

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](mailto:Rob.Hyndman@monash.edu)

Corresponding author

**Yijun Zeng**

Monash University

**Han Lin Shang**

Macquarie University

17 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. Our results suggest that the Australian pension age should be increased to 68 by 2028 and 69 by 2034 in order to maintain the old-age dependency ratio at the 2019 level, which is 22%. This approach can easily be extended to other target levels of the old-aged dependency ratio and to 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 sex-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 sex-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 ?? 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 4.1. 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 pharmaceuticals<sup>1</sup>. The means test ensures that retirees with enough personal savings will

---

<sup>1</sup>The 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 die<sup>2</sup>. 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

---

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



**Figure 1:** *Australian pension age policies.*

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 Demographic growth-balance equations

The data used are age- and sex-specific birth and death numbers of each calendar year, age- and sex-specific population members on 1 January of each year, and age- and sex-specific exposures to risk (i.e., the population of age  $x$  at 30 June) for each year. Let

$B_t(x)$  = 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 year is denoted by  $t = 1, 2, \dots, T$ , 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 migration<sup>3</sup> (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:

$$\begin{aligned} G_t(x, x+1) &= P_{t+1}(x+1) - P_t(x) + D_t(x, x+1) \quad \text{for } x = 0, 1, 2, \dots, p-2, \\ G_t(p-1^+, p^+) &= P_{t+1}(p^+) - P_t(p^+) - P_t(p-1) + D_t(p-1^+, p^+), \\ G_t(B, 0) &= P_{t+1}(0) - B_t + D_t(B, 0), \end{aligned}$$

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.

### 2.3 Old-Age Dependency Ratio

Using this notation, the OADR in year  $t$  is given by

$$O_t = \frac{\sum_{x=15}^{\lfloor a_t \rfloor - 1} P_t(x) + r_t P_t(\lfloor a_t \rfloor)}{\sum_{x \geq \lfloor a_t \rfloor} P_t(x) - r_t P_t(\lfloor a_t \rfloor)}, \quad (2)$$

where  $a_t$  is the pension age in year  $t$  and  $r_t = a_t - \lfloor a_t \rfloor$  is the non-integer part of  $a_t$ . This assumes that the pension age can only be adjusted at the beginning of a year, and that the birthday of the population is uniformly distributed over the year.

### 2.4 Forecasting demographic components

Variants of the Lee-Carter method (Lee & Carter 1992) are frequently used for demographic forecasting (see, e.g., Shang, Booth & Hyndman 2011). 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 Hyndman-Ullah method (2007) is a generalization of the Lee-Carter method and is used as the basis for forecasting demographic components in this study. In addition, the product-ratio method (Hyndman, Booth & Yasmeen 2013) is used

---

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

to eliminate the potential long-term divergence between female and male mortality rates and net-migration rates.

### Lee-Carter method

The Lee-Carter (LC) method has been widely used to forecast mortality rates across various countries, including Australia. The model proposed by Lee & Carter (1992) (but using our notation) is

$$\log m_t(x) = \mu(x) + \phi(x)\beta_t + \varepsilon_{t,x}, \quad (3)$$

where  $m_t(x)$  is the central death rate for age  $x$  in year  $t$ . This model involves fitting a matrix of log central death rates with sets of age-specific constants,  $\{\mu(x)\}$  and  $\{\phi(x)\}$ , and a time-varying index  $\beta_t$ .  $\phi(x)$  is a set of parameters that tell us the rate at which each age responds to changes in  $\beta_t$ .  $\beta_t$  measures the general level of the log death rates;  $\mu(x)$  is the general pattern across the age of the log mortality rate and  $\varepsilon_{t,x}$  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. Therefore, the constraints

$$\sum_{t=1}^T \beta_t = 0, \quad \text{and} \quad \int_x \phi(x) = 1 \quad (4)$$

are imposed, and a singular value decomposition is used to find the least squares solution of (3) (Good 1969). The estimates of  $\phi(x)$  and  $\beta_t$  are therefore the first principal component and scores of  $X$  with  $(i, t)$ th element  $[\log(m_t(x)) - \hat{\mu}(x)]$ , where  $\hat{\mu}(x)$  is the sample mean of  $\log(m_t(x))$  over  $t$ . The log transformation introduces some bias, so the LC method adjusts  $\beta_t$  such that for each year, the implied number of deaths will be equal to the actual number of deaths. The adjusted  $\beta_t$  is then extrapolated using a random walk with drift model.

### 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 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 Hyndman-Ullah (HU) model can be written as

$$y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \quad (5)$$

$$s_t(x) = \mu(x) + \sum_{j=1}^J \beta_{t,j}\phi_j(x) + e_t(x), \quad (6)$$

where  $y_t(x)$  may be log mortality, log fertility or net migration (Hyndman & Booth 2008) in year  $t$  and age  $x$ . Here  $s_t(x)$  is a smoothed version of the observed data  $y_t(x)$ , smoothed over age but not time, and estimated by constrained weighted penalized regression splines; and  $\{\varepsilon_{t,x}\}$  are independent and identically distributed standard normal error terms. Equation (6) has a similar structure to the LC method, except that it is applied to smoothed data and incorporates more than one principal component and score. The  $e_t(x)$  is the model error having mean zero and no serial correlation, while  $\mu(x)$  is estimated as the average of  $s_t(x)$  across years. Estimates of  $\phi_j(x)$  and  $\beta_{t,j}$  are the  $j^{th}$  principal components and scores of  $s_t(x) - \mu(x)$  respectively. Following Hyndman & Booth (2008), we select  $J = 6$  which we have found to be large enough for demographic applications.

By 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  $y_{T+h}(x)$  are given by

$$\hat{y}_{T+h|T}(x) = \hat{s}_{T+h|T}(x) = E[\hat{s}_{T+h|T}(x) | \mathbf{I}, \mathbf{B}] = \hat{\mu}(x) + \sum_{j=1}^J \phi_j(x) \hat{\beta}_{T+h|T,j}, \quad (7)$$

where  $\hat{\beta}_{T+h|T,j}$  denotes the  $h$ -step-ahead forecast of  $\beta_{T+h|T,j}$ . 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 selected using the automated algorithm of Hyndman & Khandakar (2008), and the parameters are estimated using maximum likelihood estimation.

### Product-ratio method

When we apply the Hyndman & Ullah (2007) model to data on males and females, we need to constrain the resulting forecasts so they do not diverge. In the product-ratio method (Hyndman, Booth & Yasmeeen 2013), the HU model is applied to the geometric mean of male and female log mortality data, and the ratio of male and female log mortality data, rather than to each series independently. That is, define

$$p_t(x) = \sqrt{s_t^M(x)s_t^F(x)}, \quad (8)$$

$$r_t(x) = \sqrt{\frac{s_t^M(x)}{s_t^F(x)}}. \quad (9)$$

Then, the HU method is applied to these two new variables:

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

where time series  $\{\beta_{t,j}\}$  are fitted with ARIMA models while  $\{\gamma_{t,l}\}$  are fitted by stationary ARFIMA models (i.e., the order of differencing is constrained to  $0 < d < \frac{1}{2}$ ). It is the stationary constraint on  $\gamma_{t,l}$  that ensures the forecasts of male and female do not diverge. Then, forecasts of sex-specific data are obtained using

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

The method for forecasting net migration by sex is similar to the above method used on mortality rates except no log scale is applied to net migration, and a sum and difference is used rather than a product and ratio.

## 2.5 Forecasting OADR

Following the general procedure proposed in Hyndman & Booth (2008), we model each of the five components of population change: female mortality rates, male mortality rates, female net migration, male net migration and fertility rates, and simulate these five components into the future. We use the HU method applied to log-fertility, and the product-ratio method applied to log-mortality and net-migration.

From these models, sample paths are simulated for age-specific fertility rates, age- and sex-specific mortality rates and age- and sex-specific net migration. Then future population values are simulated using the demographic growth-balance equations. 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. We bootstrap the residuals of the net-migration model to simulate future net-migration numbers (see Hyndman & Booth (2008) for the details). Hence, we can simulate  $h$ -step-ahead age- and sex-specific population numbers based on simulated births, deaths and net-migrants.

This simulation was repeated for years  $t = T + 1, \dots, T + h$  to obtain a sample path of the age-specific population. Many simulated population paths were obtained in this way. For each simulated population path, and each proposed set of pension age changes, we can simulate future OADR values using (2).

There are three types of randomness in the simulation: (1) randomness from the second equation in the Hyndman-Ullah specification (i.e.,  $e_t(x)$  in (5)); (2) randomness from the ARIMA model fitted to principal component scores; and (3) randomness from the Poisson distribution for births and deaths, or the bootstrap for migrants.

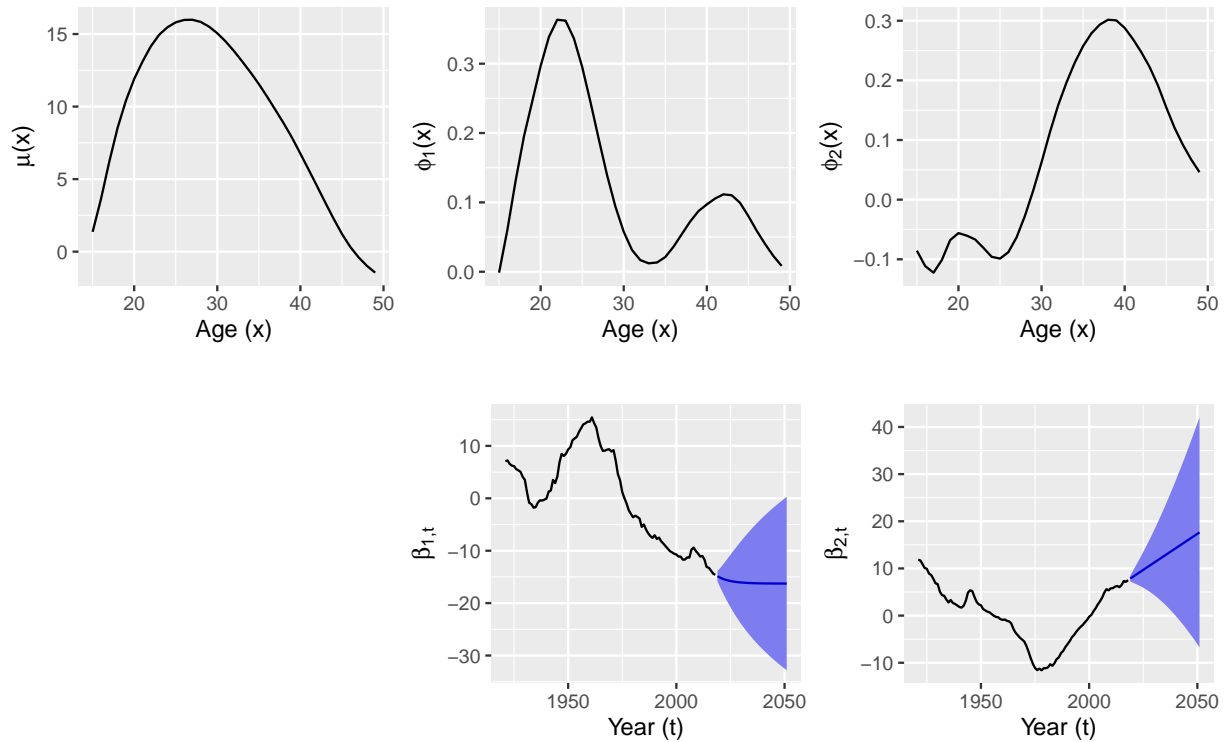
## 3 Results

### 3.1 Data

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 (2020). 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 (Hyndman & Booth 2008). Since the model for this study is valid under the assumption that extrapolative methods can be used, it seemed best 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 rates were also truncated to the same period.

### 3.2 Component forecasts

We illustrate the component model forecasts using fertility rates. Figure 2 shows the estimated mean age pattern  $\hat{\mu}(x)$ , the first two basis functions,  $\hat{\phi}_1(x)$  and  $\hat{\phi}_2(x)$ , and the first two principal component scores,  $\hat{\beta}_{t,1}$  and  $\hat{\beta}_{t,2}$  for log fertility rates. The first term,  $\hat{\phi}_1(x)\hat{\beta}_{t,1}$ , accounts for 63.7% of the variation in the log fertility rates. The first principal component scores  $\hat{\beta}_{t,1}$ , shown in black for 1950–2018, display 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 boom, while the dramatic decrease from 1960 was due to the availability of contraception pills. A noticeable further reduction of fertility rates in the late 1960s may be due to a change in the abortion laws in some states.



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

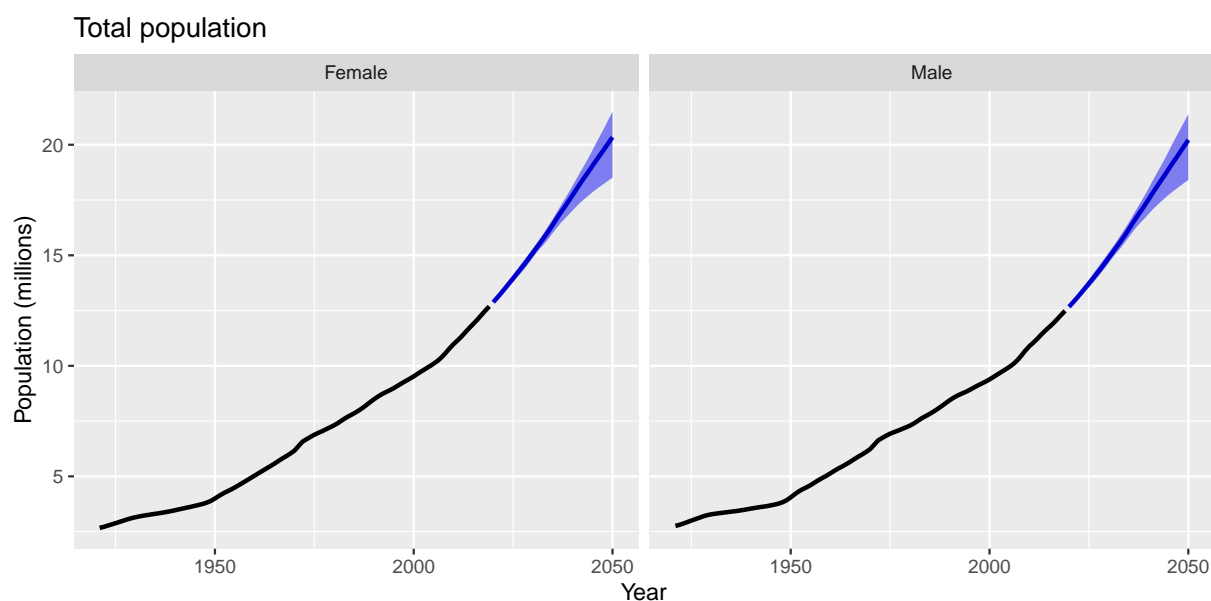
The first basis function  $\hat{\phi}_1(x)$  measures the different effects of the time-trend across ages. It indicates the general pattern of fertility has little effect on females of age 15, most likely because such pregnancies are almost always accidental and usually prior to the mother taking contraception. The two “bumps” in the first basic function suggests that the historical trend has a distinct impact on females aged around their 20s and 40s. The second basis function, accounting for 31.5% of the variation, largely concerns mothers aged over 30. The forecasts of future coefficients are shown as the blue lines in the lower panels with their 80% pointwise prediction intervals shown as blue shaded regions. These forecasts are obtained using automatically selected ARIMA models. In practice, we use six basis functions, but the four not shown have a much smaller effect on the resulting fertility forecasts, only account for 4.6% of the variation, and are much harder to interpret.

Similar models are used for the product and ratio of mortality rates, and sum and different of net migration rates.

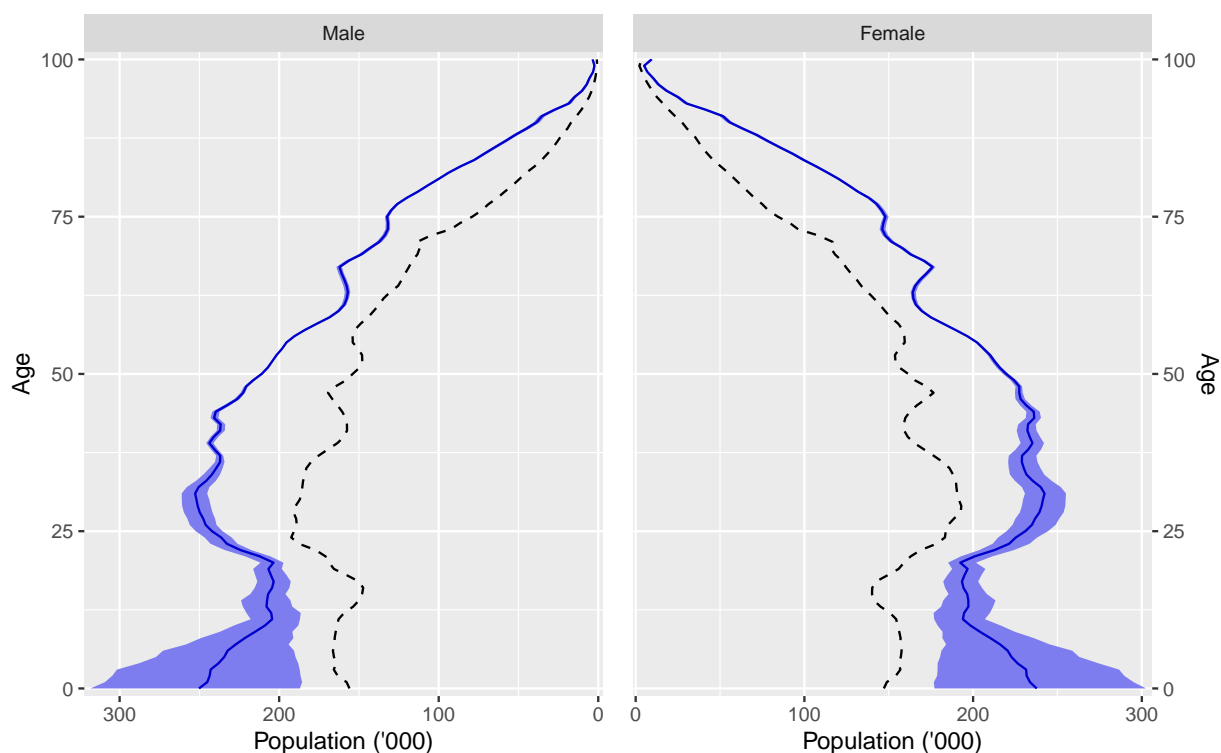
### 3.3 Population forecasts

Figure 3 shows the forecast of the total population by sex over the prediction horizon, while Figure 4 shows the 20-year forecast of population age structure in 2038. Figure 3 indicates that both the female and male population are expected to increase to around 27 million by 2068.

```
## `summarise()` regrouping output by 'Sex' (override with `.groups` argument)
```



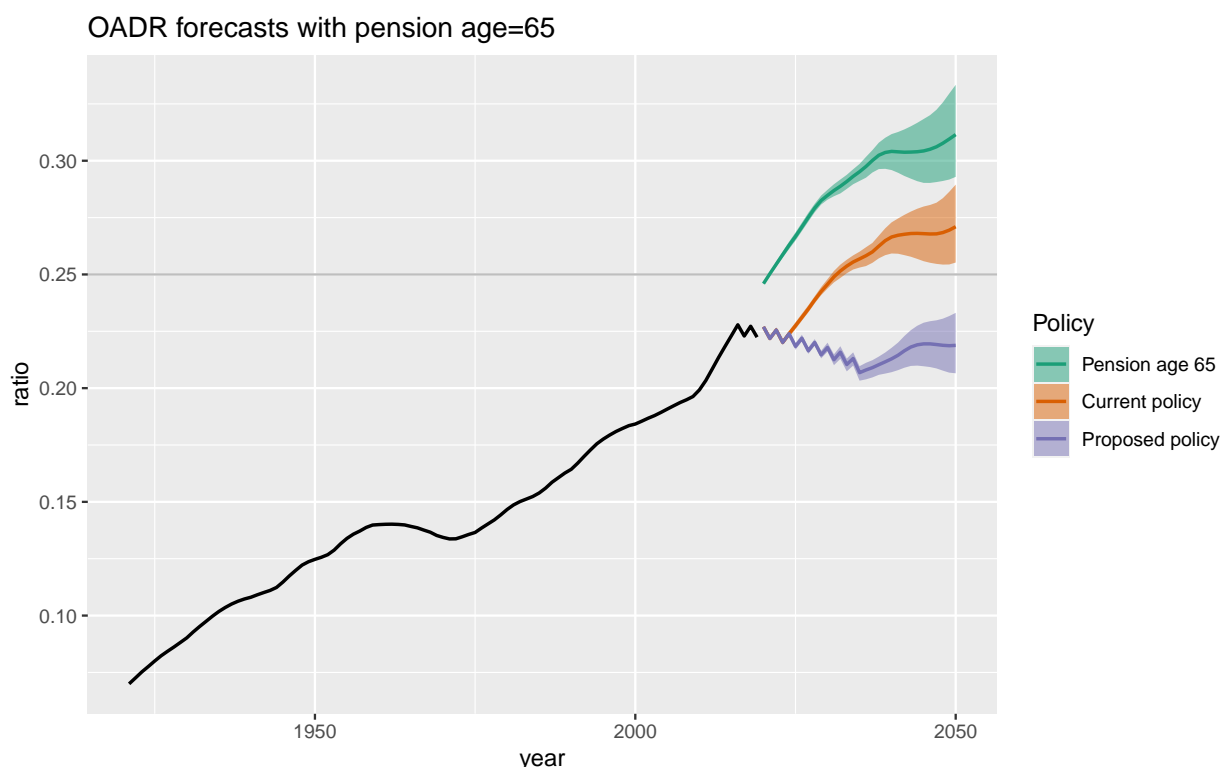
**Figure 3:** Fifty-year forecasts of the total population for each sex, along with 80% prediction intervals.



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

Figure 4 shows the mean and the 80% pointwise prediction interval of simulated population paths for the year 2038 (a 20-year forecast), along with the 2019 base population. The uncertainty of the population below aged 20 is largely due to higher forecast variation of fertility rates compared to other components.

Historical OADR values are shown in black in Figure 5 for 1921–2019. The strong positive trend indicates an increasing tax burden on workers, consistent with the prediction of the intergenerational reports in 2002 (Costello 2002) and 2007 (Costello 2007). The increased trend from 2010 is due to the retirements of “Baby Boomers”, born after WW2. We also show forecasts for three pension schemes: (1) a fixed pension age of 65; (2) the current scheme with the pension age rising to age 67 in 2023; and (3) the proposed pension policy under which the pension ages continues to rise by 6 months every two years until it reaches 70 in 2035.



**Figure 5:** Fifty-year forecasts of the OADR associated with a pension age equal to 65 (green), the current approved pension age scheme (orange), and the proposed scheme with the pension age rising to 70 in 2035 (purple). The actual OADRs for 1921–2019 are shown in black.

Once the pension age reaches 67, the OADR under the current policy grows at nearly the 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. Even the proposal to raise the pension age incrementally to 70 by 2035 leads to a similar increasing trend beyond 2035. Figure 5 demonstrates that the pension age scheme requires constant review to ensure a stable OADR over the long term.



## 4 Sustainable pension age scheme

Our aim is to determine a pension age scheme that would lead to the OADR being below a desired threshold, thus imposing a sustainability constraint.

Let  $a_t$  denote the pension age in year  $t$ , and we denote a pension age scheme over the prediction horizon by  $\mathbf{P} = [a_{T+1}, \dots, a_{T+H}]'$ , with the corresponding OADR values given by  $O_{T+1}, \dots, O_{T+H}$ . Our goal is to find the minimum pension age in each year that gives an OADR lower than the desired OADR threshold. We impose the constraint  $0 \leq a_t - a_{t-1} < 1$  to prevent two undesirable situations: (a) years where no-one is able to retire; and (b) years where retired people become ineligible for the pension. We also assume that the adjustment unit of the pension age is one month, for easier implementation.

### 4.1 Finding the target pension age scheme

Let  $O^*$  denote the desired OADR threshold. Because we do not observe the population in future years, we must estimate  $O_{T+h}$  from the simulated populations. Then we compute the pension age scheme  $\mathbf{P}$  using the following algorithm. Starting with  $h = 1$ :

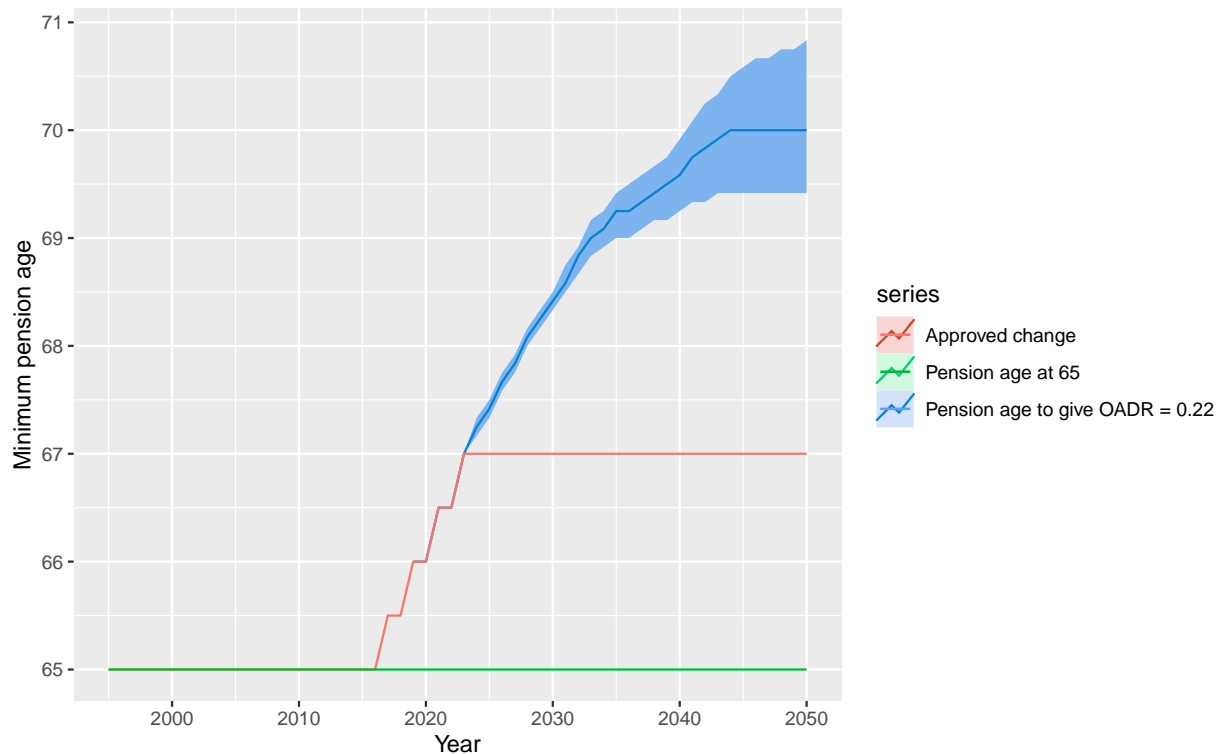
1. Set  $a_{T+h} = a_{T+h-1}$ .
2. Increment  $a_{T+h}$  by one-month intervals until either  $\hat{O}_{T+h|T} > O^*$  or  $a_{T+1} - a_T = 1$  month, where  $\hat{O}_{T+h|T}$  denotes the mean of the simulated  $O_{T+h|T}$  values.

We repeat these steps for  $h = 2, \dots, H$ .

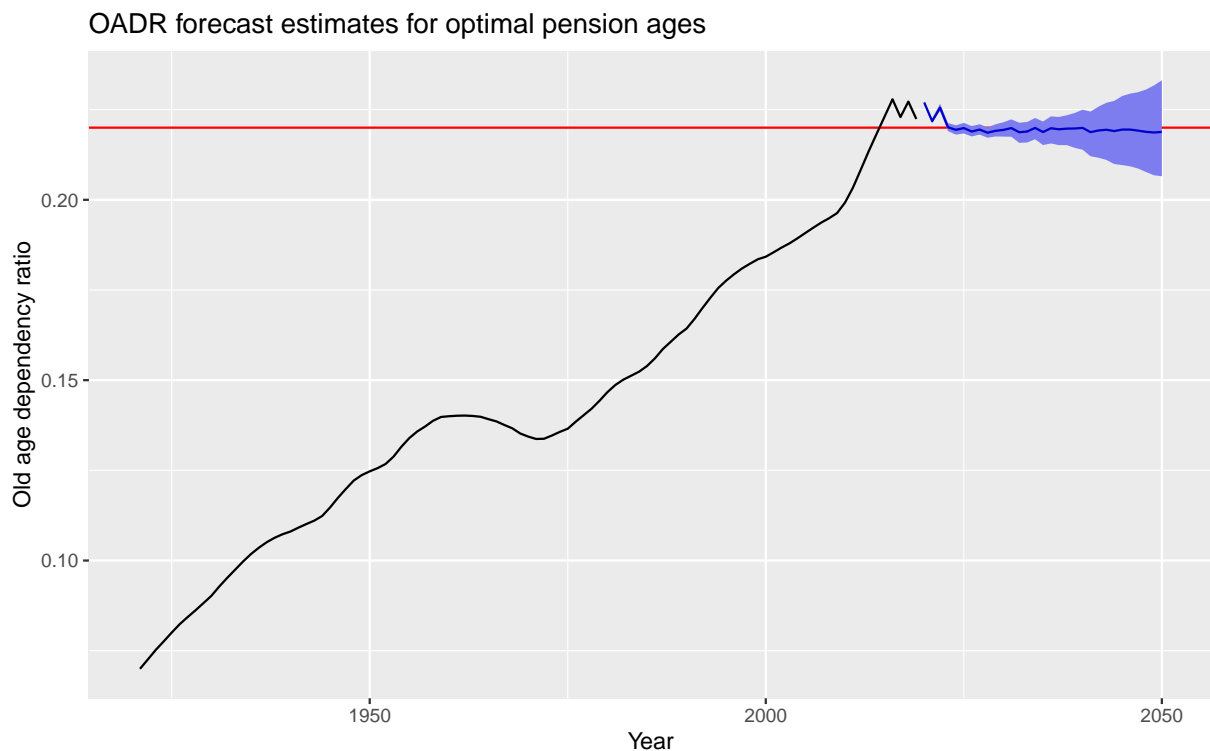
In addition to finding the target pension age based on the mean value of OADR, we also find “plausible” pension age schemes that could give the desired OADR level. Specifically, we find a range of pension age schemes for which  $O^*$  is contained within the 80% prediction intervals of the simulated  $O_{T+H|T}$  values. Because of the monotonic relationship between pension age and OADR, we need only 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 an 80% confidence interval of pension age schemes where we are 80% sure that they will give the desired outcome.

The algorithm to find the upper and lower boundary of plausible pension age schemes is the same as above, but mean of  $O_{T+H|T}$  is replaced by the 10% and 90% quantiles.

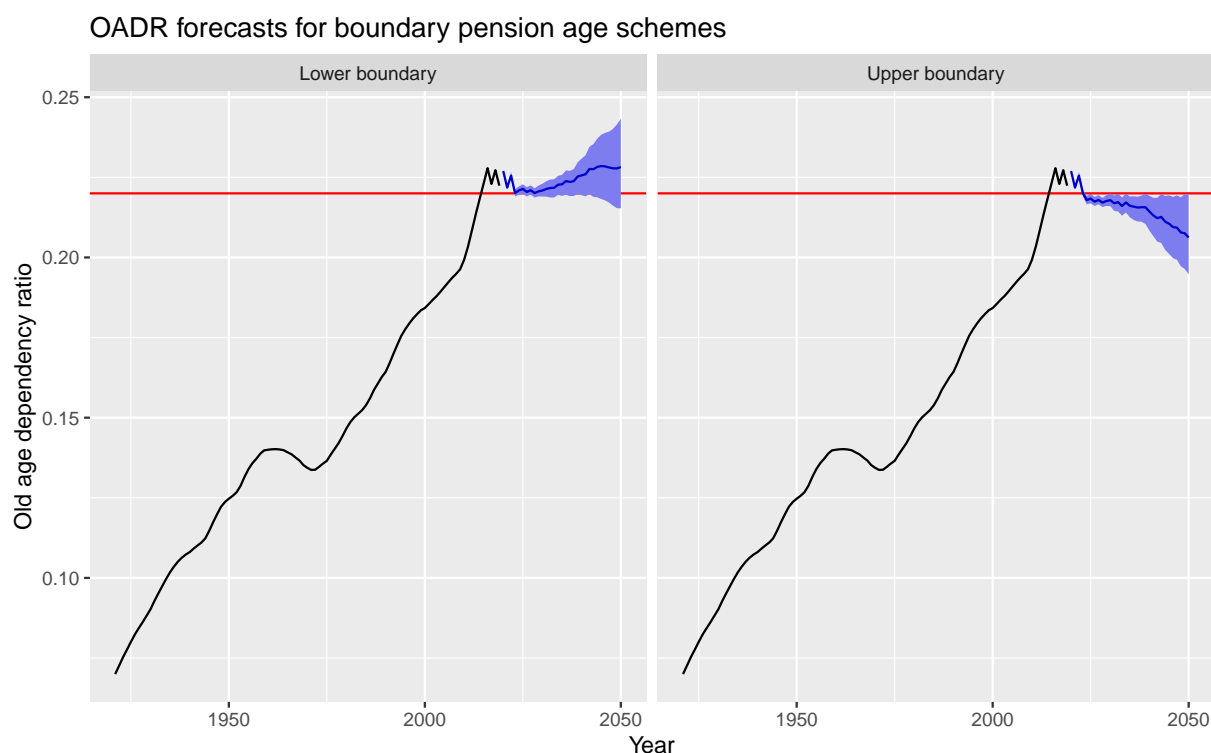
Table 1 shows the target pension age scheme, which is properly defined in section ?? . Figure 7 provides a visible comparison of the desired OADR level of 21% and the mean forecast of the



**Figure 6:** 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.



**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).

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.

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 6 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 scheme<sup>4</sup>. 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

---

<sup>4</sup>Since 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.

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 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.

## 6 Appendix: 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 forecast<sup>5</sup>, we are interested in the  $H$  prediction horizon. Then we are allowed to repeat the test  $T - H - k + 1$  times, where  $T$  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' OADR<sup>6</sup>. 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  $\omega = 1, 2, \dots, T - k - H + 1$ . Then, we can summarize the information into a matrix as follows:

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

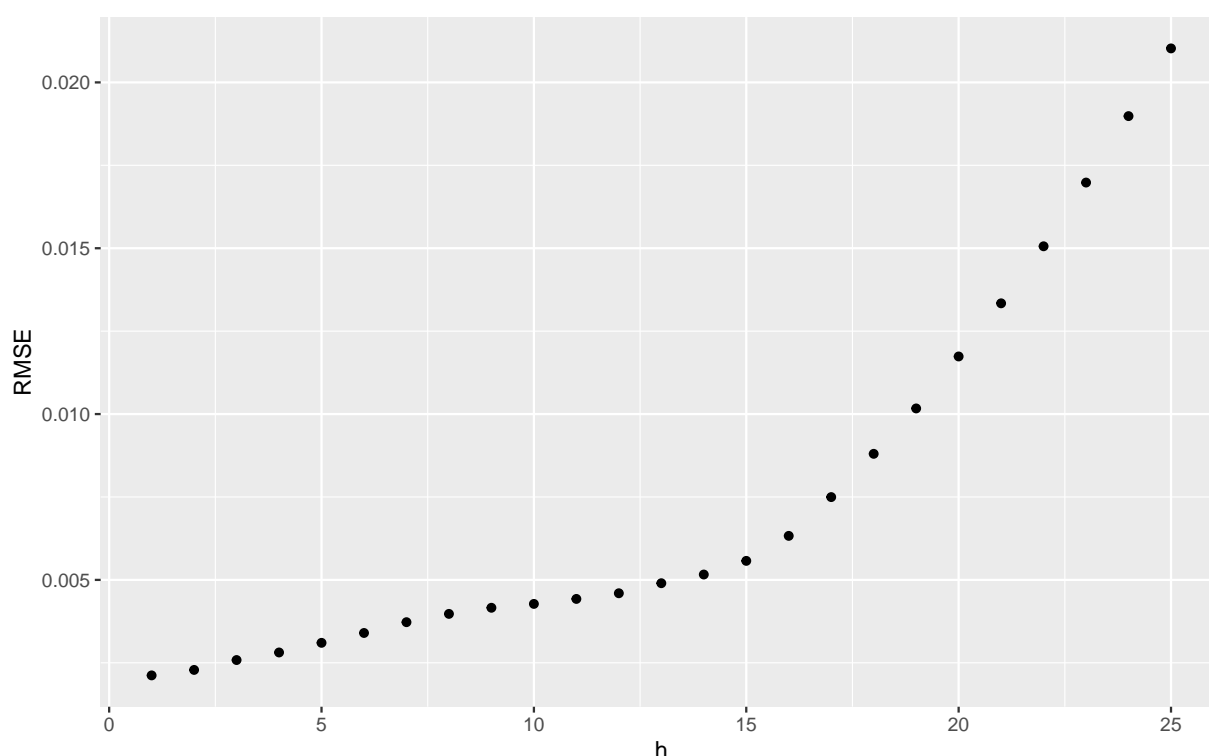
where  $h$  denotes for the forecast horizon and  $AE_{h,\omega} = |\text{OADR}_{h,\omega}^{\text{mean}} - \text{OADR}_{h,\omega}^{\text{real}}|$ . This process is also known as "rolling forecasting origin", because the "origin" ( $k + \omega - 1$ ) at which the forecast is based rolls forward in time. 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,  $T = 2018 - 1950 + 1 = 69$  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 25%.

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

<sup>5</sup> $k$  can't be too small

<sup>6</sup>The 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.

## 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 (2020). *Births, Australia: 2018*. Date accessed: 14 September 2020. Canberra. <https://bit.ly/2ZyU6vx>.
- Costello, P (2002). *Intergenerational Report 2002–2003*. Canberra: Commonwealth of Australia. <http://www.budget.gov.au/2002-03/bp5/html/>.



- 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](http://www.mortality.org) or [www.humanmortality.de](http://www.humanmortality.de) (data downloaded on 12/June/2017).
- Hyndman, RJ & G Athanasopoulos (2014). *Forecasting: principles and practice*. Melbourne, Australia: OTexts. [OTexts.org/fpp](http://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 Yasmeeen (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](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](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](http://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>.
- 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.