

# Author Manuscript

## National population growth rate, its components, and subnational contributions: A research note

Vladimir Canudas-Romo<sup>1</sup>, Tianyu Shen, and Collin Payne

Australian National University

### Abstract

A population's current growth rate is determined jointly by changes in fertility, mortality, and migration. This overall growth rate is also the average of age-specific growth rates, which can be decomposed into the result of historical changes in fertility, mortality and migration. However, doing so requires over 100 years of historical data, meaning that such analyses are possible only in select few populations. In this research note, we propose an adapted version of the variable- $r$  model to measure contributions to the population growth rate for countries with shorter demographic series. In addition, we extend this model to explore the contribution of subnational changes to the national population growth rate. Our results demonstrate that the age-specific growth rates obtained from short historical series, say 25 years, closely match those of the longer series. These abbreviated age-specific growth rates closely resemble the growth rate at birth of their respective cohorts, which is the major determinant of population growth, except at older ages where mortality becomes the main explanatory element. Exploring subnational populations, we find considerable heterogeneity in the age-profile of the components of growth, and find that the most populous regions tend to have an outsized impact on national-level growth.

*Keywords:* Variable- $r$ , Population growth, Subnational

---

<sup>1</sup> Corresponding to Vladimir Canudas-Romo, [Vladimir.Canudas-Romo@anu.edu.au](mailto:Vladimir.Canudas-Romo@anu.edu.au)

## Introduction

Population growth (increases and declines) remains one of the key demographic concerns worldwide. The world's population will continue growing from 8 billion today to a peak at nearly 10.4 billion around 2080 (UN 2022). The simple and elegant demographic balance equation allows researchers to estimate the combined contribution of births, deaths and migration to those changes in population (Preston et al. 2001). The components of the balance equation can be seen as a prospective perspective of population growth. However, counts of deaths and births are predetermined by the population's composition—for example, aging populations experience a high number of deaths and few births. Alternatively, a retrospective view towards population growth explains growth as the result of historical demographic events that have shaped the current population composition. This research note builds on previous efforts to disaggregate the demographic history of current population growth by Canudas-Romo et al. (2022), extending this method to shorter time series and disaggregated to subnational populations.

The current population growth rate is the average of the age-specific growth rates, weighted by the population composition (Horiuchi 1991). Those age-specific growth rates are in turn determined by the contribution of historical changes in fertility, mortality and migration. The variable- $r$  model separates the past contribution of each of these demographic components from the age-specific growth rates (Arthur and Vaupel 1984; Lee and Zhou 2017; Murphy 2017; Preston and Coale 1982; Preston and Stokes 2012; Preston and Vierboom 2021). However, traditional variable- $r$  methods require extremely long demographic series (birth, mortality and migration) to explore growth at all ages, with particularly high data demands at oldest-old ages. Using an alternative variable- $r$  approach it is possible to accommodate shorter demographic series and explore the historical demographic contribution between two given years (Canudas-Romo et al. 2021). In this research note, we adapt this approach to the population growth rate. Thus, we derive a direct disaggregation of the population growth rate at each age as the population growth at an earlier age and the historical cohort survival and migration between the two that produced today's population.

Historical trends in fertility, mortality and migration vary across countries, but substantial variability is also observed within countries' subnational populations (Nisén et al. 2021; Alexander et al. 2017; Bozick 2021). As a consequence, considerable heterogeneity can be hidden within national population growth rates, which may overlook diversity in growth

patterns within the subnational populations. To further explore how subnational population changes contribute to national-level growth, we propose an extension of the population growth rate to account for the regional/states contribution.

## Data

Our analysis uses age- and sex-specific population counts and death rates for national populations from the Human Mortality Database (HMD), and for subnational populations (departments/prefectures/provinces/states/territories) from the available subnational mortality databases: Australian (AHMD), Canadian (CHMD), French (FHMD), Japanese (JMD), and the US Mortality Databases (USMDB). These databases use the same methods protocol from the HMD, meaning that similar derivations of life table measures are performed for all populations, allowing comparability across countries and subpopulations. The use of this HMD protocol was the main criteria for selecting these countries, which represent a combination of growing and declining (Japan) populations during the studied period. The current HMD methods protocol (version 6) is a result of several iterations experienced during the development of the database. Similar scrutiny and possible subnational-specific methodology will need to be developed for these subnational databases. This is particularly critical for the assumptions on migration which could create bias in results on person-years and, thus, life table measures. Additionally, the information from the United Kingdom countries (England and Wales, Northern Ireland and Scotland) from the HMD were also included in our subnational study. Population counts were obtained from these databases; 2010 to 2020 was the most recent time available across all subnational and national populations and was selected as our time range of analysis. The details of the available information for each of the subnational databases are included in Table A1 in the Supplementary Online Material (SOM) Appendix 1.

## Method

The population growth rate at time  $t$ , or  $r(t)$ , is the average of the age-specific growth rates weighted by the population composition at each age as

$$r(t) = \int_0^{\omega} r_x(t) c_x(t) dx, \quad (1)$$

where  $r_x(t)$  and  $c_x(t)$  correspond to the growth rate and population composition at age  $x$  and time  $t$  respectively.

To decompose the population growth between two points in time we apply an extension of variable- $r$  method (Canudas-Romo et al. 2021). In this truncated variable- $r$  decomposition the population growth rate at age  $x$  and time  $t$ ,  $r_x(t)$ , for each cohort is separated into the growth at an earlier age  $y$ , or  $r_y(t - x + y)$ , and the cohort survival and net migration changes between the two ages as

$$r_x(t) = r_y(t - x + y) + \Delta S^{t-x}(y, x) + \Delta M^{t-x}(y, x), \quad (2)$$

where  $\Delta S^{t-x}(y, x)$  and  $\Delta M^{t-x}(y, x)$  are the changes in cohort survival probabilities and net migration probabilities from a younger age  $y$  to the older age  $x$ . The change in net migration  $\Delta M^{t-x}(y, x)$  is estimated from the residuals of Eq. (2) as in previous studies (Preston and Stokes 2012; Preston and Vierboom 2021). Since our aim is to decompose the population growth between two periods, the age-specific growth rates  $r_x(t)$  were calculated from ages 0 to 100. Thus, equations (1) and (2) combine to produce the contribution of growth at an earlier age (refer to  $r$ -past), survival and net migration to the total population growth.

Eq. (1) can also be applied to calculate subnational growth rates, or  $r_i(t)$  for subnational population  $i$ , defined as  $r_i(t) = \int_0^\omega r_{x,i}(t) c_{x,i}(t) dx$ , where  $c_{x,i}(t) = \frac{N_{x,i}(t)}{N_i(t)}$  is the population composition at age  $x$  and time  $t$  in the subnational population  $i$ . To study the contribution of each subnational population  $i$  growth rate to the national population growth there is another average,  $r(t) = \sum_i r_i(t) c_i(t)$ , where  $c_i(t) = \frac{N_i(t)}{N(t)}$  corresponds to population composition at time  $t$  in subnational population  $i$  with respect to the total population. With substitution, we can derive the formula

$$r(t) = \sum_i \int_0^\omega r_{x,i}(t) c_{x,i}(t) c_i(t) dx, \quad (3)$$

where  $r_{x,i}(t)$  can be separated with Eq. (2) further into the contribution of survival and migration as well as the previous growth rate. Further details of these equations and their derivations presented here in continuous are found in the SOM Appendix 2. All calculations were conducted in R software (R Core Team 2023), and the R codes of the programs used in the analysis are included in the GitHub repository: <https://github.com/tyaSHEN/subnational>.

## Results

This section focuses on demonstrating that the growth rate decomposition introduced in Canudas-Romo et al (2022) can be generalized and applied to shorter time series and subnational populations. We show results decomposing the rate of population growth between 2010 and 2020 for the six selected countries (Australia, Canada, France, Japan, the UK, and the US) and their subnational populations. This rate of growth is separated into three terms: the portion resulting from changes in cohort survival probabilities, the portion resulting from changes in cohort net migration probabilities, and the portion resulting from the growth rate at an earlier age ( $r$ -past). We test several different time periods of historical information to understand how sensitive our substantive results are to differing time horizons (25, 50, 75, and 100 years). Finally, we explore the drivers of changes in subnational growth rates, focusing on understanding the contribution of the subnational growth rate components to population changes at the national level.

Figure 1 depicts the age-contribution to the growth rate from Eq. (1) for the selected countries. The peaks and valleys in the age trends coincide in all nations: high values in ages 5-10, 25-35, and 55-75; and low values in ages 12-25, 40-50, and after age 75. These growth rates contributions correspond to these countries' baby-boomer cohorts (aged 55-70), the generation X, or "echo-boomers", found between ages 25 and 35, and the "baby bust" between 40 and 50. It is worth noting that these age-patterns are not universal, fluctuations in the growth rate contributions in Japan are the opposite of the age-specific trends in other countries, excepting the high growth rate above age 70 (Figure A2 in SOM Appendix 3 includes Japan).

[Figure 1 about here]

The cross-sectional representation of the population change between 2010-2020 shown in Figure 1 can be further decomposed into its historical components as presented in Eq. (2). Figure 2 demonstrates this decomposition by focusing on the US female population growth rate between the years 2010 and 2020, based on four different periods of historical information: 25, 50, 75 and 100 years. Each sub-Figure includes the age-patterns of the components of the age-specific growth rates, namely the product of the population composition, or  $c_x(t)$ , with each component: changes in cohort survival probabilities and cohort net migration probabilities, as well as the growth rate at an earlier age ( $r$ -past) as described in Eq. (2) and further explained in SOM Appendix 2. Comparing across these four periods of historical information in Figure

2, we find a remarkable similarity between the age-contributions. This suggests that short lengths of time are sufficient to estimate the age-contributions of survival, net migration and past-growth to the overall growth rate. The consistency in age-specific growth rates over age for any given cohort, except at old ages, is the main reason for the similarity in age-contributions. A cohort's growth rate starts at a level set by the growth rate at birth and, although it changes at subsequent ages due to survival and migration, it remains similar to this initial growth. This similarity between age-specific growth rates and their corresponding growth rate at birth allows us to use shorter series of 25-50 years and still obtain similar age-patterns describing the cohorts demographic history.

Although deviations arise between the age-specific growth rates at older ages and their growth rate at birth, these differences play a minor role in the overall growth rate given that they are weighted by the structure of the population. Older adults represent a smaller proportion of the total population, meaning that the contribution of those aged 85 and above to overall growth is small. This is the case even when age-specific growth rates are the highest at older ages (e.g. the highest growth rates for American females between years 2010 and 2020 are found at ages 70 and 96, see unweighted Figure A3 in SOM Appendix 3). At these advanced ages, growth is primarily driven by increases in survival. The contributions to growth vary substantially across cohorts in the US, ranging from the high growth of baby-boomers (aged 55 to 75) and the "echo-boomers" (aged 25 to 35) to the low growth rates of the "baby bust" between ages 40 and 50.

[Figure 2 about here]

Figure 3 further extends the observation of Figure 2 comparing the components of the growth rate based on 25, 50, 75 or 100 years of historical data. The similarity in growth rate at an earlier age,  $r$ -past, is confirmed regardless of the length of the historical series (similar Figures are included in SOM Appendix 3 for the other five countries in this study). However, small errors at each age accumulate, and can result in deviations in aggregate results. Table A2, in the SOM Appendix 3, shows that using shorter time-series of historical information leads to the  $r$ -past term increasing in explanatory power, while using longer time-series leads to an increase in the importance of survival and net migration to the overall growth rate. Thus, there are clear gains on precision by having a longer time series of information, particularly at the age-aggregated level. In other words, using a shorter demographic history leads to more information being contained in the  $r$ -past term, while using a longer time-series sees this

information distributed among the other two components. This is the case even when small deviations are observed at the age-specific level of the components of population growth.

[Figure 3 about here]

Figure 4 presents the national and subnational age-specific growth rates for the selected countries. Also highlighted are the subnational populations with the highest and lowest overall growth rates ( $r_i(t) = \int_0^{\omega} r_{x,i}(t)c_{x,i}(t)dx$ ). All subnational populations' age-specific growth rates follow the same age-pattern as the national trend, except for Washington DC which has higher growth at young ages (0 to 5 and 25 to 35) than the rest of the nation. As expected, the subnational populations with the highest overall growth rates ( $r_i$ ), also correspond with the highest values of age-specific growth rates. Opposing that, are the subpopulations with the lowest growth rates, which also have low age-specific values. For the United Kingdom, England and Wales corresponds to 89% of the population, so this subpopulation and the national age-specific growth rates have practically the same age-pattern.

[Figure 4 about here]

Figure 5 shows the age-contribution of the components of US population growth between the years 2010 and 2020, as well as the contribution to this growth from California, Florida and Texas. These are the most populous states in the nation representing 12%, 8% and 6% of the total population respectively, and corresponding to the  $c_i(t)$  term in Eq. (3). While for the nation the observed values are the product of the population composition, or  $c_x(t)$ , with each component, for the state results there is the additional product with the state composition in the nation,  $c_i(t)$  term in Eq. (3). As such, the national growth is dominated by the populous states and age-groups, as seen in Figure 4 for the UK. The same scale (vertical axis) is used for the age-components for the three states, and although these differ from the national-level results in scale, the age-patterns are similar. California and Florida closely follow the trends in age-specific growth of the national level described in Figure 2, with the waves of baby boomers and their echo cohorts present. However, at ages 20 to 40 the growth rate of the past ( $r$ -past) exceeds the growth rate of 2010 to 2020, resulting from negative net migration at these ages.

Patterns in Texas are less pronounced than for the other states--at younger ages (the late teens and early twenties) growth is positive, as opposed to the negative age-specific growth observed in the other states and the nation.

[Figure 5 about here]

The analysis at the subnational level in Figure 5 is extended in Table 1 to include the populations from Australia, Canada, France, Japan, the UK and the US. The three most populous subnational populations in each country and their contribution to the nation's growth rate between 2010 and 2020 are also included in the Table. The contribution to the national growth of each of the subnational populations depends substantially on the proportion of the population in those populations. For example, England and Wales which corresponds to 89% of the population in the UK explains 93% of their nation population growth of  $r = 6.22\%$ . Similarly, in Australia and Canada the three major states cover more than 75% of the population of the country, and explain 82% and 73% of their population growth respectively. In Japan and the US the three most populous states correspond to around a quarter of the population. However, while for the US they explain 51% of the population growth, in Japan they are among the few prefectures experiencing population expansion, while at the national level the population size is declining. In most countries the  $r$ -past component is the dominant component, excepting the UK where cohort survival is the primary driver of change. The only uniform trend within each country is that of survival, which consistently contributes to increased growth, while  $r$ -past and net migration differ both with respect to the national trend and between subnational populations.

[Table 1 about here]

## Discussion

The national population growth rate, as a single demographic measure, hides the historical demographic contribution of the components of change: fertility, mortality and migration (Canudas-Romo et al. 2022). In this research, we integrate the past contribution of these factors to overall population change, even for populations that do not have long demographic series of information. This is possible because age-specific growth rates are very similar to the growth rate at birth of their corresponding cohort (Preston and Vierboom 2021), except at older ages. We find consistency in age-specific trends of the components of the population growth, irrespective of the number of years of observation in all analyzed countries. In other words, at young ages the mortality and migration effects have not disrupted the age-specific growth rates enough from their initial growth rate at birth. Nevertheless, excess cohort mortality and/or migration can cause this trend to be disrupted. For example, in every single population at older ages the accumulation of mortality improvements in successive cohorts disrupts the similarity between growth rates at old ages and their corresponding growth rate at birth (SOM Appendix 3). However, this disparity between growth rates (at old ages and birth) makes only a minor contribution to the national population growth rate, as these advanced ages comprise a small proportion of the total population.

Our proposed extension to the variable- $r$  model has both strengths and weaknesses. First, this variable- $r$  extension can be widely applied to countries irrespective of the length of their available historical demographic data. However, the current model requires country-specific assessment for deciding when there is enough data to apply it, as opposed to having a rule of thumb for all populations. Secondly, the current model calculations ensure that the elements of the variable- $r$  decomposition sum to the total growth rate. However, this means that the net migration factor, domestic and international, is calculated as the residual of the other available (population, births, and deaths) information. Although this issue arises largely from a lack of high-quality data input on migration, the variable- $r$  extension results can be influenced by this lack of information, particularly in subnational populations. Thirdly, subnational data can return surprising results, such as the growth rates for the largest prefectures in Japan opposing the national trend, mainly due to the country's internal migration. However, this is in most cases unusual and the national results are largely determined by those subnational areas with the biggest population share. Finally, an important consideration is that the extended variable- $r$  model can adapt to first-, second- or third-order divisions of countries, although the findings will also vary by the geographies used.

How changes at the subnational level relate to the national-level demographic change is key to understanding the overall growth rate and other population metrics. These subnational populations experience great variability in growth (Pezzulo et al. 2017), with substantial implications for national-level trends. In our illustrations, the most populous subnational populations were the drivers of much of the observed growth at the national level, except for Japan where they worked against the national population decline. The latter result emphasizes the need to study both changes in subnational populations, as well as the contribution of those changes to the national population growth. It is particularly relevant to do a unifying analysis of the subnational contributions to national-level growth in cases such as Japan, where some subnational populations have increasing growth while others are declining. A second illustration of subnational trends that highlight the use of the variable- $r$  methodology is observed for the US state-level age-patterns of the growth rate components in Figures 4 and 5. For those states, unusual historical demographic trends that have an effect on current population growth were identified for Washington DC (Figure 4) and Texas (Figure 5), which differed from national and other state trends. These methods could also be used to explore the contribution of non-geographic subnational groups (such as those defined by educational attainment, race/ethnicity, or migration status) to overall population change. Furthermore, other demographic variables, e.g. life expectancy or total fertility rate, can benefit from studying changes at the subnational level, which can have differential contributions to national results.

## **Conclusion**

Even as we age, a photograph from the past shows the sustained traits of an individual that exist in the present day. Similarly, a population's age-specific growth rates resemble the levels of their corresponding growth rate at birth and can be used instead of the birth values in variable- $r$  applications. Our proposed method develops this relationship, allowing researchers to use shorter periods of historical data and account for subnational contributions when exploring changes to the national growth rate.

## Reference

- Australian Human Mortality Database (AHMD). The Australian National University. Available at: [demography.cass.anu.edu.au/research/australian-human-mortality-database](http://demography.cass.anu.edu.au/research/australian-human-mortality-database)
- Alexander, M., Zagheni, E. & Barbieri, M. (2017). A flexible Bayesian model for estimating subnational mortality. *Demography*, 54(6):2025-2041.
- Arthur, W. B., & Vaupel, J. W. (1984). Some general relationships in population dynamics. *Population Index*, 50(2), 214-226.
- Bozick, R. (2021). Age, period, and cohort effects contributing to the great American migration slowdown. *Demographic Research*, 45:1269-1296.
- Burch, T. (2022). On the use of engineering notation in demography: a technical note. Demotrends blog. [demotrends.org/2022/01/13/](https://demotrends.org/2022/01/13/)
- Canudas-Romo, V., Shen, T., & Payne, C. F. (2021). The role of reductions in old-age mortality in old-age population growth. *Demographic Research*, 44:1073-1084.
- Canudas-Romo, V., Shen, T., & Payne, C. F. (2022). The components of change in population growth rates. *Demography*, 59(2):417-431.
- Canadian Human Mortality Database (CHMD). Université de Montréal (Canada), Max Planck Institute for Demographic Research (Germany), and University of California, Berkeley (USA). Available at: [www.bdlc.umontreal.ca/CHMD/index.htm](http://www.bdlc.umontreal.ca/CHMD/index.htm)
- French Human Mortality Database (FHMD). French Institute for Demographic Studies and the Paris School of Economics (France). Available at: [frdata.org/fr/french-human-mortality-database/](http://frdata.org/fr/french-human-mortality-database/)
- Human Mortality Database (HMD). Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at: [www.mortality.org](http://www.mortality.org)
- Japanese Mortality Database (JMD). National Institute of Population and Social Security Research. Available at: [www.ipss.go.jp/p-toukei/JMD/index-en.asp](http://www.ipss.go.jp/p-toukei/JMD/index-en.asp)
- Lee, R., & Zhou, Y. (2017). Does fertility or mortality drive contemporary population aging? The revisionist view revisited. *Population and Development Review*, 43:285–301.
- Horiuchi, S. (1991). Assessing the effects of mortality reduction on population ageing. *Population Bulletin of the United Nations* 31/32:38–51.
- Murphy, M. (2017). Demographic determinants of population aging in Europe since 1850. *Population and Development Review*, 43(2), 257-283.

- Nisén, J., Klüsener, S., Dahlberg, J., Dommermuth, L., Jasilioniene, A., Kreyenfeld, M., Lappegård, T., Li, P., Martikainen, P., Neels, K. & Riederer, B. (2021). Educational differences in cohort fertility across sub-national regions in Europe. *European Journal of Population*, 37(1):263-295.
- Pezzulo, C., Hornby, G.M., Sorichetta, A., Gaughan, A.E., Linard, C., Bird, T.J., Kerr, D., Lloyd, C.T. & Tatem, A.J. (2017). Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. *Scientific data*, 4(1):1-15.
- Preston, S. H., & Coale, A. J. (1982). Age structure, growth, attrition, and accession: a new synthesis. *Population Index*, 48(2), 217-259. doi: 10.2307/2735961
- Preston, S., P. Heuveline, and M. Guillot. (2001). *Demography: Modelling and measuring population processes*. Oxford: Blackwell.
- Preston, S. H., & Stokes, A. (2012). Sources of population aging in more and less developed countries. *Population and Development Review*, 38:221–236.
- Preston, S. H., & Vierboom, Y. C. (2021). The changing age distribution of the United States. *Population and Development Review*, 47:527–539.
- R Core Team. 2023. R: A Language and Environment for Statistical Computing (Version 4.3.0). Vienna, Austria: R Foundation for Statistical Computing.
- United Nations. (2022). World Population Prospects 2022. [www.population.un.org/wpp/](http://www.population.un.org/wpp/)
- United States Mortality Database (USMDB). Centers for Disease Control (CDC, US), Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at: [usa.mortality.org](http://usa.mortality.org)

**Table 1.** Female population growth rate between 2010 and 2020 for selected countries, based on historical information on survival, net migration and past growth rate of 25 years as in Eq. (2), and the contribution in those components of the most populous subnational populations as in Eq. (3).

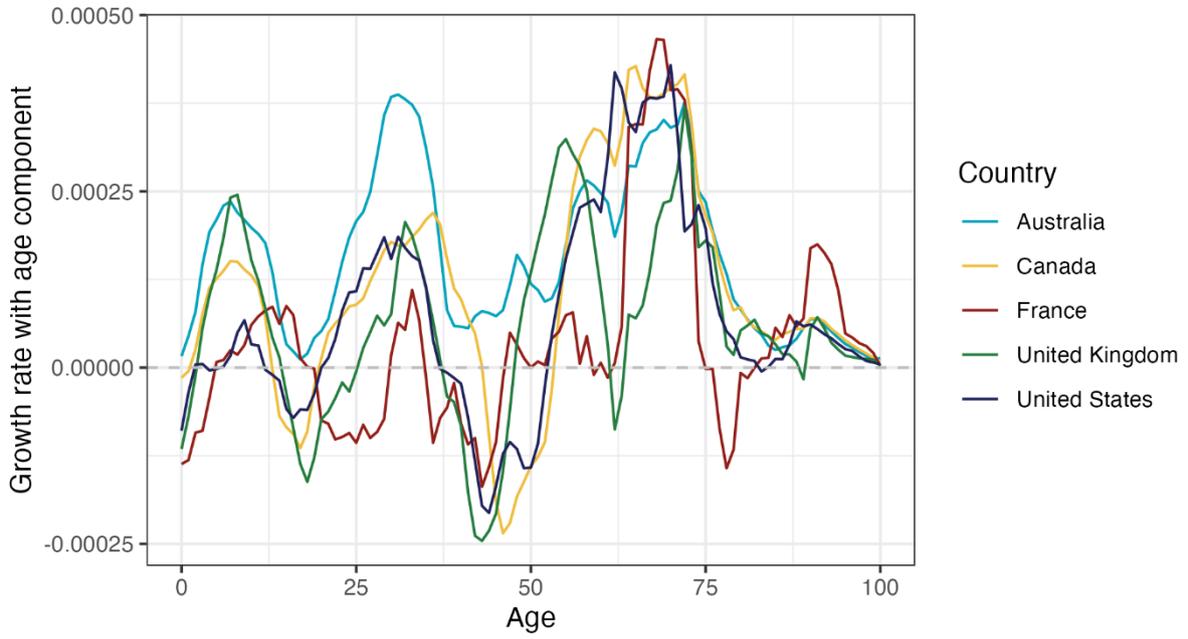
Country/State	Population %	Growth (%)	$r$ -past	Survival	Net migration
<b>Australia</b>	100	15.93	9.82	1.15	4.95
New South Wales	32	4.38 (27)	2.45	0.40	1.53
Victoria	25	5.23 (33)	1.68	0.31	3.24
Queensland	20	3.45 (22)	3.79	0.21	-0.55
Sum	78	13.06 (82)	7.92	0.92	4.22
<b>Canada</b>	100	10.85	6.40	1.04	3.40
Ontario	39	4.28 (39)	3.77	0.45	0.05
Quebec	23	1.66 (15)	0.54	0.27	0.85
British Columbia	13	1.98 (18)	2.00	0.15	-0.16
Sum	75	7.92 (73)	6.31	0.87	0.74
<b>United States</b>	100	6.50	6.37	0.89	-0.75
California	12	0.54 (8)	1.08	0.12	-0.66
Texas	8	1.44 (22)	1.21	0.06	0.16
Florida	6	1.37 (21)	1.95	0.11	-0.69
Sum	27	3.35 (51)	4.24	0.29	-1.19
<b>United Kingdom</b>	100	6.22	1.52	1.63	3.09
England and Wales	89	5.80 (93)	1.71	1.44	2.65
Scotland	8	0.29 (5)	-0.24	0.14	0.38
Northern Ireland	3	0.13 (2)	0.07	0.04	0.02
Sum	100	6.22 (100)	1.54	1.62	3.05
<b>France</b>	100	3.80	1.86	1.29	0.65
Île-de-France	19	0.82 (22)	0.63	0.21	-0.02

Auvergne-Rhône-					
Alpes	12	0.77 (20)	0.41	0.16	0.20
Nouvelle-Aquitaine	9	0.04 (1)	-0.12	0.13	0.03
Sum	40	1.63 (43)	0.92	0.50	0.21
Japan	100	-1.92	-3.84	1.49	0.44
Tokyo	10	0.63 (-33)	-0.87	0.15	1.36
Osaka Prefecture	7	0.00 (0)	-0.40	0.11	0.29
Kanagawa Prefecture	7	0.17 (-9)	0.16	0.09	-0.08
Sum	24	0.80 (-42)	-1.11	0.35	1.57

Sources: authors' calculations based on data described in Table A1 in the SOM.

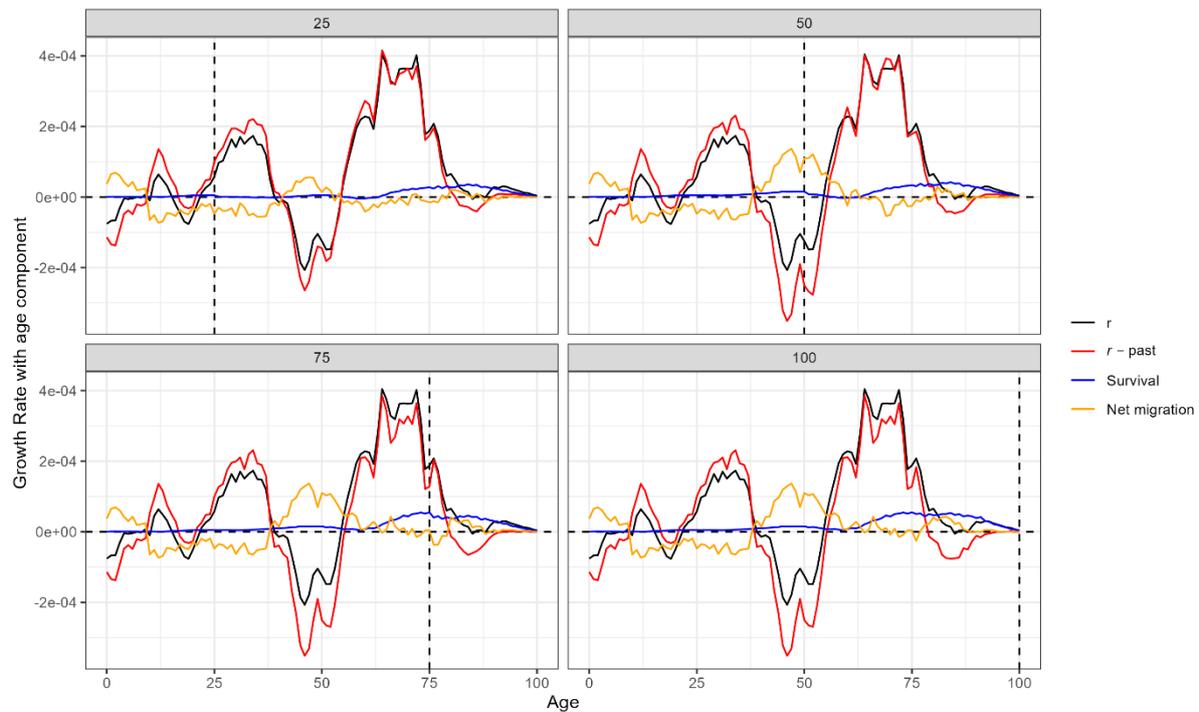
Notes: Values in columns 2 to 5 have been multiplied by 1000. Countries ordered by national growth rate from the highest in Australia (15.93%) to the lowest in Japan (-1.92%). Percentage subnational population and contributions to national growth rates are presented in the first and second columns, the latter in parenthesis. The UK values do not include the overseas territories which explains the remaining gaps to the total.

**Figure 1.** Age-contribution to the female population growth rate between 2010 and 2020,  $r_x(t)$ , for selected countries.



Source: authors' calculations based on HMD (2023) data.

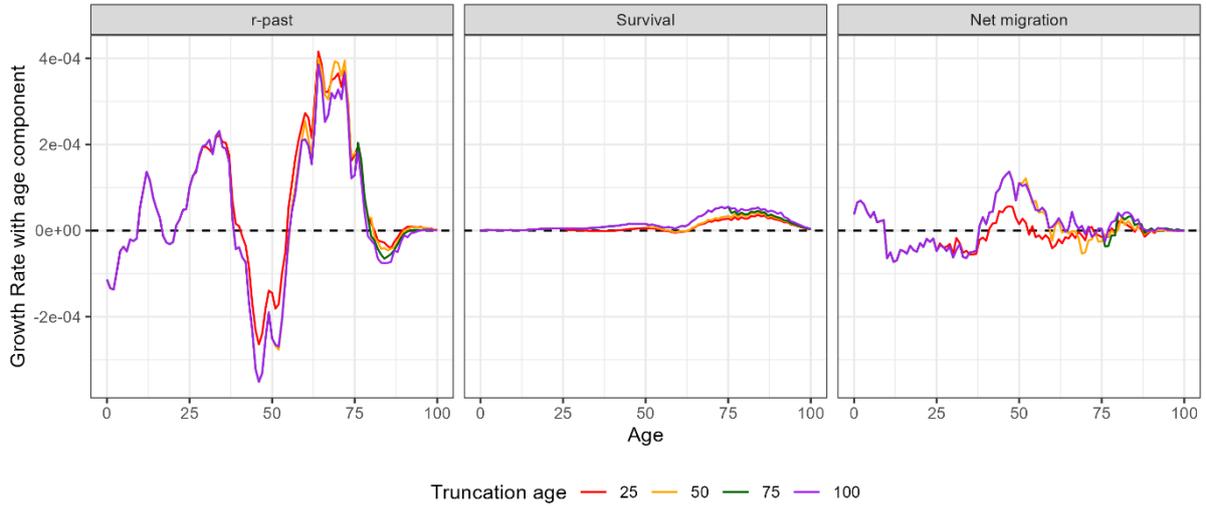
**Figure 2.** US female age-specific growth rates between 2010 and 2020,  $r_x(t)$ , and its components in Eq. (2) with 25, 50, 75 and 100 years of historical data: Age-specific growth rate at a younger age  $y$ , or  $r$ -past, cohort survival change and cohort net migration change between ages  $y$  and  $x$ .



Source: authors' calculations based on data described in Table A1 in the SOM.

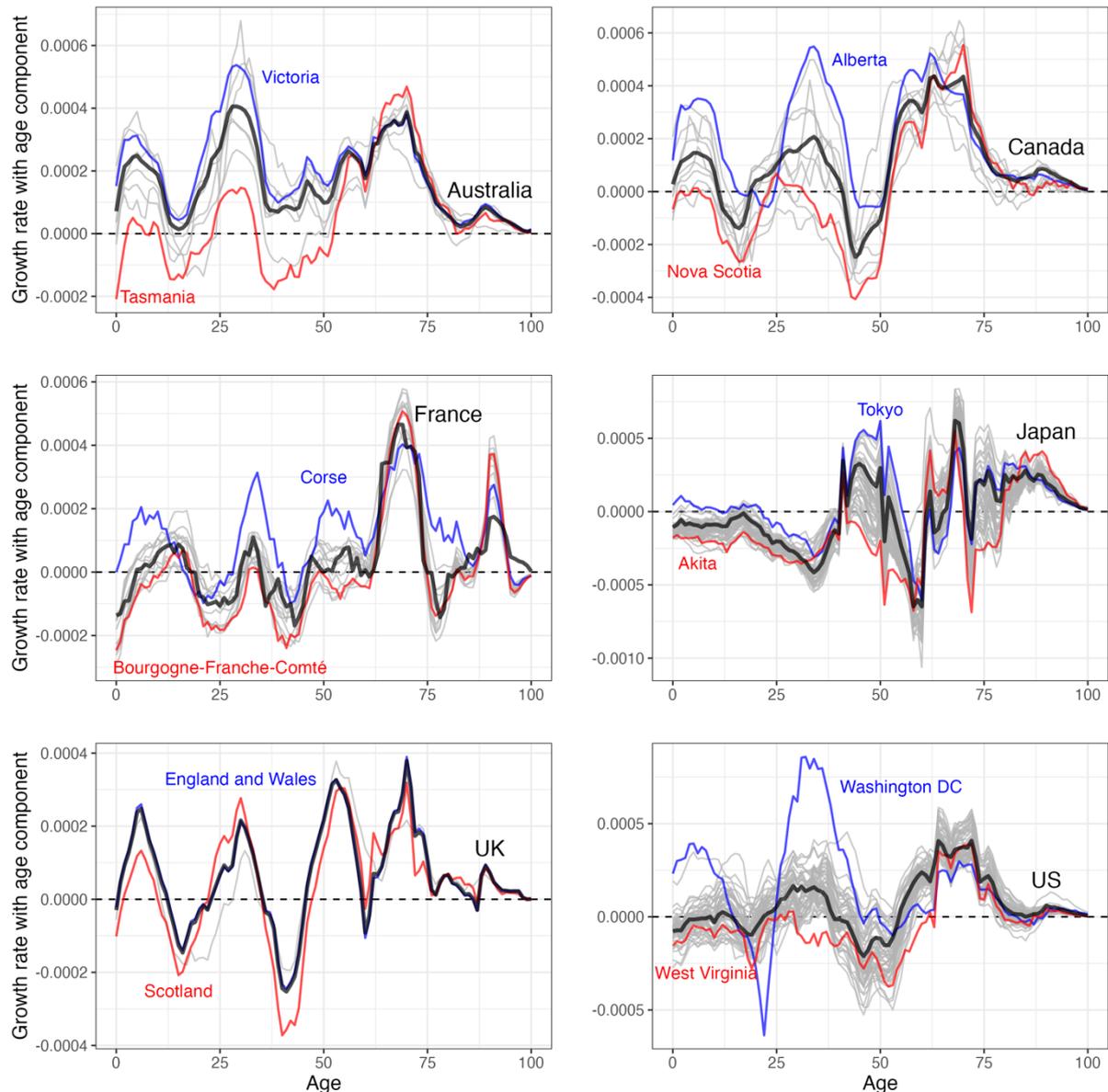
Notes: the vertical dashed-line in Figure 2 separates the young ages to the left as those with full cohort historical information and those on the right without that (see SOM Appendix 2 for more details). Scientific notation is used on the vertical-axis of Figures to highlight small numbers (Burch 2022).

**Figure 3.** Components of the US female age-specific growth rates between 2010 and 2020, as in Eq. (2), with 25, 50, 75 and 100 years of historical data: Age-specific growth rate at a younger age  $y$ , or  $r$ -past, cohort survival change and cohort net migration change between ages  $y$  and  $x$ .



Source: authors' calculations based on data described in Table A1 in the SOM.

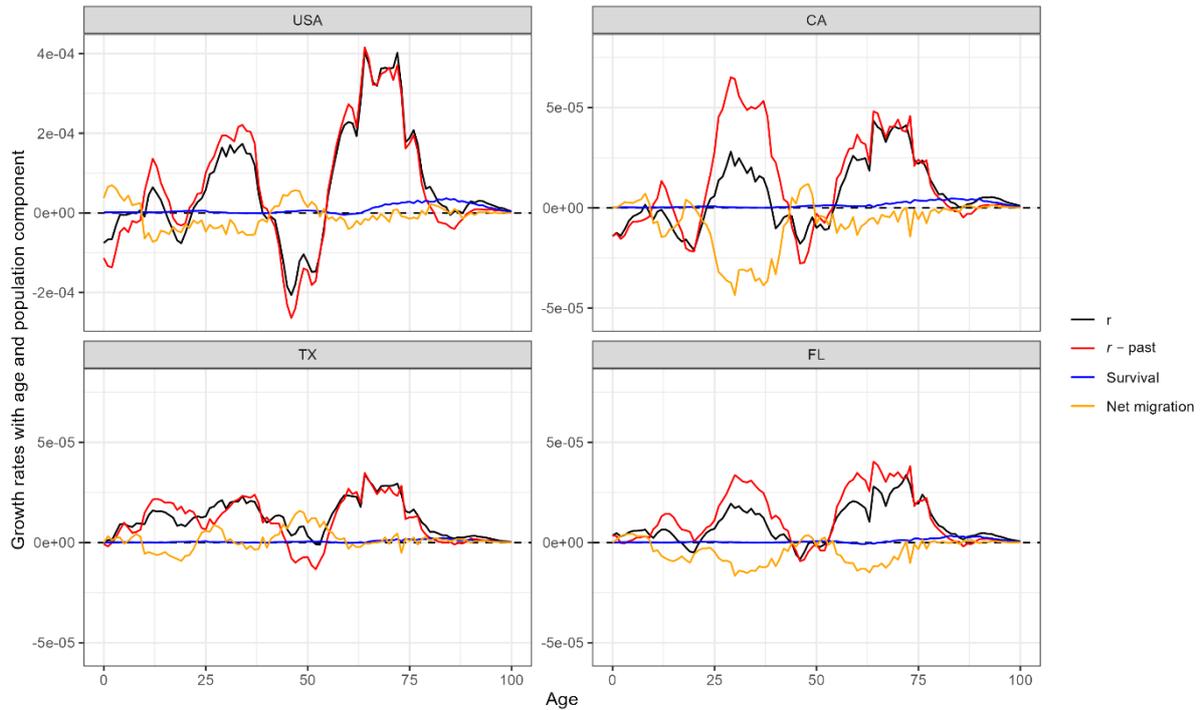
**Figure 4.** National and subnational female age-specific growth rates (multiplied by population composition) for selected HMD countries. The subnational populations with the highest and lowest subnational overall growth rate are highlighted, 2010-2020.



Source: authors' calculations based on data described in Table A1 in the SOM.

Note: For the United Kingdom, England and Wales corresponds to 89% of the population, so this subpopulation and the national age-specific growth rates have practically the same age-pattern.

**Figure 5.** US female age-specific growth rates between 2010 and 2020, based on 25 years of historical information on survival, net migration and past growth rate ( $r$ -past) as in Eq. (2), and the contribution to those components by the most populous states as in Eq. (3): California (CA), Florida (FL) and Texas (TX).



Source: authors' calculations based on data described in Table A1 in the SOM.

Notes: National vertical-axis scale is different than for the state, since for the state results there is the additional product with the state composition in the nation,  $c_i(t)$  term in Eq. (3).

## **Supplementary Online Material**

### *Content*

*Appendix 1. National and subnational databases used*

*Appendix 2. Methodological notes*

*Appendix 3. Tables and Figures*

Author Manuscript

### **Appendix 1. National and subnational databases used**

Table A1 presents the available Human Mortality Database information at the national level, as well as the subnational HMD data for the countries studied. All these databases result from collaborations between country specialists with the HMD team to produce country specific information at the subnational level. All these populations use the same methods protocol as the HMD for calculating life tables. In other words, each subnational population is treated as an independent population from the rest of the subnational and national data. This could create problematic trends, since coherence between trends is not taken into account, as studies of forecasting do (Pascariu et al. 2018; Bergeron-Boucher et al. 2018). Although, the subnational population counts are not affected by such issues, further assessment is needed for the death rates. Within the subnational data there were also different counts of populations and age-specific death rates which had to be matched, and the shortest of the two series was taken here. Population counts for the subnational populations in the French FHMD and American USMDB were obtained directly from the developers of those collections, namely: Florian Bonnet and Magali Barbieri respectively. We also sourced US Census Bureau (2020) for the US subnational population from 2019 to 2020 and Eurostat (2024) for French subnational population with five-year age group from 2016 to 2020.

**Table A1.** National and subnational databases used in the analysis.

<b>Country</b>	<b>National (HMD)</b>	<b>Subnational population</b>	<b>Subnational death rates</b>
Australia (AHMD)	1921-2021	1972-2021	1971-2021
Canada (CHMD)	1921-2021	1950-2021	1950-2021
France (FHMD)	1816 - 2021	1946-2020	1901-2021
Japan (JMD)	1947-2021	1975-2021	1975-2021
UK (HMD)	1922-2021	1922-2021	1922-2021
US (USMDB)	1933-2021	1980-2020	1959-2021

Note: The latest year for the Canadian territory of Yukon is 2016 and it was not included in the analysis.

## Appendix 2. Methodological notes

The population growth rate is the relative derivative of population counts with respect to time, denoted as  $r(t) = \frac{\dot{N}(t)}{N(t)}$ , where a dot on top of a variable represents the derivative with respect to time of the population count at time  $t$ ,  $N(t)$ . The age-specific population growth rate can hence be represented in the same manner as  $r_x(t) = \frac{\dot{N}_x(t)}{N_x(t)}$ . The population growth rate is the weighted average

$$r(t) = \int_0^{\omega} r_x(t) c_x(t) dx, \quad (\text{A1})$$

where  $c_x(t) = \frac{N_x(t)}{N(t)}$  corresponds to the population composition at age  $x$  and time  $t$ .

To decompose the population growth between two times we apply an extension of variable- $r$  method (Canudas-Romo et al. 2021). In the truncated variable- $r$  decomposition for each cohort the population growth rate at age  $x$  and time  $t$ ,  $r_x(t)$ , are separated into the growth at an earlier age  $y$  and time  $t - x + y$ , or  $r_y(t - x + y)$ , and the cohort and net migration changes between the two ages as

$$r_x(t) = r_y(t - x + y) + \Delta S^{t-x}(y, x) + \Delta M^{t-x}(y, x), \quad (\text{A2})$$

where  $\Delta S^{t-x}(y, x) = \frac{\dot{s}^{t-x}(y, x)}{s^{t-x}(y, x)}$  and  $\Delta M^{t-x}(y, x) = \frac{\dot{m}^{t-x}(y, x)}{m^{t-x}(y, x)}$  represent the relative derivatives in cohort survival and net migration probabilities from the younger age  $y$  to the older age  $x$ . The change in net migration  $\Delta M^{t-x}(y, x)$  is estimated as the residuals of Eq. (A2) similar to previous studies (Preston and Stokes 2012; Preston and Vierboom 2021). In this paper, since we aim to decompose the population growth between two time periods, we need to calculate the age-specific growth rates  $r_x(t)$  from ages 0 to 100 using Eq. (A1-2) and maintaining the age gap  $x-y$  constant over all ages. For the young ages with values less than the interval of study,  $x < (x - y)$ , Eq. (2) is modified by taking only the small range of information and the growth rate at birth for those cohorts as

$$r_x(t) = r_B(t - x) + \Delta S^{t-x}(0, x) + \Delta M^{t-x}(0, x), \quad (\text{A3})$$

where  $r_B(t - x) = \frac{\dot{B}(t-x)}{B(t-x)}$  is the growth rate at birth at time  $t-x$ .

Figure A1 presents a visual representation of the information used in the analysis for three different ages (20, 50 and 100) and the growth rate components of those ages as presented in

Eq. A2 and A3: the age-specific growth rates between 2010 and 2020, the age-specific growth rates of earlier years 1985-1995, and the survival between those years.

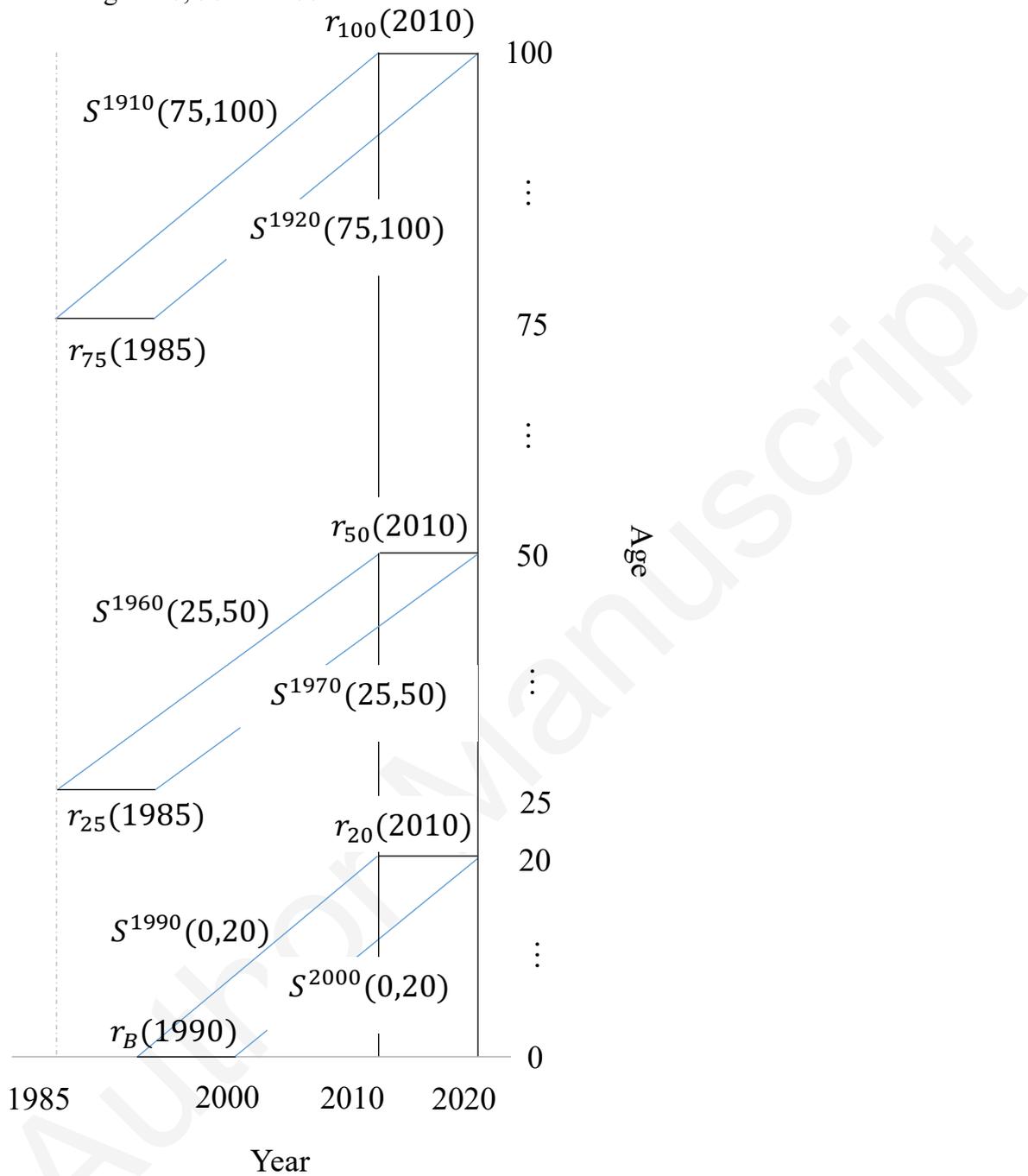
For example, at age 50 the growth rate  $r_{50}(2010)$  is calculated over 10 years between 2010 and 2020, and following Eq.(A2) used the growth  $r_{25}(1985)$ , and the survival changes  $\Delta S^{1960}(25,50)$ , or changes between  $S^{1960}(25,50)$  and  $S^{1970}(25,50)$ .

However, at ages younger than 25 we used instead Eq. (A3). For example, at age 20 the growth rate  $r_{20}(2010)$  following Eq.(A3) used the growth rate at birth  $r_B(1990)$ , and the survival changes  $\Delta S^{1990}(0,20)$ , or changes between  $S^{1990}(0,20)$  and  $S^{2000}(0,20)$ . As such, younger cohorts include all their cohort information – in our case younger cohort were those *less than*  $t_2 - t_1 - 10 = 2020 - 1985 - 10 = 25$ , with the 10 corresponding to the number of years used for the calculations of growth rates.

Finally, for all these cases the migration component was calculated as the residual term of the difference between the growth rate in 2010, minus that in 1985, and minus the survival between them. For example, for age one hundred we set  $\Delta M^{1910}(75,100)$  as equal to  $r_{100}(2010)$  minus the younger growth rate  $r_{75}(1985)$  and the survival changes  $\Delta S^{1910}(75,100)$ .

For the examples presented in this study and the schematic application in Figure A1, the current and earlier growth rates, as well as the growth rate at birth, are all of the same length—10 calendar years. However, as emphasized in this study, the number of years of historical data used can vary from 25 to 100 years, or other chosen lengths of time.

Figure A1. Depiction of the population growth rate between 2010 and 2020 and its components for three ages: 20, 50 and 100.



***Approximating Continuous Change***

Our equations in the main text and in Appendix 2, are in continuous, however data is found only annually and approximations are needed for: the relative derivative, midpoint calculations, and derivatives. We implemented standard approximations for those terms (Vaupel and Canudas-Romo 2003; Preston et al. 2001).

Given a demographic function  $v(x,t)$  measured at two time points  $t$  and  $t+h$ , we approximated the relative derivative assuming a constant rate of change over time as

$$\frac{\dot{v}(x,t+h/2)}{v(x,t+h/2)} \approx \frac{\ln\left[\frac{v(x,t+h)}{v(x,t)}\right]}{h}. \quad (\text{A4})$$

We approximate the midpoint for a function using

$$v(x, t + h/2) \approx [v(x, t)v(x, t + h)]^{\frac{1}{2}}, \quad (\text{A5})$$

and the derivative with respect to time by

$$\dot{v}(x, t + h/2) = \left[\frac{\dot{v}(x,t+h/2)}{v(x,t+h/2)}\right] v(x, t + h/2). \quad (\text{A6})$$

**Appendix 3. Tables and Figures**

Table A2 presents the results for the US female population growth rate between years 2010 and 2020, when using different historical information: 25, 50, 75 and 100 years. Although the population growth rate is the same irrespective of the length of the historical information, the contribution of each of the components changes with the amount of data. These results serve as sensitivity analysis of the length of the demographic series and its impact on the survival, net migration and past-growth components as described in Eq. (2). As observed in Table A2, with the increase in demographic data the  $r$ -past reduces, and the survival and net migration increase in their contribution. In other words, using a shorter demographic history leads to more information being contained in the  $r$ -past term, while using a longer time-series sees this information distributed among the other two components.

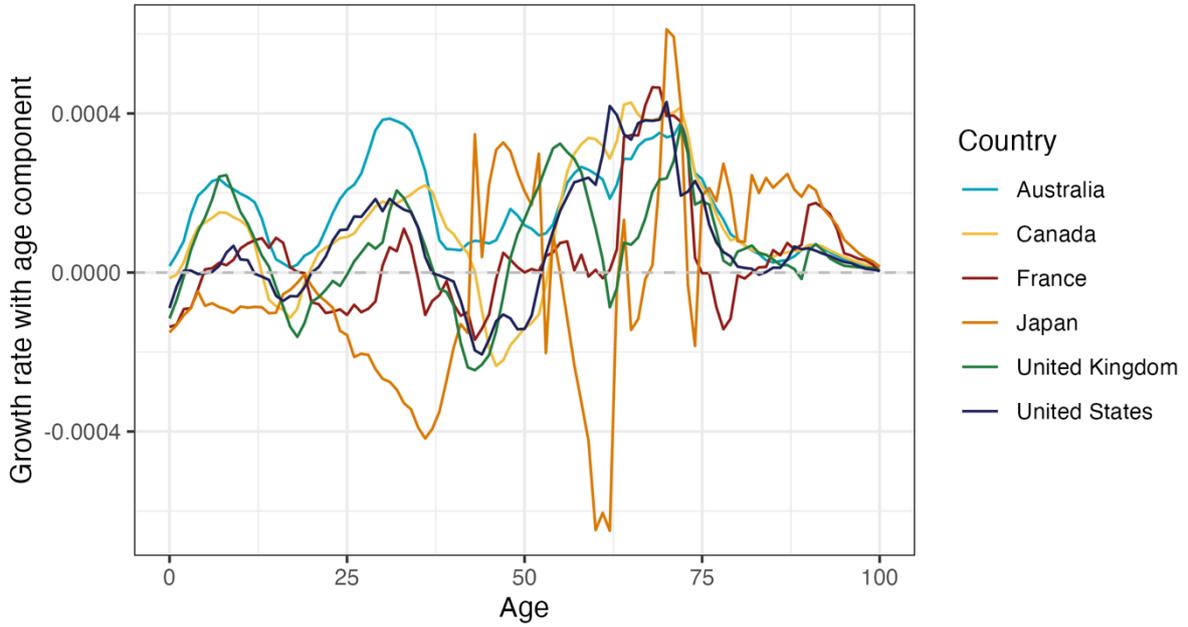
**Table A2.** US female population growth rate between 2010 and 2020, based on historical information on survival, migration and past growth rate of 25, 50, 75 and 100 years.

<b>Historical information (years)</b>	<b>Growth rate <math>r(t)</math></b>	<b><math>r</math>-past</b>	<b>Survival</b>	<b>Net migration</b>
25	6.50	6.37	0.89	-0.76
50	6.50	4.67	1.26	0.57
75	6.50	3.84	1.68	0.98
100	6.50	3.52	1.85	1.13

Source: authors' calculations based on data described in Table A1 in the SOM.

Note: Values have been multiplied by 1000.

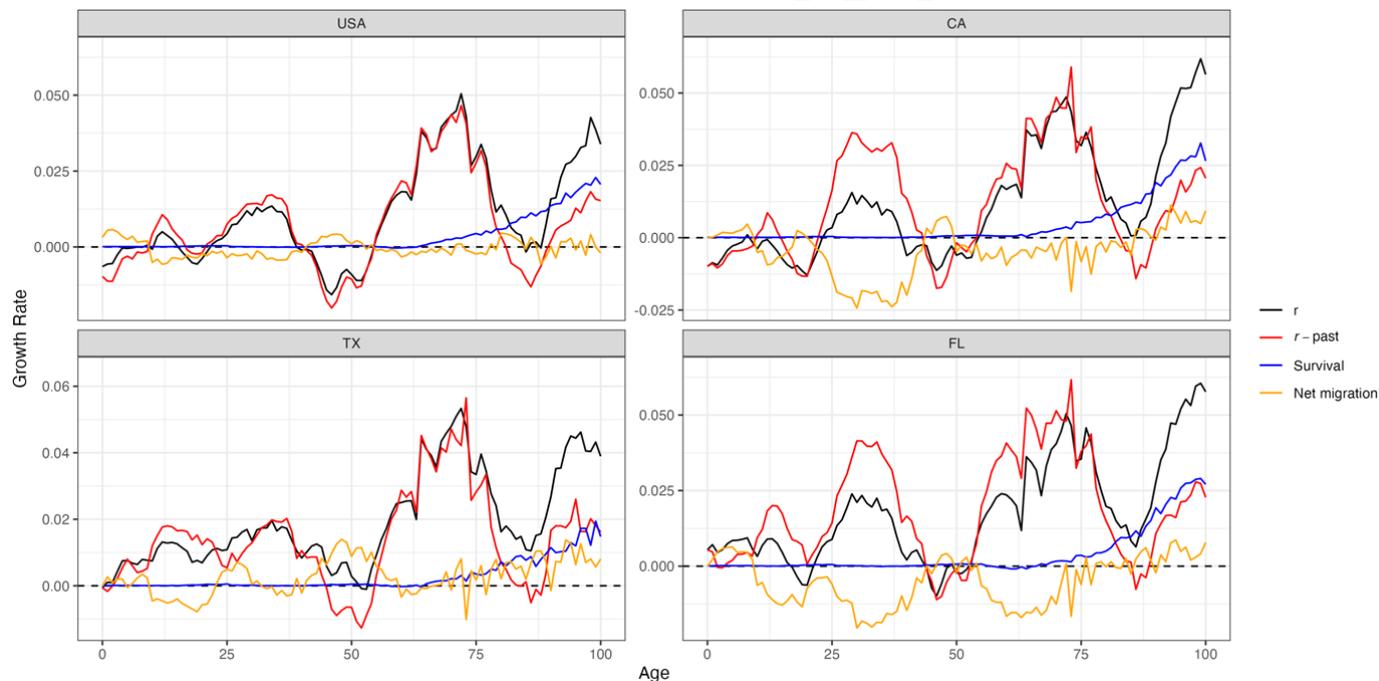
**Figure A2.** Age-specific female population growth rate between 2010 and 2020,  $r_x(t)$ , for selected countries.



Source: authors' calculations based on HMD (2023) data.

Figure A3 is analogous to Figure 4 in the main text but without the age composition,  $c_x(t)$ , from Eq. (2), nor the state composition  $c_i(t)$  from Eq. (3). The high growth rates at old ages and the high contribution of improvements in survival at those ages are the two most evident contrasts between Figure A3 and the composition weighted Figure 4. The latter Figure 4 takes into account the small population composition at older ages, respect to the total, and weakens the contribution of the progress in survival in age-specific growth rates observed in Figure A3.

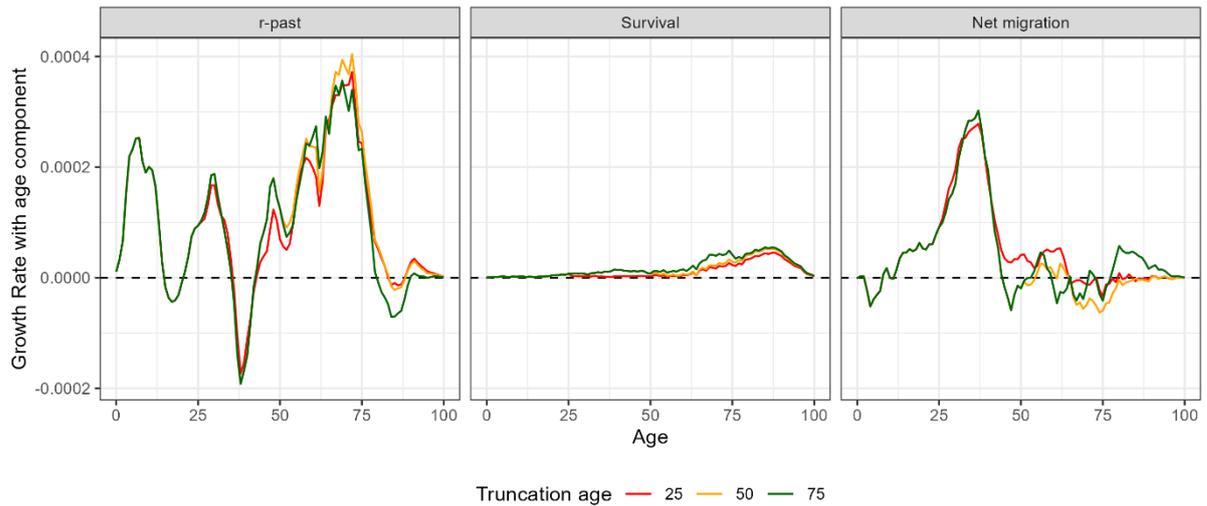
**Figure A3.** US, California, Texas and Florida female age-specific growth rates between 2010 and 2020, based on 25 years of historical information on survival, migration and past growth rate ( $r$ -past), but without the age composition,  $c_{x,i}(t)$ , from Eq. (2) or Eq. (3).



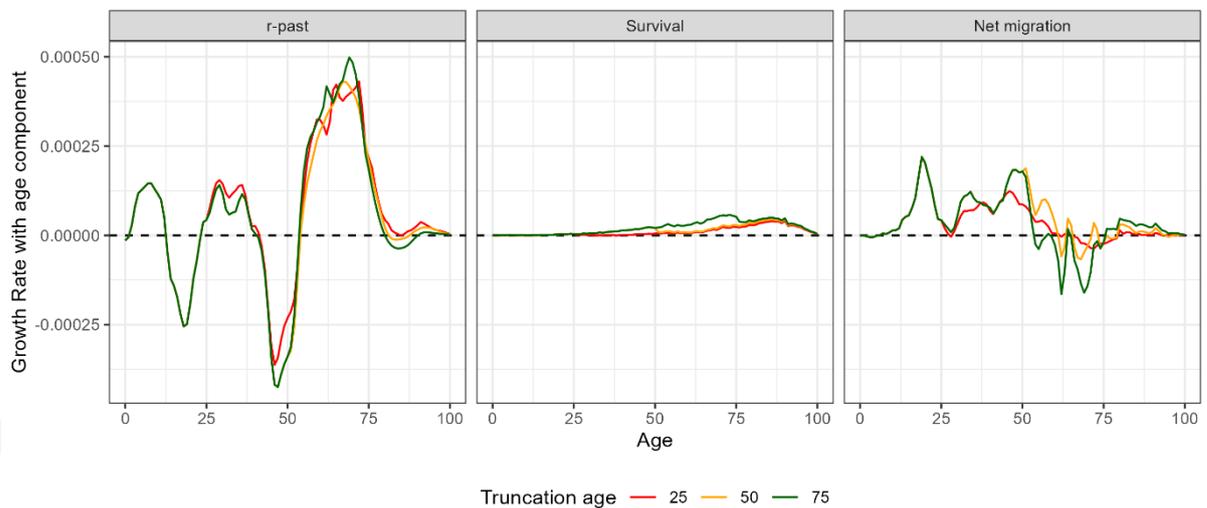
Source: authors' calculations based on data described in Table A1 in the SOM.

**Figure A4.** Components of the female age-specific growth rates between 2010 and 2020 for five countries, as in Eq. (2), with 25, 50, 75 and 100 years of historical data (depending on availability of historical information): Age-specific growth rate at a younger age  $y$ , or  $r$ -past, cohort survival change and cohort net migration change between ages  $y$  and  $x$ .

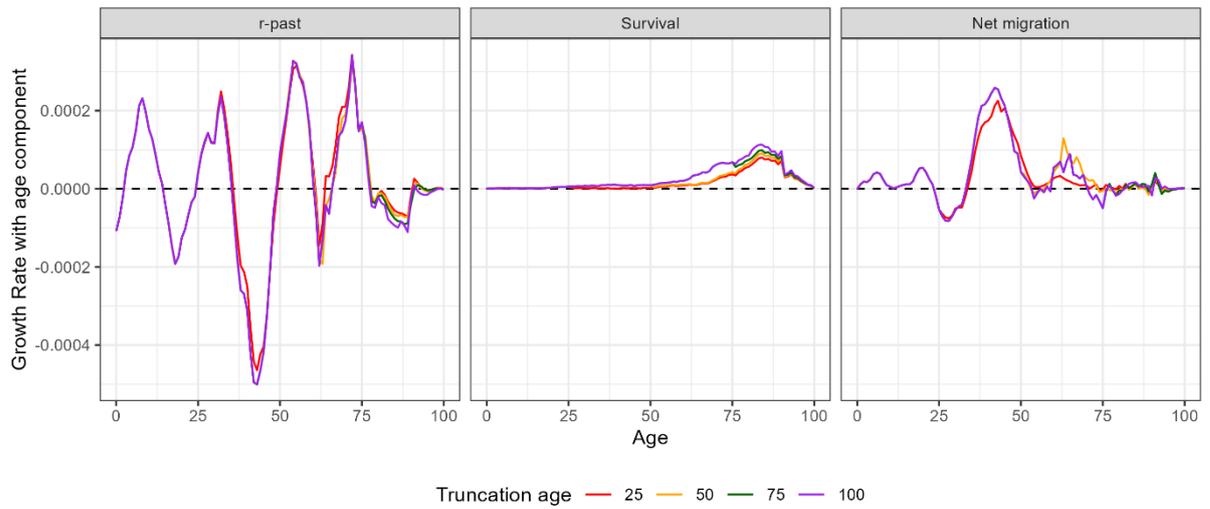
**a. Australia**



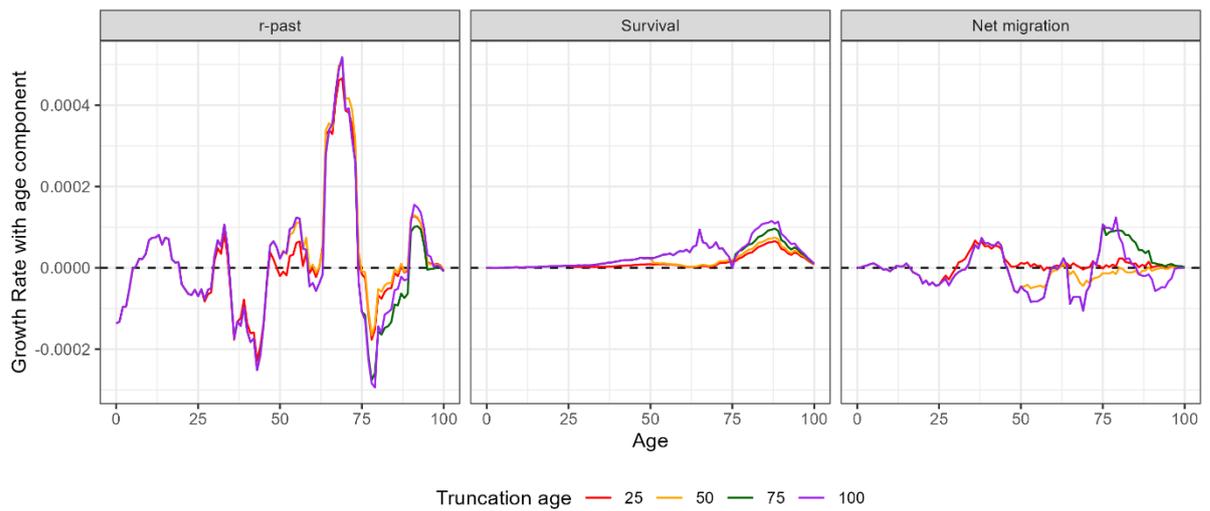
**b. Canada**



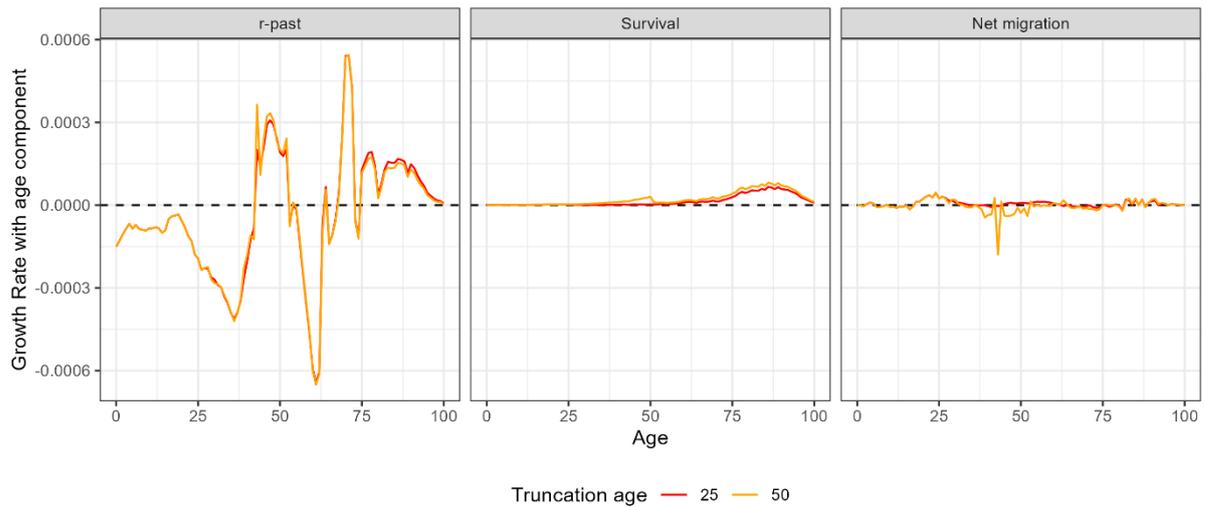
**c. England and Wales**



**d. France**



**e. Japan**



## References

Australian Human Mortality Database (AHMD). The Australian National University.

Available at: [demography.cass.anu.edu.au/research/australian-human-mortality-database](https://demography.cass.anu.edu.au/research/australian-human-mortality-database)

Bergeron-Boucher, M.P., Canudas-Romo, V., Pascariu M. and Lindahl-Jacobsen, R. (2018).

Modelling and forecasting sex differences in mortality: a sex-ratio approach. *Genus*, 74(20).

Canadian Human Mortality Database (CHMD). Université de Montréal (Canada), Max Planck Institute for Demographic Research (Germany), and University of California, Berkeley (USA). Available at: [www.bdlc.umontreal.ca/CHMD/index.htm](http://www.bdlc.umontreal.ca/CHMD/index.htm)

Canudas-Romo, V., Shen, T., & Payne, C. (2021). The role of reductions in old-age mortality in old-age population growth. *Demographic Research*, 44:1073-1084.

Eurostat (2024). *Population on 1 January by age group, sex and NUTS 3 region*. Available at: [https://doi.org/10.2908/DEMO\\_R\\_PJANGRP3](https://doi.org/10.2908/DEMO_R_PJANGRP3)

French Human Mortality Database (FHMD). French Institute for Demographic Studies and the Paris School of Economics (France).

Available at: [frdata.org/fr/french-human-mortality-database/](http://frdata.org/fr/french-human-mortality-database/)

Human Mortality Database (HMD). Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at: [www.mortality.org](http://www.mortality.org)

Japanese Mortality Database (JMD). National Institute of Population and Social Security Research. Available at: [www.ipss.go.jp/p-toukei/JMD/index-en.asp](http://www.ipss.go.jp/p-toukei/JMD/index-en.asp)

- Pascariu, M., Canudas-Romo, V. and Vaupel, J.W. (2018). The Double-Gap Life Expectancy Forecasting Model. *Insurance: Mathematics and Economics*, 78:339–350.
- Preston, S., P. Heuveline, and M. Guillot. (2001). *Demography: Modelling and measuring population processes*. Oxford: Blackwell.
- Preston, S. H., & Stokes, A. (2012). Sources of population aging in more and less developed countries. *Population and Development Review*, 38:221–236.
- Preston, S. H., & Vierboom, Y. C. (2021). The changing age distribution of the United States. *Population and Development Review*, 47:527–539.
- Vaupel, J.W. and Canudas-Romo V. (2003). Decomposing Change in Life Expectancy: A Bouquet of Formulas in Honor of Nathan Keyfitz's 90th Birthday. *Demography*, 40(2): 201-216.
- US Census Bureau. (2021). *State Population by Characteristics: 2010-2020*. Available at: <https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-state-detail.html>
- United States Mortality Database (USMDB). Centers for Disease Control (CDC, US), Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at: [usa.mortality.org](http://usa.mortality.org)