

Department of Econometrics and Business Statistics

<http://monash.edu/business/ebs/research/publications>

Forecasting the old-age dependency ratio to determine the best pension age

Rob J Hyndman, Yijun Zeng, Han Lin Shang

September 2020

Working Paper no/yr

Forecasting the old-age dependency ratio to determine the best pension age

Rob J Hyndman

Department of Econometrics & Business Statistics
Monash University, Clayton VIC 3800, Australia
Email: Rob.Hyndman@monash.edu
Corresponding author

Yijun Zeng

Monash University

Han Lin Shang

Macquarie University

10 September 2020

JEL classification: C10,C14,C22

Forecasting the old-age dependency ratio to determine the best pension age

Abstract

We forecast the old-age dependency ratio for Australia under various pension age proposals, and estimate a pension age scheme that will provide a stable old-age dependency ratio at a specified level. Our approach involves a stochastic population forecasting method based on coherent functional data models for mortality, fertility and net migration, which we use to simulate future age-structures of the population. We apply these ideas to Australian historical demographic data and the results suggest that the pension age should be increased to 67 by 2021, 70 by 2032 and 75 by 2058 in order to maintain the old-age dependency ratio at the 2018 level, which is 21%. This approach can easily be extended to other target levels of the old-aged dependency ratio and other countries.

Keywords: coherent forecasts, demographic components, functional time series model, pension and insurance

Adjust numbers in abstract as per Figure 9

1 Introduction

Australia's ageing population is a result of three factors: (1) significant technological advances in medical care leading to longer life spans; (2) high fertility rates after World War II resulting in the "baby-boomer" generation; and (3) a large decline in fertility rates over the past twenty or so years (Fehr, Jokisch & Kotlikoff 2008). Consequently, there are many older people leaving the workforce, and fewer workers replacing them. Similar demographic challenges face many other countries. One way of measuring this problem is via the old-age dependency ratio (OADR): the number of people post-retirement to the number of people of working age (Walker 1982). The OADR has been increasing for many decades, but recently the rate of increase has jumped markedly, and we show that it will continue to increase at a rapid rate if there are no age-related policy changes.

This is consistent with the Australian Intergenerational Reports (Costello 2002, 2007; Swan 2010) that explored the extent to which existing age-related policies would affect the burden of future working generations. To date, the Australian government has addressed the problem through a variety of approaches including modifications to compulsory superannuation policies, instigating work incentives for older people, and increasing the pension age (defined as the age at which eligible residents can access the age pension for the first time). In 2009, the federal government announced an increase in pension age from 65 to 67 by 2023 (Nielson 2010). In May 2014, the Federal Treasurer proposed a further increase in the pension age to 70 by 2035. This new proposal is in line with the recommendations from the Productivity Commission (2013) and the National Commission of Audit (2014). In this paper, we investigate the effect of these changes in the pension age on the forecast OADR.

The research presented in this paper contributes to the discussion by forecasting the age structure of the population, quantifying the effects of changing the pension age on financial burdens, and hence finding an appropriate pension age scheme which is economically viable for Australia. The analysis of the economic impact of increasing pension age can be complicated. In this paper, we will focus on only one aspect: the impact on the burden of future working-age generations.

In this paper, we use the OADR as a measure of the financial burden on the Australian workforce because a rise in the OADR will reduce the ability of working-age population to finance pensions and health costs (Apps, Rees & Wood 2007; Costello 2002, 2007). The OADR is defined as the ratio of the number of people aged over the pension age to the number of people of working age (Apps, Rees & Wood 2007). This can be expressed using the formula below:

$$\text{OADR} = \frac{\text{number of people aged over pension age}}{\text{number of people aged 15 to pension age}} * 100\% \quad (1)$$

We provide forecasts of future population age structures over the next 50 years. We then use those forecasts to compute the OADR associated with pension age both before and after the adjustments. One contribution of the current study is the combined use of the Hyndman-Ullah forecasting method (Hyndman & Ullah 2007) and the product-ratio method (Hyndman, Booth & Yasmeen 2013) to forecast a population age structure. Forecasts of population structure have been built on forecasts of age- and gender-specific mortality rates, fertility rates and net migration using this method. These forecasts presuppose that the historical trends of past years will continue. Historical data of those rates have thus been used as input to simulate future paths of age- and gender-specific population structure.

Using these predicted population paths, we can calculate the average OADR and predict the levels of old-age dependency that correspond to different pension age schemes. By testing different pension age schemes, we aim to find the pension age with the desired OADR outcome. Therefore, other contributions of this research are (1) to examine the effect of adjusting pension ages by comparing OADR before and after adjustment; and (2) to suggest a suitable pension age scheme to ensure that the economic burden will remain relatively constant at the desired level in the future. It should be noted that the desired level has been set at 21%, the estimated level of OADR in 2018. Whether this OADR level is the most appropriate level for Australia remains to be examined and is not the focus of this paper. This paper concentrates on providing a framework for policymakers to find a target pension age for a given desired level of OADR.

Some may argue that using OADR as a measure of a financial burden for the next generation could be biased, with a large number of retirees able to rely on self-financed retirement income, such as superannuation. However, 80% of Australians over pension age rely fully or partly on the age pension (Power 2014). Therefore, the proportion of people over working age is still a good representation of the financial burden on future working generations. Nevertheless, using OADR as a single measurement of financial burden may overlook some important economic factors, and Hu & Yang (2012) have suggested the use of an adjusted OADR that takes account of some of these factors in analyzing pension finance. We do not consider these here. Our approach is intended to illustrate how population forecasting can be used to explore the pension age issue in a simple context, and it could be extended to take into account other variables.

The remainder of this paper proceeds as follows. Section 2 reviews key literature on population forecasting and the pension age in Australia. Section 3 describes the data and modelling framework we use in this study. We show how our framework can be used to estimate a pension age scheme that leads to a target OADR in Section 4. Finally, Section 5 gives concluding remarks as well as identifies the scope for future research.

2 Background

2.1 The age pension in Australia

Age pension in Australia is income support and provides access to a range of concessions for older Australians who have passed means-testing. For example, it provides single home-owners with a real income of \$23,889 p.a and additional allowances to cover expenses such as rent or pharmaceuticals¹. The means test ensures that retirees with enough personal savings will

¹The pension amount is adjusted every 6 months for inflation

become ineligible for the public provision by reducing the pension at fixed levels of pay as personal wealth increases. The age pension in Australia is generous, even with the means test. For instance, a couple holding assets of more than one million dollars is still eligible for a partial age pension.

In 2006, the age pension was the primary source of income for 70–80% of Australian residents over 64 years (Australian Bureau of Statistics 2006). Even though the government has promoted self-financed retirement by introducing the Superannuation Guarantee in 1992, the Superannuation Guarantee accumulations are still modest, with the average less than \$100,000 at retirement (ASFA 2007). It has been estimated by Petrichev & Thorp (2008) that a retiree will need to have \$450,000 at 65 in order to support a pension-equivalent income and this is more than eight times of the average retirement accumulation of Australians. The required wealth to comply with personal-funded retirement savings requirements is too high for most Australians, suggesting that retirees will continue to rely on the age pension in the coming decades.

The ageing trend in Australia, along with retirees' dependence on the age pension, has dramatically increased the fiscal cost of the age pension. For instance, the fiscal cost of the age pension is currently around at 2.5% of Gross Domestic Profit (GDP), but is expected to increase to 4.4% by the 2050s (Australian government, 2007). The Productivity Commission (2013) also indicates the government will need to raise taxes by 21% to pay for extra aged care and health costs, which is caused by longer life expectancies in newer generations.

The issue is not only a problem for the government, but also for future retirees. Given children born today are expected to live up to their 90s on average, the Productivity Commission (2013) suggests that Australians won't be able to afford to spend 35 years in retirement (i.e., retire around 65). They have predicted that a third of baby boomers are expected to spend all their money and assets before they die². This problem might eventually deteriorate the sustainability of the age pension, although the welfare system works well at the moment, because demographic changes often have delayed effects (Costello 2002).

The Australian government is trying to take early intervention in relation to the pension age to balance between providing social support and fiscal cost. The pension age for males has been set at 65 for over a century. While the pension age for females was 60 years until 1994, it has steadily increased to 65 years, in line with that of males. The previous Labor government announced in 2009 that the pension age would increase six months every two years until 2017, capping at 67 in July 2023. This change is expected to have a significant impact on labor force

²Baby Boomers are people born during the Post-World War II baby boom between the years 1946 and 1964.

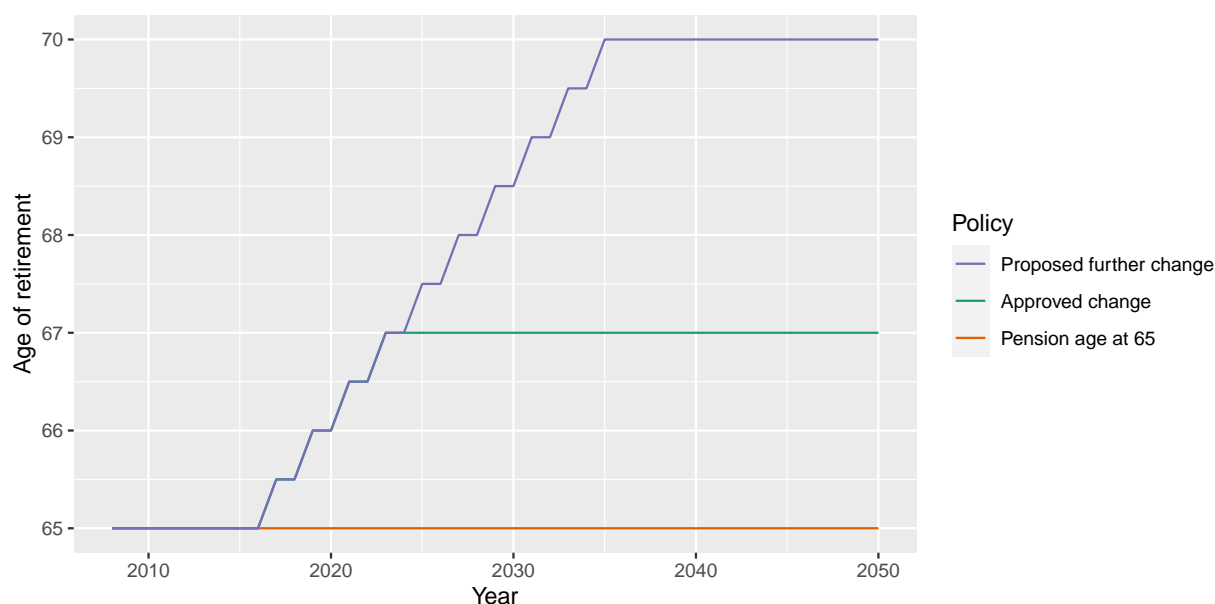


Figure 1: *Government's adjustments on pension age scheme.*

participation and to decrease the retirement duration of residents because the pension age is regarded as a normal age for retirement in Australia. However, the Productivity Commission (2013) proposed further lifting of the pension age to 70 on the grounds that increasing the pension age to 70 would save taxpayers \$150 billion in welfare costs and health spending, and older people are capable of continuing work. The National Commission of Audit (NCOA) (2014) holds the same opinion, and recommended lifting the pension age to 70 from 2053. The former Treasurer Joe Hockey announced that the government would increase the pension age to 70 from 2035, which is proposed in the Federal Budget Statement for 2014–2015 to be confirmed by the Australian Federal Government. Figure 1 illustrates the process of the adjustments on pension age scheme. This announcement received overwhelmingly negative responses in the media. Guest (2014) argues from an economic perspective that Australia currently has much less age pension pressure than other OECD countries, because spending on the age pension in Australia is less than half of the OECD average. He also claims that Australia has a higher consumption of all other goods and services, and so why are they not allowed to have more leisure time. Further, Power (2014) contends that many Australians are not physically able to work full-time to 70, so that recommendations made by the PC and NCOA are simplistic solutions to a more complex social problem.

All supportive arguments for increasing the pension age in Australia have only justified the need to increase the pension age. No justification on the pace or degree of the change has been provided. Also, the government has not disclosed the method it used to determine the new pension age. It is therefore not surprising that the adjustment of the pension age has been

questioned by the public. A consistent method to determine the appropriate pension age would be more persuasive in general. For example, it is preferred by OECD countries that pension age should be linked to life expectancy (OECD 2012). There is a practical issue as to how to link life expectancy with pension ages given that there is no information on future mortality rates (Productivity Commission 2013). Knox (2007, p7) proposed directly linking pension age with life expectancy in the way that “pension age should increase by approximately 50% of any increase in life expectancy”. However, using a “rule of thumb” approach might not be ideal, as 50% of the change in life expectancy may not be a good adjustment to pension age in the long term. Our study suggests an approach, linking the pension age with the aged dependency ratio, based on forecasts of future mortality and fertility rates. Our approach has the advantage that it can be used to justify the change in pension age and directly links the pension age with the goal of a stable financial burden for the public.

2.2 Population forecasting

In population forecasting, two different approaches — a deterministic approach (sometimes called a quasi-statistical approach) and a stochastic approach (a purely statistical approach) — are frequently used. For instance, the Australian Bureau of Statistics (ABS) is using a deterministic approach for its demographic projection, while some other countries or institutes forecast demographic trends based on purely statistical approaches such as Statistics New Zealand and the United States Census Bureau. In the deterministic approach, experts project future demographic changes using judgement-based scenarios. It ignores the non-demographic factors such as major pandemics or wars, although historically they are important factors that affect demographic behaviour (Productivity Commission 2013). In stochastic approaches, the Lee-Carter (LC) method (Lee & Carter 1992) is the most prominent demographic forecasting method. It now has several variants and extensions that relax some of the assumptions of the basic Lee-Carter method (see, e.g., Shang, Booth & Hyndman 2011). Although statistical forecasting models are rigorous, they still leave a basis for judgment-based adaptation, as they sometimes produce unrealistic projections of demographic factors such as long-run projections of mortality rates (Productivity Commission 2013).

Therefore, some official agencies decide to mix the two approaches in their forecasts in order to provide a supplement to both approaches. For instance, the most recent forecast of the population in Australia by the Productivity Commission (2013) was based on basic Lee-Carter methods plus some judgement-based adaptations to address the characteristics of demographic trends that the Lee-Carter method failed to capture. The approach of forecasting used in this study is similar to the Productivity Commission (2013). The Hyndman-Ullah method (2007)

is used as a basic forecasting approach, which is a comprehensive extension of the Lee-Carter method. In addition, an adapted product-ratio method is used to eliminate the potential long-term divergence between female and male mortality rates. A more detailed review of forecasting methods is given below.

Deterministic methods used by the ABS

The deterministic method is judgment-based. The Australian Bureau of Statistics formulates assumptions on the basis of demographic trends, in conjunction with advice from individual experts or government departments. The projections are not predictions or forecasts, “but are simply illustrations of the growth and change in population which would occur if certain assumptions about future levels of fertility, mortality, internal migration and overseas migration were to prevail over the projection period” (Australian Bureau of Statistics 2012, p.2). In addition, they do not attempt to take into account of non-demographic factors, such as major government policy changes and catastrophes. Therefore, they normally report multiple (usually three) series based on different levels of assumptions of demographic factors to allow for a variety of population projections.

Lee-Carter method

The Lee-Carter method (Lee & Carter 1992) is a significant milestone in demographic forecasting. It has been widely used to forecast mortality rates across various countries, including Australia.

The model proposed by Lee & Carter (1992) is

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

where $m_{x,t}$ is the central death rate (properly defined in Section 3) for age x in year t . This model involves fitting a matrix of log central death rates with sets of age-specific constants, $\{a_x\}$ and $\{b_x\}$, and a time-varying index k_t . b_x is a set of parameters that tell us the rate at which each age responds to changes in k_t . k_t measures the general level of the log death rates; a_x is the general pattern across the age of the log mortality rate and $\varepsilon_{x,t}$ is an error term with zero mean and constant variance, reflecting the randomness that can't be captured by the model.

The model is over-parameterised. For example, if the vector a, b, k is one solution, then for any scalar c , $a, b/c, kc$ or $a - bc, b, k + c$ also must be two possible solutions. Therefore, Lee and Carter (1992) impose the following constraints:

$$\sum_{t=1}^n k_t = 0, \quad \sum_{x=x_1}^{x_m} b_x = 1 \quad (3)$$

Lee & Carter (1992) use a singular value decomposition (SVD) method to find the least squares solution of (2) (Good 1969). The estimates of b_x and k_t are therefore the first principal component and scores of X with (i, t) th element $[\log(m_{x_i,t}) - a_{x_i}]$, where a_{x_i} is the sample mean of $\log(m_{x_i,t})$ over t . The log transformation introduces some bias, so the LC method adjusts k_t such that for each year, the implied number of deaths will be equal to the actual number of deaths. The adjusted k_t is then extrapolated using a random walk with drift model, which is equivalent to an ARIMA (0, 1, 0) model with a drift parameter.

Hyndman-Ullah method

Hyndman & Ullah (2007) proposed a nonparametric method for demographic modelling and forecasting that extended the LC model in four ways:

1. The demographic rates are smoothed using penalized regression splines (Booth, Hyndman & Tickle 2014) before modeling.
2. It uses functional principal components analysis, which is a continuous version of PCA.
3. It uses more than one principal component, which addresses the main weakness of the LC method.
4. The forecasting models for the principal component scores are more complex than the random walk with drift model used in the LC model.

The model proposed by Hyndman & Ullah (2007) is as follows:

$$y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \quad (4)$$

$$s_t(x) = \mu(x) + \sum_{j=1}^J \beta_{t,j}\phi_j(x) + e_t(x) \quad (5)$$

Equation (4) describes the smoothing process. $y_t(x)$ is the observed data in year t and age x , and $s_t(x)$ is the smoothed function of $y_t(x)$, estimated by constrained weighted penalized regression splines (Booth, Hyndman & Tickle 2014). $\sigma_t(x)$ allows the variance to be a function of time and age; and $\varepsilon_{t,x}$ is independent and identically distributed standard normal error terms. Equation (5) has a similar structure to the LC method, except that it incorporates more than one principal component and score. It describes the dynamics of $s_t(x)$ over time. The $\phi_j(x)$ for every j and $\mu(x)$ are sets of age-specific constants, and $\beta_{t,j}$ for every j is a time-varying index. The $e_t(x)$ is the model error, which is the residual function with mean zero and no serial correlation. The $\mu(x)$ is estimated by the average of $s_t(x)$ across years, which measures the general shape of the observed data $y_t(x)$. The estimates of $\phi_j(x)$ and $\beta_{t,j}$ are the j^{th} principal component and scores of $s_t(x) - \mu(x)$ respectively. In other words, we take singular value decomposition of

$s_t(x) - \mu(x)$, then (5) now becomes

$$s_t(x) - \mu(x) = \phi_1(x)\beta_{t,1} + \phi_2(x)\beta_{t,2} + \cdots + \phi_J(x)\beta_{t,J}$$

We choose J in a way that we believe the majority of information for $s_t(x) - \mu(x)$ has been captured by the first J components. Following Hyndman & Booth (2008), we select $J = 6$ which has been found empirically to be large enough for demographic applications.

Ty conditioning on the observed data $\mathbf{I} = \{y_1(x), \dots, y_n(x)\}$ and the set of functional principal components $\mathbf{B} = \{\phi_1(x), \phi_2(x), \dots, \phi_J(x)\}$, the h -step-ahead forecast of log rate $y_{n+h}(x)$ are given by

$$\hat{y}_{n+h|n}(x) = \hat{s}_{n+h|n}(x) = E[\hat{s}_{n+h|n}(x) | \mathbf{I}, \mathbf{B}] = \hat{\mu}(x) + \sum_{j=1}^J \phi_j(x) \hat{\beta}_{n+h|n,j}, \quad (6)$$

where $\hat{\beta}_{n+h|n,j}$ denotes the h -step-ahead forecast of $\beta_{n+h|n,j}$. Because $\{\beta_{t,j}\}$ is time-varying, it controls the dynamics of the process. Each component series $\{\beta_{t,j}\}$ is modelled with an ARIMA(p_j, d_j, q_j) process, where p_j, d_j, q_j are determined by minimizing the Akaike Information Criterion (Hyndman & Khandakar 2008), and the parameters are estimated using maximum likelihood estimation.

3 Data

We use the same notation and similar procedures to construct data as in Hyndman & Booth (2008), because the fundamental model for forecasting population for this study is the Hyndman & Ullah (2007) method. The data used to construct population age structure included: age- and gender-specific birth and death numbers of each calendar year, age- and gender-specific population members on 1 January of each year, and age- and gender-specific exposures to risk (i.e., the population of age x at 30 June) for each year. The notations are as follows:

$B_t(x)$ = Births in calendar year t to females of age x

$D_t(x)$ = Deaths in calendar year t of persons of age x

$P_t(x)$ = Population of age x at 1 January of year t

$E_t(x)$ = Population of age x exposed to risk at 30 June of year t

where $x = 0, 1, 2, \dots, p-1, p^+$ denotes age with p^+ the open-ended upper age group, set to 100^+ in this study. The years is denoted by $t = 1, 2, \dots, n$, and superscripts M and F denote male and female respectively.

We can obtain historical mortality rates $m_t(x)$ (also called central death rates) and fertility rates $f_t(x)$ as follows:

$$m_t(x) = \frac{D_t(x)}{E_t(x)} = \text{age-sex-specific central death rates in calendar year } t$$

$$f_t(x) = \frac{B_t(x)}{E_t^F(x)} = \text{age-specific fertility rates in calendar year } t.$$

In addition, the net migration³ (denoted G) is estimated using demographic growth-balance equations, which summarize the relationship between population change and three factors: net migration, births and deaths. The equations are expressed as:

$$G_t(x, x+1) = P_{t+1}(x+1) - P_t(x) + D_t(x, x+1) \quad \text{for } x = 0, 1, 2, \dots, p-2,$$

$$G_t(p-1^+, p^+) = P_{t+1}(p^+) - P_t(p^+) - P_t(p-1) + D_t(p-1^+, p^+),$$

$$G_t(B, 0) = P_{t+1}(0) - B_t + D_t(B, 0),$$

where $G_t(x, x+1)$ refers to net migration in calendar year t of persons aged x at the beginning of year t , $G_t(p-1^+, p^+)$ refers to migration in calendar year t of persons aged $p-1$ and older at the beginning of year t , and $G_t(B, 0)$ refers to deaths in calendar year t of births during year t ; and similarly for deaths, $D_t(x, x+1)$, $D_t(p-1^+, p^+)$, and $D_t(B, 0)$. The deaths are estimated using the standard life table approach of population projection. It should be noted that the estimated net migration includes errors in data recording.

Historical central death rates and start-year and mid-year populations of residents in Australia by sex and age in single years for age group 0–99 and 100+ have been sourced from the Human Mortality Database (2020) for 1921–2018. The start-year population for 2019 is also used to estimate net migration. Data for age-specific annual fertility rates by single years of age for 15–49 over the period 1950–2018 was obtained from the Australian Bureau of Statistics (2012). Even though annual mortality rates are available from 1921 to 2018, only data collected after 1950 is used for two reasons. First, the period between 1921–1950 has different mortality patterns from later years because of wars and epidemics (Hyndman & Booth 2008). Since the model for this study is valid under the assumption that extrapolative methods can be used, it seemed best

³Net migration is the difference between immigration and emigration in a certain area during a specified time frame.

to delete the less relevant data, which might influence estimations of parameters if we include them. Second, fertility rates are only available from 1950, so the mortality rate was cut to fit the same period.

3.1 Forecasting population age structure and OADR

In this study, three steps were taken to simulate OADR associated with a specific pension age scheme. A stochastic functional data model was applied to five components of population change: female mortality rates, male mortality rates, female net migration, male net migration and fertility rates to simulate these five components over the prediction horizon (which is 50 years in this study). Age- and gender-specific population paths over the prediction horizon were simulated using the demographic growth-balance equation. The Monte Carlo simulation method was used consistently in this research. Finally, calculation of OADR associated with a specific pension age scheme was performed by applying the OADR formula. Even though the process seems complicated, it is relatively easy to implement using *R* (R Core Team 2020). The detailed explanation of each step is as follows. The corresponding R code is attached as a supplement.

Step 1

The main procedure for forecasting the population age structure follows Hyndman & Booth (2008), except that the data are updated, and we incorporate the product-ratio method (Hyndman, Booth & Yasmeen 2013) to improve the accuracy of the forecast. The Hyndman & Ullah (2007) model, introduced in section 2.2, is applied to the mortality rate, fertility rate and net migration separately, based on the assumption that they behave independently. Therefore, to fit the model to historical fertility rates, let $y_t(x)$ in equation (4) denote the observed historical fertility data. However, for the mortality rate by sex, the product-ratio method is used to ensure the forecast of females and males are non-divergent or coherent (Hyndman, Booth & Yasmeen 2013). In the product-ratio method, after smoothing the female and male mortality rates, two other series were created: product series (denoted by $p_t(x)$) and ratio series (denoted by $r_t(x)$), which are defined as follows:

$$p_t(x) = \sqrt{s_t^M(x)s_t^F(x)}, \quad (7)$$

$$r_t(x) = \sqrt{\frac{s_t^M(x)}{s_t^F(x)}}. \quad (8)$$

Then, Hyndman & Ullah (2007) was applied to these two new variables instead of male and female mortality rate as follows:

$$\begin{aligned}\log[p_t(x)] &= \mu_p(x) + \sum_{j=1}^J \beta_{t,j} \phi_j(x) + e_t(x) \\ \log[r_t(x)] &= \mu_r(x) + \sum_{l=1}^L \gamma_{t,l} \psi_l(x) + z_t(x)\end{aligned}$$

where time series $\{\beta_{t,j}\}$ are fitted with an ARIMA model while $\{\gamma_{t,l}\}$ are fitted by stationary ARFIMA model (i.e., d in $\text{ARIMA}(p, d, q)$ constrained to $0 < d < \frac{1}{2}$). It is the stationary constraint on $\gamma_{t,l}$ that ensures the forecast of male and female do not diverge. Then, simulating h -step-ahead forecast values of product variable and ratio variable was performed (denoted by $p_{n+h|n}(x)$ and $r_{n+h|n}(x)$ respectively) according to (7) and (8). From here, obtaining h -step-ahead forecast of sex-specific coherent mortality rates (denoted $m_{n+h|n,M}(x)$ and $m_{n+h|n,F}(x)$) was performed using a back-transformation:

$$\begin{aligned}m_{n+h|n}^M(x) &= p_{n+h|n}(x) r_{n+h|n}(x) \\ m_{n+h|n}^F(x) &= p_{n+h|n}(x) / r_{n+h|n}(x)\end{aligned}$$

The method for simulating h -step-ahead values of net migration by sex is similar to the above method used on mortality rates except no log scale is applied to net migration. A stationary constraint was imposed on the ARIMA model for net migration to make sure the difference between sexes did not diverge.

Step 2

After simulating paths for age-specific fertility rate, age- and gender-specific mortality rate and age- and gender-specific net migration, future population value was simulated using the demographic growth-balance equations over the prediction horizon. The future deaths and births in this equation are assumed to follow a Poisson distribution, with parameters as a function of future mortality and fertility rates (see Hyndman and Booth, 2008 for the details of the functions). Hence, we can simulate h -step-ahead age- and gender-specific births and deaths based on simulated mortality and fertility rates. According to the demographic growth-balance equation, one-step-ahead net migrations, deaths, births and population in the last year were simulated, allowing for the generation of one-step-ahead population predictions by sex and age. This simulation was repeated for years $t = n + 1, \dots, n + h$ to obtain a path of population. The process was repeated and simulated i paths of population based on Monte Carlo simulation

method were created (we set $i = 1000$ in this research). It should be noted that there are three types of randomness in the simulation: randomness from the second equation in the Hyndman-Ullah specification (i.e., $e_t(x)$ in (4)), randomness from the ARIMA model fitted to principal component scores and randomness from the Poisson distribution. It should be noted that the simulated population will be an array with three dimensions: age, year and simulation.

Step 3

Even though the OADR is a simple function of the population age structure and pension age, the formula can't be directly applied to the simulated population array because the pension age changes over time. To overcome this problem, we assume that the pension age can only be adjusted at the beginning of a year for these calculations. Therefore, one pension age is applicable for a whole year and hence OADR is generated at the beginning of each year to show the trend. We will show the process of obtaining a one-step-ahead OADR (i.e., forecast of OADR in the year 2019), given a specific pension age. The same method can be applied to all years over the prediction horizon, given their corresponding pension age. Further, the fractional pension age causes practical issues in calculating OADR, since ages are recorded as integers. We will also show how this issue was handled in the example below.

Suppose the pension age for $n + 1$ is $a.r$ where a is the integer and r is the remainder of pension age. Recall that for $h = 1$, we can express the simulated population paths as a matrix with age in the row and simulation in the column as follows:

$$\mathbf{p}_{n+1} = [p_{n+1,m,i}] = \begin{bmatrix} p_{1,1} & \cdots & p_{1,1000} \\ \vdots & \ddots & \vdots \\ p_{100+,1} & \cdots & p_{100+,1000} \end{bmatrix}.$$

When $r = 0$, then the simulation of a population aged 15 to pension age is a subset of the above matrix from 15^{th} row to $a - 1^{th}$ row, which is:

$$\begin{bmatrix} p_{15,1} & \cdots & p_{15,1000} \\ \vdots & \ddots & \vdots \\ p_{a-1,1} & \cdots & p_{a-1,1000} \end{bmatrix}. \quad (9)$$

To calculate the total number of working-age, we take a column sum of the trimmed matrix (9). The simulations of the total number of working-age become a vector:

$$\mathbf{p}_{n+1}^{\text{work}} = \left[\sum_{m=15}^{a-1} p_{m,1}, \sum_{m=15}^{a-1} p_{m,2}, \dots, \sum_{m=15}^{a-1} p_{m,1000} \right] \quad (10)$$

Using the same method, the simulations of the total number of aged pensioners in the population can be shown as:

$$\mathbf{p}_{n+1}^{\text{aged}} = \left[\sum_{m=a}^{100^+} p_{m,1}, \sum_{m=a}^{100^+} p_{m,2}, \dots, \sum_{m=a}^{100^+} p_{m,1000} \right] \quad (11)$$

When $r \neq 0$, the vector of the total number of working-age population in (10) is added, and the vector of the total number of aged population in (11) is subtracted with an adjustment:

$$\text{adjustment} = 0.r \times 100\% [p_{a,1}, p_{a,2}, \dots, p_{a,1000}] \quad (12)$$

This adjustment is based on the assumption that the birthday of the population is uniformly distributed over the year. Then, $r \times 100\%$ of the population with a recorded integer age of a at beginning of the year is actually younger than $a.r$, so they are still under the pension age. In other words, $(1 - r \times 100\%)$ of the population with recorded integer age a at the beginning of the year is older than $a.r$, so they should be counted as the aged population. The simulation vectors of total work age population and total aged population should then be:

$$\begin{aligned} \mathbf{p}_{n+1}^{\text{work}} &= [p_{n+1,1}^{\text{work}}, p_{n+1,2}^{\text{work}}, \dots, p_{n+1,1000}^{\text{work}}] \\ &= \left[\sum_{m=15}^{a-1} p_{m,1}, \sum_{m=15}^{a-1} p_{m,2}, \dots, \sum_{m=15}^{a-1} p_{m,1000} \right] + 0.r \times 100\% [p_{a,1}, p_{a,2}, \dots, p_{a,1000}] \\ \mathbf{p}_{n+1}^{\text{aged}} &= [p_{n+1,1}^{\text{aged}}, p_{n+1,2}^{\text{aged}}, \dots, p_{n+1,1000}^{\text{aged}}] \\ &= \left[\sum_{m=a}^{100^+} p_{m,1}, \sum_{m=a}^{100^+} p_{m,2}, \dots, \sum_{m=a}^{100^+} p_{m,1000} \right] - 0.r \times 100\% [p_{a,1}, p_{a,2}, \dots, p_{a,1000}] \end{aligned}$$

After the adjustment, we can generate simulation vectors of the one-step-ahead OADR by dividing every element of $\mathbf{p}_{n+1}^{\text{work}}$ by the corresponding element of $\mathbf{p}_{n+1}^{\text{aged}}$, i.e.:

$$\begin{aligned} \text{OADR}_{n+1} &= [\text{OADR}_{n+1,1}, \text{OADR}_{n+1,2}, \dots, \text{OADR}_{n+1,1000}] \\ &= \left[\frac{p_{n+1,1}^{\text{work}}}{p_{n+1,1}^{\text{aged}}}, \dots, \frac{p_{n+1,1000}^{\text{work}}}{p_{n+1,1000}^{\text{aged}}} \right] \end{aligned}$$

The average value of elements of OADR_{n+1} is the mean prediction of one-step-ahead OADR.

The same algorithm is applied to the simulation matrix of the population age structure to obtain h -step-ahead simulation vectors of OADR for $h = 1, 2, \dots, H$, where $H = 50$ for this study (i.e., a prediction horizon of 2019 to 2068). Simulations of OADR over a 50 year prediction horizon by

the matrix is summarized below, with the forecast year in rows and simulation in the columns:

$$\text{OADR}^{\text{forecast}} = [\text{OADR}_{h,i}] = \begin{bmatrix} \text{OADR}_{n+1,1} & \cdots & \text{OADR}_{n+1,1000} \\ \vdots & \ddots & \vdots \\ \text{OADR}_{n+50,1} & \cdots & \text{OADR}_{n+50,1000} \end{bmatrix}.$$

3.2 Results

We use the fitted model of fertility rates as an example to interpret the results of fitted models. It should be noted that models fitted to product and ratio series are less likely to be interpreted in an economic sense. Figure 2 shows the estimated mean age pattern $\hat{\mu}(x)$, the first two estimated basis functions, $\hat{\phi}_1(x)$ and $\hat{\phi}_2(x)$, and first two estimated coefficients (i.e., estimated principal component scores $\hat{\beta}_{t,1}$ and $\hat{\beta}_{t,2}$) for fertility rates. The first term ($\hat{\phi}_1(x) \times \hat{\beta}_{t,1}$) accounts for the most variation in fertility, so a relatively sensible interpretation can be given to the first term. In the plot of the first coefficient, estimated historical first principal component scores $\hat{\beta}_{t,1}$ are shown in black for 1950–2018. It captures the general trend of fertility change over time. It indicates a sharp increase between 1950–1960, followed by a rapid decline in fertility rates from 1961. This increase of fertility is linked to the Post-World War II baby boom⁴ in Australia, resulting from the recovery from war and the Great Depression⁵. The dramatic decrease from 1960 was due to the availability of contraception pills in Australia. A noticeable further reduction of fertility rates in the late 1960s may be due to a change in the law allowing abortion in some states in Australia.

The first estimated basis function $\hat{\phi}_1(x)$ measures the different effects of the trend over time across ages. It indicates the general pattern of fertility has little effect on females of age 15. This makes intuitive sense because most early pregnancies are accidental. Therefore, any of the previously discussed events has no effect on early pregnancies. In addition, there are two spikes in the first basic function, which suggests that the historical trend has a distinct impact on females aged around their 20s and 40s. Furthermore, the blue line is the fitted value of future coefficients (i.e., $\hat{\beta}_{n+h,1}$, for $h = 1, 2, \dots, 50$) and the yellow shading shows their 80% pointwise prediction interval according to the fitted ARIMA model of first coefficients $\{\hat{\beta}_{t,1}\}$. Other terms further explain variation that has not been captured by the first term, but less economic explanations can be given for them.

⁴Between the years 1946 and 1964.

⁵The Great Depression was a severe economic depression in the decade before World War II.

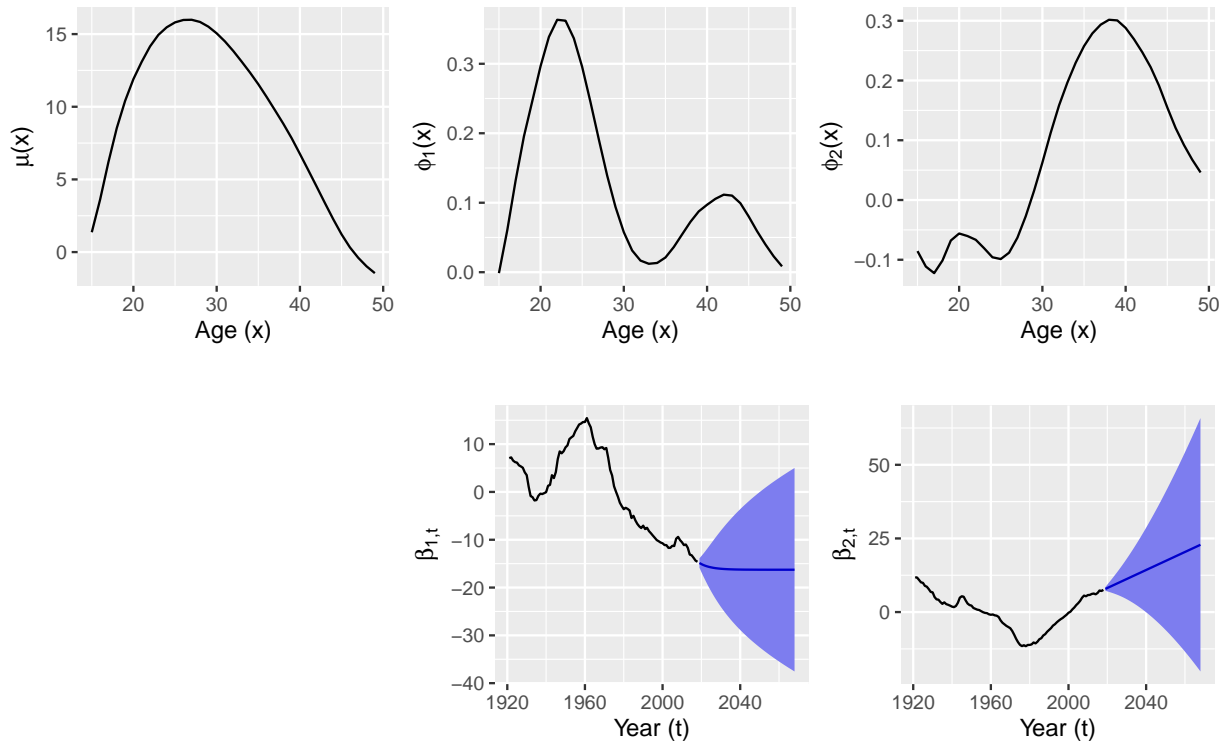


Figure 2: Fitted basis functions and coefficients for Australian age-specific fertility rates.

Figures 3 and 4 show the forecast results of population. Figure 3 shows the forecast of the total population by gender over the prediction horizon, while Figure ?? shows the forecast of population age structure in a specific year, 2038. Figure 3 indicates that both the female and male population are expected to increase to around 27 million by 2068. Since the fertility rate is expected to remain low, the possible reason for the significant increase in population size is net migration. We can also observe that the further the projected year is, the bigger the forecast variation of a population. This increase in variation shows that we have less relevant information for further apart prediction horizons.

Figure 4 shows the mean and the 80% pointwise prediction interval of simulated population paths for the year 2038, along with the 2019 base population. It can be observed that the overall population size is expected to increase substantially. The uncertainty of the infant population is much higher than other ages, which reflects the higher forecast variation of fertility rates than mortality rates. The uncertainty of the young age population is mainly due to migration (Hyndman & Booth 2008).

The analysis was performed by first generating an OADR associated with a fixed pension age of 65 and an OADR associated with pension age scheme with the approved change, gradually increasing the pension age from 65 to 67 by mid-2023. Even though the government might further increase the pension age in future, we kept the pension age at 67 from 2024 to see

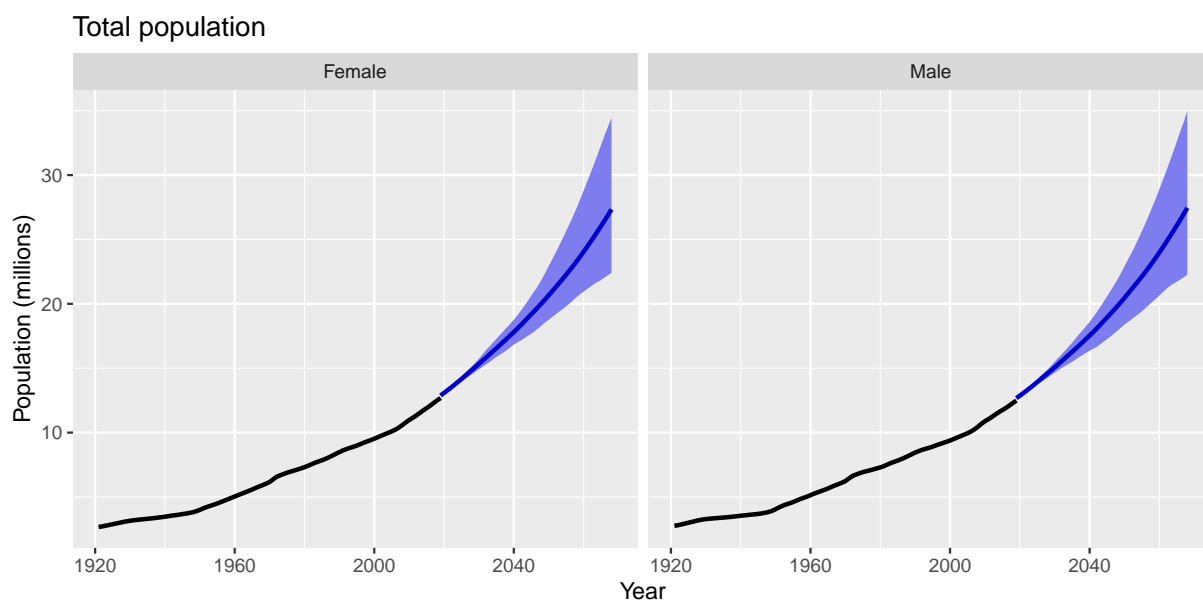


Figure 3: Fifty-year forecasts of the total population for each sex, along with 80% prediction intervals. The actual population from 1950 to 2019 is shown using circles.

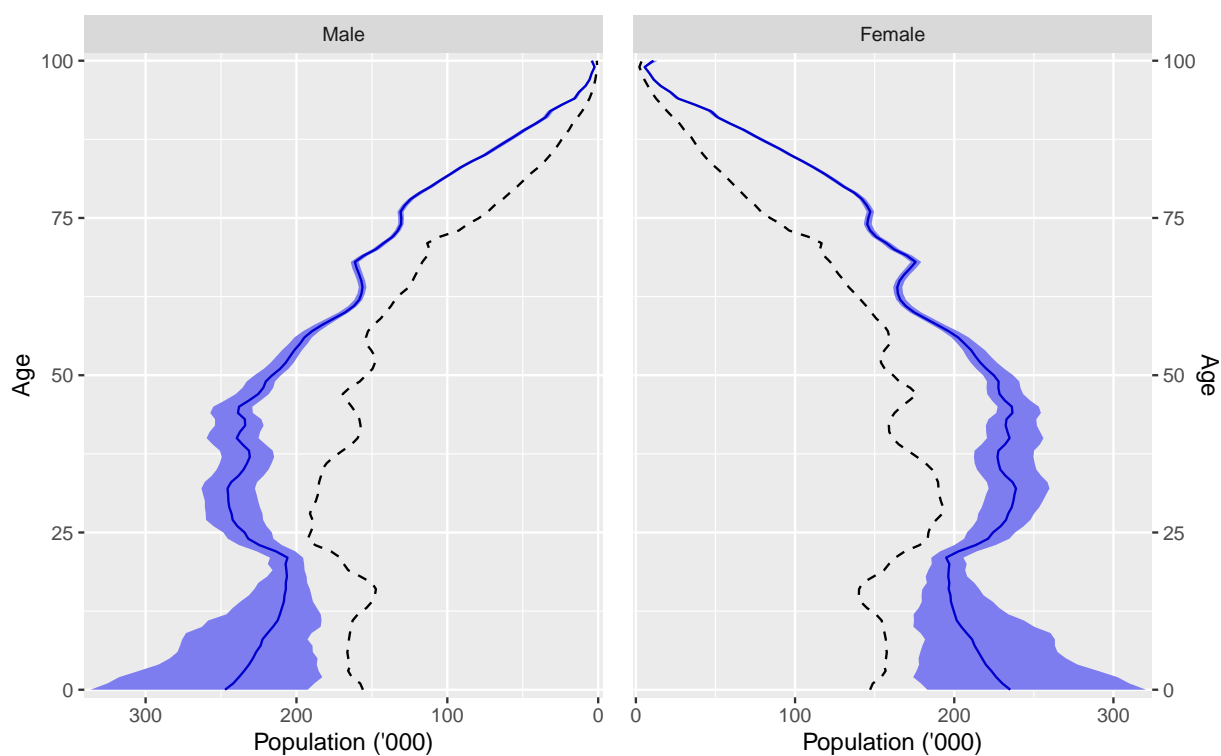


Figure 4: Forecast population pyramid for 2038, along with 80% pointwise prediction intervals. The actual pyramid for 2019 is shown using dashed lines.

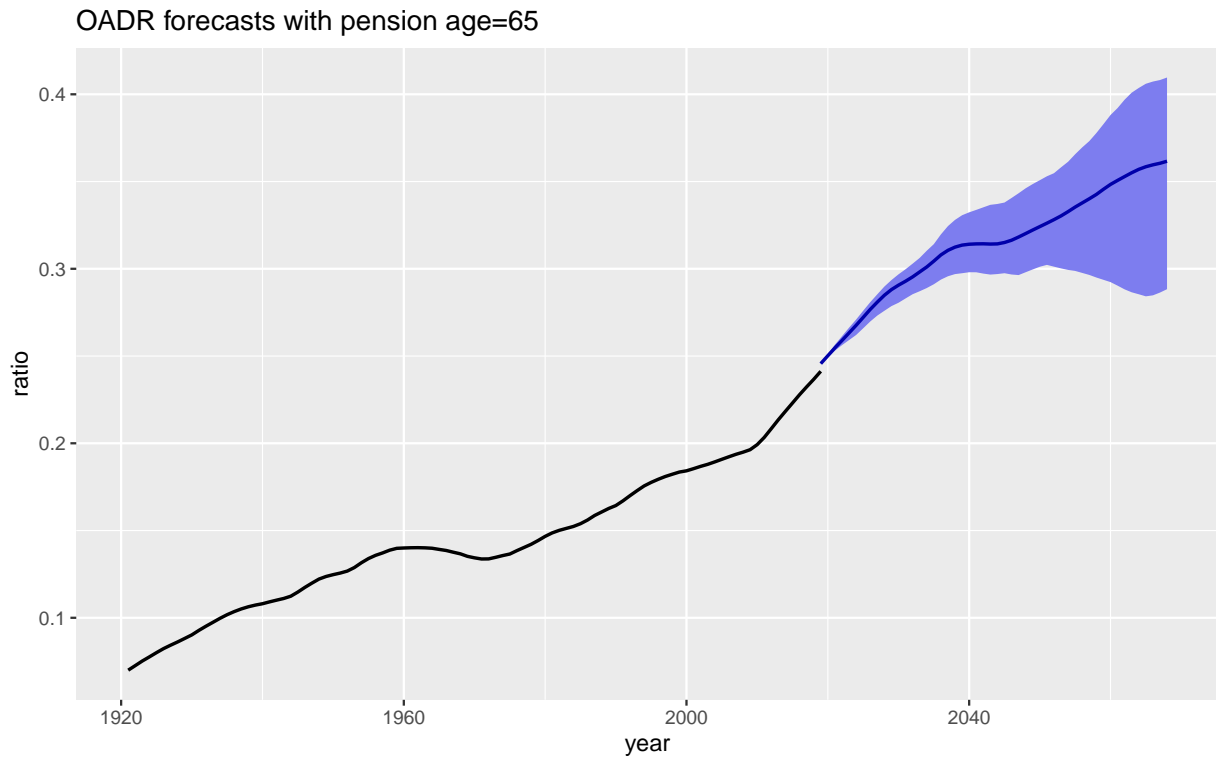


Figure 5: Fifty-year forecasts of the OADR with the assumption of pension age equal to 65 for both sexes, along with 80% pointwise prediction interval.

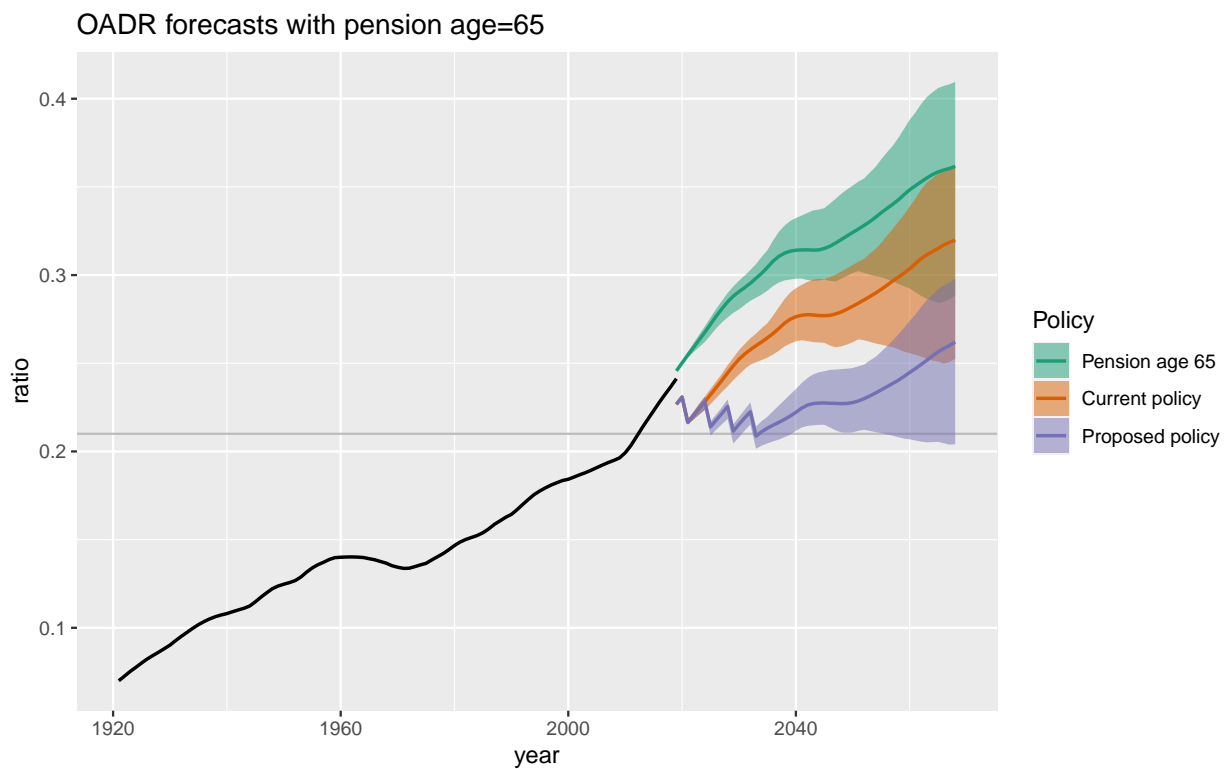


Figure 6: Fifty-year forecasts of the OADR associated with a pension age equal to 65 (green line) and equal to the pension age scheme with approved change (orange line) respectively, along with the desired OADR level in purple. The actual OADRs for 1921–2018 are shown in black.

what would happen if no further action was taken. Figure 5 shows the mean forecast and pointwise prediction interval of the OADR with a pension age of 65, along with the historical OADR for 1921–2018. A strong positive trend was observed, which indicates an increasing burden on the next working-age generation if the pension age is fixed at 65. This is consistent with the prediction by inter-generation reports in 2002 (Costello 2002) and 2007 (Costello 2007). According to former Australian Treasurer Costello (2004), this could be a result of better health care and low fertility rates. These two reasons could be regarded as permanent change and therefore a persistent trend. In addition, there is a big jump from actual data to predicted data. This jump is possibly a result of the Baby Boomers' retirement. Those born in 1946, the start of the baby boom, would begin retiring in 2011 when the pension age was 65.

Figure 6 presents a comparison of the mean forecast of OADR with the pension age equal to 65 and a pension age equal to the pension age scheme with the approved change, along with a comparison of both forecasts with the desired level of OADR, which has been set at 21%. It should be noted that OADR corresponding to pension age scheme with approved change is approaching the desired level, which suggests that the approved change does achieve the aim of providing an affordable financial burden, at least for a moment. However, once the pension age reaches 67, the OADR of the pension scheme with approved adjustment grows at the nearly same rate as the OADR of a pension age equal of 65. This is consistent with the idea that there is a persistent positive trend in OADR associated with a constant pension age, due to low fertility rates and better health care. Therefore, Figure 6 demonstrates that the pension age scheme requires constant review to ensure a stable OADR over the long term. However, the questions of when the action needs to be taken and how much of an adjustment needs to be made remain to be answered. This answer will allow us to respond to the recent proposal of further increasing the pension age to 70 by 2035.

3.3 Evaluating forecast accuracy

Since all the results and implications are dependent on the forecast accuracy, it is important to evaluate the accuracy of our forecast. It is not valid to assess the model's fit in terms of historical data. We are concerned with how well a model performs on out-of-sample data, which is known but not used to fit the model. Therefore, we will use a cross-validation procedure (Hyndman & Athanasopoulos 2014) to evaluate the accuracy of the models used in this paper.

The logic of a cross-validation test is to use a portion of available data for fitting, and use the remaining data to test. Suppose we need k years' observations to produce a reliable forecast⁶, we

⁶ k can't be too small

are interested in the H prediction horizon. Then we are allowed to repeat the test $n - H - k + 1$ times, where n is the number of total years of observation. The procedure is as follows:

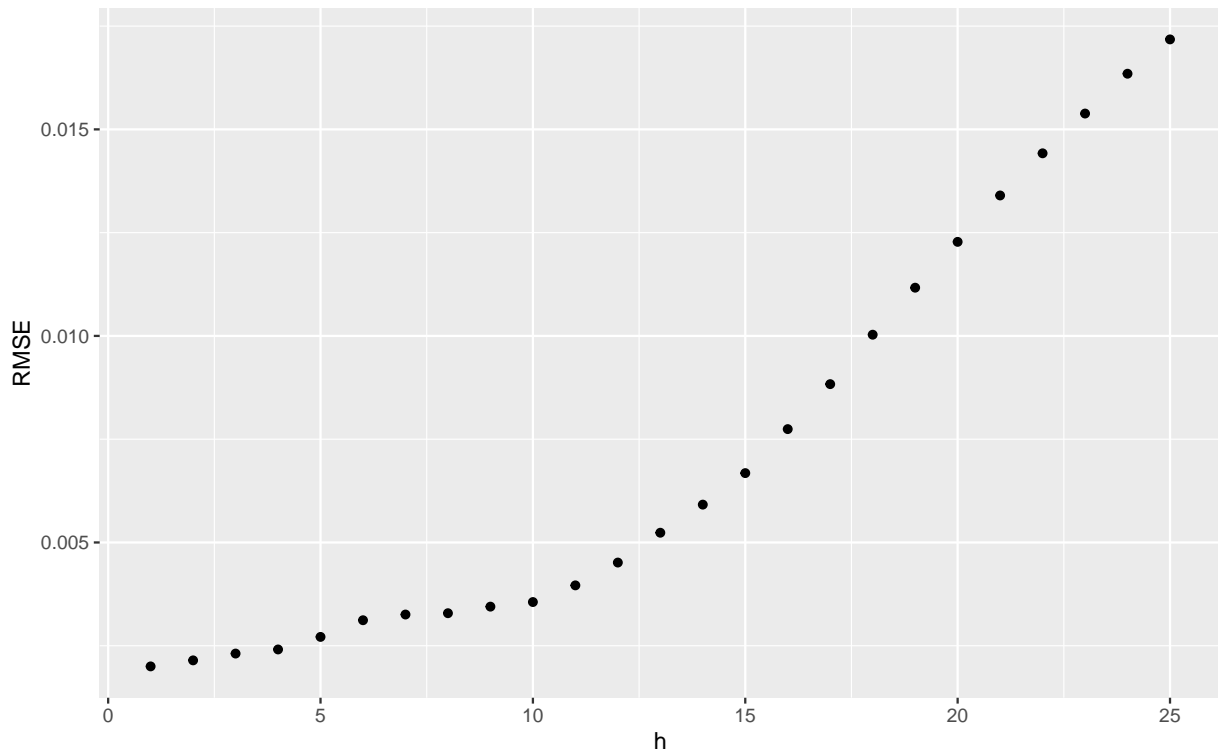
Let $\omega = 1$, we first select historical data for a time interval of $[\omega, \omega + 1, \dots, \omega + k - 1]$ as “in sample data” to fit the model, and then forecast H years’ OADR⁷. Then, we can calculate the difference of forecasted OADR over the period $[\omega + k, \omega + k + 1, \dots, \omega + k + H - 1]$ from the real OADR for the same period, which is known as absolute error. We repeat the following steps for $\omega = 1, 2, \dots, n - k - H + 1$. Then, we can summarize the information into a matrix as follows:

$$\text{Absolute Error} = [AE_{h,\omega}] = \begin{bmatrix} AE_{1,1} & \cdots & AE_{1,n-k-H+1} \\ \vdots & \ddots & \vdots \\ AE_{H,1} & \cdots & AE_{H,n-k-H+1} \end{bmatrix}$$

where h denotes for the forecast horizon and $AE_{h,\omega} = |OADR_{h,\omega}^{\text{mean}} - OADR_{h,\omega}^{\text{real}}|$. This process is also known as “rolling forecasting origin”, because the “origin” ($k + \omega - 1$) at which the forecast is based rolls forward in time. We take the average of each row to obtain an absolute mean error of h -step-ahead forecast of OADR where $h = 1, 2, \dots, H$.

In our sample, $n = 2018 - 1950 + 1 = 69$ and we set $k = 25$ and $H = 25$ to perform the test. The mean absolute error of h -step-ahead forecast of OADR for $h = 1, 2, \dots, 25$ is shown in Figure ???. From this Figure, we observe that the mean absolute error is increasing as h increases, which is reasonable as we have less relevant information to make a forecast for further apart OADR. Overall, the mean absolute error is small. For example, the mean absolute error for 25-step-ahead forecast is 1.6%, which is small if compared to the desired level of OADR, which is at 21%.

⁷The OADR is generated based on a pension age equal to 65 for the purpose of test.



It would normally be ideal for comparing the mean absolute error of the forecast from our model to other available models to measure the accuracy. For example, suppose we can take the population projections from ABS reports over the last few decades, their projection of population age structure reported in different years can be used to mimic the rolling origin test of OADR, because we now have the real OADR for those projected years. However, we are not able to perform it because full projection data is no longer available online and ABS projections have not used a consistent and replicate methodology.

4 Finding the target pension age scheme

The goal of this research is to keep the pension age low, as long as OADR is lower than or equal to the desired OADR level. According to the formula of OADR, OADR at a given time can be written as a function of pension age, when the population age structure is given. Suppose we denote pension age (PA) schemes over the prediction horizon by $PA = [PA_{n+1}, PA_{n+2}, \dots, PA_{n+H}]$. The OADR function of pension age over the prediction horizon is denoted as $OADR_h(PA_{n+h})$ for $h = 1, 2, \dots, H$. There is a negative relationship between pension age and OADR. The negative relationship implies that the higher the pension age, the lower the OADR. Therefore, our goal becomes finding the minimum pension age which has an OADR lower than desired OADR level. Several practical issues need to be considered to achieve this goal. First, only simulated forecasting values of future OADR are available, so it is assumed that the mean value of the

simulated OADR is a good estimate of true underlying OADR. Second, pension age cannot be increased by more than one within a year or decreased because either one of the two actions will force some retirees back to the workforce, which is unreasonable in practical applications. Therefore, the target pension age scheme PA^{target} can be defined in the following way: for $h = 1, 2, \dots, H$

$$\text{minimise } PA_{n+h} \text{ such that } OADR_h(PA_{n+h}) \leq OADR^{\text{desired}}$$

$$\text{subject to the constraint } 0 \leq PA_{n+h} - PA_{n+h-1} \leq 1$$

In order to achieve this maximization subject to constraint, these calculations assume that the adjustment unit of pension age is one month for practical convenience. It is also supposed that the set starting value of pension age is at 65 for all years⁸, which is:

$$PA = [PA_{n+1}, PA_{n+2}, \dots, PA_{n+H}] = [65, 65, \dots, 65]$$

We take the following algorithm: For $h = 1, 2, \dots, H$, we keep increasing pension age vector $[PA_{n+h}, PA_{n+h+1}, \dots, PA_{n+H}]$ by a month at a time, unless we find that

$$OADR_h^{\text{mean}}(PA_{n+h}) > OADR^{\text{desired}} \quad \text{or} \quad PA_{n+h} - PA_{n+h-1} \geq 1$$

where $OADR_h^{\text{mean}}(PA_{n+h})$ is the mean value of h -step-ahead forecast of OADR with a pension age at the time $n + h$ equal to PA_{n+h} .

In addition to finding the target pension age based on the mean value of OADR, the aim is to find plausible pension age schemes that could give the desired OADR level. To be more specific, “plausible”, means finding a range of pension age schemes that could lead to the desired OADR, contained by 80% prediction intervals of simulated OADR. Because of the monotonic relationship between pension age and OADR, we only need to find the upper (lower) boundary of plausible pension age schemes, at which the upper (lower) limit of 80% prediction intervals of OADR is equal to the desired OADR level. Hence, any pension age scheme within those two boundaries is plausible. The range of plausible pension age schemes can be thought of as a 80% confidence interval of pension age scheme where we are 80% sure that there will be the desired outcome. In other words, any pension age scheme outside the range is undesirable to

⁸Even though the pension age for males and females is 65.5 in mid-2017, the same pension age for males and females was set to simplify the calculations. This simplification is reasonable because the pension age will be same in mid-2017 onward

Table 1: Fifty-year target pension age scheme with stable OADR around desired OADR. Y denotes years and M denotes months.

Year	Age	Year	Age	Year	Age	Year	Age	Year	Age
2019	65Y	2029	68Y+1M	2039	70Y+1M	2049	70Y+1M	2059	71Y+1M
2020	65Y	2030	68Y+1M	2040	70Y+1M	2050	70Y+1M	2060	71Y+1M
2021	65Y	2031	68Y+1M	2041	70Y+1M	2051	70Y+1M	2061	71Y+1M
2022	65Y	2032	68Y+1M	2042	70Y+1M	2052	70Y+1M	2062	71Y+1M
2023	65Y+12M	2033	68Y+1M	2043	70Y+1M	2053	70Y+1M	2063	71Y+1M
2024	66Y+12M	2034	68Y+1M	2044	70Y+1M	2054	70Y+1M	2064	71Y+1M
2025	67Y+1M	2035	68Y+1M	2045	70Y+1M	2055	70Y+1M	2065	72Y+1M
2026	67Y+1M	2036	69Y+1M	2046	70Y+1M	2056	70Y+1M	2066	72Y+1M
2027	67Y+1M	2037	69Y+1M	2047	70Y+1M	2057	71Y+1M	2067	72Y+1M
2028	67Y+1M	2038	69Y+1M	2048	70Y+1M	2058	71Y+1M	2068	72Y+1M

Table 2: Fifty-year lower and higher boundary of pension age scheme with 80% confidence. Y denotes years and M denotes months.

Year	Lo	Hi	Year	Lo	Hi	Year	Lo	Hi	Year	Lo	Hi	Year	Lo	Hi
2019	65Y	65Y	2029	68Y+1M	68Y+1M	2039	69Y+1M	70Y+1M	2049	69Y+1M	70Y+1M	2059	69Y+1M	71Y+1M
2020	65Y	65Y	2030	68Y+1M	68Y+1M	2040	69Y+1M	70Y+1M	2050	69Y+1M	70Y+1M	2060	69Y+1M	71Y+1M
2021	65Y	65Y	2031	68Y+1M	68Y+1M	2041	69Y+1M	70Y+1M	2051	69Y+1M	70Y+1M	2061	69Y+1M	71Y+1M
2022	65Y	65Y	2032	68Y+1M	68Y+1M	2042	69Y+1M	70Y+1M	2052	69Y+1M	70Y+1M	2062	69Y+1M	71Y+1M
2023	65Y+12M	65Y+12M	2033	68Y+1M	68Y+1M	2043	69Y+1M	70Y+1M	2053	69Y+1M	70Y+1M	2063	69Y+1M	71Y+1M
2024	66Y+12M	66Y+12M	2034	68Y+1M	68Y+1M	2044	69Y+1M	70Y+1M	2054	69Y+1M	70Y+1M	2064	69Y+1M	71Y+1M
2025	67Y+1M	67Y+1M	2035	68Y+1M	68Y+1M	2045	69Y+1M	70Y+1M	2055	69Y+1M	70Y+1M	2065	69Y+1M	72Y+1M
2026	67Y+1M	67Y+1M	2036	69Y+1M	69Y+1M	2046	69Y+1M	70Y+1M	2056	69Y+1M	70Y+1M	2066	69Y+1M	73Y+1M
2027	67Y+1M	67Y+1M	2037	69Y+1M	69Y+1M	2047	69Y+1M	70Y+1M	2057	69Y+1M	71Y+1M	2067	69Y+1M	73Y+1M
2028	67Y+1M	67Y+1M	2038	69Y+1M	69Y+1M	2048	69Y+1M	70Y+1M	2058	69Y+1M	71Y+1M	2068	69Y+1M	74Y+1M

use, because we are 80% certain that the corresponding OADR will deviate from the desired level of OADR.

The algorithm to find the upper and lower boundary of plausible pension age schemes is similar to the algorithm to find the target pension age scheme, except that the mean value of OADR is replaced by the boundary of the 80% prediction interval. Mean is replaced by the upper (lower) limit of the 80% prediction interval to find the upper (lower) boundary for pension age schemes.

4.1 Plausible pension age schemes

Table 1 shows the target pension age scheme, which is properly defined in section 3.2.2. It is assumed that it is impractical to change the pension age during 2011–2018, so the pension age has to be set at 65 for those years. Figure 7 provides a visible comparison of the desired OADR level of 21% and the mean forecast of the OADR corresponding to the target pension age. Note that the OADR associated with target pension age scheme has nearly overlapped the desired OADR level, which indicates that we have successfully found the pension age with stable OADR at the desired levels.

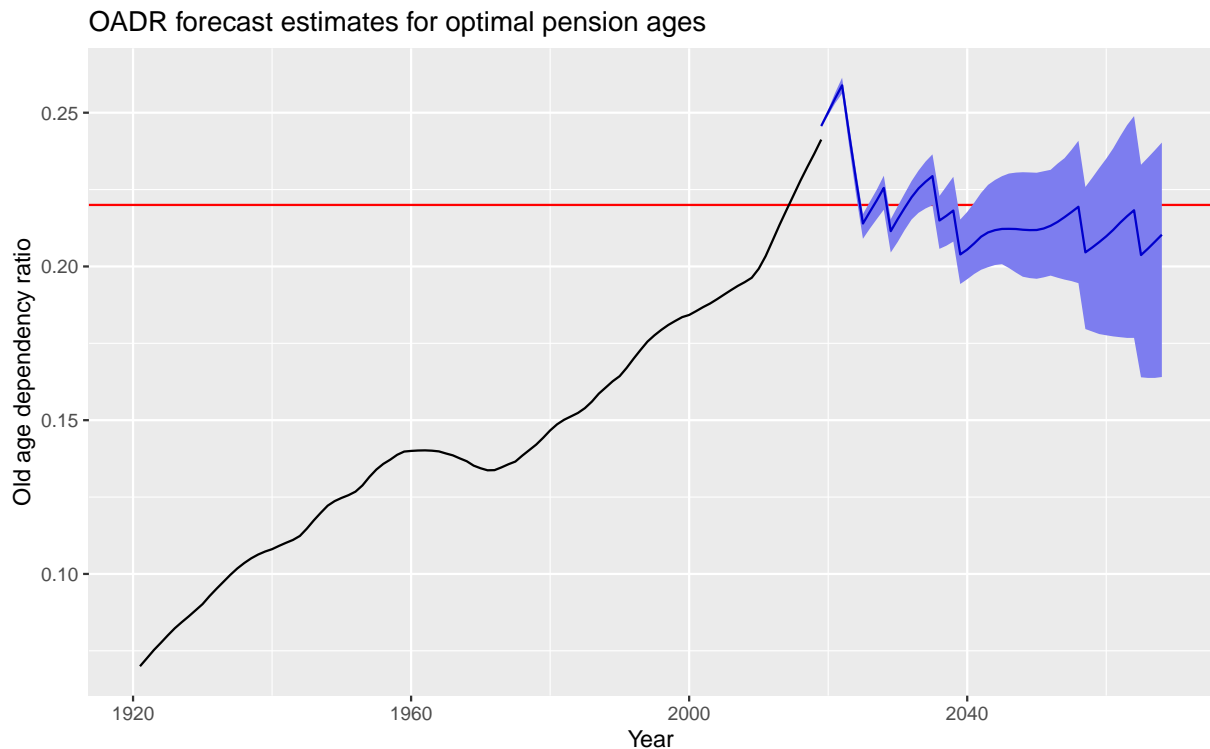


Figure 7: The mean forecast of OADR associated with target pension age scheme (listed in Table 1) shown in blue, along with the desired OADR level in red (i.e., mean forecast of OADR in 2018). The 80% prediction interval of the forecast of OADR is shown as blue shading.

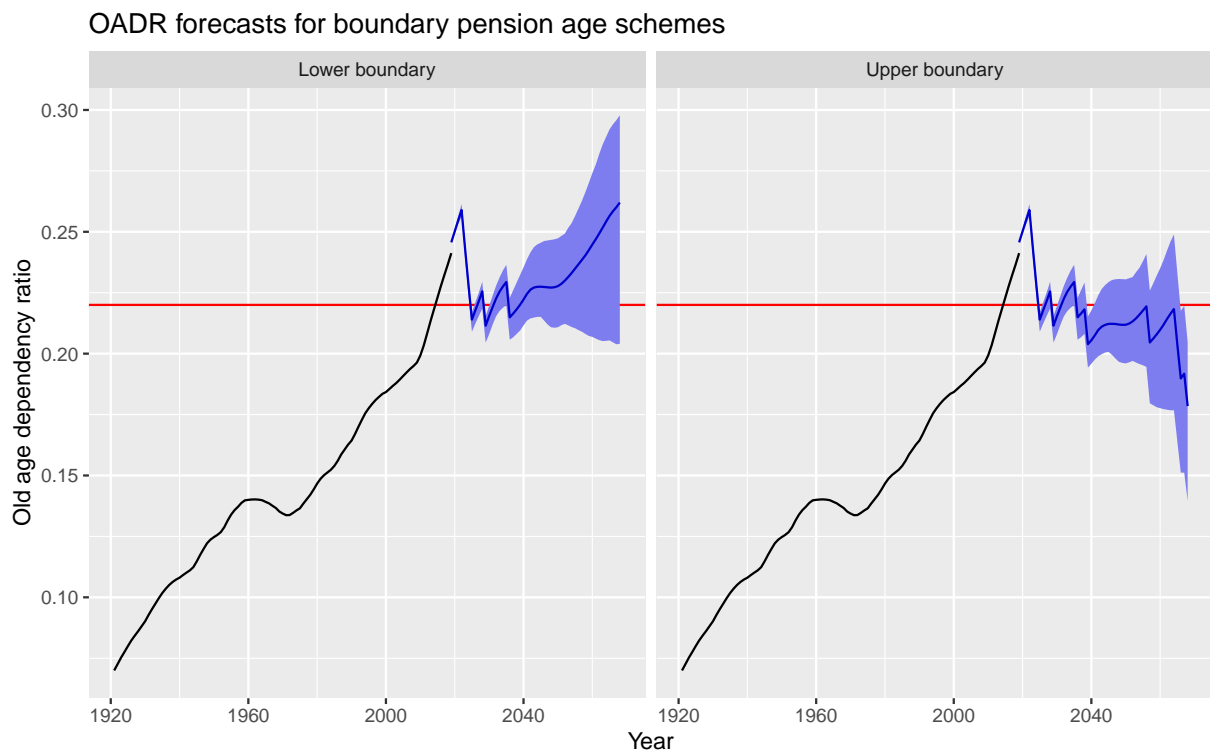


Figure 8: The 80% prediction interval forecast of OADR (in blue) associated with the upper and lower boundary of plausible pension age schemes (listed in Table 2), along with desired OADR level (in red).

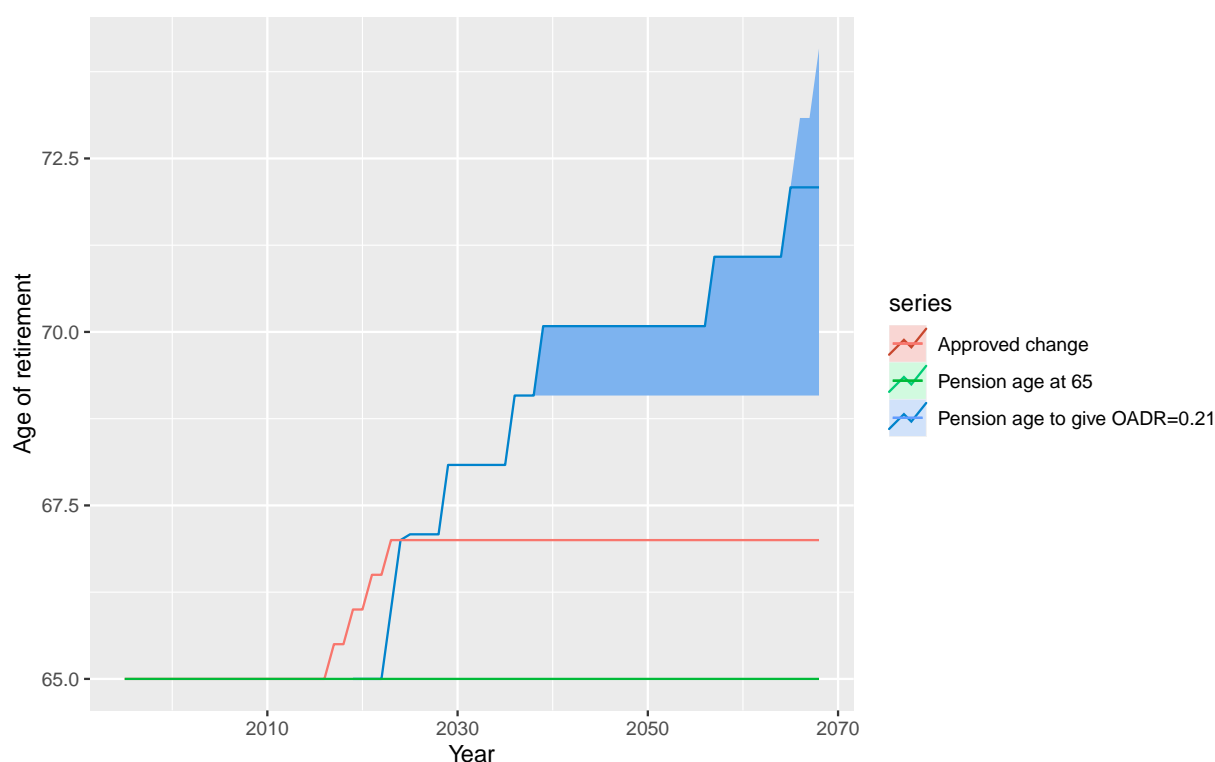


Figure 9: The comparison between the target pension age scheme (in green) and real pension age scheme that has been set or proposed by the government, along with the 80% confidence interval of the target pension age shown in the green shade. The solid red line indicates the changes to the pension age from 65 to 67, announced in 2009.

The upper and lower boundaries of the plausible pension age schemes (also called the 80 % confidence interval of the target pension age scheme in this paper) are reported in Table 2, and their corresponding OADR simulation results are shown in Figure 8. It can be seen that the lower (upper) limit of the 80% prediction interval of OADR associated with the lower (upper) boundary of the 80% confidence interval pension age scheme is entirely stable around the desired level. Note the negative relationship between OADR and pension age. It suggests, on the one hand, that any pension age higher than the upper boundary in Table 2 is deemed to be undesirable because the age is too high to allow the age pension to provide enough welfare support or insurance to the aged. On the other hand, any pension age scheme lower than the lower boundary of the pension age reported in Table 2 will also be regarded as unreasonable, because it sacrifices the interests of the next generation to provide a generous age pension immediately.

Figure 9 compares the target pension age scheme and its 80% confidence interval with the pension age scheme with proposed further change. It was found that the target pension age increases at a higher rate than the proposed new pension age scheme. Reported in Table 1, the target pension age is supposed to reach 67 by 2021 and 70 by 2032, while pension age has been

proposed to change to 67 by 2023 and 70 by 2035. Moreover, the proposed pension age scheme is below the lower boundary of plausible pension age schemes, which means we are 80% sure that it will cause an OADR higher than the desired level. From this result, it is concluded that the government will need to adjust pension age slightly quicker than the proposed pace to ensure we have a stable OADR around the desired level at 21%.

Note that the proposed pension age scheme does not diverge too far from the plausible pension age schemes, suggesting proposed new pension age scheme may be appropriate if we can afford slightly higher OADR. Therefore, it is worth performing a sensitivity test on the desired level here.

Using the approach suggested in this paper, the target pension age associated with the desired OADR equal to current level plus 1% has been found. We have shown the comparison of it and its 80% confidence level with a real pension age scheme in Figure ???. When we allow 1% more in desired OADR, the real pension age becomes plausible. Figure @ref(fig:OADR_realpensionage) shows simulated OADR associated with proposed pension age scheme⁹. It confirms that OADR is stable around a level higher than current level of OADR before 2035. This results suggest that pension age scheme with proposed change could be suitable if the government are targeting a slightly higher financial pressure than current level.

code
are
missing
for two
figures

5 Conclusions

The ageing of the population is expected to become a critical issue in Australia due to the increase in life expectancy and the retirement of baby boomers. The fiscal cost of the age pension is projected to sharply increase as a result. For the purpose of sustainability of the welfare system in Australia, the government has announced an increase in the pension age to 67 in 2023 and 70 in 2035 to reduce the financial burden on society.

In this paper, we have suggested a statistical and demographic approach to address the issue. We attempt to quantify the effects of changes in the pension age on financial burdens using an OADR as a measure of the financial burden on Australian society. The quantification process involves applying stochastic models to forecast future population age structure and then using those future age structure projections to find the most appropriate pension age scheme that will ensure stable OADR around the desired level. To forecast the population age structure, three demographic components of the population were modeled separately using functional

⁹Since we don't know the change rate set by government for 2023–2035, we assume it follows increase rate for 2017–2023, which increase half year pension age in every two years.

time series method (Hyndman & Ullah 2007), with an extension of the product-ratio method (Hyndman, Booth & Yasmeen 2013). Next, future age structures were used to simulate OADR associated with different pension age schemes until the minimum pension age scheme was found, which ensures the OADR no greater than the desired level. This approach is proposed to find a target pension age scheme based on a given desired level of OADR.

We undertake the approach based on assuming the desired level of OADR is at the current level of 21%. Using the historical data from 1950 to 2018, the forecast of future populations suggests a substantial increase in total population, and confirms the inherent ageing trend. It was found that pension age should not be fixed, but instead should grow at a rate that will ensure a stable OADR at a level which can be affordable to Australians of working age. Findings indicate that the pension age scheme with proposed change was not too high, as it is still lower than the target pension age scheme and outside its confidence interval. It was concluded that the pension age set by the government should be adjusted slight quicker to be in line with our target pension age, which ensures a stable OADR of around 21%.

The empirical analysis presented in this paper shows the adjustment of the pension age announced in 2009 was effective in reducing the financial burden to the desired level, which was set at 21%. However, a one-time adjustment is not sufficient to ensure a stable OADR in the long term, due to ongoing changes in the population age structure. Therefore, we recommend consistently reviewing the pension age according to the presented statistical approach. This approach has been used to calculate the target pension age from now until 2068 and its 80% confidence interval. The target pension age found in this research increases to 67 by 2023, 70 by 2032 and 75 by 2058. This generates a stable OADR of around 21%. The rate of growth in the pension age as proposed by the government in 2014 is slightly lower than the target pension age found in this study, which suggests that the government may need to slightly accelerate the growth of the pension age to maintain an OADR at the current level.

Although our results are promising, there are still grounds for further research. First, our research results are based on forecasts from one particular model. Further studies can apply the same approach but with different models of population forecasting, which allows for enhanced credibility of results. Moreover, our research addresses the issue from a purely demographic perspective. It is possible to include an economic perspective. For example, the economic implications of retaining older workers in the workforce can be considered. Second, Our research uses an OADR as a single measurement of the financial burden. Further research should allow the effect of other economic factors to be taken into account. For instance, the

effect of growth in superannuation can be considered in a way that the effects of high OADR can be offset by the growth in superannuation.

Colophon

This report was generated on 2020-09-10 18:53:15.

The current Git commit details are:

```
## Local:      master /home/robjhyndman/git/Research/PensionAge
## Remote:     master @ origin (git@github.com:robjhyndman/PensionAge.git)
## Head:       [6fd19ec] 2020-09-10: proselint and styler
```

We used the following computational environment and dependencies:

```
## - Session info -----
## setting value
## version R version 4.0.2 (2020-06-22)
## os      Ubuntu 20.04.1 LTS
## system  x86_64, linux-gnu
## ui      X11
## language en_AU:en
## collate en_US.UTF-8
## ctype   en_US.UTF-8
## tz      Australia/Melbourne
## date    2020-09-10
##
```

```
## - Packages -----
## ! package          * version  date      lib source
## P abind             1.4-5    2016-07-21 [?] CRAN (R 4.0.0)
## P assertthat        0.2.1    2019-03-21 [?] CRAN (R 4.0.0)
## P backports         1.1.6    2020-04-05 [?] CRAN (R 4.0.0)
## P bookdown          0.18     2020-03-05 [1] CRAN (R 4.0.0)
## P callr             3.4.3    2020-03-28 [?] RSPM (R 4.0.0)
## P cli               2.0.2    2020-02-28 [?] CRAN (R 4.0.0)
## P cluster           2.1.0    2019-06-19 [?] CRAN (R 4.0.0)
## P cobs              1.3-4    2020-01-24 [?] CRAN (R 4.0.0)
## P colorspace        1.4-1    2019-03-18 [?] CRAN (R 4.0.2)
## P crayon            1.3.4    2017-09-16 [?] CRAN (R 4.0.2)
## P curl              4.3      2019-12-02 [?] CRAN (R 4.0.0)
## P demography        * 1.22     2019-04-22 [?] CRAN (R 4.0.0)
```

##	P desc	1.2.0	2018-05-01	[?]	CRAN	(R 4.0.0)
##	devtools	2.3.0	2020-04-10	[1]	CRAN	(R 4.0.0)
##	P digest	0.6.25	2020-02-23	[?]	CRAN	(R 4.0.0)
##	P dplyr	* 0.8.5	2020-03-07	[?]	CRAN	(R 4.0.0)
##	P ellipsis	0.3.0	2019-09-20	[?]	CRAN	(R 4.0.0)
##	P evaluate	0.14	2019-05-28	[?]	CRAN	(R 4.0.0)
##	P fansi	0.4.1	2020-01-08	[?]	CRAN	(R 4.0.0)
##	P fda	5.1.4	2020-04-20	[?]	CRAN	(R 4.0.0)
##	P forecast	* 8.12	2020-03-31	[?]	CRAN	(R 4.0.2)
##	P fracdiff	1.5-1	2020-01-24	[?]	CRAN	(R 4.0.0)
##	P fs	1.4.1	2020-04-04	[?]	CRAN	(R 4.0.0)
##	P ftsa	5.6	2020-04-23	[?]	CRAN	(R 4.0.0)
##	P geometry	0.4.5	2019-12-04	[?]	CRAN	(R 4.0.0)
##	P ggplot2	* 3.3.0	2020-03-05	[?]	CRAN	(R 4.0.0)
##	P glue	1.4.0	2020-04-03	[?]	CRAN	(R 4.0.0)
##	P gtable	0.3.0	2019-03-25	[?]	CRAN	(R 4.0.0)
##	P hdrcke	3.3	2018-12-21	[?]	CRAN	(R 4.0.0)
##	P htmltools	0.4.0	2019-10-04	[?]	CRAN	(R 4.0.0)
##	P KernSmooth	2.23-17	2020-04-26	[?]	CRAN	(R 4.0.0)
##	P knitr	1.28	2020-02-06	[?]	CRAN	(R 4.0.0)
##	P ks	1.11.7	2020-02-11	[?]	CRAN	(R 4.0.0)
##	P lattice	0.20-41	2020-04-02	[?]	CRAN	(R 4.0.2)
##	P lifecycle	0.2.0	2020-03-06	[?]	CRAN	(R 4.0.0)
##	P lmtest	0.9-37	2019-04-30	[?]	CRAN	(R 4.0.0)
##	P magic	1.5-9	2018-09-17	[?]	CRAN	(R 4.0.0)
##	P magrittr	1.5	2014-11-22	[?]	CRAN	(R 4.0.0)
##	P MASS	7.3-51.6	2020-04-26	[?]	CRAN	(R 4.0.0)
##	P Matrix	1.2-18	2019-11-27	[?]	CRAN	(R 4.0.0)
##	P MatrixModels	0.4-1	2015-08-22	[?]	CRAN	(R 4.0.0)
##	P mclust	5.4.6	2020-04-11	[?]	CRAN	(R 4.0.0)
##	memoise	1.1.0	2017-04-21	[1]	CRAN	(R 4.0.0)
##	P mgcv	1.8-31	2019-11-09	[?]	CRAN	(R 4.0.0)
##	MonashEBSTemplates	0.2	2020-05-04	[1]	Github	(robjhyndman/MonashEBSTemplates)
##	P munsell	0.5.0	2018-06-12	[?]	CRAN	(R 4.0.0)

##	P mvtnorm	1.1-0	2020-02-24	[?]	RSPM	(R 4.0.0)
##	P nlme	3.1-147	2020-04-13	[?]	RSPM	(R 4.0.0)
##	P nnet	7.3-14	2020-04-26	[?]	CRAN	(R 4.0.0)
##	P patchwork	* 1.0.0	2019-12-01	[?]	CRAN	(R 4.0.0)
##	P pcaPP	1.9-73	2018-01-14	[?]	CRAN	(R 4.0.0)
##	P pdfCluster	1.0-3	2018-12-04	[?]	CRAN	(R 4.0.0)
##	P pillar	1.4.3	2019-12-20	[?]	CRAN	(R 4.0.0)
##	P pkgbuild	1.0.7	2020-04-25	[?]	CRAN	(R 4.0.0)
##	P pkgconfig	2.0.3	2019-09-22	[?]	CRAN	(R 4.0.0)
##	P pkgload	1.0.2	2018-10-29	[?]	CRAN	(R 4.0.0)
##	P prettyunits	1.1.1	2020-01-24	[?]	CRAN	(R 4.0.0)
##	P processx	3.4.2	2020-02-09	[?]	CRAN	(R 4.0.0)
##	P ps	1.3.2	2020-02-13	[?]	CRAN	(R 4.0.0)
##	P purrr	0.3.4	2020-04-17	[?]	RSPM	(R 4.0.0)
##	P quadprog	1.5-8	2019-11-20	[?]	CRAN	(R 4.0.0)
##	P quantmod	0.4.17	2020-03-31	[?]	RSPM	(R 4.0.0)
##	P quantreg	5.55	2020-04-01	[?]	CRAN	(R 4.0.0)
##	P R6	2.4.1	2019-11-12	[?]	CRAN	(R 4.0.0)
##	P rainbow	3.6	2019-01-29	[?]	CRAN	(R 4.0.0)
##	P Rcpp	1.0.4.6	2020-04-09	[?]	CRAN	(R 4.0.0)
##	P remotes	2.1.1	2020-02-15	[?]	CRAN	(R 4.0.0)
##	renv	0.9.3	2020-02-10	[1]	CRAN	(R 4.0.0)
##	rlang	0.4.6	2020-05-02	[1]	CRAN	(R 4.0.0)
##	P rmarkdown	2.1	2020-01-20	[?]	CRAN	(R 4.0.0)
##	P rprojroot	1.3-2	2018-01-03	[?]	CRAN	(R 4.0.0)
##	P scales	1.1.0	2019-11-18	[?]	CRAN	(R 4.0.0)
##	P sde	2.0.15	2016-04-13	[?]	CRAN	(R 4.0.0)
##	sessioninfo	1.1.1	2018-11-05	[1]	CRAN	(R 4.0.0)
##	P SparseM	1.78	2019-12-13	[?]	RSPM	(R 4.0.0)
##	P stringi	1.4.6	2020-02-17	[?]	CRAN	(R 4.0.0)
##	P stringr	1.4.0	2019-02-10	[?]	CRAN	(R 4.0.0)
##	P testthat	2.3.2	2020-03-02	[?]	CRAN	(R 4.0.2)
##	P tibble	3.0.1	2020-04-20	[?]	CRAN	(R 4.0.0)
##	P tidyr	* 1.0.2	2020-01-24	[?]	CRAN	(R 4.0.0)

```
## P tidyselect      1.0.0    2020-01-27 [?] CRAN (R 4.0.0)
## P timeDate        3043.102 2018-02-21 [?] CRAN (R 4.0.0)
## P tseries         0.10-47  2019-06-05 [?] CRAN (R 4.0.0)
## P TTR             0.23-6   2019-12-15 [?] CRAN (R 4.0.0)
## P urca            1.3-0    2016-09-06 [?] CRAN (R 4.0.0)
##   usethis         1.6.1    2020-04-29 [1] CRAN (R 4.0.0)
## P vctrs           0.2.4    2020-03-10 [?] CRAN (R 4.0.0)
## P withr           2.2.0    2020-04-20 [?] CRAN (R 4.0.0)
## P xfun            0.13     2020-04-13 [?] CRAN (R 4.0.0)
## P xts             0.12-0   2020-01-19 [?] CRAN (R 4.0.0)
## P yaml            2.2.1    2020-02-01 [?] CRAN (R 4.0.0)
## P zoo             1.8-7    2020-01-10 [?] CRAN (R 4.0.0)
##
## [1] /home/robjhyndman/git/Research/PensionAge/renv/library/R-4.0/x86_64-pc-linux-gnu
## [2] /tmp/RtmpL12d86/renv-system-library
##
## P -- Loaded and on-disk path mismatch.
```

References

- Apps, P, R Rees & M Wood (2007). Population ageing, taxation, pensions and health costs. *Australian Journal of Labour Economics* **10**(2), 79–97.
- ASFA (2007). *Pre-budget Submission for 2007–2008 Federal Budget*. Sydney: Association of Superannuation Funds of Australia.
- Australian Bureau of Statistics (2006). *Population projections, Australia: 2004–2101*. Canberra: AGPS.
- Australian Bureau of Statistics (2012). *Population projections, Australia: 2012–2101*. Canberra: AGPS.
- Booth, H, RJ Hyndman & L Tickle (2014). “Prospective Life Tables”. In: *Computational Actuarial Science with R*. Ed. by A Charpentier. Chapman & Hall/CRC, pp.323–348.
- Costello, P (2002). *Intergenerational Report 2002-2003*. Canberra: Commonwealth of Australia. <http://www.budget.gov.au/2002-03/bp5/html/>.
- Costello, P (2004). *Australia's demographic challenges*. Canberra: Commonwealth of Australia Treasury Paper.
- Costello, P (2007). *Intergenerational Report 2007*. Canberra: Commonwealth of Australia. <http://www.treasury.gov.au/PublicationsAndMedia/Publications/2007/Intergenerational-Report-2007>.
- Fehr, H, S Jokisch & LJ Kotlikoff (2008). Fertility, mortality and the developed world's demographic transition. *Journal of Policy Modeling* **30**(3), 455–473.
- Good, IJ (1969). Some applications of the singular decomposition of a matrix. *Technometrics* **11**(4), 823–831.
- Guest, R (2014). *The argument for changing the age pension doesn't stack up*. <http://theconversation.com/the-argument-for-changing-the-age-pension-doesnt-stack-up-26046>. Accessed: 2014-5-11. <http://theconversation.com/the-argument-for-changing-the-age-pension-doesnt-stack-up-26046>.
- Hu, N & Y Yang (2012). The Real Old-Age Dependency Ratio and the Inadequacy of Public Pension Finance in China. *Journal of Population Ageing* **5**(3), 193–209.
- Human Mortality Database (2020). *University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)*. Available at www.mortality.org or www.humanmortality.de (data downloaded on 12/June/2017).
- Hyndman, RJ & G Athanasopoulos (2014). *Forecasting: principles and practice*. Melbourne, Australia: OTexts. OTexts.org/fpp.

- Hyndman, RJ & H Booth (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting* **24**(3), 323–342.
- Hyndman, RJ, H Booth & F Yasmeen (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography* **50**(1), 261–283.
- Hyndman, RJ & Y Khandakar (2008). Automatic time series forecasting : the forecast package for R. *Journal of Statistical Software* **26**(3).
- Hyndman, RJ & S Ullah (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics & Data Analysis* **51**(10), 4942–4956.
- Knox, D (2007). *Pensions for longer life: Linking Australia's pension age with life expectancy*. CEDA Information Paper 89. Committee for Economic Development of Australia. <https://trove.nla.gov.au/work/35010227?selectedversion=NBD42351550>.
- Lee, RD & LR Carter (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* **87**(419), 659–671.
- National Commission of Audit (2014). *Towards Responsible Government (Phase One)*. Canberra: Commonwealth of Australia. http://www.ncoa.gov.au/report/docs/phase_one_report.pdf.
- Nielson, L (2010). *Chronology of superannuation and retirement income in Australia*. Canberra: Commonwealth of Australia. http://www.aph.gov.au/binaries/library/pubs/bn/eco/chron_superannuation.pdf.
- OECD (2012). *OECD Pensions Outlook 2012*. Paris: Organisation for Economic Co-operation and Development, OECD Publishing.
- Petrichev, K & S Thorp (2008). The private value of public pensions. *Insurance: Mathematics and Economics* **42**, 1138–1145.
- Power, T (2014). Alert! Commission of audit targets age pension age, super access, CSHC, PBS and more. *SuperGuide*. www.superguide.com.au.
- Productivity Commission (2013). *An ageing Australia: preparing for the future*. Commission Research Paper. Canberra, Australia. <https://www.pc.gov.au/research/completed/ageing-australia>.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Shang, HL, H Booth & RJ Hyndman (2011). Point and interval forecasts of mortality rates and life expectancy : a comparison of ten principal component methods. *Demographic Research* **25**(5), 173–214.

Swan, W (2010). *Australia to 2050: Future challenges*. Canberra: Commonwealth of Australia.

<http://archive.treasury.gov.au/igr/igr2010/>.

Walker, A (1982). Dependency and old age. *Social Policy & Administration* **16**(2), 115–135.