International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: repository@iiasa.ac.at
Web: www.iiasa.ac.at

Working Paper **WP-16-010**

Expectation of Life at Old Age: Revisiting Horiuchi-Coale and Reconciling with Mitra

Dalkhat M. Ediev (ediev@iiasa.ac.at)

Approved by

Wolfgang Lutz

Program Director, World Population Program

August 2016

Contents

Abstract

Horiuchi & Coale (1982) have suggested a simple formula to improve the estimates of the expectation of life at old age, a formula that was questioned as unrealistic by Mitra (1984; 1985), whose alternative approach was in return criticized for being prone to biases due to age exaggeration (Coale 1985). The discussion between the researchers have not been resolved. This lack of agreement and the simplifying assumptions behind both approaches prevented their integration in the practical toolkit of demographic analysis. This paper shows that both methods are useful in drastically reducing the estimation errors in life expectancy. The disputes were largely due to inconsistent inputs rather than methodological contradictions. Furthermore, the methods produce by far better results as compared to the popular alternative, the extrapolation of the death rates. In combination the two alternative methods may even lead to a better outcome.

Acknowledgments

We thank Prof. Shiro Horiuchi, Dr. Patrick Gerland and participants of the Colloquium at the Vienna Institute of Demography for their valuable comments. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no ERC2012-AdG 323947-Re-Ageing.

About the Author

Dalkhat Ediev is a Research Scholar with the World Population (POP) Program at the Wittgenstein Centre (IIASA, VID/ÖAW, WU), International Institute for Applied Systems Analysis, and Institute Director and Professor with the North-Caucasian State Humanitarian-Technological Academy, Institute for Applied Mathematics and Information Technologies. His research interests include population change and projections, population ageing, formal demography, demographic methods and models. Dr. Ediev graduated from Moscow Institute of Physics and Technology in 1993, where he also obtained his PhD in 1999. He was granted the Docent degree in mathematics (2002) and the degree of Doktor Nauk in Physical-Mathematical Sciences (2008) by the Russian Highest Attestation Commission.

Expectation of Life at Old Age: Revisiting Horiuchi-Coale and Reconciling with Mitra

Dalkhat M. Ediev

1. Introduction

The life table model, which describes the survival of individuals throughout their life course, is an important tool in studying mortality and its many implications such as insurance policies, social policies and population projections. It builds upon age-specific death rates and produces various indicators of mortality, survival and longevity for the age groups. A limitation of the model appears at the older age where data scarcity or deficiency forces statisticians to disregard age details and aggregate the available data into a single "open age interval".

Typical data problems that prevent the extension of the life table to an older age are age exaggeration and other types of age misreporting, such as poorly documented return migration following retirement and missing death records. Currently, different approaches come into operation to solve or circumvent this hitch. The World Health Organization (WHO) (Mathers & Ho 2014) used to extrapolate the death rates above the age of 85 by assuming a logistic mortality model. The United Nations (1982) construct life tables for developing countries by fitting the Gompertz-Makeham model at younger ages and closing at age 85+. The Human Mortality Database (HMD) (Wilmoth et al. 2007) also corrects the original data at ages 80+. A recent study by Randall and Coast (2016) suggests that the data quality at ages 60+ in low-income countries is yet "very rough" with only a limited improvement over time in African countries. Even in developed countries with generally good mortality data, incompleteness of migrant registration or misreported ages at death may bias the mortality estimates for the elderly (Preston et al. 1996; Khlat & Courbage 1996; Kibele et al. 2008). In regional demography and demography of social groups, small population size may be another source of data limitation that demands lowering the age at the start of the open age interval, so that to include at least some death events (Scherbov & Ediev 2011).

Choosing a broader open age interval that begins at younger age may help mitigate some of the data quality problems, including the problems of age exaggeration and small population size. Lowering the age at the beginning of the open age interval, however, may itself have severe consequences for the quality of life table estimates because of departures from stationarity which is conventionally assumed for the age composition at the open age interval (Preston et al. 2001). In a stationary population, life expectancy at the beginning of the open age interval is inverse to the death rate:

$$e_a = {M_{a+}}^{-1}, \qquad (1)$$

hereinafter, a denotes the starting age of the open age interval; e_a is the life expectancy at age a; and M_{a+} is the death rate in the open age interval. In this paper, we call (1) *the classical estimate (approach).*

The classical approach is not the only one available for cases with age reporting problems. The most common alternative method to calculate the last 'problematic' age range of the life table is *extrapolation* of death rates based on their change at younger ages in combination with a mortality model (United Nations 1982; Wilmoth et al. 2007; Mathers & Ho 2014; Missov et al. 2016). Extrapolation does not assume population stationarity or any other population model. Such extrapolation, however, ignores the original empirical data pertaining to the open age interval. Furthermore, as we demonstrate in this paper, the extrapolation method tends to be less accurate in terms of life expectancy than the other methods considered here.

Two other alternatives to the classical approach also work with the open age interval, just like the classical method, but relax the stationarity assumption presumed in (1). Indeed, populations are rarely stationary. Improving survival, changing fertility and immigration modify the population age composition. Growing populations are typically of younger age composition, with more weight on younger ages with lower mortality (Preston et al. 2001); their death rate in the open age interval is lower than in a stationary population with similar mortality. As a result, the classical method is prone to overestimate life expectancies for such populations. A step forward as compared to the zero-growth stationary population model is the stable population model that assumes, generally speaking, that allows for population growth and produces population structures younger or older than the corresponding stationary populations (Preston et al. 2001). This model – in combination with the Gompertz mortality model – was used by Horiuchi and Coale (1982) to develop the following simple formula to improve the classical estimate:

$$e_a = {M_{a+}}^{-1} e^{-\beta_a r {M_{a+}}^{-\alpha_a}}, \qquad (2)$$

here, *r* is the annual growth rate of the population in the open age interval; α_a and β_a are the model parameters (for numerical values, see Horiuchi & Coale 1982 or Table 1).

Adjustment (2) was challenged by Mitra (1984; 1985) who also assumed population stability and derived a closed form solution:

$$e_a = {M_{a+}}^{-1} e^{-r\left[{M_{a+}}^{-1} - \left(1 + r{M_{a+}}^{-1}\right)(\bar{x} - a)\right]}, \qquad (3)$$

where $\bar{x}$ stands for the mean age of the population in the open age interval. Using his method, Mitra came to estimates fairly close to the classical ones and suggested that the adjustment by Horiuchi-Coale was unrealistic. Mitra's approach was, however, criticised for being prone to biases due to age exaggeration (Coale 1985). Indeed, having age exaggeration as a problem in the first place, one would cautiously use the empirical mean population age $\bar{x}$ in an adjustment procedure.

Partly because the discussion between Horiuchi-Coale and Mitra was never resolved, but also due to the strong assumptions used in both approaches, their methods have not made it to a wider practical use by demographers and population statisticians.

Two developments since the time of discussion between Horiuchi-Coale and Mitra call for revisiting their results. First, the empirical basis for mortality studies and computational resources have advanced substantially, as it is reflected in the rich collection of high quality data in the HMD (2016). This enables the development and testing of methods on a sounder empirical base, and relaxes some of the original assumptions. Second, life expectancy in many countries has systematically advanced since 1980. On the one hand, better survival to old age boosts the importance of the open age interval for life table estimates. On the other hand, higher (and improving) life expectancies call for testing if the old methodology works well on current data. Our paper is a response to these needs and introduces some useful modifications to the original adjustment formulas. We also aim at reconciling the dispute between Horiuchi-Coale and Mitra and combining both alternative methods for a better outcome.

Table 1. Original parameters of the Horiuchi-Coale model (Alfa, Beta), the Beta-parameter of the model re-estimated on the Human Mortality Database data (Beta.hmd), and the coefficients (estimated on HMD data) of the regression model (5) for the mean population age in the open age interval (C, k_1, k_2).

Sex	a	Alfa	Beta	Beta.hmd	C	k_1	k_2
Female	40	1.0	0.283	0.321	50.045	0.241	-4.918
Female	55	1.1	0.207	0.241	61.025	0.303	-4.503
Female	65	1.4	0.095	0.100	69.200	0.335	-3.670
Female	75	1.4	0.095	0.109	77.701	0.380	-2.676
Female	85	1.4	0.095	0.104	86.460	0.470	-1.883
Female	95	1.4	0.095	0.062	95.591	0.626	-0.867
Male	40	1.0	0.283	0.330	50.924	0.196	-3.919
Male	55	1.1	0.207	0.236	61.406	0.269	-3.722
Male	65	1.4	0.095	0.102	69.229	0.318	-3.180
Male	75	1.4	0.095	0.108	77.563	0.379	-2.398
Male	85	1.4	0.095	0.102	86.355	0.482	-1.863
Male	95	1.4	0.095	0.058	95.633	0.609	-0.914
Total	40	1.0	0.283	0.308	50.839	0.206	-3.849
Total	55	1.1	0.207	0.234	61.115	0.293	-4.030
Total	65	1.4	0.095	0.099	69.117	0.335	-3.324
Total	75	1.4	0.095	0.108	77.583	0.387	-2.481
Total	85	1.4	0.095	0.102	86.405	0.477	-1.803
Total	95	1.4	0.095	0.061	95.518	0.658	-0.929

Note: a = starting age of the open age interval.

2. Testing the models of expectation of life at old age on empirical data

To examine the biases in estimated life expectancy for selected estimation methods at different open age intervals, we use all period life tables, single-year mortality data, and population exposures contained in the HMD (2016)[1]. Altogether, the database (data downloaded on 12.02.2016) contains 4436 country-calendar years for each gender (males, females, total), for 46 countries/populations, spans over the years 1751 to 2014, and life expectancies at birth from 16.7 to 86.4 years. For each of the 3x4436 database entries we recalculate its life table assuming alternative open age intervals (the threshold age a spanning from 55 to 95 years) and various estimation methods.

For each open age interval, we calculate the aggregated death rate for the open age interval using the death rates and population exposures from the HMD:

$$M_{a+} = \frac{\sum_{x=a}^{\omega} M_x P_x}{\sum_{x=a}^{\omega} P_x}, \quad (4)$$

here M_x and P_x are the HMD death rate and the population exposure at age x; and ω is the maximum attainable age group (110+ in HMD). After obtaining the death rate for the open age interval, we apply alternative life table methods and compare the results to the life table that is based on the full age scale spanning up to the age 110+.

First, we apply the classical life table method (1) and use estimates (4) and the death rates at ages below age a to calculate a new "truncated" life table for each HMD entry. Life expectancy from the truncated life table is compared to life expectancy from the full life table; where the "full" life table stands for the life table with the maximum possible, a=110. The difference between the two gives the estimation error in the classical life table method due to the truncation of the data at the selected open age interval.

In a similar fashion, we examine the estimation errors in the methods of Horiuchi-Coale (2) and Mitra (3), and in the popular method where death rates are extrapolated into the open age interval based on their rate of increase at younger age. For extrapolations, we use the Gompertz model (Gompertz 1825; Doray 2008) where the force of mortality increases exponentially with age. After some experimentation, we opt for the marginally better extrapolation without jump of the death rate at age a; with parameters fit on 20-years-of-age-long age intervals below the open age interval. This extrapolation is close to the common practices for low-quality data cases. The Gompertz model is also useful as a bridge to the work by Horiuchi-Coale who relied on the model in developing their own method.

Results for life expectancy at birth in the classical life table method for open age intervals 55+, 65+, 75+, 85+, and 95+ are presented in Figure 1 and Table 2. On the

[1] HMD modifies (smooths and extrapolates) the original data at some ages above age 80 using the logistic model. This might have distorted our findings for the highest open age intervals, especially for the extrapolation method. Yet, after re-running our calculations on the raw, not modified, death rates also provided in the HMD, we came to results similar to those presented in this paper. Root mean squared errors (RMSE) obtained on raw data differ only in the second digit after comma from the RMSEs presented in Table 2 for all methods, except for the extrapolation. Even for the extrapolation method, there was only one case when RMSE based on the raw data differed in the first digit after comma from the results presented in Table 2 and was substantial in relative terms: females, life expectancy at birth range 40-50 years, a= 85 (RMSE on the smoothed death rates: 0.04; RMSE on the raw data: 0.16).

horizontal scale of the figure, we put the life expectancy at birth from the full life table, and the vertical scale represents the difference between life expectancies at birth in the truncated and full life tables.

The estimation errors in the classical method are predominantly positive because of the worldwide growth of elderly population in the course of demographic transition, which produced population age structures younger (of lower mortality) than the structure of stationary populations assumed in the method. The biases were relatively small (yet, quite substantial) for periods with shorter life expectancy at birth and soared to high levels as life expectancy grew. This comes well in agreement with the formal derivations by Horiuchi-Coale and Mitra. Currently, closing the life table at age 65, both sexes combined, would produce an upward bias in life expectancy at birth as high as 10 years in some countries and more than 2.5 years in many other cases. These errors, with the secular trend of life expectancy increasing by about two years per decade (Oeppen & Vaupel 2002; White 2002), correspond to gains of life expectancy over 12-50 years. If data permits closing the life table at age 75 or 85, the estimation errors are down to under 4 or 0.5 years respectively. Closing the life tables at age 95 yields only minor errors in e_0, while closing at age 55 drives the errors to unacceptably high levels. The errors are generally higher when considering only the female populations or both sexes together.

In Figure 2 and Table 2, we show the estimation errors in life expectancy e_a at the beginning of the open age interval as percent of the life expectancy from the corresponding full life table. Our interest in this kind of estimation errors is driven by the usual practice in population projections, where the population of the open age interval is projected on the basis of the life table death rate for that age group (Preston et al. 2001). As Figure 2 illustrates, substituting the actual population change in the open age interval by its estimate, based on the life table death rate, may lead to annual downward biases of dozens of percent in the numbers of deaths in the open age interval.

Numerical results for the adjustments by Horiuchi-Coale and Mitra as well as the extrapolation method are presented in Figure 3 and Table 2 (as a refernce, we also put the results for the classical method into Figure 3). We improved the stability of the Horiuchi-Coale and Mitra formulas by using the population growth rates averaged over 10-year periods prior to the estimaton year. If we used the annual rates (results not presented here), the adjusted life expectancies would contain more outliers, especially in the Mitra method.

In all procedures, errors tend to increase as life expectancy grows. Interestingly, all methods tend to err more often to the positive side, i.e. they overestimate life expectancy, although the non-classical methods are free from the classical method's sources of error that is nested in the stationarity assumption. The reasons for the positive errors are different among the methods. The extrapolation method produces positive or negative errors depending on wether the death rates increase steeper above or below the minimal age of the open age interval. Mortality acceleration at younger old ages (Horiuchi 1997; Horiuchi & Wilmoth 1997), which is more typical in female populations, explains positive biases in the extrapolation method as seen in Figure 3 (notice the stronger pattern of positive biases for females). At the same time, mortality deceleration at older ages (Horiuchi & Wilmoth 1997; Horiuchi & Wilmoth 1998; Horiuchi et al. 2003) may explain somewhat more prevalent negative biases of the extrapolation method at the open age interval 85+ and the tendency of the method to produce negative errors for the open age interval 95+ (results not shown here). The Horiuchi-Coale and Mitra methods, in their turn, tend to produce positive errors, because protracted periods of mortality decline at old age, as observed in many countries, produce population age structures even younger than the stable populations assumed in the two methods (Horiuchi & Preston 1988; Guillot 2003; Ediev 2014).

The Horiuchi-Coale formula provides a remarkable improvement over the classical method in terms of estimation errors, although the parameters for the formula were estimated back in the 1980s. The vast reduction of estimation errors, after applying the adjustment, indicates that the method is rather robust to violations of its underlying assumptions (the Gompertzian death rates and stable population age strcuture). Table 2 also includes results for the Horiuchi-Coale formula where we kept the original values for the parameter α_a but re-estimated the other parameter β_a based on our database ("H.-C. (hmd)" columns of Table 2; see Table 1 for the re-estimated parameter values). Updating the model parameters provides only a marginal improvement in terms of root-mean squared error (RMSE). The method's RMSEs may perhaps be further reduced by fitting the model to more homogeneous data (to groups of populations with similar mortality dynamics and growth histories).

The Mitra formula is generally more accurate than the Horiuchi-Coale method (Table 2). However, the method appeared to be prone to produce outliers, especially overestimates of life expectancy (Figure 3; similar results for other open age intervals are not shown here). The formula involves the population mean age in the open age interval, $\bar{x}$, an indicator easy to calculate for populations with good-quality data, such as the HMD populations, but problematic for populations with age exaggeration. Hence, we checked, if the formula remains accurate after substituting $\bar{x}$ by its prediction based on the regression involving the growth rate and the observed death rate:

$$\bar{x} = C + k_1 {M_{a+}}^{-1} + k_2 r {M_{a+}}^{-1} \quad (5)$$

(see Table 1 for model parameters estimated on HMD data). Absolute and percentage RMSEs of estimate (5) are presented in the last two columns of Table 2. Results for the Mitra formula with the approximate mean age (5) are shown in Table 2, in the "Mitra (regr.)" columns. Substituting the true mean age by its indirect estimate only marginally increases RMSEs. Even based on indirect mean age estimates, the method remains more accurate than the Horiuchi-Coale method. The differences between the two methods, however, are minor as compared to the errors in the classical method.

The extrapolation method, surprisingly, does not improve over the classical method in terms of bias and is rather unstable. It shows higher RMSEs as compared to the classical life table method and is by far less acurrate than both the Horiuchi-Coale and Mitra approaches. Note that we used the Gompertz model that assumes exponential growth of the death rates as a function of age. If we used a more optimistic logistic-type model (e.g., the Kannisto model (Thatcher et al. 1998) fits better the pattern of mortality decelaration at oldest old age and is used by the WHO and the HMD), the upward biases in life expectancy estimates would be even higher (selected results for the Kannisto model are presented in section 3).

Figure 1. Estimation errors in life expectancy at birth obtained by the classical life table method for selected open age intervals (starting at ages shown in the right-hand side of the panes), men, women and both sexes combined (in years). Author's calculations based on the Human Mortality Database (HMD 2016).

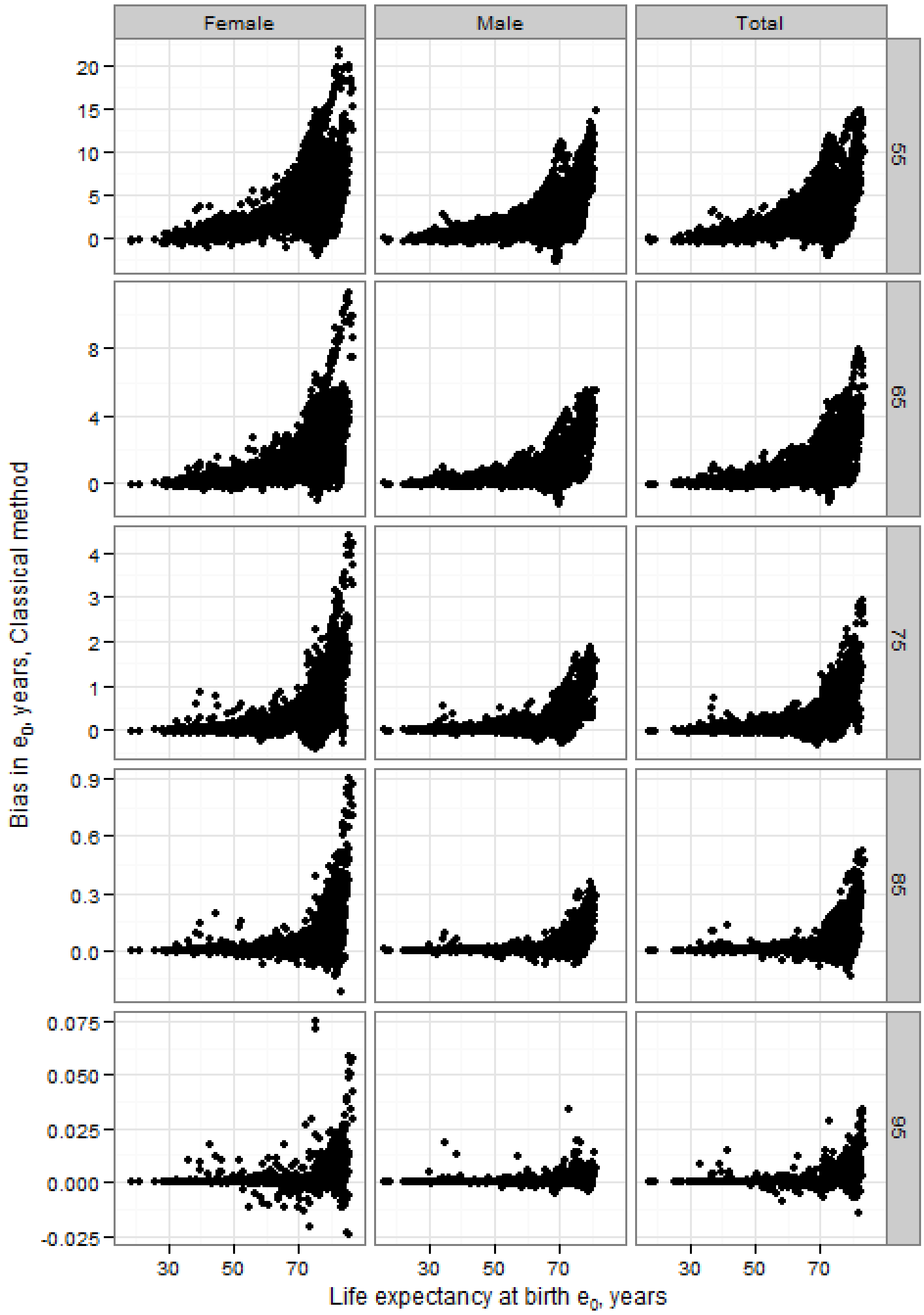


Figure 2. Estimation errors in life expectancy at the starting age of the open age interval for the classical life table method (starting ages of the selected open age intervals are shown in the right-hand side of the panes), men, women and both sexes combined (in percent). Author's calculations based on the Human Mortality Database (HMD 2016).

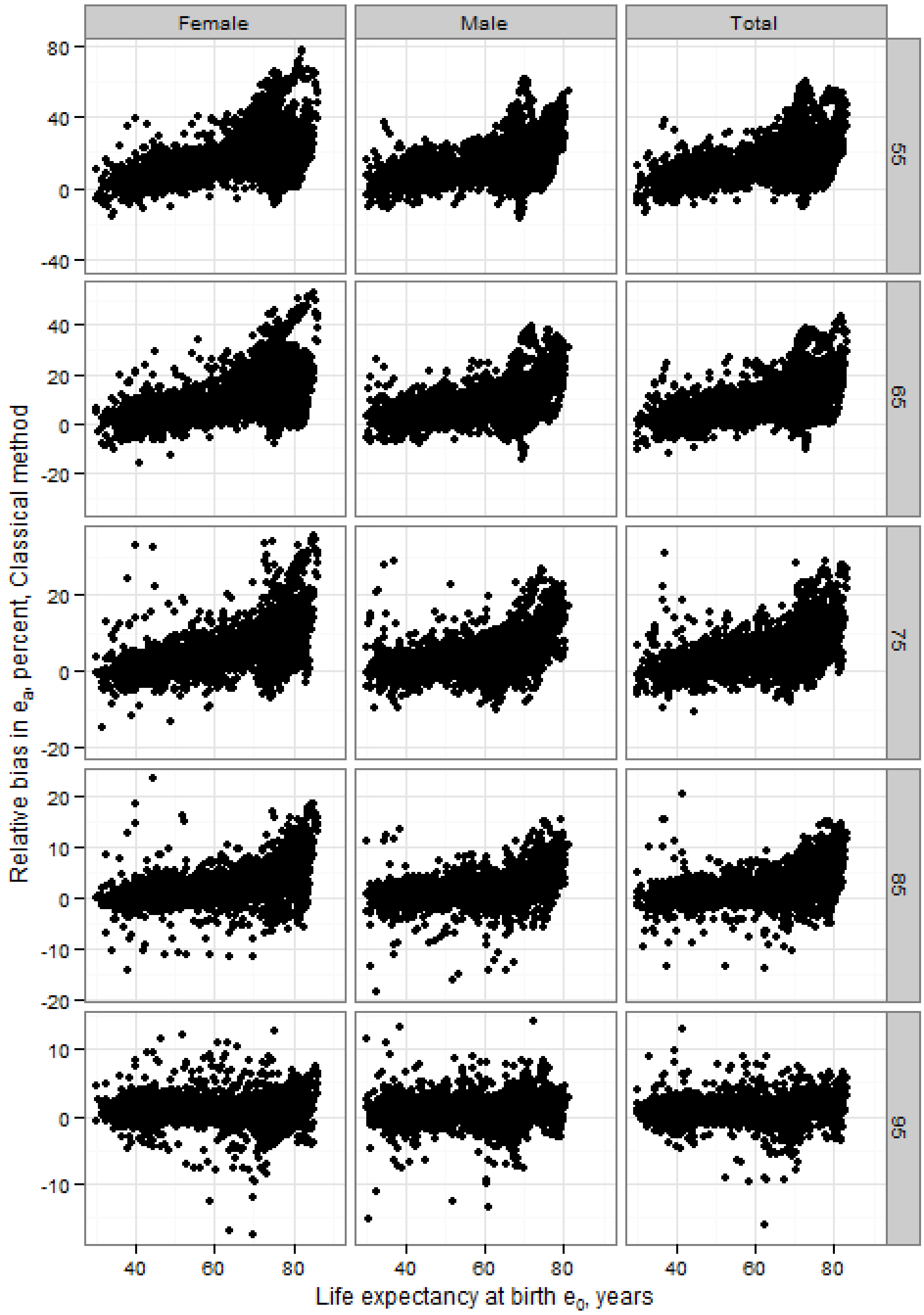


Figure 3. Estimation errors in life expectancy at birth obtained by selected methods (methods indicated in the right-hand side of the panes) for the open age interval 75+, men, women and both sexes combined (in years). Author's calculations based on the Human Mortality Database (HMD 2016).

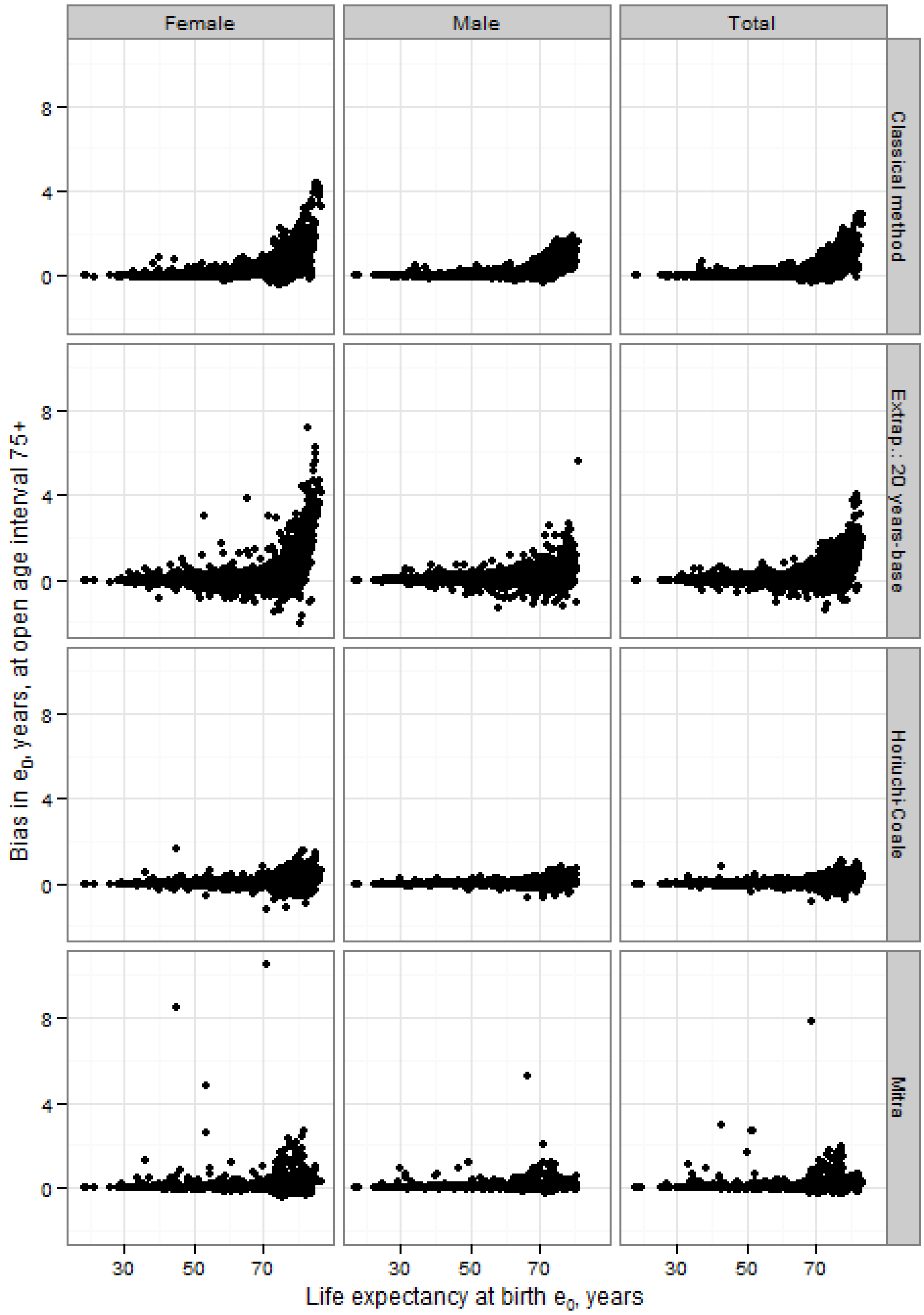


Table 2. Root mean squared errors (RMSE) in life expectancy at birth e_0, percentage of RMSEs in life expectancy in the open age interval e_a (by method) and RMSEs in the mean population age in the open age interval (model (5)): by sex, level of life expectancy at birth and open age interval (a+).

			RMSE in e_0 by method (years)						Percentage RMSE in e_a by method (percent)						RMSE in the mean pop. age, years	RMSE in the mean pop. age, percent
Sex	e_0 range	a	Classical	Extrapol.	H.-C.	H.-C. (hmd)	Mitra	Mitra (regr.)	Classical	Extrapol.	H.-C.	H.-C. (hmd)	Mitra	Mitra (regr.)		
Female	40-50	55	0.75	6.60	0.40	0.41	0.47	0.46	8.05	73.76	4.54	4.64	5.33	5.22	0.58	0.88
Female	40-50	65	0.25	0.70	0.17	0.17	0.23	0.22	5.17	15.61	3.67	3.71	4.83	4.72	0.43	0.59
Female	40-50	75	0.06	0.16	0.05	0.05	0.05	0.05	3.40	8.42	2.78	2.85	2.93	2.88	0.37	0.47
Female	40-50	85	0.01	0.03	0.01	0.01	0.02	0.02	2.21	9.50	2.29	2.36	3.03	2.65	0.32	0.36
Female	50-60	55	1.53	10.23	0.57	0.54	0.38	0.48	11.86	81.82	4.54	4.31	3.14	3.90	0.57	0.86
Female	50-60	65	0.53	1.45	0.23	0.23	0.21	0.24	7.65	21.53	3.41	3.38	3.07	3.51	0.42	0.58
Female	50-60	75	0.12	0.25	0.07	0.08	0.08	0.08	4.54	12.93	2.82	2.85	2.73	2.89	0.37	0.46
Female	50-60	85	0.01	0.04	0.01	0.01	0.01	0.01	2.60	7.94	2.00	2.04	2.08	2.09	0.24	0.27
Female	60-70	55	3.57	18.78	0.79	0.77	0.43	0.57	21.21	114.15	4.63	4.59	2.52	3.34	0.52	0.79
Female	60-70	65	1.23	1.56	0.29	0.29	0.19	0.25	12.99	17.34	3.11	3.12	1.99	2.65	0.39	0.53
Female	60-70	75	0.25	0.29	0.09	0.09	0.06	0.09	6.55	9.15	2.54	2.53	1.77	2.30	0.47	0.58
Female	60-70	85	0.02	0.06	0.01	0.01	0.01	0.01	3.11	8.14	1.95	1.95	1.81	1.97	0.44	0.48
Female	70-80	55	5.38	3.32	1.36	1.19	0.99	1.08	24.14	15.00	5.99	5.29	4.53	4.93	0.57	0.85
Female	70-80	65	2.33	2.18	0.49	0.48	0.37	0.42	16.87	15.72	3.53	3.48	2.70	3.06	0.37	0.50
Female	70-80	75	0.67	0.56	0.16	0.15	0.11	0.13	10.10	8.25	2.56	2.35	1.75	2.14	0.29	0.36
Female	70-80	85	0.09	0.10	0.03	0.03	0.02	0.03	4.78	7.20	1.70	1.72	1.34	1.57	0.22	0.24
Female	80-90	55	8.08	23.67	1.90	1.61	0.96	1.17	28.64	81.77	6.86	5.81	3.50	4.22	0.56	0.81
Female	80-90	65	3.78	4.71	0.91	0.89	0.64	0.74	20.07	25.14	4.91	4.79	3.45	3.95	0.37	0.48
Female	80-90	75	1.42	2.11	0.34	0.27	0.20	0.27	13.80	20.10	3.35	2.72	2.04	2.60	0.25	0.30
Female	80-90	85	0.26	0.16	0.09	0.08	0.06	0.08	7.43	4.46	2.54	2.37	1.67	2.33	0.26	0.29
Male	40-50	55	0.71	3.47	0.26	0.24	0.20	0.24	8.26	45.83	3.27	3.07	2.59	3.01	0.45	0.70
Male	40-50	65	0.22	0.62	0.10	0.09	0.09	0.10	5.24	15.58	2.57	2.56	2.47	2.73	0.35	0.48
Male	40-50	75	0.04	0.09	0.02	0.02	0.02	0.02	3.05	7.71	2.05	2.04	1.89	2.12	0.29	0.37
Male	40-50	85	0.00	0.02	0.00	0.00	0.00	0.00	1.73	9.38	1.58	1.60	1.53	1.58	0.23	0.26
Male	50-60	55	1.49	4.23	0.46	0.43	0.32	0.39	12.30	34.85	4.03	3.80	2.71	3.38	0.47	0.72
Male	50-60	65	0.50	0.99	0.17	0.17	0.13	0.16	8.05	16.32	2.97	2.90	2.25	2.74	0.37	0.51
Male	50-60	75	0.10	0.18	0.05	0.05	0.04	0.05	4.65	9.42	2.56	2.51	2.13	2.51	0.38	0.47
Male	50-60	85	0.01	0.03	0.01	0.01	0.00	0.01	2.53	7.78	1.90	1.91	1.63	1.82	0.27	0.31
Male	60-70	55	2.45	2.76	0.55	0.58	0.47	0.48	15.32	17.23	3.49	3.71	3.00	3.06	0.43	0.66

Sex	e_0 range	a	RMSE in e_0 by method (years) Classical	Extrapol.	H.-C.	H.-C. (hmd)	Mitra	Mitra (regr.)	Percentage RMSE in e_a by method (percent) Classical	Extrapol.	H.-C.	H.-C. (hmd)	Mitra	Mitra (regr.)	RMSE in the mean pop. age, years	RMSE in the mean pop. age, percent
Male	60-70	65	0.82	0.68	0.21	0.22	0.16	0.18	9.58	8.36	2.52	2.60	1.89	2.17	0.36	0.49
Male	60-70	75	0.17	0.20	0.06	0.06	0.05	0.06	5.43	7.20	2.07	2.12	1.63	1.92	0.32	0.39
Male	60-70	85	0.02	0.03	0.01	0.01	0.01	0.01	2.74	6.19	1.71	1.74	1.60	1.66	0.28	0.31
Male	70-80	55	5.34	3.02	1.30	1.06	0.66	0.77	24.10	13.85	6.00	4.95	3.17	3.62	0.43	0.64
Male	70-80	65	2.19	1.30	0.51	0.46	0.36	0.40	16.35	9.65	3.83	3.59	2.79	3.06	0.29	0.39
Male	70-80	75	0.66	0.53	0.16	0.13	0.11	0.13	10.44	8.39	2.66	2.23	1.81	2.09	0.22	0.28
Male	70-80	85	0.09	0.06	0.03	0.03	0.02	0.03	5.22	4.47	1.88	1.80	1.21	1.72	0.20	0.23
Total	40-50	55	11.01	9.56	1.91	1.75	0.65	1.00	41.98	36.36	7.34	6.73	2.48	3.82	0.49	0.72
Total	40-50	65	4.30	1.54	0.51	0.37	0.27	0.51	25.39	9.18	2.96	2.18	1.58	2.95	0.38	0.51
Total	40-50	75	1.06	1.97	0.33	0.25	0.19	0.22	12.10	21.67	3.82	2.90	2.24	2.51	0.17	0.21
Total	40-50	85	0.19	0.09	0.06	0.05	0.03	0.05	7.12	3.48	2.13	1.86	1.21	1.77	0.16	0.19
Total	50-60	55	0.72	4.80	0.33	0.31	0.32	0.34	8.09	57.95	3.83	3.76	3.89	4.02	0.52	0.79
Total	50-60	65	0.22	0.65	0.13	0.13	0.18	0.18	4.93	15.28	3.13	3.15	4.10	4.07	0.38	0.53
Total	50-60	75	0.05	0.09	0.03	0.03	0.03	0.03	3.05	6.24	2.24	2.27	2.14	2.28	0.32	0.40
Total	50-60	85	0.01	0.02	0.01	0.01	0.01	0.01	1.88	8.65	1.77	1.80	1.89	1.83	0.24	0.27
Total	60-70	55	1.47	5.74	0.48	0.45	0.30	0.39	11.81	47.86	4.00	3.78	2.53	3.26	0.47	0.72
Total	60-70	65	0.50	1.16	0.18	0.17	0.13	0.16	7.58	17.95	2.80	2.74	2.01	2.56	0.35	0.49
Total	60-70	75	0.10	0.16	0.05	0.05	0.05	0.06	4.26	7.56	2.40	2.40	2.22	2.43	0.33	0.41
Total	60-70	85	0.01	0.03	0.01	0.01	0.01	0.01	2.31	6.72	1.58	1.60	1.38	1.58	0.22	0.25
Total	70-80	55	3.12	3.93	0.66	0.62	0.48	0.54	18.72	25.25	4.01	3.79	2.96	3.29	0.49	0.74
Total	70-80	65	1.08	0.94	0.25	0.25	0.18	0.22	11.69	10.59	2.78	2.79	2.02	2.48	0.36	0.49
Total	70-80	75	0.23	0.23	0.08	0.08	0.06	0.07	6.32	6.88	2.35	2.36	1.81	2.16	0.39	0.48
Total	70-80	85	0.02	0.04	0.01	0.01	0.01	0.01	2.87	6.50	1.76	1.79	1.77	1.71	0.35	0.39
Total	80-90	55	4.54	2.52	1.21	1.06	0.77	0.82	20.94	11.87	5.47	4.87	3.72	3.86	0.48	0.71
Total	80-90	65	1.95	1.25	0.45	0.44	0.34	0.37	14.76	9.39	3.40	3.31	2.56	2.78	0.34	0.45
Total	80-90	75	0.59	0.53	0.15	0.13	0.10	0.12	9.36	8.21	2.40	2.15	1.63	1.97	0.25	0.31
Total	80-90	85	0.08	0.07	0.03	0.03	0.02	0.03	4.76	4.88	1.66	1.65	1.18	1.55	0.21	0.24

Note: e0= life expectancy at birth; a=starting age of the open age interval; “Classical”= classical life table method; “Extrapol.”=extrapolation of the death rates using the Gompertz model; “H.-C.”=Horiuchi-Coale formula; “H.-C. (hmd)”=Horiuchi-Coale formula with the Beta parameter re-estimated on the HMD data; “Mitra”=Mitra formula; “Mitra (regr.)”=Mitra formula with the mean population age in the open age interval substituted from the indirect estimate (5); “RMSE” = root-mean squared estimation errors; “Percentage RMSE” = root-mean squared estimation errors as percent of true values “mean pop. age”= mean population age in the open age interval indirectly estimated using model (5).

3. Reconciling Horiuchi-Coale and Mitra formulas. Conditioned extrapolation of death rates

Our above results show that both the Horiuchi-Coale and Mitra formulas perform well in reducing life expectancy estimation errors caused by aggregating data for the open age interval. The Mitra method is marginally more accurate but less stable. Its reliance on the possibly exaggerated mean population age may be overcome by using the indirect estimates of the mean age (5). Both methods are by far superior to the classical and the extrapolation methods. It is a pity that the two methods did not make it to a wider practical use and rather surprising that their authors came to contradicting results in their papers.

It is worthwhile to note, however, that a large part of numerical differences between Horiuchi-Coale and Mitra were, in fact, due to different inputs used in their calculations rather than methodological differences. In particular, the largest discrepancy in the original papers was for e_{65} for El Salvador in 1961: Mitra's estimates were larger by 3.12 years for women and 3.02 years for men. When we recalculated the life expectancies using similar inputs in both approaches (Mitra 1984, pp.11–12), we found that the two approaches are more consistent: Mitra's formula gives estimates by 1.69 and 1.27 years larger, respectively. After our recalculation, the estimates become much closer also for Canada, Japan, Switzerland, UK, Mexico, and Malaysia. Altogether, the two methods differ by more than one year in only two cases, El Salvador and Puerto Rico, out of 13. If one takes into account that the Mitra method was relying on potentially biased official estimates of the mean population age in the open age interval, it becomes clear that the authors were more consistent than they concluded.

One way of utilising both approaches might be to use the Horiuchi-Coale method to adjust the conventional life expectancy estimates as a first step and, in the second step, use the Mitra formula to obtain the estimate of the mean population age in the open age interval (deriving from Eq. (3)):

$$\bar{x} = a + \frac{1+\frac{M_{a+}}{r} ln(e_a M_{a+})}{M_{a+}+r}. \tag{6}$$

Here, life expectancy e_a may be substituted by the Horiuchi-Coale estimate (2). This way, it might be possible not only to compensate the estimates of life expectancy for biases caused by data aggregation but also to assess the magnitude of age exaggeration in the data (an approach alternative to our estimate (5)). Testing model (6) on HMD data (detailed results not shown here), we found that the method is inferior to the regression method (5). RMSEs of estimates (6) were, in principle, acceptably low (shorter than one year) but much higher (by up to several times) as compared to the RMSE of the regression approach. Nonetheless, one should be encouraged to check, which approach of assessing the mean age is better fitted to a specific country case.

Another way of combining the two approaches might be to implement them in parallel (substituting the indirectly estimated mean age (5) into the Mitra formula) and use the difference between the two estimates in a rough check for possible outliers.

Our results show that the extrapolation of death rates using a mortality model does not improve much over the classical model in terms of the life expectancy estimation bias. However, one may combine the more accurate adjusted estimates of life expectancy e_a with the extrapolation model by constraining the parameters of the latter to fit the life expectancy estimate. For the Gompertz and Kannisto models that contain only two parameters, we suggest fixing the model's death rate to the original death rate at the age below the open age interval, M_{a-1}, and finding the other parameter of the model by fitting them to the life expectancy e_a from either the Horiuchi-Coale or Mitra formulas. In the Gompertz model,

$$M_x = Ce^{b(x-a+1)}, \tag{7}$$

one might set $C = M_{a-1}$ and fit b to the e_a estimate; while in the Kannisto model,

$$M_x = \frac{Ce^{b(x-a+1)}}{1+Ce^{b(x-a+1)}}, \tag{8}$$

one would set $C = \frac{M_{a-1}}{1-M_{a-1}}$ and fit b to the e_a estimate. In Figure 4 we present extrapolations (conventional and conditioned on e_a from the Horiuchi-Coale method) applying the two mortality models to the death rates in Japan in 2012, at selected open age intervals (a=65, 75, or 85 years). In all cases the constrained extrapolations fit the actual death rates better than the unconstrained ones, although the improvement was small in the case of males, open age interval 65+. In the worst case (females, younger open age interval), the conventional extrapolations may mislead producing the death rates by up to ten times lower than the actual rates at old age. Generally, the Kannisto model fits the age pattern of the death rates better than the Gompertz model, although the two models behave rather similar at younger ages. At open age interval starting at least at 75 years, the Kannisto model conditioned on the life expectancy estimate produces the best outcomes, staying very close to the empirical rates.

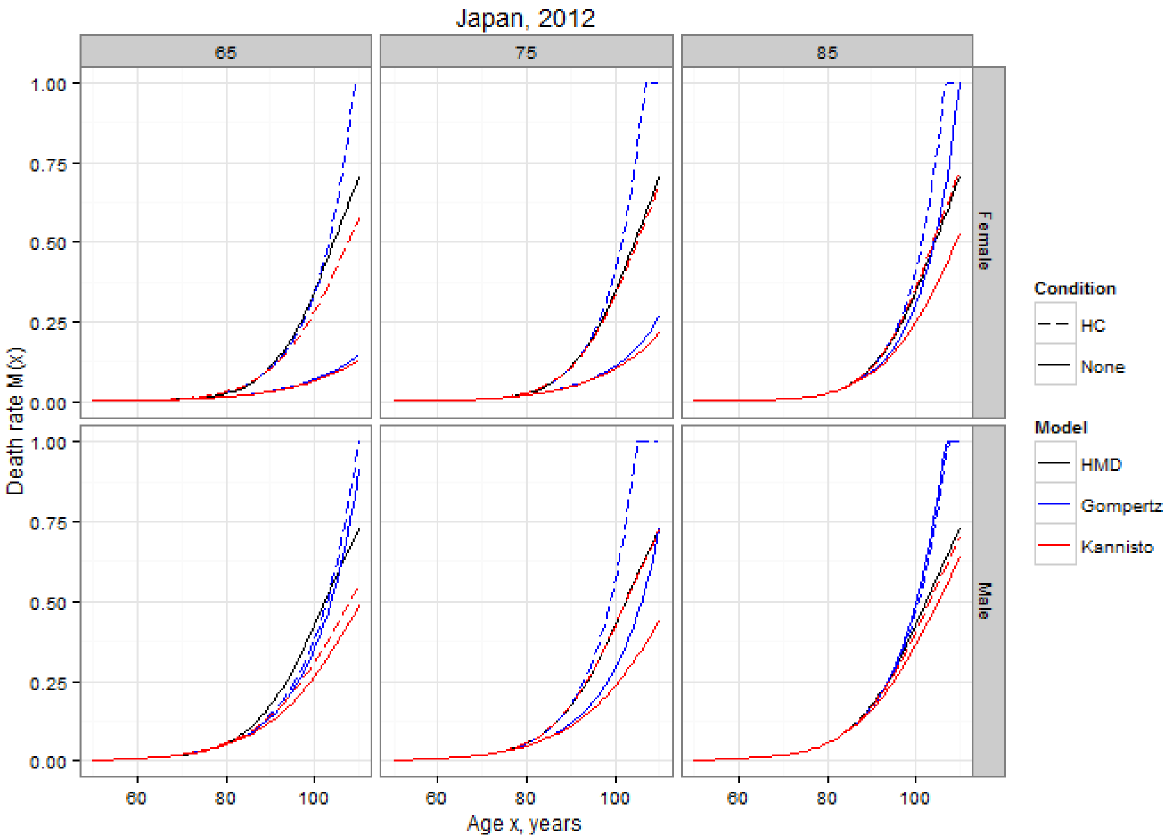
4. Discussion

Our results show that the violation of the stationary population assumption of the classical life table method has strong consequences for the accuracy of life expectancy estimates. The errors in the classical method increase as mortality declines. For a currently low-mortality population, closing the life table at age 65 would produce an HMD-average upward bias of more than 3 years and even larger RMSEs in the life expectancy at birth calculated by the classical method. Continuing increases in life expectancy will drive the biases of the classical method to even higher levels.

The methods developed by Horiuchi-Coale and Mitra drastically reduce estimation errors in the expectation of life as compared to both the classical and the extrapolation methods. Wider usage of these methods should be encouraged for populations where data on old-age mortality are missing (for example, due to small population size) or corrupt by age exaggeration or age misreporting in general. The methods may also be combined to study the extent of age exaggeration within data, although our alternative (5) seems to provide a better service.

The Horiuchi-Coale and Mitra adjustments appear to be less sensitive to the specific features of the pattern of death rates at old age. We have checked the sensitivity of the methods by applying them to the raw death rates in the HMD that were not modified using a logistic model (according to the HMD protocol, the HMD death rates used in our calculations above are smoothed and modified at old age by applying a logistic model). Our results based on the raw death rates were similar to those based on the smoothed rates from HMD. In the context of sensitivity, the good performance of the Horiuchi-Coale method (with parameters estimated on Gompertz mortality schedules back in the 1980s) on the up-to-date empirical data is worthwhile noting. Departures from a stationary population age composition is the main driver of estimation biases and differences between the estimates of life expectancy. This does not come as a big surprise if one takes into account that for a stationary population, irrespective of the age pattern of mortality, both the classical and the adjusted life expectancies are identical and equal to the true life expectancy.

Figure 4. Constrained and unconstrained extrapolations of the death rates for Japan, 2012, men and women, at various ages at the start of the extrapolation (*a*=65, 75, or 85 years as shown in the upper parts of the panels).



Note: 'HC'= extrapolation constrained on the life expectancy e_a from the Horiuchi-Coale method; 'None'=no constraints imposed (model parameters are estimated on the death rates for 20-years-long age frame below age *a*); 'HMD'=original death rates from the Human Mortality Database (HMD 2016); 'Gompertz'=death rates extrapolated using the Gompertz model at ages *a*+; 'Kannisto'= death rates extrapolated using the Kannisto model at ages *a*+.

The lack of usage of the Horiuchi-Coale and Mitra methods in demographic analysis is even more surprising and upsetting if one notes their perfect fit to a number of popular indirect demographic methods. The Brass Growth Balance method, the Preston and Coale method, the Hill Generalized Growth Balance method, and the Bennett and Horiuchi Synthetic Extinct Generations method, all involve estimates of the population growth rate (Nations 1983; Moultrie et al. 2013). These methods provide ready inputs for the Horiuchi-Coale and Mitra formulas.

The method of extrapolating the death rates into the open age interval does not appear to work better than the classical method. The WHO and the HMD use an S-shaped model for extrapolating the death rates as alternative to the J-shaped Gompertz model used here. Indeed, logistic-type models were shown to better fit the deceleration of mortality at oldest old ages (Thatcher et al. 1998; Missov et al. 2016). However, our results for the younger open age intervals where all mortality models fit close to each other, suggest that the conclusion about the inferiority of extrapolation to the Horiuchi-Coale and Mitra methods is in general applicable to any mortality models other than the Gompertz model. In fact, a logistic-type model might even accentuate the upward biases in life expectancy estimates. In many applications, however, it is important to extend the age profile of the death rates into the open age interval. Although our results discourage from using the popular extrapolations, we suggest to improve those extrapolations by conditioning the life expectancy form on the more accurate Horiuchi-Coale and Mitra methods. Our preliminary results indicate prominence of the Kannisto model conditioned on the life expectancy estimate. Such improved extrapolations of the death rates into the older ages may be valuable for population projections and reconstructions.

As an important consequence of the discrepancy between death rate and life expectancy for the open age interval, the traditional approach of projecting the population in the open age interval (Preston et al. 2001) may lead to overestimates by dozens of percent of deaths in the open age interval. The adjustment formulas considered here as well as related conditioned extrapolations of the death rates may be used to compensate for this projection bias.

One may further improve life expectancy estimates by fitting the models on populations with closer history of growth and mortality reduction (e.g., of regions of a given country). Another direction of improvement might be considering population models more advanced than the stable population. For example, one may consider effects of the changing growth rate and mortality on the population age composition (Brouard 1986; Horiuchi & Preston 1988; Guillot 2003; Ediev 2014).

Life expectancy adjustments and conditioned mortality extrapolations may be useful in estimating multi-state life tables that would, typically, suffer from data quality issues. Another promising area for application might be the construction of life tables for cohorts that may depart from stationary age composition in the presence of migration, state transitions or other types of attrition. Indeed, growth rate and stable population concepts may not be directly applied to the birth cohort. Therefore, the Horiuchi-Coale method would be difficult to use in estimating the cohort life tables. However, one may consider using the Mitra formula where the growth parameter r is substituted by its approximation based on the observed death rate and the mean population age in the open age interval (an approach similar to Eq. (5)).

5. References

Brouard, N. 1986. Structure et dynamique des populations. La pyramide des années à vivre, aspects nationaux et exemples régionaux. Espace, Populations, Sociétés 4(2): 157–168.

Coale, A.J. 1985. Estimating the expectation of life at old ages: Comments on the article by Mitra. *Population Studies* 39(3): 507–509.

Doray, L.G. 2008. *Inference for Logistic-type Models for the Force of Mortality*. Paper presented at the Living to 100 Symposium, January 7-9, 2008, Orlando, Fla. (USA): Society of Actuaries.

Ediev, D.M. 2014. Why increasing longevity may favour a PAYG pension system over a funded system. *Population Studies* 68(1): 95–110.

Gompertz, B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 115: 513–583.

Guillot, M. 2003. The cross-sectional average length of life (CAL): A cross-sectional mortality measure that reflects the experience of cohorts. *Population Studies* 57(1): 41–54.

HMD. 2016. *Human Mortality Database*. Online database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). http://www.mortality.org/

Horiuchi, S. et al. 2003. Differential patterns of age-related mortality increase in middle age and old age. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 58(6): 495–507.

Horiuchi, S. 1997. Postmenopausal acceleration of age-related mortality increase. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 52(1): B78-92.

Horiuchi, S. and A.J. Coale. 1982. A Simple Equation for Estimating the Expectation of Life at Old Ages. *Population Studies* 36(2): 317.

Horiuchi, S. and S.H. Preston. 1988. Age-Specific Growth Rates: The Legacy of Past Population Dynamics. *Demography* 25(3): 429.

Horiuchi, S. and J.R. Wilmoth. 1998. Deceleration in the age pattern of mortality at older ages. *Demography* 35(4): 391–412.

Horiuchi, S. and J.R.R. Wilmoth. 1997. Age patterns of the life table aging rate for major causes of death in Japan, 1951-1990. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 52(1): B67-77.

Khlat, M. and Y. Courbage. 1996. Mortality and causes of death of Moroccans in France, 1979-91. *Population* 8: 59–94.

Kibele, E., R. Scholz, and V.M. Shkolnikov. 2008. Low migrant mortality in Germany for men aged 65 and older: Fact or artifact? *European Journal of Epidemiology* 23(6): 389–393.

Mathers, C. and J. Ho. 2014. *WHO Methods for Life Expectancy and Healthy Life Expectancy*. Geneva, Switzerland: WHO Department of Health Statistics and Information Systems.

Missov, T.I., L. Németh, and M.J. Dańko. 2016. How much can we trust life tables? Sensitivity of mortality measures to right-censoring treatment. *Palgrave Communications* 2: 15049.

Mitra, S. 1984. Estimating the expectation of life at older ages. *Population Studies* 38(2): 313.

Mitra, S. 1985. On estimating the expectation of life at old ages: Reply to Professor Coale. *Population Studies* 39(3): 511–512.

Moultrie, T. et al. 2013. *Tools for Demographic Estimation*. Paris, France: International Union for the Scientific Study of Population.

Oeppen, J. and J.W. Vaupel. 2002. Demography. Broken limits to life expectancy. *Science* 296(5570): 1029–1031.

Preston, S.H. et al. 1996. African-American mortality at older ages: results of a matching study. *Demography* 33(2): 193–209.

Preston, S.H., P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford, UK: Blackwell Publishers.

Randall, S. and E. Coast. 2016. The quality of demographic data on older Africans. *Demographic Research* 34: 143–174.

Scherbov, S. and D. Ediev. 2011. Significance of life table estimates for small populations: Simulation-based study of standard errors. *Demographic Research* 24: 527–550.

Thatcher, A.R., V. Kannisto, and J.W. Vaupel. 1998. *The Force of Mortality at Ages 80-120. Monographs on Population Aging*. Odense, Denmark: Odense University Press.

United Nations. 1982. *Model Life Tables for Developing Countries*. Population Studies 77. New York, USA: United Nations Department of International Economic and Social Affairs.

United Nations. 1983. *Manual X: Indirect Techniques for Demographic Estimation.* Population Studies 81. New York, USA: United Nations Department of International Economic and Social Affairs.

White, K.M. 2002. Longevity advances in high-income countries, 1955-96. *Population and Development Review* 28(1): 59–76.

Wilmoth, J.R. et al. 2007. Methods Protocol for the Human Mortality Database. http://www.mortality.org/Public/Docs/MethodsProtocol.pdf.