Han Lin, and Yang Yang. 2021. Forecasting Australian subnational age-specific mortality rates. *Journal of Population Research* 38: 1–24. [\[CrossRef\]](#)
- Shi, Jiaxin, and Martin Kolk. 2022. How does mortality contribute to lifetime pension inequality? Evidence from five decades of Swedish taxation data. *Demography* 59: 1843–71. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wickramasuriya, Shanika L. 2021. Properties of point forecast reconciliation approaches. *arXiv* arXiv:2103.11129.
- Wickramasuriya, Shanika L., George Athanasopoulos, and Rob J. Hyndman. 2019. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association* 114: 804–19.
- Zellner, Arnold, and Justin Tobias. 2000. A note on aggregation, disaggregation and forecasting performance. *Journal of Forecasting* 19: 457–65. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.