On the sources of instability of the Mitra model for years of life at old-age Dalkhat M. Ediev

(<u>dalkhat@hotmail.com; ediev@iiasa.ac.at</u>) North-Caucasian State Academy (Russia), Lomonosov Moscow State University (Russia), International Institute for Applied Systems Analysis (Austria)

We explore sources of this instability of the Mitra (1984) model for estimating the remaining life expectancy at old age by revisiting Mitra's own derivation and keeping the terms of higher order necessary for our inquiry. Dropping the higher-order terms and errors in estimates of the mean population age appear to lead to only small errors of up to three percent of the true remaining life expectancy. The growth parameter's inadequate estimates, however, lead to large estimation errors and are the source of the outliers of the model. Our results confirm that Mitra model's biases are typically on the side of overestimating the life expectancy. This provides formal support for an earlier method (Ediev 2018) based on combining the Mitra and the classical life table models. Based on the performed analysis, we also propose several alternative models of estimating the life expectancy at old age in stable and non-stable population contexts. Our iterative method of estimating the expectation of life at old age does not rely on approximate relations assumed in Mitra and other traditional models and performs as good as the Mitra model overall, yet, being free from outliers and skewedness of estimation errors characterizing the Mitra model.

Acknowledgements

The work was supported by the Russian Foundation for Basic Research (grant N_{2} 18-01-00289 «Mathematical models and methods of correcting the distortions of the age structure and mortality rates of elderly population»).

1. Introduction

Distorted data on population age distribution (Bennett and Horiuchi 1984; Shryock and Siegel 1973), particularly the age exaggeration (Coale and Kisker 1990; Duthé et al. 2010; Khlat and Courbage 1996; Kibele, Scholz, and Shkolnikov 2008; Preston et al. 1996), undermines quality of life table at older ages and necessitates improving models for estimating the remaining life expectancy in the open age interval. Horiuchi and Coale (1982) – assuming population stability and using regression approximations –

suggested the conventional estimate of the remaining life expectancy in the open age interval is distorted for non-stationary populations and proposed the correction formula:

$$e_a = M_{a+}{}^{-1} e^{-\beta_a r M_{a+}{}^{-\alpha_a}},\tag{1}$$

here M_{a+} is the death rate in the open age interval a+ beginning at age a, r is the population growth rate, α_a and β_a are model parameters. The first multiplier in (1) represents the conventional estimate used in constructing life tables (Preston, Heuveline, and Guillot 2001):

$$e_a = M_{a+}^{-1},$$
 (2)

that assumes stationarity of population age composition. Indeed, in a (stationary) life table population with deaths above age *a* equal to the survival by that age l_a and population size above age *a* equal to the life table person-years T_a , the death rate in the open age interval is in clear inverse relationship (2) to the life expectancy: $M_{a+} = \frac{l_a}{T_a} = e_a^{-1}$.

Mitra (1984) used a more rigorous formal-demographic approach and suggested alternative model:

$$e_{a} = M_{a+}^{-1} e^{-r[M_{a+}^{-1} - (1 + rM_{a+}^{-1})(\bar{x} - a)]},$$
(3)

here \bar{x} is the mean population age in the open age interval. This has led to debate (Coale 1985; Mitra 1985) that was not resolved until recently.

More recent extensive empirical test of models (1)-(3) as compared also to the conventional extrapolation of death rates to older ages (Ediev 2018) revealed that models (1) and (3) are consistent with each other and substantially improve over the conventional approaches (the life table model (2) and the extrapolation). The original disagreement between these authors appears to be caused by using different inputs for the same countries rather than by using different models. In the empirical tests, the Mitra model appeared to be most accurate, although the difference in accuracy between models (1) and (3) was of smaller magnitude than improvement of either of the models over the conventional approaches. However, practical usability of the Mitra model appears to be limited, because it is also the least stable among the studied approaches due to frequent outliers produced by the model. Furthermore, Mitra model's biases were predominantly positive in the empirical tests.

Neither the original works by Mitra, nor the empirical tests provide a clue as to what might have been causing the model instability and biasedness. Here, we address this problem by re-deriving the Mitra formula and studying its sensitivity to violations of the underlying assumptions.

2. Revisiting Mitra's formula

In this section, we re-derive Mitra's formula by and large following the same steps as proposed by Mitra (1984). We re-derive the equation, because the original derivation by Mitra was not straightforward, led to various side-results and partly relied on results from the stable population theory. Furthermore, Mitra omitted higher-order terms before arriving to the result, which is an obstacle in finding out if these were the neglected higher-order terms that led to models' instability and biasedness.

Consider the number of deaths in the open age interval
$$a$$
+ in a stable population:

$$D_{a+} = \int_{a}^{\omega} D(x) dx = \int_{a}^{\omega} N(x) \mu(x) dx = N(0) \int_{a}^{\omega} \mu(x) l(x) e^{-rx} dx =$$

$$-N(0) \int_{a}^{\omega} \frac{dl(x)}{dx} e^{-rx} dx = -N(0) [l(x) e^{-rx} |_{a}^{\omega} + r \int_{a}^{\omega} l(x) e^{-rx} dx] =$$

$$N(0) [l(a) e^{-ra} - r \int_{a}^{\omega} l(x) e^{-rx} dx],$$
(4)

1

here, l(x) is the survival (probability to survive from birth to age x), N(x) = $N(0)l(x)e^{-rx}$ is the age structure of the stable population with r being the growth parameter (Keyfitz and Caswell 2005), $\mu(x) = -\frac{dl(x)}{l(x)dx}$ is the force of mortality, and ω is the lifespan $(N(x), l(x) \equiv 0 \text{ at } x \geq \omega)$.

Population in the open age interval equals:

$$N_{a+} = \int_{a}^{\omega} N(x) dx = N(0) \int_{a}^{\omega} l(x) e^{-rx} dx,$$
(5)

that is, the open age interval death rate equals

$$M_{a+} = \frac{D_{a+}}{N_{a+}} = \frac{l(a)e^{-ra} - r\int_a^{\omega} l(x)e^{-rx}dx}{\int_a^{\omega} l(x)e^{-rx}dx} = \frac{l(a)e^{-ra}}{\int_a^{\omega} l(x)e^{-rx}dx} - r.$$
(6)

In the stationary case (r=0), (6) implies that the conventional estimate (2) is accurate and no correction is needed:

$$M_{a+}^{-1} = \frac{\int_{a}^{\omega} l(x)dx}{l(a)} = e(a),$$
(7)

Regrouping terms in (6) leads to an auxiliary equation that will be used further:

$$\frac{1}{M_{a+}+r} = \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx.$$
(8)

The remaining life expectancy at age *a* may be approximated as follows:

$$e(a) = \frac{1}{l(a)} \int_{a}^{\omega} l(x) dx = \frac{e^{r\bar{x}}}{l(a)} \int_{a}^{\omega} l(x) e^{-rx} e^{r(x-\bar{x})} dx \approx \frac{e^{r\bar{x}}}{l(a)} \left(\int_{a}^{\omega} l(x) e^{-rx} dx + r \int_{a}^{\omega} (x-\bar{x}) l(x) e^{-rx} dx + \frac{r^2}{2} \int_{a}^{\omega} (x-\bar{x})^2 l(x) e^{-rx} dx \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r \frac{\int_{a}^{\omega} (x-\bar{x}) l(x) e^{-rx} dx}{\int_{a}^{\omega} l(x) e^{-rx} dx} + \frac{r^2}{2} \frac{\int_{a}^{\omega} (x-\bar{x})^2 l(x) e^{-rx} dx}{\int_{a}^{\omega} l(x) e^{-rx} dx} \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r(\bar{x}-\bar{x}) + \frac{r^2}{2} \sigma^2 \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r(\bar{x}-\bar{x}) + \frac{r^2}{2} \sigma^2 \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r(\bar{x}-\bar{x}) + \frac{r^2}{2} \sigma^2 \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r(\bar{x}-\bar{x}) + \frac{r^2}{2} \sigma^2 \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + r(\bar{x}-\bar{x}) + \frac{r^2}{2} \sigma^2 \right) = e^{r(\bar{x}-a)} \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx \left(1 + \frac{r^2}{2} \sigma^2 \right), \tag{9}$$

here

$$\bar{x} = \frac{\int_a^\omega x N(x) dx}{\int_a^\omega N(x) dx} = \frac{\int_a^\omega x e^{-rx} l(x) dx}{\int_a^\omega e^{-rx} l(x) dx}$$
(10)

is the mean population age in the open age interval and

$$\sigma^{2} = \frac{\int_{a}^{\omega} (x - \bar{x})^{2} N(x) dx}{\int_{a}^{\omega} N(x) dx} = \frac{\int_{a}^{\omega} (x - \bar{x})^{2} e^{-rx} l(x) dx}{\int_{a}^{\omega} e^{-rx} l(x) dx}$$
(11)

is the variance of age of individuals in the open age interval in the stable population.

Using the identity $l(a) = -\int_{a}^{\omega} \frac{dl(x)}{dx} dx$, the survival is approximated as: $l(a) = -\int_{a}^{\omega} \frac{dl(x)}{dx} dx = -e^{r\overline{x_{D}}} \int_{a}^{\omega} e^{r(x-\overline{x_{D}})} e^{-rx} \frac{dl(x)}{dx} dx \approx -e^{r\overline{x_{D}}} \left(\int_{a}^{\omega} e^{-rx} \frac{dl(x)}{dx} dx + r \int_{a}^{\omega} (x-\overline{x_{D}}) e^{-rx} \frac{dl(x)}{dx} dx + \frac{r^{2}}{2} \int_{a}^{\omega} (x-\overline{x_{D}})^{2} e^{-rx} \frac{dl(x)}{dx} dx \right) = -e^{r\overline{x_{D}}} \int_{a}^{\omega} e^{-rx} \frac{dl(x)}{dx} dx \left(1 + \frac{r^{2}}{2} \sigma_{D}^{2} \right),$ (12)

where:

$$\overline{x_D} = \frac{\int_a^{\omega} xD(x)dx}{\int_a^{\omega} D(x)dx} = \frac{\int_a^{\omega} xN(x)\mu(x)dx}{\int_a^{\omega} N(x)\mu(x)dx} = \frac{\int_a^{\omega} xe^{-rx}\frac{dl(x)}{dx}dx}{\int_a^{\omega} e^{-rx}\frac{dl(x)}{dx}dx}$$
(13)

is the mean age at death in the open age interval and

$$\sigma_D^2 = \frac{\int_a^\omega (x - \overline{x_D})^2 D(x) dx}{\int_a^\omega D(x) dx} = \frac{\int_a^\omega (x - \overline{x_D})^2 e^{-rx} \frac{dl(x)}{dx} dx}{\int_a^\omega e^{-rx} \frac{dl(x)}{dx} dx}$$
(14)

is the variance of age at death in the open age interval in the stable population.

Using (8), one gets:

$$\int_{a}^{\omega} e^{-rx} \frac{dl(x)}{dx} dx = e^{-rx} l(x) |_{a}^{\omega} + r \int_{a}^{\omega} l(x) e^{-rx} dx = -e^{-ra} l(a) + r \int_{a}^{\omega} l(x) e^{-rx} dx = -e^{-ra} l(a) \left(1 - r \int_{a}^{\omega} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r \int_{a}^{\infty} \frac{l(x)}{l(a)} e^{-r(x-a)} dx\right) = -e^{-ra} l(a) \left(1 - r$$

Substituting (15) into (12):

$$l(a) \approx e^{r(\overline{x_D} - a)} l(a) \frac{M_{a+}}{M_{a+} + r} \left(1 + \frac{r^2}{2} \sigma_D^2 \right), \tag{16}$$

which implies:

$$M_{a+} + r \approx e^{r(\overline{x_D} - a)} M_{a+} \left(1 + \frac{r^2}{2} \sigma_D^2 \right).$$
(17)

Substituting (17) into (9):

$$e(a) \approx \frac{1}{M_{a+}} e^{r(\bar{x} - \overline{x_D})} \left(1 + \frac{r^2}{2} (\sigma^2 - \sigma_D^2) \right).$$
(18)

To conclude the derivation, Mitra suggests relating the mean age at death to the mean population age using Eqs. (8), (10) and (13):

$$\overline{x_D} = \frac{\int_a^{\omega} xe^{-rx} \frac{dl(x)}{dx} dx}{\int_a^{\omega} e^{-rx} \frac{dl(x)}{dx} dx} = \frac{\int_a^{\omega} xe^{-rx} \frac{dl(x)}{dx} dx}{-e^{-ra} l(a) \frac{M_{a+}}{M_{a++r}}} = \frac{xe^{-rx} l(x)|_a^{\omega} + r \int_a^{\omega} xl(x)e^{-rx} dx - \int_a^{\omega} l(x)e^{-rx} dx}{-M_{a+} \int_a^{\omega} l(x)e^{-rx} dx} = \frac{a(M_{a+}+r)}{M_{a+}} + \frac{1-r\bar{x}}{M_{a+}} = a + \frac{1-r(\bar{x}-a)}{M_{a+}}.$$
(19)

Substituting (19) into (18) and rearranging terms:

$$e(a) \approx \frac{1}{M_{a+}} e^{r \frac{(\bar{x}-a)(M_{a+}+r)-1}{M_{a+}}} \left(1 + \frac{r^2}{2} (\sigma^2 - \sigma_D^2) \right).$$
(20)

Mitra's formula (3) is obtained from (20) by dropping the higher order terms.

3. Sources of instability and biasedness of the Mitra model

As follows from the above derivation of the Mitra formula, there are three possible sources of biases in the Mitra model: neglecting the higher order terms in (20), erroneous estimates of the mean population age \bar{x} , or deviation of the population age composition from the (stable) one assumed in the model derivation. Within the model's own framework, the error in the growth parameter may be taken as a proxy for the latter source of model biases. Following, we investigate contribution of each of these sources to the estimation errors.

To assess the contribution of neglecting the higher order terms when moving from (20) κ (3), note that the magnitude of the growth parameter *r* does not usually exceed 0,05 (Keyfitz and Flieger 1990) and the variance of age at death is of the same magnitude as the variance of the population age distribution (but smaller at younger onset age *a* of the open age interval). Both the standard deviation of age at death and of population age distribution are of magnitude of several years, i.e., the quadratic term in (20) may be roughly assessed as no more than $\frac{0.05^2}{2}5^2 \approx 0.03$, i.e. three percentage points. Hence, the quadratic terms dropped in the Mitra model may not have caused the outliers observed in empirical tests. Note, however, that the quadratic terms addressed here are only the first ones in a series of non-linear terms neglected in the Mitra formula. In fact, linearization assumed in the Mitra formula may have profound consequences for the estimation accuracy when combined with large errors in the input value of the growth parameter (see the discussion further down in the next section).

Sensitivity to estimation errors of the mean population age \bar{x} might have been another source of instability in the Mitra model. Indeed, Coale (1985) has pointed to this parameter of the Mitra model as a problematic one when dealing with data subject to age exaggeration. Although, it was shown that the mean population age may be efficiently estimated even in the case of age exaggeration or in the absence of data on actual population age composition in the open age interval (Ediev 2017, 2018), such (regression-based) estimates may nonetheless be subject to substantial errors and lead to Models' overall instability. To explore what might have been the effect of biased estimation of the mean population age in the Mitra model, consider sensitivity of the life expectancy (3) to this parameter:

$$\frac{dln[e_M(a)]}{d\bar{x}} = \frac{d}{d\bar{x}} \left[r \frac{(\bar{x}-a)(M_{a+}+r)-1}{M_{a+}} \right] = r \frac{(M_{a+}+r)}{M_{a+}} \sim r,$$
(21)

The magnitude of errors in the mean population age may roughly be assessed as not exceeding 0.5 years based on standard errors of the regression model proposed for the parameter (Ediev 2016, table 2). Taking this into account, we get from (21), at r < 0.05:

 $\delta ln[e_M(a)] \sim 0.5r < 0.025, (22)$

i.e., approximation errors in the mean age \bar{x} contribute to the relative error of $e_M(a)$ no more than about 3 percentage points.

To assess the contribution of the third possible source of instability of the Mitra model, let us consider sensitivity of (3) to the growth parameter r:

$$\frac{dln[e_M(a)]}{dr} = \frac{d}{dr} \left[r \frac{(\bar{x}-a)(M_{a+}+r)-1}{M_{a+}} \right] = (\bar{x}-a) \left(1 + \frac{1}{M_{a+}} \right) + r \frac{(\bar{x}-a)}{M_{a+}} - \frac{1}{M_{a+}}.$$
(23)

Noting that $\frac{1}{M_{a+}}$ equals, roughly, the remaining life expectancy in the open age interval, while the age difference $(\bar{x} - a)$ is, roughly, half the remaining life expectancy, we get: $\frac{dln[e_M(a)]}{dr} \approx \frac{e^2(a)(1+r)-e(a)}{2} \approx \frac{e^2(a)}{2}.$ (24)

The remaining life expectancy in the open age interval is of magnitude of years to dozens of years while the error of the growth parameter might be up to a percentage point. Therefore, it follows from (24) that erroneous growth parameter may lead to enormous deviations of the estimated life expectancy from the actual one. Taking more broadly, violations of the Mitra model's assumption about stability of the age structure of the population in the open age interval appear to be the prime source of the model's instability.

Noting the quadratic term in (3), one may explain the empirical observation of predominantly positive biases, especially of the large ones, of the Mitra model (Ediev 2017, 2018). To this end, consider finite increments of the growth parameter that, as we have established above, drives most the biases of the Mitra model. Let us assume the simplest (yet, characteristic) case when the true growth parameter equals zero, i.e. the population is actually a stationary one, and the correcting multiplier in (3) equals one. In that case, the relative error of the model when assuming, erroneously, non-zero growth parameter r, equals:

$$e^{r\frac{(\bar{x}-a)(M_{a+}+r)-1}{M_{a+}}} - 1 \approx e^{0.5re(a)(re(a)-1)} - 1.$$
(25)

The lower limit of expression (25) is attained at re(a) = 0.5 and equals -0.11. Hence, the model may underestimate the remaining life expectancy is by no more than about 11 percent. On the overestimation side, however, the errors have no upper limit and may easily lead to outliers, as the power in (25) is a parabolic function that opens up. For example, at re(a) = 2, the relative bias (25) reaches 1,72 (model error of 172 percent) and at re(a) = -2, the model error equals 1909 percent.

Discussion. On the combined and iterative methods of estimating the remaining life expectancy at old age

We have revisited Mitra's derivation while streamlining the derivation and keeping the higher-order terms in the formula. That enabled us to conduct a comprehensive analysis of the sources of the biases and of instability of the results of the Mitra model. Our analysis shows that the leading source of instability of the model is inadequacy of the growth parameter used in the model or, by implication, the violation of the assumption of stability of the population age composition.

We have also confirmed an earlier hypothesis (Ediev 2017, 2018) that upward biasedness of the model was not a coincidental observation related to particular data used but is, rather, an essential feature of the model. This latter observation provides a formal ground to the earlier proposed combined method (Ediev 2018) that relies on using the minimum of the Mitra and of the classical estimates of the remaining life expectancy in the open age interval:

$$e_a = \min(e_a^{Clas.}, e_a^{M.}) \tag{26}$$

Taking positive difference between the Mitra and the classical estimates as indication of an outlier of the Mitra model, one may extend the combined method by including more than two alternative estimates:

$$e_a = \begin{cases} e_a^{M} \text{ if } e_a^{Clas.} \ge e_a^{M} \\ e_a^{\text{alt.}} \text{ if } e_a^{Clas.} < e_a^{M} \end{cases}$$

$$(27)$$

where $e_a^{\text{alt.}}$ is an alternative estimate of the remaining life expectancy to be used when an outlier of the Mitra model is detected. With the classical estimate used as such an alternative, $e_a^{\text{alt.}} = e_a^{Clas.}$, (27) turns into (26). More accurate alternative estimates, such as the Horiuchi-Coale estimate (Ediev 2018; Horiuchi and Coale 1982), may be used to improve accuracy of the combined method. Even better perhaps, one may consider models that do not assume population stability at all to form the alternative estimates $e_a^{\text{alt.}}$. A prominent choice of that kind might be the regression model relating the remaining life expectancy to the death rate at a given age, because of the model's independence of population stability and the lack of correlation between the estimates based on the regression model and on models by Horiuchi-Coale and Mitra (Ediev 2019).

We found no strong effects of dropping the quadratic terms in the Mitra formula. This, presumably, indicates that the effects of neglecting other higher-order terms should be of the same magnitude (inducing errors of several percentage points) or less. Despite that, combined effect of all non-linear terms dropped in the Mitra formula might well be sizable, especially in combination with large errors in the input value of the growth parameter. Indeed, note that the Mitra formula suggests non-monotone (log-quadratic) relationship of the observed death rate in the open age interval to the growth parameter: $M_{a+} \sim e^{r\frac{(\bar{x}-a)(M_{a+}+r)-1}{M_{a+}}}$. This, however, is merely a reflection of inadequacy of the linear approximation used in deriving the formula, because in reality the open age interval death rate in the open age interval is a weighted average of age-specific death rates in the interval with weights equal to the population size in respective age. When growth parameter is higher, the stable population age composition is weighted heavier at younger ages (with lower mortality) and the open age interval death rate becomes lower.

To address the Mitra model's inadequacy at extreme values of the growth parameter, we suggest making use of the constrained mortality extrapolation model (Ediev 2017) that enables projecting, rather accurately, the death rates to old ages given the reasonable estimate of the remaining life expectancy at old age. Having the death rates $\mu(x)$ projected to old ages and assuming the growth parameter, one may form the life table functions and the stable population age structure $N(x) \sim l(x)e^{-rx}$. Now, the idea is to use this stable population age structure directly, without relying on indirect regressions as in the Horiuchi-Coale method or on linear approximations as in the Mitra model, in order to derive the correction coefficient necessary to estimate the remaining life expectancy from the observed death rate in the open age interval:

$$e_{a} = \frac{\int_{a}^{\omega} l(x)dx}{l(a)} = \frac{\int_{a}^{\omega} l(x)dx}{l(a)} \frac{M_{a+}}{M_{a+}} = M_{a+}^{-1} \cdot \frac{\int_{a}^{\omega} l(x)dx}{l(a)} \frac{\int_{a}^{\omega} \mu(x)l(x)e^{-rx}dx}{\int_{a}^{\omega} l(x)e^{-rx}dx} = M_{a+}^{-1} \cdot K, \quad (28)$$

here

$$K = \frac{\int_a^{\omega} l(x)dx}{l(a)} \frac{\int_a^{\omega} \mu(x)l(x)e^{-rx}dx}{\int_a^{\omega} l(x)e^{-rx}dx} = \frac{\int_a^{\omega} l(x)dx}{\int_a^{\omega} l(x)e^{-rx}dx} \frac{\int_a^{\omega} d(x)e^{-rx}dx}{\int_a^{\omega} d(x)dx}$$
(29)

is the correction coefficient alternative to the ones used in Horiuchi-Coale or Mitra methods, $d(x) = \mu(x)l(x)$ is the deaths density function. The latter expression for the correction coefficient shows the coefficient to be a robust function of the stable population age composition. Hence, the idea is to find such an estimate for the remaining life expectancy e_a that, in combination with the constrained mortality extrapolation model and relation (28), is consistent with the observed death rate M_{a+} .

The estimation procedure for the stable population model (28) combined with the constrained extrapolation may efficiently be organized in an iterative manner, where the life expectancy at each step is used to generate the correction coefficient K and improve the life expectancy estimate in the next step. We have empirically evaluated this iterative method on the Human Mortality Database (University of California, Berkeley and Max Planck Institute for Demographic Research (Rostock) 2018) using the same testing methodology as in earlier works (Ediev 2018). The empirical tests show that the iterative method differs from the Mitra model in two important ways. Firstly, the iterative method appears not to be prone to produce outliers and is close, in this respect, to the Horiuchi-Coale model. Secondly, unlike in both the Mitra and Horiuchi-Coale models, estimation errors of the iterative method are not skewed towards positive biases. These differences confirm that the origins of outliers and of predominantly positive biases of the Mitra model lie in the log-quadratic form of approximation used in the Mitra formula. This approximation serving perfectly in cases of small absolute values of the growth parameter performs increasingly inadequately when the growth parameter gets higher in absolute terms. This is a problematic feature of the model in cases of low-mortality populations that experience rapid ageing and expansions of elderly populations.

Indeed, the very stable population model (28) that forms the basis for the Horiuchi-Coale and Mitra models as well as for our iterative model may appear inadequate for cases of rapid mortality decline. That is because the population age composition is the context of changing mortality is formed by cohort survival profiles and not by the period life tables assumed in the mentioned models. In that case, the cross-sectional average length of life (Brouard 1986; Guillot 2003) or other tempoadjusted indicators of lifespan (Ediev 2008) may be used instead of the life expectancy indicators in generating inputs for the correction coefficient (29). Alternatively, one may modify our iterative model by applying it to a period of time covering the cohorts forming the currently elderly population. In this way, after estimating the life expectancy and the death rates in each iteration, one could calculate the cohort survival and use it in estimating the correction coefficients (29) for the next step. Yet another prominent improvement to our model might be to estimate both the life expectancy and the growth parameter within the iteration procedure. The growth parameter may be estimated based on the ratio of the population size in the open age interval to the population size in the age groups below the open age interval. These improvements to our model need more elaboration and will be considered in further research.

References

- Bennett, Neil G. and Shiro Horiuchi. 1984. "Mortality Estimation from Registered Deaths in Less Developed Countries." *Demography* 21(2):217.
- Brouard, Nicolas. 1986. "Structure et Dynamique Des Populations. La Pyramide Des Années à Vivre, Aspects Nationaux et Exemples Régionaux." *Espace*, *Populations, Sociétés* 4(2):157–68.
- Coale, Ansley J. 1985. "Estimating the Expectation of Life at Old Ages: Comments on the Article by Mitra." *Population Studies* 39(3):507–9.
- Coale, Ansley J. and EE Kisker. 1990. "Defects in Data on Old-Age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages." *Asian and Pacific Population Forum* 4:1–31.
- Duthé, Géraldine, Irina Badurashvili, Karine Kuyumjyan, France Meslé, and Jacques Vallin. 2010. "Mortality in the Caucasus: An Attempt to Re-Estimate Recent Mortality Trends in Armenia and Georgia." *Demographic Research* 22(23):691–732.

- Ediev, Dalkhat M. 2008. On the Theory of Distortions of Period Estimates of the Quantum Caused by the Tempo Changes. 03. Vienna, Austria.
- Ediev, Dalkhat M. 2016. *Expectation of Life at Old Age: Revisiting Horiuchi-Coale and Reconciling with Mitra*. WP-16-010. Laxenburg.
- Ediev, Dalkhat M. 2017. "Constrained Mortality Extrapolation to Old Age: An Empirical Assessment." *European Journal of Population*.
- Ediev, Dalkhat M. 2018. "Expectation of Life at Old Age: Revisiting Horiuchi-Coale and Reconciling with Mitra." *Genus* 74(1).
- Ediev, Dalkhat M. 2019. "On the Statistical Dependence of Life Expectancy and Mortality Rate at a Given Age." *Voprosy Statistiki* 26(6):39–46.
- Guillot, Michel. 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.
- Horiuchi, S. and Ansley J. Coale. 1982. "A Simple Equation for Estimating the Expectation of Life at Old Ages." *Population Studies* 36(2):317–26.
- Keyfitz, Nathan and Hal Caswell. 2005. *Applied Mathematical Demography*. Springer.
- Keyfitz, Nathan and Wilhelm. Flieger. 1990. *World Population Growth and Aging* : *Demographic Trends in the Late Twentieth Century*. University of Chicago Press.
- Khlat, Myriam and Youssef Courbage. 1996. "Mortality and Causes of Death of Moroccans in France, 1979-91." *Population. English Selection* 8:59–94.
- Kibele, Eva, Rembrandt D. Scholz, and Vladimir M. Shkolnikov. 2008. "Low Migrant Mortality in Germany for Men Aged 65 and Older: Fact or Artifact?" *European Journal of Epidemiology* 23(6):389–93.
- Mitra, S. 1984. "Estimating the Expectation of Life at Older Ages." *Population Studies* 38(2):313–19.
- Mitra, S. 1985. "On Estimating the Expectation of Life at Old Ages: Reply to Professor Coale." *Population Studies* 39(3):511–12.
- Preston, Samuel H., I. T. Elo, I. Rosenwaike, and M. Hill. 1996. "African-American Mortality at Older Ages: Results of a Matching Study." *Demography* 33(2):193– 209.

- Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishers.
- Shryock, H. S. and Jacob S. Siegel. 1973. *The Methods and Materials of Demography*. Washington D.C.: United States Bureau of the Census.
- University of California, Berkeley and Max Planck Institute for Demographic Research (Rostock). 2018. "Human Mortality Database. Online Database Sponsored by University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)." Retrieved May 15, 2018 (www.mortality.org).