Forecasting the old-age dependency ratio to determine a sustainable pension age

Rob J Hyndman, Yijun Zeng, Han Lin Shang

March 2021

Forecasting the old-age dependency ratio to determine a sustainable 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
Shanghai, China

Han Lin Shang
Macquarie University, Sydney, Australia

15 March 2021

Forecasting the old-age dependency ratio to determine a sustainable 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 the future age-structure of the population. Our results suggest that the Australian pension age should be increased to 68 by 2030, 69 by 2036, and 70 by 2050, in order to maintain the old-age dependency ratio at 23%, just above the 2018 level. Our general 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, pension age

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 thirty years (Fehr, Jokisch & Kotlikoff 2008). Consequently, there are many older people leaving the workforce, and fewer workers replacing them. Many other countries face similar demographic challenges. 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, in line with the recommendations from the Productivity Commission (2013). We investigate the effect of these changes in the pension age on the forecasts of the OADR.

The OADR can be expressed as

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

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). We provide forecasts of future population age structures over the next 50 years, and use those forecasts to compute the OADR that would result under various pension age schemes.

We then invert the analysis by finding a pension age scheme that would lead to a specified OADR value. For our analysis, the specified OADR value has been set at 23%, just above the 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. Instead, we provide a methodology for policymakers to find a target pension age that leads to a given OADR value.

Forecasts of population structure depend on forecasts of age- and sex-specific mortality rates, fertility rates and net migration. To this end, we combine the functional demographic models of Hyndman & Ullah (2007) and the product-ratio method of Hyndman, Booth & Yasmeen (2013) within the general stochastic population forecasting framework of Hyndman & Booth (2008), to forecast a population age structure. 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.

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, approximately two-thirds of Australians over pension age rely fully or partly on the age pension (Australian Institute of Health and Welfare 2019), and this figure has remained relatively constant over the last 20 years. 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 if required.

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

The age pension in Australia provides income support and access to a range of concessions and additional allowances for older Australians who have passed a means test.

In 2018, 67% of Australian residents aged 65 and over received income support payments (Australian Institute of Health and Welfare 2019). Even though the government has promoted self-financed retirement by introducing the Superannuation Guarantee in 1992, the Superannuation Guarantee accumulations are still modest, with many retirees having no superannuation or very modest balances (Stephenson et al. 2017).

The ageing trend in Australia, along with retirees' dependence on the age pension, means the cost of the age pension is currently around at 2.7% of Gross Domestic Product. With an ageing population, the pension age will need to increase or the fiscal burden will become unsustainable.

The pension age for males was 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 2014, in line with that of males. From July 2017, pension eligibility ages for both males and females have been increasing by six months every two years and will reach 67 in July 2023. 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 former Treasurer Joe Hockey announced that the
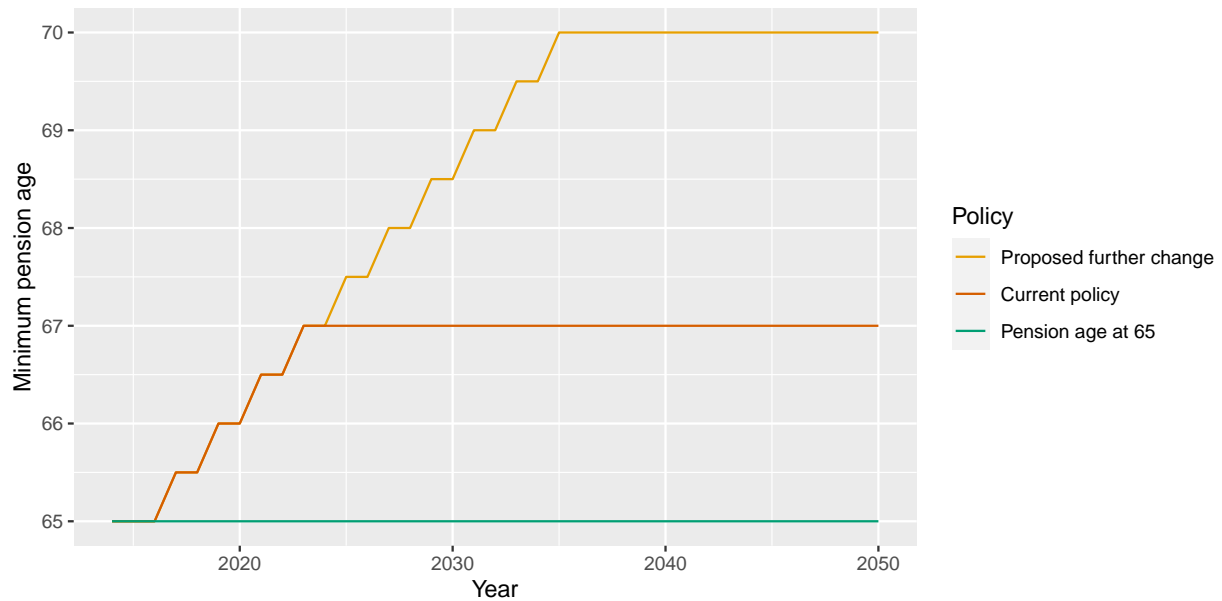
Figure 1: *Australian pension age policies.*

government would increase the pension age to 70 by 2035, and this was proposed in the Federal Budget Statement for 2014–2015, although it has never been legislated by the Australian Federal Government. Figure 1 illustrates the process of the adjustments on pension age scheme.

The government has not disclosed the method it used to determine the new pension age. 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. We propose that the pension age should be linked to the old age dependency ratio, and 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 required for our proposal 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

$$B_t(x) = \text{Births in calendar year } t \text{ to females of age } x,$$

$$D_t(x) = \text{Deaths in calendar year } t \text{ of persons of age } x,$$

$$P_t(x) = \text{Population of age } x \text{ at 1 January of year } t,$$

$$E_t(x) = \text{Population of age } x \text{ exposed to risk at 30 June of year } t,$$

where $x = 0, 1, 2, \ldots, p - 1, p^+$ denotes age with p^+ the open-ended upper age group, set to 100^+ in this study. The year is denoted by $t = 1, 2, \ldots, T$, 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[1] (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:

$$G_t(x, x+1) = P_{t+1}(x+1) - P_t(x) + D_t(x, x+1) \qquad \text{for } x = 0, 1, 2, \ldots, p - 2,$$
$$G_t(p - 1^+, p^+) = P_{t+1}(p^+) - P_t(p^+) - P_t(p - 1) + D_t(p - 1^+, p^+),$$
$$G_t(B, 0) = P_{t+1}(0) - B_t + D_t(B, 0),$$

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 migration in calendar year t of persons born 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 (Preston, Heuveline & Guillot 2000). 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

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

---

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

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 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

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

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 β_t. $\phi(x)$ is a set of parameters that tell us the rate at which each age responds to changes in β_t. β_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

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

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 β_t are therefore the first principal component and scores of X with (i, t)th element $[\log(m_t(x)) - \widehat{\mu}(x)]$, where $\widehat{\mu}(x)$ is the sample mean of $\log(m_t(x))$ over t. The log transformation introduces some bias, so the LC method adjusts β_t such that

for each year, the implied number of deaths will be equal to the actual number of deaths. The adjusted β_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

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

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

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 jth 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 $\boldsymbol{I} = \{y_1(x), \ldots, y_n(x)\}$ and the set of functional principal components $\boldsymbol{B} = \{\phi_1(x), \phi_2(x), \ldots, \phi_j(x)\}$, the h-step-ahead forecast of $y_{T+h}(x)$ are given by

$$\widehat{y}_{T+h|T}(x) = \widehat{s}_{T+h|T}(x) = \mathrm{E}[\widehat{s}_{T+h|T}(x) \mid \boldsymbol{I}, \boldsymbol{B}] = \widehat{\mu}(x) + \sum_{j=1}^{J} \phi_j(x)\widehat{\beta}_{T+h|T,j}, \tag{7}$$

where $\widehat{\beta}_{T+h|T,j}$ 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 & Yasmeen 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

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

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

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

$$\log[p_t(x)] = \mu_p(x) + \sum_{j=1}^{J} \beta_{t,j} \phi_j(x) + e_t(x)$$

$$\log[r_t(x)] = \mu_r(x) + \sum_{l=1}^{L} \gamma_{t,l} \psi_l(x) + z_t(x)$$

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

$$m_{T+h|T}^{\text{M}}(x) = p_{T+h|T}(x) r_{T+h|T}(x)$$

$$m_{T+h|T}^{\text{F}}(x) = p_{T+h|T}(x) / r_{T+h|T}(x)$$

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. We used weighted penalized splines (Hyndman & Ullah 2007) when estimating the log-fertility and log-mortality curves, where the weights account for Poisson variation in deaths and births. We use loess when estimating the net-migration curves.

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 $t = T + 1, \ldots, T + h$ 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 Application to Australia

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 $\widehat{\mu}(x)$, the first two basis functions, $\widehat{\phi}_1(x)$ and $\widehat{\phi}_2(x)$, and the first two principal component scores, $\widehat{\beta}_{t,1}$ and $\widehat{\beta}_{t,2}$) for log fertility rates. The first term, $\widehat{\phi}_1(x)\widehat{\beta}_{t,1}$, accounts for 63.7% of the variation in the log fertility rates. The first principal component scores $\widehat{\beta}_{t,1}$, 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.
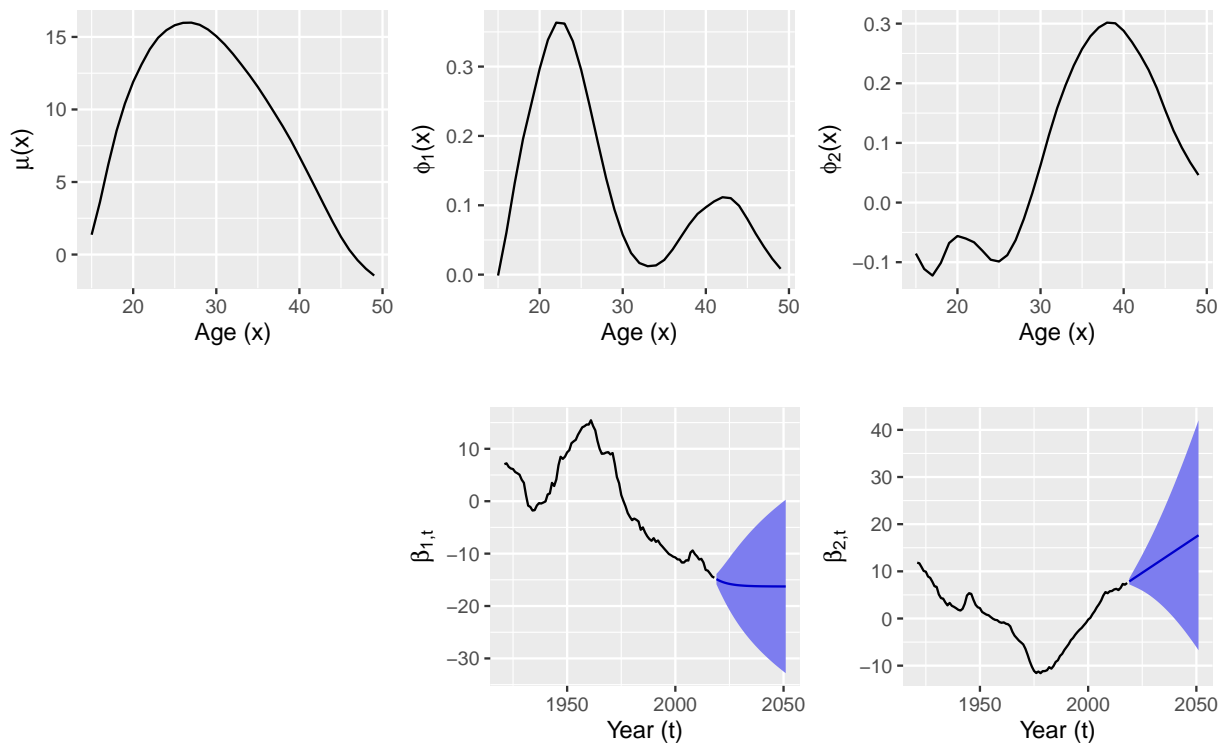
Figure 2: *Fitted basis functions and coefficients for Australian age-specific fertility rates.*

The first basis function $\widehat{\phi}_1(x)$ 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.
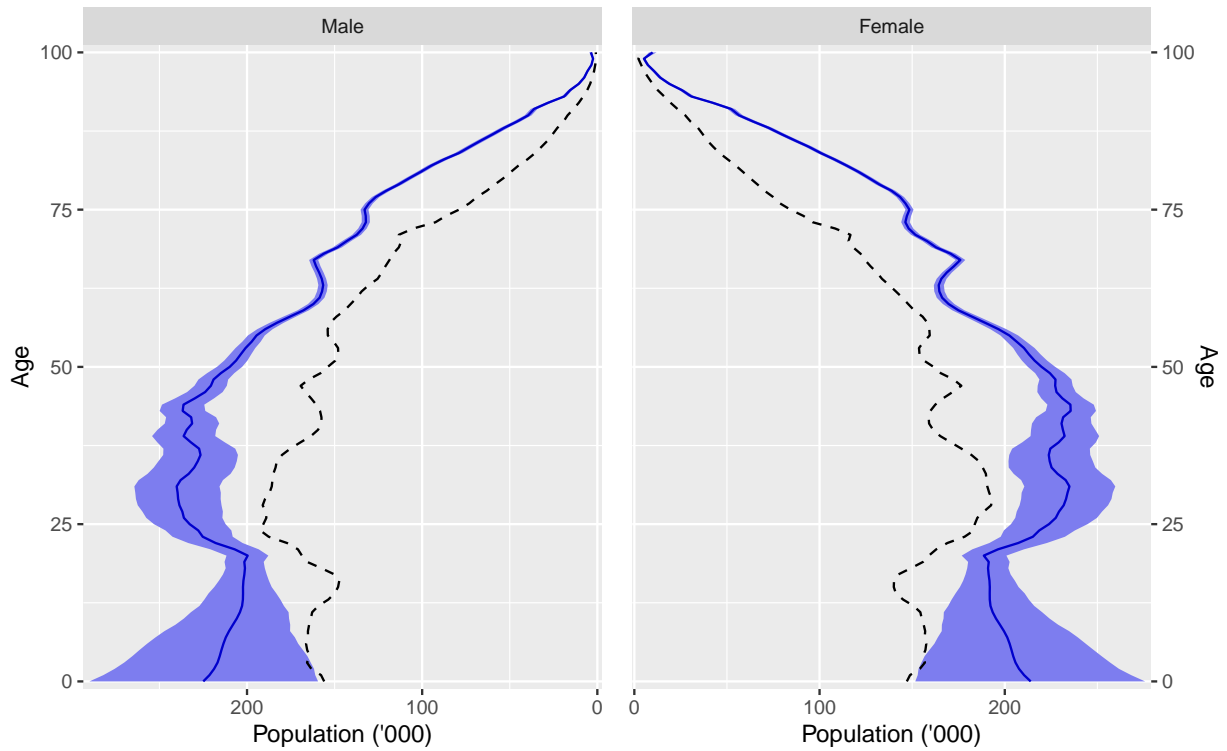
3.3 Population forecasts

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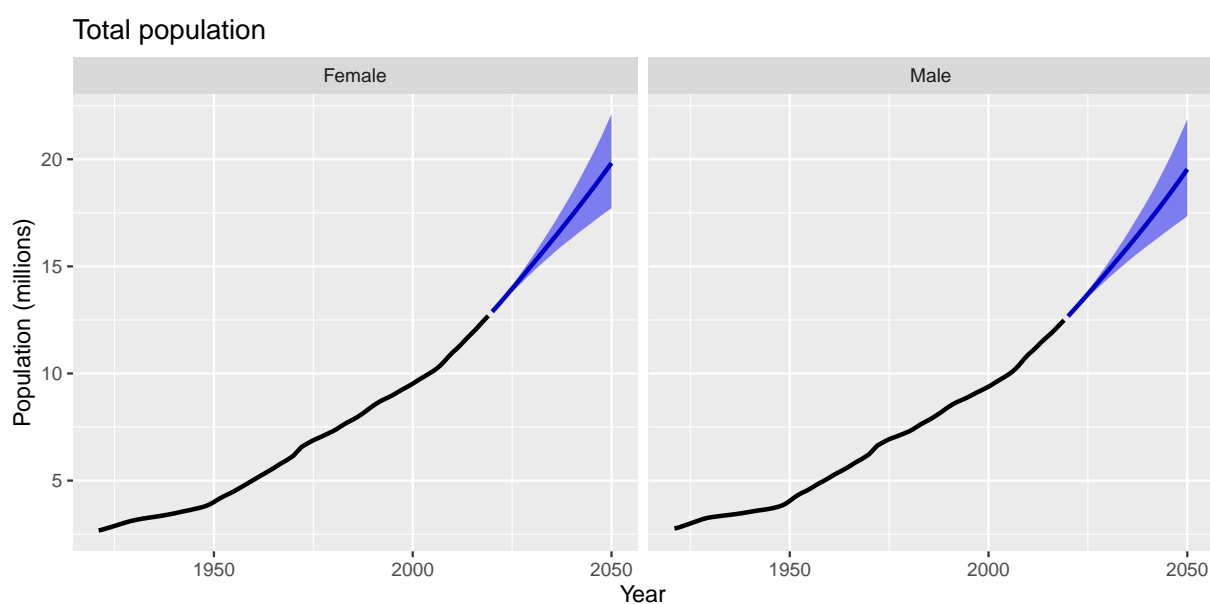
Figure 3: *Fifty-year forecasts of the total population for each sex, along with 80% prediction intervals.*

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

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

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.

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. This proposal also leads to a slow decline in the OADR over the next 15 years. Figure 5 demonstrates that the pension age scheme requires constant review to ensure a stable OADR over the long term.

4 Sustainable pension age schemes

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 $\boldsymbol{P} = [a_{T+1}, \ldots, a_{T+H}]'$, with the corresponding OADR values given by $O_{T+1}, \ldots, O_{T+H}$. Our goal is to find the minimum pension age in each year that gives an OADR

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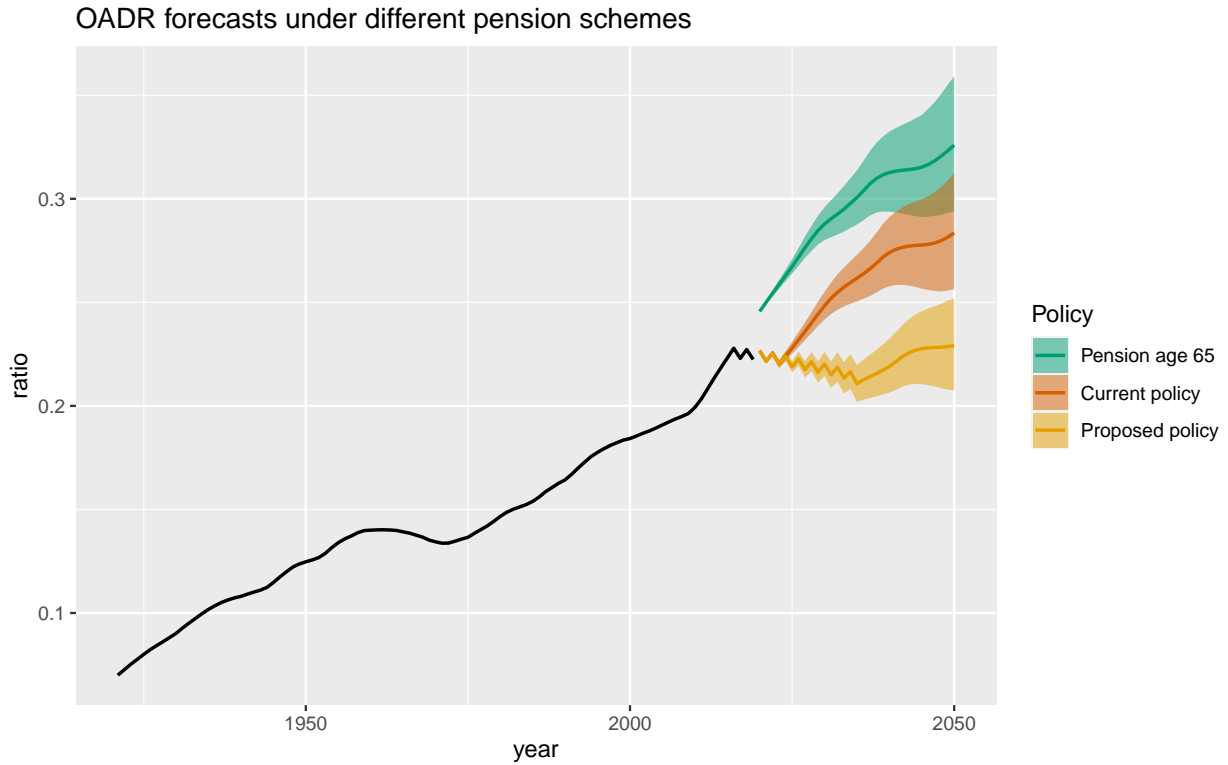
Figure 5: *Fifty-year forecasts of the OADR associated with a pension age equal to 65 (green), the current approved pension age scheme (orange), and the proposed scheme with the pension age rising to 70 in 2035 (purple). The actual OADRs for 1921–2019 are shown in black.*

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 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 $\widehat{O}_{T+h|T} \leq O^*$ or $a_{T+h} - a_{T+h-1} = 12$ months, where $\widehat{O}_{T+h|T}$ denotes the mean of the simulated $O_{T+h|T}$ values.

We repeat these steps for $h = 2, \ldots, H$. The second limit in Step 2 arises because the pension age should not jump more than 12 months between years, to prevent pension-age people losing their pension.

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 the mean of $O_{T+H|T}$ is replaced by the 10% and 90% quantiles.

Figure 6 shows (in blue) the pension scheme produced by this algorithm, along with 80% confidence intervals. In comparison, the previous (green), current (red) and proposed (orange) schemes are also shown.
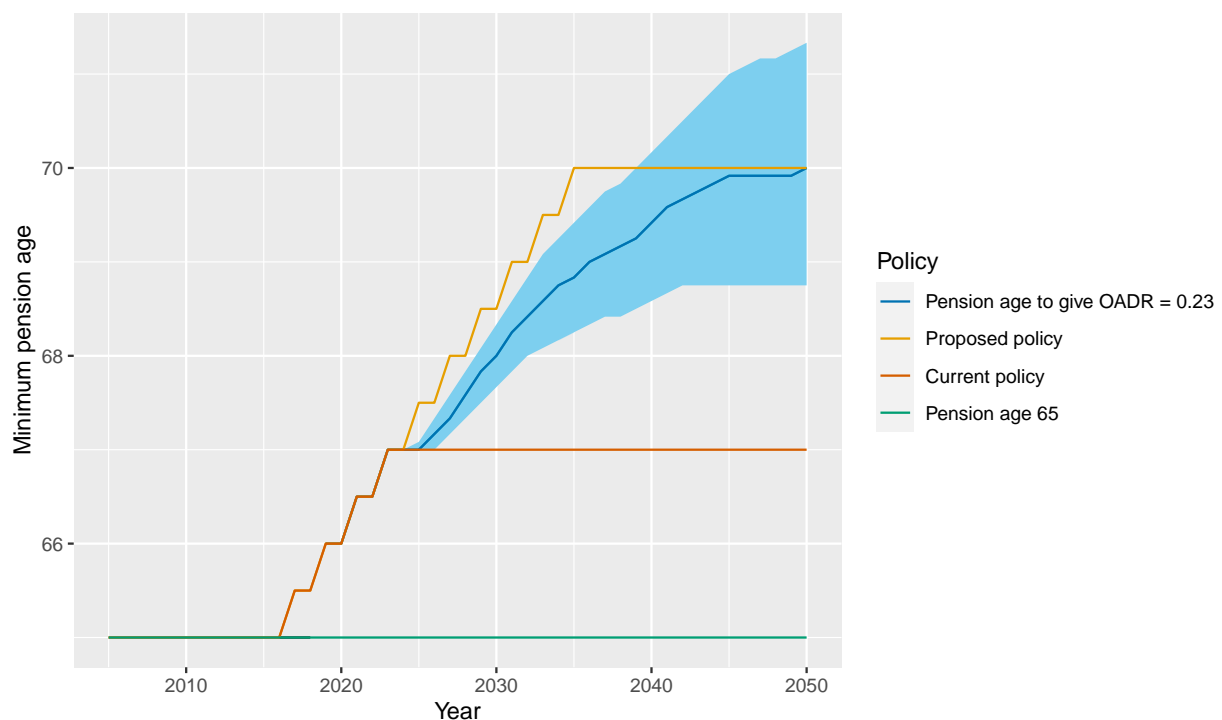

Figure 6: *The target pension age scheme (in blue) along with the pension age schemes that have been set and proposed by the government. The blue shaded region shows 80% confidence intervals for the target pension age.*

Figure 7 shows the OADR forecasts as a result of using the target age pension scheme shown in Figure 6. The target level of 23% and the mean forecast of the OADR corresponding to the target pension age almost overlap. Figure 8 shows the OADR forecasts corresponding to the upper

OADR forecast estimates for the target pension age scheme
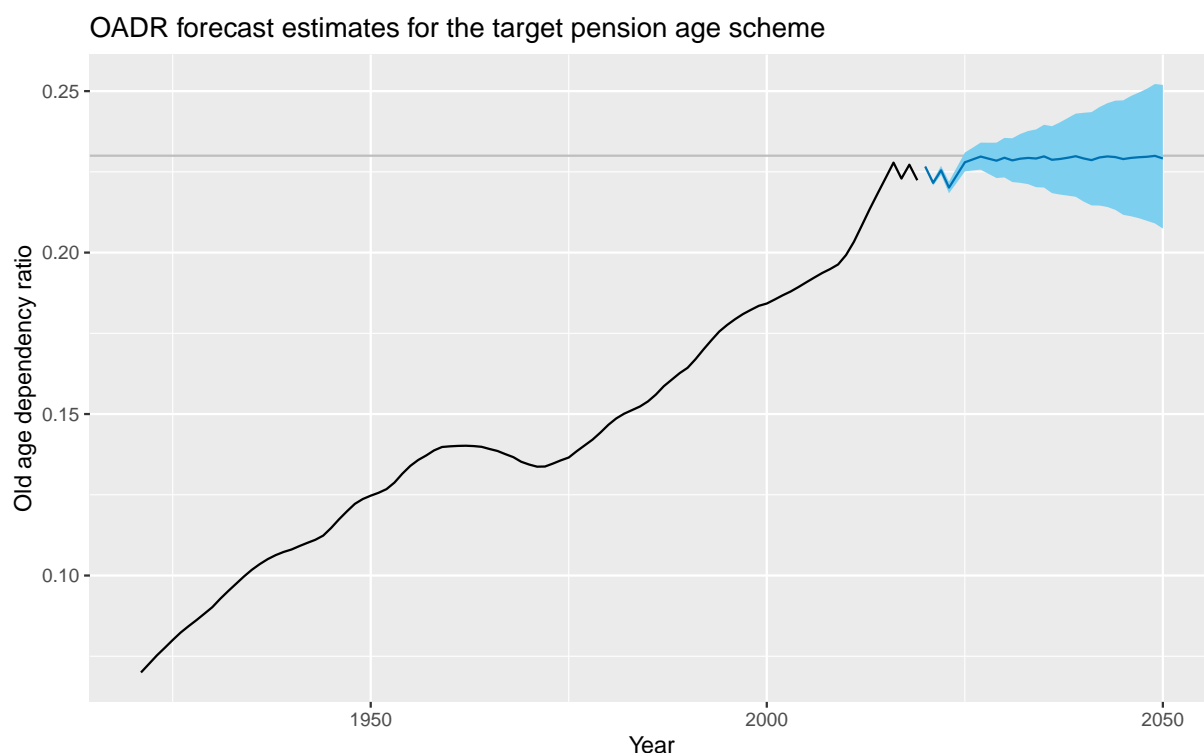
Figure 7: *The mean forecast and 80% prediction intervals for the OADR associated with the target pension age scheme shown in blue, along with the desired OADR level in gray.*

OADR forecasts for boundary pension age schemes
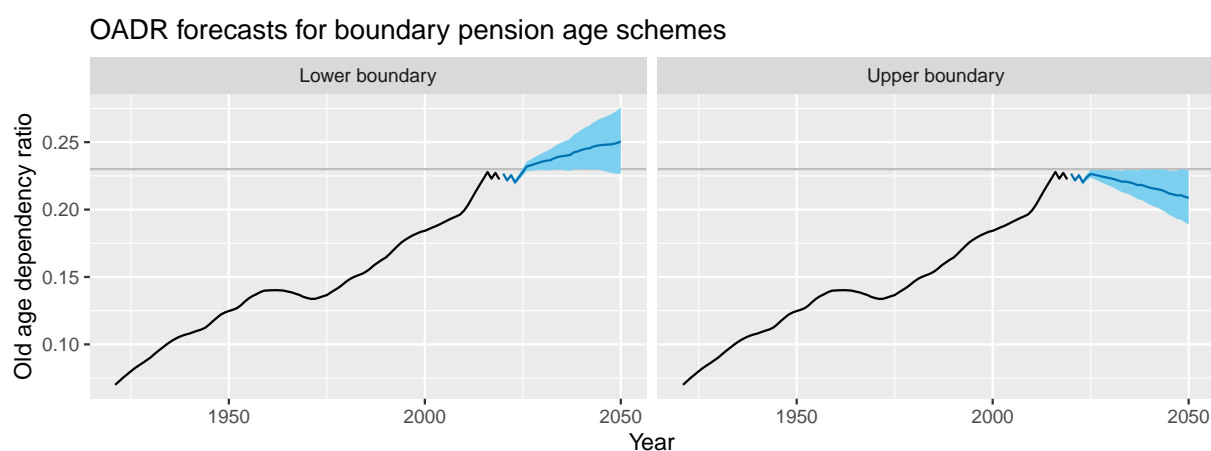
Figure 8: *The 80% prediction intervals of OADR (in blue) associated with the upper and lower boundary of plausible pension age schemes, along with desired OADR level in gray.*

and lower limits of the 80% confidence intervals shown in Figure 6, demonstrating that they correspond to having OADR forecast prediction interval boundaries that lie very close to the target OADR value.

Thus any pension age higher than the upper boundary shown in Figure 6 will give an OADR value that is significantly lower than the target value, while any pension age scheme lower than the lower boundary of the pension age reported in Figure 6 will give an OADR value that is significantly higher than the target value.

Figure 6 also compares the target pension age scheme and its 80% confidence interval with the pension age scheme that has been proposed but not yet enacted. Clearly, it is outside the 80% confidence interval and so is too high for our targeted OADR value of 23%. Under our proposed scheme, the Australian pension age should be increased to 68 by 2030, 69 by 2036, and 70 by 2050, in order to maintain the old-age dependency ratio at 23%.

5 Conclusions

The ageing of the population is an increasingly critical issue in Australia (and many other countries) due to the increase in life expectancy and the retirement of baby boomers. The fiscal cost of the age pension has increased as a result, and is likely to continue to be a financial burden unless there are changes in age pension schemes. Under current Australian government policy, the pension age will rise to 67 in 2023, and there is a proposal to further increase it to 70 by 2035. Our analysis suggests that this increase is too rapid unless the aim is to reduce the OADR to below 2018 values.

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 the OADR, which we take as a crude measure of the financial burden of the age pension. We have applied stochastic models to forecast the future population age structure and then used those future age structure projections to find the most appropriate pension age scheme that will ensure a 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 methods (Hyndman & Ullah 2007), along with an extension of the product-ratio method (Hyndman, Booth & Yasmeen 2013). Then the framework of Hyndman & Booth (2008) was used to simulate future populations by age and sex. These were then used to simulate OADR values associated with different pension age schemes. Finally, we inverted the process to find a pension age scheme that leads to a specified OADR value.

We illustrate the approach based on assuming the desired level of OADR is around the 2018 level of 23%. Using historical data from 1950 to 2019, the forecasts 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.

The empirical analysis presented in this paper shows the adjustment of the pension age to age 67 has been effective in stopping the increasing trend in the OADR values. 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 proposed statistical approach. This approach has been used to calculate the target pension age from now until 2050 along with its 80% confidence interval. The target pension age found in this research increases to 67 by 2023, 68 by 2030, 69 by 2036 and 70 by 2050. This generates a stable OADR of around 23%.

This research addresses the pension age issue from a purely demographic perspective. Further analysis should include an economic perspective as well including the impact of the superannuation guarantee and other economic measures to offset the fiscal burden of the age pension. There are several issues to be considered here. First, not all people aged above the pension age will receive the pension, and the ratio of those receiving the pension may decrease as more retiring people can access superannuation. Second, labour force participation is probably related to age as self-funded retirees may opt out of paid employment before they reach pension age. Hence, raising the pension age may not have as large an effect on OADR as we have estimated. Third, we have not considered the effect of changing per capita GDP which helps fund the pension. If per capita GDP declines, the fiscal burden of the pension may increase even with a stable OADR.

We have also not allowed for the impact of COVID-19 in this analysis as it is too early to properly estimate its impact on mortality rates and future age structures. It is known that the mortality rates from COVID-19 are greater amongst older people, so it should lead to a decrease in the OADR. However, the relatively small number of deaths in Australia mean that the effect of the pandemic will probably not substantially change our recommendations.

Reproducibility

The data and code to fully reproduce our results is available at https://github.com/robjhyndman/PensionAge. We used R (R Core Team 2020), along with the drake (Landau 2018), tidyverse (Wickham et al. 2019), demography (Hyndman et al. 2019) and forecast (Hyndman et al. 2020) packages.

References

Apps, P, R Rees & M Wood (2007). Population ageing, taxation, pensions and health costs. *Australian Journal of Labour Economics* **10**(2), 79–97.

Australian Bureau of Statistics (2020). *Births, Australia: 2018*. Date accessed: 14 September 2020. Canberra. `bit.ly/2ZyU6vx`.

Australian Institute of Health and Welfare (2019). *Income support payments for older people*. Accessed on 20 September 2020. `www.aihw.gov.au/reports/australias-welfare/income-support-payments-for-older-people`.

Costello, P (2002). *Intergenerational Report 2002-2003*. Downloaded 10 September 2020. Canberra: Commonwealth of Australia. `archive.budget.gov.au/2002-03/bp5/bp5_combined_pdf.zip`.

Costello, P (2007). *Intergenerational Report 2007*. Accessed 20 September 2020. Canberra: Commonwealth of Australia. `treasury.gov.au/publication/intergenerational-report-2007-2`.

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.

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)*. Data downloaded on 18 September 2020. `mortality.org`.

Hyndman, R, G Athanasopoulos, C Bergmeir, G Caceres, L Chhay, M O'Hara-Wild, F Petropoulos, S Razbash, E Wang & F Yasmeen (2020). *forecast: Forecasting functions for time series and linear models*. R package version 8.13. `https://pkg.robjhyndman.com/forecast/`.

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, L Tickle & J Maindonald (2019). *demography: Forecasting Mortality, Fertility, Migration and Population Data*. R package version 1.22. `https://CRAN.R-project.org/package=demography`.

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. Accessed on 20 September 2020. Committee for Economic Development of Australia. `trove.nla.gov.au/work/35010227?selectedversion=NBD42351550`.

Landau, WM (2018). The drake R package: a pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software* **3**(21). `https://doi.org/10.21105/joss.00550`.

Lee, RD & LR Carter (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* **87**(419), 659–671.

Nielson, L (2010). *Chronology of superannuation and retirement income in Australia*. Background note. Accessed on 20 September 2020. Canberra: Commonwealth of Australia. `www.aph.gov.au/binaries/library/pubs/bn/eco/chron_superannuation.pdf`.

Preston, SH, P Heuveline & M Guillot (2000). *Demography: measuring and modeling population processes*. Oxford: Wiley-Blackwell.

Productivity Commission (2013). *An ageing Australia: preparing for the future*. Commission Research Paper. Accessed on 20 September 2020. Canberra, Australia. `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.

Stephenson, A, Z Zhu, P Toscas, A Reeson, B Koo, C O'Hare & D Cox (2017). *Household Assets Among Australian Age Pensioners: A preliminary analysis of data from the Department of Human Services*. Report EP174179. CSIRO Risklab Australia.

Swan, W (2010). *Australia to 2050: Future challenges*. Accessed 20 September 2020. Canberra: Commonwealth of Australia. `treasury.gov.au/sites/default/files/2019-03/IGR_2010_Overview.pdf`.

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

Wickham, H, M Averick, J Bryan, W Chang, LD McGowan, R François, G Grolemund, A Hayes, L Henry, J Hester, M Kuhn, TL Pedersen, E Miller, SM Bache, K Müller, J Ooms, D Robinson, DP Seidel, V Spinu, K Takahashi, D Vaughan, C Wilke, K Woo & H Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software* **4**(43), 1686.