Population Research and Policy Review (2025) 44:7
https://doi.org/10.1007/s11113-024-09926-y

ORIGINAL RESEARCH

Optimized Random-Combinations of Total Fertility Rates and Life Expectancies at Birth for Probabilistic Population Projections

Man Li · Shanwen Zhu · Zhenglian Wang · Qiushi Feng · Junni Zhang · Fengqing Chao, et al. ***[full author details at the end of the article]***

Received: 8 December 2023 / Accepted: 4 December 2024

Abstract

Many studies indicate that Total fertility rates (TFR(t)) are negatively correlated with life expectancies at birth (e0(t)). We found that complete random-combinations of TFR(t) and e0(t) would result in about 24% and 22.2% of improbable combinations in probabilistic population projections (PPPs) for developing and developed countries, respectively, namely, high (or low) TFR(t) combined with high (or low) female e0(t). Thus, we propose optimized random-combinations of probabilistically projected TFR(t) and e0(t) for PPPs, and we use different strategies of the optimized randomcombinations across developing and developed countries due to different empirical patterns observed. As illustrative applications, we conducted PPPs for 11 developing countries (Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand, Viet Nam), and 6 developed countries (Canada, France, Italy, Japan, the United Kingdom, the United States), using our proposed optimized random-combinations of probabilistically projected TFR(t) and e0(t). We found that optimized random-combinations largely reduce percentages of improbable combinations of TFR(t) and e0(t) and substantially narrow the prediction intervals width compared to complete random-combinations in both developing countries and developed countries. This is important in a real-world practical sense since it would substantially improve the accuracy of PPPs, which are useful for socioeconomic planning. The present study is part of our ongoing research program on probabilistic households and living arrangement projections (PHPs) that builds upon and is consistent with the UNPD PPPs. The PHPs are useful for various studies of healthy aging and sustainable development.

Keywords Probabilistic population projections · Probabilistic households and living arrangement projections · Correlations between total fertility rates and life expectations at birth · Optimized random-combinations · Complete random-combinations

Man Li and Shanwen Zhu co-first authors.

Published online: 08 January 2025

Introduction

The United Nations Population Division (UNPD) made a great historical breakthrough by publishing probabilistic population projections (PPPs) for each country worldwide since June 2014 in the World Population Prospects (United Nations, 2013; Alkema et al., 2015). The UNPD PPPs are based on probabilistic projections of total fertility rates (TFR(t)) in year t and life expectancy at age 0 (e0(t)) in year t, as well as the complete random-combinations of TFR(t) and e0(t) (Alkema et al., 2011; Gerland et al., 2014; Raftery et al., 2014; Ševčíková et al., 2013). The term "complete random-combinations" (or "completely randomly combined") in this article refers random-combinations without any scope strategy of assembling expert options and judgements.

The research question to be addressed in the present article is: "To what extent the complete random-combinations of the probabilistically projected TFR(t) and e0(t) may result in improbable combinations of TFR(t) and e0(t) in future years?" The authors believe that the correlations between TFR(t) and female e0(t) need to be considered when conducting PPPs. This is because that if complete random-combinations of the projected trajectories of TFR(t) and e0(t) in the future years are adopted, very low (or high) trajectories of TFR(t) and very low (or high) trajectories of e0(t) may be randomly combined, which are inconsistent with empirical observations that have been reported in numerous previous studies and may accordingly reduce the accuracy in the outcomes of PPPs.

Using the UNPD-released probabilistically projected TFR(t) and female e0(t) in the future years, this study investigates the percentages of the improbable combinations if the projected TFR(t) and female e0(t) are completely randomly combined. We explore different optimized random-combinations of probabilistically projected 2,000 trajectories of TFR(t) and 2,000 trajectories of female e0(t) for each of the 11 developing countries (Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand, and Viet Nam) and each of the 6 developed countries (Canada, France, Italy, Japan, the United Kingdom and the United States (countries are listed in alphabetical order) to substantially reduce the percentages of improbable combinations of TFR(t) and female e0(t). The term "optimized random-combinations" in this article refers to random-combinations with appropriate expert judgements.

In the Discussion, we also describe and discuss our ongoing research program of probabilistic household and living arrangement projections (PHPs) for informed decision-making and sustainable development research, which is built upon and consistent with the probabilistic population projections of UNPD.

Background

The basic ideas underlying the UNPD PPPs are to first estimate prior probability distributions of TFR(t) and female e0(t) based on their historical trends from 1950 to the last year of the available estimates for almost all countries (Alkema, et al., 2011; Raftery et al., 2013) and relevant demographic transition theories

(Lee, 2003). The second step uses Bayesian hierarchical models to estimate posterior probability distributions of each country s trajectories of TFR(t) and female e0(t) in 2021–2100, combining information from prior probability distributions and country-specific demographic information. Different statistical models, such as the double logistic model and the first-order auto-regression model, are used to model the parameters at various stages of the demographic transition. The third step is to use a Markov chain Monte Carlo (MCMC) algorithm to sample the parameters from their posterior probability distributions and simulate the probabilistically projected 2,000 trajectories of TFR(t) and 2,000 trajectories of female e0(t) independently. Since male e0(t) and female e0(t) are highly correlated, the UNPD PPPs project probabilistically the female e0(t) first and then project the male e0(t) according to the historically observed associations between male e0(t) and female e0(t) and the projected gender gap in e0(t). In the fourth step, the 2,000 trajectories of TFR(t) and the 2,000 trajectories of female e0(t) of each country are completely randomly combined to produce the PPPs for each of the countries worldwide in 2021–2100 (UNPD, 2022a, 2022b).

Each of the 2,000 trajectories of TFR(t) is combined with standard schedules of age-specific fertility rates to obtain a projected trajectory of age-specific fertility rates. The sex-age-specific mortality rates are calculated to be consistent with the 2,000 trajectories of projected e0(t) for each year, with year-specific age patterns of death rates for each country separately projected using model life tables, a modified Lee-Carter method, or other models. The trajectories of age-specific fertility rates and the trajectories of sex-age-specific mortality rates are then combined with deterministically projected sex-age-specific numbers of net migration. Together with the baseline population by age and sex, these trajectories are input of the cohort-component method for projecting populations to produce 2,000 probabilistically projected trajectories of future populations by age and sex. Various population and demographic indicators can be summarized with 80% and 95% prediction intervals. This statistical framework for probabilistic projections and estimations of fertility, mortality and population for all countries has also been embodied in a computer program package named "BayesPop" written in R and publicly available on the Comprehensive R Archive Network (CRAN) (Ševčíková, H., and Raftery, A. E.,2016).

Analyses of empirical data from different countries show that high (or low) TFR(t) is associated with low (or high) e0(t). Prior studies have found a negative correlation between TFR(t) and e0(t) based on either country-specific time series data (Jannat, 2022; Jafrin et al., 2021; Onwube et al., 2021; Abbas and Awan, 2015; Delavari et al., 2016; Sarlak & Savari, 2016; Summoogum & Fah, 2016; Singariya, 2013) or cross-sectional data from multiple countries (Barlow & Vissandjée, 2011; Girum et al., 2018; Husain, 2002; Kabir, 2008; Kalita, 2018; Low et al., 2013; Mondal & Shitan, 2013, 2014; Mondal et al., 2015). Statistically significant negative correlations between TFR(t) and e0(t) have been found in Bangladesh (Jannat, 2022), in 5 countries of the South Asian Association for Regional Cooperation (Jafrin, 2021), in Iran (Delavari, 2016), in Oman (Sarlak & Savari, 2016), in India (Singariya, 2013), in Nigeria (Onwube, 2021), and in one developed and three developing countries in Asia (Summoogum & Fah, 2016).

Multi-country studies also show evidence of the negative correlation between TFR(t) and e0(t). Girum (2018) found a statistically significant negative linear relationship between TFR(t) and e0(t) using aggregates of health indicator data from the WHO, World Bank, UNPD, and UNICEF databases in 83 countries with low and medium human development index scores. This finding is consistent with the studies conducted by Low (2013) in 177 countries, in 91 developing countries by Husain (2002), Kabir (2008), Mondal and Shitan (2013, 2014), and Barlow and Vissandjée (2011) in 77 countries. Within-country studies also support the negative relationship between TFR(t) and e0(t) (Krupp, 2012; Nettle, 2011).

Data and Methods

Data

The dada analyzed in this article are derived from the UNPD World Population Prospects report (United Nations, 2022), including country-specific annual time series data of TFR(t) and female e0(t) in 1950–2020 in each of the 11 developing countries of Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand, and Viet Nam, and each of the 6 developed countries of Canada, France, Italy, Japan, the United Kingdom and the United States.

Measurements

To further confirm that high (or low) TFR(t) is associated with low (or high) female e0(t) as reviewed in the Background section, we estimated Kendall s Tau-b rank correlation coefficients to empirically analyze temporal covariations of the country-specific annual time series data of TFR(t) and female e0(t) in 1950–2020 in 11 developing countries and 6 developed countries as presented in the first section of Results section, using the data from the UNPD World Population Prospects report (United Nations, 2022).

Since the correlations between TFR(t) and female e0(t) observed in 1950–2020 were relatively less strong in the developed countries than that in the developing countries (see Table 1), we use different measurements to investigate the percentages of the improbable combinations of TFR(t) and female e0(t), if they are completely randomly combined. The different measurements for developing countries and developed countries are described as follows.

For the developing countries, we divide the probabilistically projected trajectories of TFR(t) and female e0(t) released by UNPD into five groups of percentiles of the highest 20%, second highest 20%, middle 20%, second lowest 20%, and the lowest 20%, according to the means of the trajectories of TFR(t) and female e0(t). Because TFR(t) is negatively correlated with female e0(t), it would result in 6 groups of improbable combinations as follows in developing countries, if TFR(t) and female e0(t) are completely randomly combined:

Table 1 Kendall s Tau-b rank correlation coefficients measuring correlations between observed TFR(t) and female e0(t) in1950-2020 in the 11 developing countries and 6 developed countries

Country	n (TFR)	n (e0)	Tau-b rank correlation coefficient	95% Confidential interval	*P* value
11 developing countries					
Brazil	72	72	−.9775	−1.000 ~ −0.952	*P* < .001
China	72	72	−.6703	−0.780 ~ −0.556	*P* < .001
Indonesia	72	72	−.8065	−0.911 ~ −0.702	*P* < .001
Madagascar	72	72	−.8816	−0.963 ~ −0.800	*P* < .001
Pakistan	72	72	−.8180	−0.901 ~ −0.692	*P* < .001
Philippines	72	72	−.9718	−1.000 ~ −0.940	*P* < .001
Saudi Arabia	72	72	−.8618	−0.964 ~ −0.759	*P* < .001
Singapore	72	72	−.8934	−0.938 ~ −0.849	*P* < .001
Sri Lanka	72	72	−.8522	−0.933 ~ −0.772	*P* < .001
Thailand	72	72	−.9370	−0.980 ~ −0.894	*P* < .001
Viet Nam	72	72	−.7452	−0.864 ~ −0.626	*P* < .001
6 developed countries					
Canada	72	72	−.7886	−0.861 ~ −0.716	*P* < .001
France	72	72	−.4558	−0.612 ~ −0.300	*P* < .001
Italy	72	72	−.5649	−0.681 ~ −0.449	*P* < .001
Japan	72	72	−.7692	−0.855 ~ −0.684	*P* < .001
United Kingdom	72	72	−.4697	−0.583 ~ −0.356	*P* < .001
The United States	72	72	−.4095	−0.579 ~ −0.226	*P* < .001

Data source: The Kendall s Tau-b rank correlation coefficients listed in this Table 1 were based on the time series data of observed TFR(t) and female e0(t) in 1950–2020 in the 17 countries from the UNPD World Population Prospects report (United Nations, 2022)

(1) The 20% highest TFR(t) being combined with the 20% highest female e0(t);
(2) The 20% highest TFR(t) being combined with the 20% second highest female e0(t);
(3) The 20% second highest TFR(t) being combined with the 20% highest female e0(t);
(4) The 20% lowest TFR(t) being combined with the 20% lowest female e0(t);
(5) The 20% lowest TFR(t) being combined with the 20% second low female e0(t);
(6) The 20% second lowest TFR(t) being combined with the 20% lowest female e0(t).

For the developed countries, we divide the probabilistically projected trajectories of TFR(t) and female e0(t) released by UNPD into three groups of percentiles of the 33.3% highest, 33.3% middle, and 33.3% lowest, according to the means of the trajectories of TFR(t) and female e0(t). It would result in 2 groups of improbable combinations as follows in developed countries, if TFR(t) and female e0(t) are completely randomly combined:

(a) The 33.33% highest TFR(t) being combined with the 33.33% highest female e0(t);
(b) The 33.33% lowest TFR(t) being combined with the 33.33% lowest female e0(t).

Methods for the optimized random-combinations of TFR(t) and female e0(t) in the PPPs.

The method for optimized random-combinations of TFR(t) and female e0(t) for the PPPs in developing countries is outlined in (O5-1) ~ (O5-5) below (abbreviated as "optimized random-combinations for developing countries" hereafter):

(O5-1) Randomly combine the highest 20% of TFR(t) and the lowest 20% of female e0(t);
(O5-2) Randomly combine the second highest 20% of TFR(t) and the second lowest 20% of female e0(t);
(O5-3) Randomly combine the middle 20% of TFR(t) and the middle 20% of female e0(t);
(O5-4) Randomly combine the second lowest 20% of TFR(t) and the second highest 20% of female e0(t);
(O5-5) Randomly combine the lowest 20% of TFR(t) and the highest 20% of female e0(t).

The method for optimized random-combinations of TFR(t) and female e0(t) for the PPPs in developed countries is outlined in (O3-1) ~ (O3-3) below (abbreviated as "optimized random-combinations for developed countries" hereafter):

(O3-1) Randomly combine the 33.3% highest of TFR(t) and the 33.3% lowest of female e0(t);
(O3-2) Randomly combine the 33.3% middle of TFR(t) and the 33.3% middle of female e0(t);
(O3-3) Randomly combine the 33.3% lowest of TFR(t) and the 33.3% highest of female e0(t).

Results

Empirical Data Analyses of the Correlations Between TFR(t) and Female e0(t)

To further confirm that high (or low) TFR(t) is associated with low (or high) female e0(t) as reviewed in the Background, we empirically analyzed the national annual time series data of TFR(t) and female e0(t) in 1950–2020 in 11 developing countries of Brazil, China, Indonesia, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand and Viet Nam, and 6 developed countries of Canada, France Italy, Japan, the United Kingdom, and the United States (countries are listed alphabetically). The data were taken from the UNPD World Population Prospects report (United Nations, 2022). We estimated Kendall s Tau-b rank correlation coefficients as a nonparametric measure of the strength of the correlation between TFR(t) and

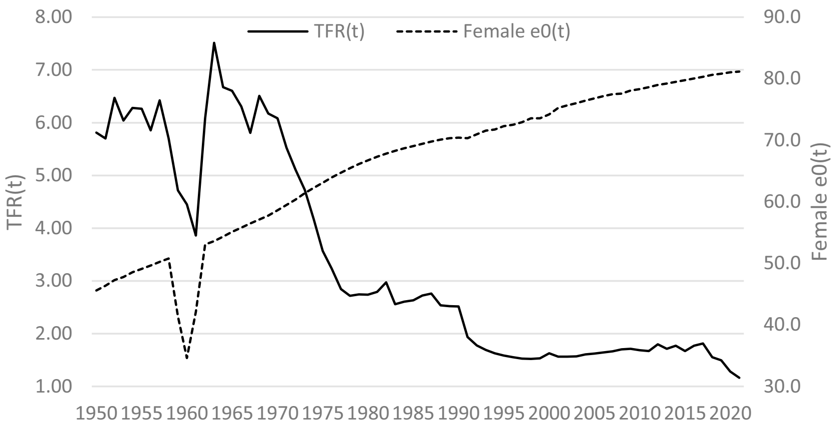


Fig. 1 Observed TFR(t) and female e0(t) in China, 1950–2020

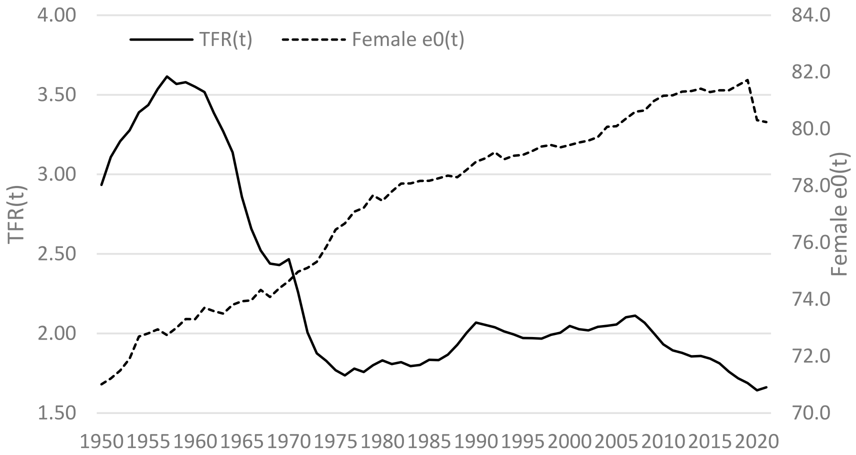


Fig. 2 Observed TFR(t) and female e0(t) in the United States, 1950–2020

female e0(t). As shown in Table 1, there are highly significant negative correlations ($P < 0.001$) between TFR(t) and female e0(t) in the 11 developing countries, as well as in the 6 developed countries. All of the p values of Kendall s Tau-b rank correlation coefficients in these 17 countries were less than 0.001 (see Table 1). The visual graphics of TFR(t) and female e0(t) in 1950–2020 for China and the United States are presented in Figs. 1, 2, and for the other 15 countries are presented in the online Supplementary Materials (SM) SM-Figs. 1–15. The Figs. 1, 2 and SM-Figs. 1–15 show the clear patterns of negative correlations between TFR(t) and female e0(t) in the past 71 years in the 17 countries.

Expectations and Empirical Verification of the Improbable Combinations of TFR(t) and Female e0(t) if they are Completely Randomly Combined

As shown in Table 1, the tau-b rank correlation coefficients of the TFR(t) and female e0(t) in 1950–2020 in the 11 developing countries, which are now in Phase II of the fertility transition (i.e., fertility rates decline and gradually being close or reach the replacement level) were mostly and substantially larger than those in the 6 developed countries, which are now in Phase III of the fertility transition (i.e., fertility rates substantially decline to below replacement level and then recover toward replacement fertility; ref. Alkema et al., 2011). Thus, our expectations of the implausible combinations of TFR(t) and female e0(t) for developing countries and developed countries are not the same, as discussed and empirically verified below.

We expect that the improbable combinations as outlined in (1) ~ (6) in the measurements sub-section of Data and Methods did not exist or were very rare in observed time series datasets of TFR(t) and female e0(t) in developing countries. To empirically assess this expectation, we sorted the time series data of TFR(t) and female e0(t) in the 71 years from 1950 to 2020 into five groups of percentiles of 20% highest, 20% second highest, 20% middle, 20% second lowest, and 20% lowest, according to the means of TFR(t) and female e0(t) in the country, for each of the 11 developing countries of Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand, and Viet Nam, respectively. We found that the percentages of the improbable combinations among the total numbers of combinations of TFR(t) and female e0(t) in 1950–2020 in each of the 11 developing countries are all 0% (see Table 2), and hence the expectation of the improbable combinations of TFR(t) and female e0(t) if they were completely randomly combined outlined in (1) ~ (6) of Sect. 2.2.1 is empirically verified by the observed time series data of the 11 developing countries.

We expect that the improbable combinations as outlined in (a) ~ (b) of the measurements sub-section of Data and Methods section did not exist or were very rare in observed time series datasets of TFR(t) and female e0(t) in developed countries. To empirically assess this expectation, we sorted the time series data of TFR(t) and female e0(t) in the 71 years from 1950 to 2020 into three groups of percentiles of 33.33% high, 33.33% middle, and 33.33% low, according to the means of TFR(t) and female e0(t) in the country, for each of the 6 developed countries. We found that the percentages of the improbable combinations among the total numbers of combinations of TFR(t) and female e0(t) in 1950–2020 in Canada, France, Italy, Japan, the United Kingdom and the United States are 0%, 0%, 0%, 0%, 0% and 1.4%, respectively (see Table 3). Hence, the expectation of the improbable combinations of TFR(t) and female e0(t) if they were completely randomly combined outlined in (a) ~ (b) is empirically verified by the observed time series data of the 6 developed countries. Note that the 1.4% of the improbable combinations of TFR(t) and female e0(t) in 1950–2020 in the United States was because TFR(t) reached its peak (2.1) in 2007 due to the economic boom in the early 2000s which resulted in a huge number of international immigrants who had high fertility rates (Camarota, 2011).

Table 2 Percentages of the improbable combinations of the observed TFR(t) and female e0(t) in 1950–2019 in 11 developing countries

Country	20% highest TFR(t) with 20% highest e0(t)	20% highest TFR(t) with 20% 2nd high e0(t)	20% 2nd high TFR(t) with 20% highest e0(t)	20% lowest TFR(t) with 20% lowest e0(t)	20% lowest TFR(t) with 20% 2nd low e0(t)	20% 2nd low TFR(t) with 20% lowest e0(t)	Total % of improbable combinations
Brazil	0%	0%	0%	0%	0%	0%	0%
China	0%	0%	0%	0%	0%	0%	0%
Indonesia	0%	0%	0%	0%	0%	0%	0%
Madagascar	0%	0%	0%	0%	0%	0%	0%
Pakistan	0%	0%	0%	0%	0%	0%	0%
Philippines	0%	0%	0%	0%	0%	0%	0%
Saudi Arabia	0%	0%	0%	0%	0%	0%	0%
Singapore	0%	0%	0%	0%	0%	0%	0%
Sri Lanka	0%	0%	0%	0%	0%	0%	0%
Thailand	0%	0%	0%	0%	0%	0%	0%
Viet Nam	0%	0%	0%	0%	0%	0%	0%

Table 3 Percentages of the improbable combinations of the observed TFR(t) and female e0(t) in 1950–2020 in 6 developed countries

Country	33.33% high TFR(t) with 33.33% highest e0(t)	33.33% lowest TFR(t) with 33.33% lowest e0(t)	Total % of the improbable combinations
Canada	0%	0%	0%
France	0%	0%	0%
Italy	0%	0%	0%
Japan	0%	0%	0%
The United Kingdom	0%	0%	0%
The United States	1.4%	0%	1.4%

Table 4 Demonstrations of the expected number of improbable combinations of TFR(t) and female e0(t) (marked by orange color in boldface type), if projected 2,000 trajectories of TFR(t) and female e0(t) are completely randomly combined in the case of developing countries

TFR(t)	Female e0(t)					
	20% Highest	20% 2nd highest	20% middle	20% 2nd lowest	20% lowest	Total
20% Highest	**80**	80	80	80	80	400
20% 2nd Highest	**80**	**80**	80	80	80	400
20% Middle	80	80	80	80	80	400
20% 2nd Lowest	80	80	80	**80**	**80**	400
20% Lowest	80	80	80	80	**80**	400
Total	400	400	400	400	400	2000

There are in total 480 pairs (marked by orange color in boldface type) of the improbable complete random-combinations of TFR(t) and female e0(t), which account for 24% (=480/2,000) of the total number of complete random-combinations, for developing countries

The Improbable Combinations of the Probabilistically Projected Trajectories of TFR(t) and Female e0(t) When they are Completely Randomly Combined

The empirical analyses of the observed time series data presented in the first subsection of the Results imply that it would result in a considerable number of improbable combinations of trajectories of TFR(t) and female e0(t) if they were completely randomly combined for PPPs in the future years. As demonstrated in Table 4, we expect that there are 80 improbable combinations in each of the 6 groups of improbable combinations among the 2,000 pairs of completely randomly combined trajectories of TFR(t) and female e0(t) in future years for a developing country. In total, it is expected that there are 480 (=80×6) improbable combinations, which account for 24% (=480 /2,000) of the 2,000 pairs of completely randomly combined TFR(t) and female e0(t) trajectories used for PPPs in the developing countries.

To empirically assess this expectation, we downloaded the 1,000 sets of trajectories of TFR(t) and female e0(t) for the PPPs of the 11 developing countries including

Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand, and Viet Nam in 2020–2100 released by UNPD at the BayesPop website of the University of Washington. Note that the BayesPop Website of the University of Washington included 1,000 sets (rather than 2,000 sets) of trajectories of TFR(t) and e0(t) in the future years, and the UNPD PPP Website does not include trajectories of TFR(t) and e0(t). We sorted them into five groups of percentiles of 20% highest, 20% second highest, 20% middle, 20% second lowest, and 20% lowest, respectively, according to the means of the trajectories of TFR(t) and female e0(t). As shown in the first section of Table 5, for example, the percentages of improbable combinations among the complete random-combinations of TFR(t) and female e0(t) in the UNPD PPPs are 22.6%-24.8% (close to 24%) in 2060 for the 11 developing countries.

The SM-Tables 1–6 present more detailed numerical results of the ranges of 20% groups of TFR(t) and 20% groups of e0(t), and percentages of the improbable combinations, when the probabilistically projected TFR(t) and e0(t) are completely randomly combined, in the PPPs of 11 developing countries in 2030–2100. For example, the complete random-combinations of TFR(t) and e0(t) for PPPs in Brazil would result in that 23.8% of the combinations are improbable in 2060, including 20% highest TFR(t) (1.82–2.28) combined with 20% highest e0(t) (87.00–89.67), 20% lowest TFR(t) (0.55–1.27) combined with 20% lowest e0(t) (80.42–84.20) (ref. SM-Table 1). Clearly, our expectation for the developing countries outlined in the first paragraph of this Section is empirically verified.

For the developed countries (such as Canada, France Italy, Japan, the United Kingdom and the United States), as demonstrated in Table 6, we expect that there are 222 improbable combinations in each of the two groups of improbable combinations among the 2,000 pairs of completely randomly combined trajectories of TFR(t) and female e0(t) in future years for a developed country. In total, it is expected that there are 444 improbable combinations, which account for 22.2% (=444/2,000) of the 2,000 pairs of completely randomly combined TFR(t) and female e0(t) trajectories used for PPPs in the developed countries in the future years.

To empirically assess this expectation, we downloaded the 1,000 sets of trajectories of TFR(t) and female e0(t) for the PPPs of the 6 developed countries in 2020–2100 released by UNPD at the BayesPop website of the University of Washington. We sorted them into three groups of percentiles of 33.3% highest, 33.3% middle, and 33.3% lowest, respectively, according to the means of the trajectories of TFR(t) and female e0(t). As shown in the second section of Table 5, for example, the percentages of improbable combinations among the complete random-combinations of TFR(t) and female e0(t) in the PPPs are 19.6%-23.8% in 2060 for Canada, France. Italy, Japan, the United Kingdom and the United States. The SM-Tables 7–9 present more detailed numerical results of the ranges of 33.3% groups of TFR(t) and 33.3% groups of e0(t), and percentages of the improbable combinations, when projected TFR(t) and e0(t) are completely randomly combined in the PPPs of the 6 developed countries in 2030–2100. For example, the complete random-combinations of TFR(t) and e0(t) for PPPs in the United States would result in 24.42% of the improbable combinations in 2100, including 33.3% highest TFR(t) (1.91–2.78) combined with 33.3% highest e0(t) (91.41–100.36) and 33.3% lowest TFR(t) (1.10–1.75) combined

Table 5 Comparisons of percentages of the improbable combinations between the optimized random- combinations and completely random-combinations of probabilistically projected TFR(t) and female e0(t) for the PPPs of the 11 developing countries and the 6 developed countries

Year	2040		2060		2080		2100	
Country	Optimized random-comb	Completely random-comb	Optimized random-comb	Completely random-comb	Optimized Random-comb	Completely random-comb	Optimized random-comb	Completely random-comb
11 developing countries								
Brazil	7.3%	24.7%	1.2%	23.8%	3.1%	23.5%	8.3%	24.1%
China	9.6%	24.6%	3.3%	22.6%	2.1%	22.8%	6.1%	22.6%
Indonesia	6.9%	24.8%	1.9%	23.2%	2.3%	25.3%	6.9%	23.7%
Madagascar	4.4%	25.7%	0.9%	24.7%	0.2%	24.2%	2.2%	26.2%
Pakistan	6.9%	24.3%	1.8%	24.7%	1.6%	24.0%	6.9%	24.2%
Philippines	7.3%	24.0%	1.5%	24.8%	1.8%	25.0%	6.0%	23.9%
Saudi Arabia	8.5%	24.3%	1.7%	23.4%	1.4%	24.3%	6.6%	23.8%
Singapore	9.0%	24.9%	2.0%	24.6%	0.8%	23.8%	2.9%	24.2%
Sri Lanka	6.0%	25.6%	2.0%	24.1%	2.8%	23.9%	8.2%	24.3%
Thailand	5.0%	25.5%	1.2%	22.7%	3.2%	21.8%	9.3%	23.1%
Viet Nam	8.2%	23.5%	1.6%	24.0%	0.9%	24.7%	4.4%	23.3%
6 developed countries								
Canada	7.7%	19.3%	2.6%	19.7%	3.6%	20.1%	9.1%	20.3%
France	8.8%	21.9%	4.5%	19.6%	3.3%	20.5%	4.5%	20.6%
Italy	7.7%	20.2%	3.0%	21.3%	2.4%	22.1%	6.5%	21.5%
Japan	9.4%	22.8%	4.1%	22.3%	3.1%	21.5%	6.1%	21.7%
The United Kingdom	10.7%	24.8%	4.2%	23.8%	4.6%	21.7%	7.2%	21.9%
The United States	8.3%	21.5%	3.3%	23.4%	3.4%	23.9%	6.5%	24.4%

More detailed numerical results of the improbable combinations of *optimized random-combinations* and completely random-combinations are presented in the SM-Tables1-18

Table 6 Demonstrations of the expected numbers of improbable combinations of TFR(t) and female e0(t) (marked by orange color in boldface type), if projected 2,000 trajectories of TFR(t) and female e0(t) are completely randomly combined in the case of developed countries

TFR(t)	Female e0(t)			
	33.3% highest	33.3% middle	33.3% lowest	Total
33.3% Highest	**222**	222	222	666
33.3% Middle	222	224	222	668
33.3% Lowest	222	222	**222**	666
Total	666	668	666	2000

There are in total 444 pairs (marked by boldface type) of the improbable completely random-combinations of TFR(t) and female e0(t), which account for 22.2% (=444/2,000) of the total number of completely random-combinations, for developed countries

with 33.3% lowest e0(t) (82.55–89.00) (ref. SM-Table 9). Clearly, our expectation for the developed countries outlined in the first paragraph of this Section is empirically verified.

The Results of Optimized Random-Combinations of the Probabilistically Projected TFR(t) and Female e0(t) in Future Years

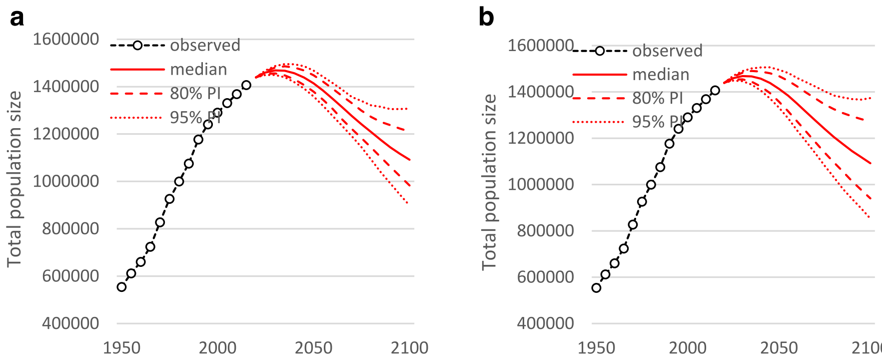
The analyses presented in the first and second sub-sections of Results section have verified that substantial percentages of combinations of trajectories of TFR(t) and female e0(t) in future years would be improbable when they are completely randomly combined. Inspired by the SCOPE strategy of assembling expert opinions in probabilistic mortality forecasting (Vaupel, 2019), we apply the approach of optimized random-combinations (i.e. appropriately stratified random-combinations) of the probabilistically projected trajectories of TFR(t) and female e0(t) in the PPPs for developing countries, as described in the third sub-section of the Data and Methods Section.

Note that using the optimized random-combinations does not only include the combinations as outlined in (O5-1) ~ (O5-5) in the third sub-section of the Methods section, but also results in other probable combinations, such as the highest 20% of TFR(t) combined with the second lowest 20% of female e0(t), second highest 20% of TFR(t) combined with the lowest 20% of female e0(t), the lowest 20% of TFR(t) combined with the second highest 20% of female e0(t), and the second lowest 20% of TFR(t) combined with the highest 20% of female e0(t). This is because each of the probabilistically projected TFR(t) trajectories and female e0(t) trajectories fluctuates up and down in future years, and the actual optimized random-combinations (according to the means) of TFR(t) and female e0(t) in each of the future years may not be exactly the same as the initial design of (O5-1) ~ (O5-5). Thus, using the optimized random-combinations may not totally avoid improbable combinations.

To test the approach as outlined in (O5-1) ~ (O5-5) in the third sub-section of the Data and Methods section for developing countries, we applied our extension of UNPD publicly released R package BayesPop to obtain the optimized

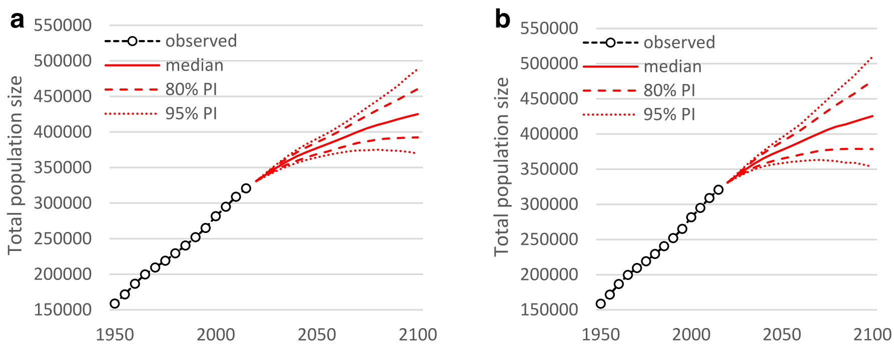
Table 7 The reductions of 80% predictive interval (PI width) and 95% PI width of projected total population size in 2040–2100, using optimized random-combinations of TFR(t) and e0(t), as compared with those using the completely random-combinations

Year	2040		2060		2080		2100	
country	Reduction of 80% PI width	Reduction of 95% PI width	Reduction of 80% PI width	Reduction of 95% PI width	Reduction of 80% PI width	Reduction of 95% PI width	Reduction of 80% PI width	Reduction of 95% PI width
11 developing countries								
Brazil	−13.45%	−9.13%	−22.99%	−20.05%	−20.79%	−16.67%	−17.58%	−13.00%
China	−23.51%	−24.02%	−43.76%	−38.85%	−43.69%	−34.89%	−31.37%	−21.51%
Indonesia	−18.01%	−22.75%	−25.11%	−28.90%	−29.64%	−28.79%	−22.73%	−22.37%
Madagascar	−15.97%	−13.44%	−18.60%	−14.48%	−22.87%	−17.73%	−22.30%	−17.08%
Pakistan	−11.73%	−8.12%	−15.21%	−15.88%	−15.94%	−16.91%	−17.68%	−20.48%
Philippines	−10.93%	−12.65%	−14.37%	−11.40%	−19.21%	−12.40%	−18.52%	−13.87%
Saudi Arabia	−11.41%	−10.69%	−25.36%	−23.54%	−25.08%	−26.57%	−21.14%	−18.25%
Singapore	−19.29%	−20.04%	−36.32%	−30.49%	−28.50%	−30.42%	−21.09%	−21.43%
Sri Lanka	−17.62%	−20.07%	−24.06%	−24.50%	−17.08%	−18.01%	−13.54%	−12.87%
Thailand	−21.74%	−17.39%	−22.14%	−25.73%	−16.05%	−20.45%	−10.26%	−16.00%
Viet Nam	−21.56%	−19.89%	−37.55%	−32.90%	−34.74%	−27.01%	−27.32%	−21.03%
6 developed countries								
Canada	−12.76%	−11.46%	−14.92%	−20.69%	−14.88%	17.74%	−9.31%	−18.76%
France	−14.57%	−12.59%	−17.78%	−17.65%	−15.92%	−12.17%	−13.76%	−8.91%
Italy	−23.48%	−23.07%	−30.96%	−33.80%	−21.03%	−27.93%	−16.31%	−18.44%
Japan	−24.86%	−23.42%	−34.79%	−36.98%	−30.56%	−28.76%	−20.83%	−20.33%
The United Kingdom	−20.90%	−19.07%	−33.63%	−28.58%	−34.35%	−27.71%	−28.75%	−21.20%
The United States	−20.90%	−19.07%	−33.63%	−28.58%	−34.35%	−27.71%	−28.75%	−21.20%



Year	2030	2040	2050	2060	2070	2080	2090	2100
Reductions of 80% PI width	-12.85%	-23.51%	-37.26%	-43.76%	-42.51%	-43.69%	-37.41%	-31.37%
Reductions of 95% PI width	-13.10%	-24.02%	-33.45%	-38.85%	-42.39%	-34.89%	-27.31%	-21.51%

Fig. 3 Total population size observed in 1950–2020 and probabilistically projected in 2021–2100, China. Summary: The reductions of 80% PI width and 95% PI width of projected total population size in 2030–2100 in China, using optimized random-combinations of TFR(t) and e0(t), as compared with that using the completely random-combinations. **a** Using the bayesPop and optimized random-combinations of projected TFR(t) and e0(t). **b** Using the bayesPop and completely random combinations of projected TFR(t) and e0(t)



Year	2030	2040	2050	2060	2070	2080	2090	2100
Reductions of 80% PI width	-7.50%	-20.90%	-28.56%	-33.63%	-35.86%	-34.35%	-31.05%	-28.75%
Reductions of 95% PI width	-5.79%	-19.07%	-25.85%	-28.58%	-27.44%	-27.71%	-25.18%	-21.20%

Fig. 4 Total population size observed in 1950–2020 and probabilistically projected in 2021–2100, the Unite States. Summary: The reductions of 80% PI width and 95% PI width of projected total population size in 2030–2100 in the United States, using optimized random-combinations of TFR(t) and e0(t), as compared with that using the completely random-combinations. **a** Using the bayesPop and optimized random- combinations of projected TFR(t) and e0(t). **b** Using the bayesPop and completely random combinations of projected TFR(t) and e0(t)

random-combinations of TFR(t) and female e(t) in the PPPs following the approach of five groups of percentiles for the 11 developing countries of Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand and Viet Nam. As shown in the first section of Table 7, the optimized random-combinations for developing countries outlined in (O5-1) ~ (O5-5) indeed substantially reduce the percentages of the improbable combinations of TFR(t) and female e0(t) as compared with the complete random-combinations. For example, the percentages of the improbable combinations of TFR(t) and female e0(t) in 2060 using the approach of five groups of percentiles of optimized random-combinations are 1.2%-3.3% for the 11 developing countries (see the first section of Table 5), which is much smaller than that of using complete random-combinations for these 11 developing countries (22.6%-24.8%).

Our illustrative applications have shown that the five groups of percentiles of optimized random-combinations of the probabilistically projected TFR(t) and female e0(t) substantially narrow the 80% prediction interval width and the 95% prediction interval width of the projections of total population sizes, as compared with the complete random-combinations of the probabilistically projected TFR(t) and female e0(t) in 2021–2100 for the 11 developing countries (ref. Figure 3 and SM-Figs. 16–25 for visual comparisons and numerical summaries). For example, the optimized random-combinations of TFR(t) and female e0(t) in China reduced the 80% prediction interval width and 95% prediction interval width by 43.76% and 38.85% in 2060 and by 31.37% and 21.51% in 2100, as compared with that using the complete random-combinations (see Table 7 and the summary underneath Fig. 3).

We applied the extended R package BayesPop and used the method for optimized random-combinations of TFR(t) and female e(t) for developed countries as outlined in (O3-1) ~ (O3-3) in the PPPs for the 6 developed countries of Canada, France Italy, Japan, the United Kingdom and the United States. As shown in the second section of Table 7, the optimized random-combinations of TFR(t) and female e(t) substantially reduce the percentages of the improbable combinations of TFR(t) and female e0(t) compared to the complete random-combinations. For example, the percentages of improbable combinations of TFR(t) and female e0(t) for the 6 developed countries using the method for optimized random-combinations are 2.6%-4.5% in 2060 (see the second section of Table 5), which are much smaller than that using complete random-combinations for these developed countries (19.6%-23.8%) presented in the second section of Table 5.

Our illustrative applications have shown that the method for optimized random-combinations of the probabilistically projected TFR(t) and female e0(t) substantially narrowed the 80% PI width and the 95% PI width of the projections of total population sizes, as compared with the complete random-combinations of the probabilistically projected TFR(t) and female e0(t) in 2021–2100 for the 6 developed countries (see the second section of Table 7 and ref. Figure 4 and SM-Figs. 35–40 for visual comparisons and numerical summaries). For example, the optimized random-combinations of TFR(t) and female e0(t) in the United States reduced the 80% PI width and 95% PI width by 33.63% and 28.58% in 2060 and by 28.75% and 21.20% in 2100, as compared with that using the complete random-combinations (see the second section of Table 7 and the summary underneath Fig. 4).

Discussion and Conclusion

Our literature review and empirical data analyses provide strong evidence that TFR(t) is negatively correlated with e0(t) in various countries worldwide. Our analyses indicated that complete random-combinations of TFR(t) trajectories and female e0(t) trajectories resulted in about 22.6%-24.8% and 19.6%-23.8% of the combinations that are improbable, namely, high (or low) TFR(t) combined with high (or low) female e0(t) in the probabilistic population projections for developing countries and developed countries, respectively. Because the correlations of the TFR(t) and e0(t) in 1950–2020 in the developed countries were less strong as compared with that in the developing countries, we proposed different approaches of the optimized random-combinations of probabilistically projected TFR(t) and female e0(t) for PPPs for developing countries and developed countries.

As illustrative applications, we conducted PPPs for the 11 developing countries of Brazil, China, Indonesia, Madagascar, Pakistan, Philippines, Saudi Arabia, Singapore, Sri Lanka, Thailand and Viet Nam and 6 developed countries of Canada, France Italy, Japan, the United Kingdom and the United States, using our proposed approaches of the optimized random-combinations of trajectories of TFR(t) and female e0(t) probabilistically projected by UNPD. Our illustrative applications for the 17 countries demonstrated that the optimized random-combinations largely reduce the percentages of improbable combinations of TFR(t) and female e0(t) and substantially narrow the 80% prediction interval width and 95% prediction interval width, as compared with the complete random-combinations. For example, in the 17 countries, the percentages of the improbable combinations of TFR(t) and female e0(t) in 2060 are 1.2%-4.5% using the optimized random-combinations, in contrast to 19.6%-24.8% using the complete random-combinations; and 80% predictive interval widths of the probabilistically projected total population size in 2060 using the optimized random-combinations of TFR(t) and e0(t) are substantially reduced by 14.4%-43.8% as compared with using the complete random-combinations of TFR(t) and e0(t). Clearly, the optimized random-combinations largely reduce percentages of improbable combinations of TFR(t) and e0(t) and substantially narrow the prediction intervals width compared to complete random-combinations in both developing countries and developed countries. This is important in a real-world practical sense since it would substantially improve the accuracy of PPPs, which are useful for socioeconomic planning.

We also tried the approach of ten groups of percentiles for optimized random-combinations of the projected trajectories of TFR(t) and female e0(t) (ref. Section SM1 of Supplementary Materials for details). We found that the use of ten groups of percentiles resulted in substantially higher percentages of improbable random-combinations of TFR(t) and e0(t) as compared with the approaches of the five groups of percentiles for developing countries and three groups of percentiles for developed countries. Furthermore, the approaches of five groups of percentiles and three groups of percentiles have middle groups, whereas the approaches of ten groups of percentiles, eight groups of percentiles, six groups of percentiles, and four groups of percentiles do not have middle groups, and they are not appropriate for the optimized random-combinations of the projected trajectories of TFR(t) and e0(t).

The present study is part of our ongoing research program on extending ProFamy cohort-component methods and software to probabilistic household and living arrangement projections (PHPs) that build upon and are consistent with the UNPD PPPs. The PHPs are useful extensions of PPPs (which do not include household information) since the outcomes of PHPs can be readily used in various studies on healthy aging and sustainable development. To prepare the inputs for PHPs, we have developed new methods for probabilistically projecting the total marriage rate (TMR(t)) and the total divorce rate (TDR(t)) based on the time series of TMR(t) and TDR(t) estimated using the BayesRates R package and commonly available demographic datasets (Zhang et al., 2024). Our illustrative applications of estimating the time series of TMR(t) and TDR(t) and probabilistically projecting TMR(t) and TDR(t) in future years for China and United States are successful (Zeng et al., 2024). We will investigate the optimized random-combinations of probabilistically projected TMR(t) and TFR(t) as well as TDR(t) and TFR(t) for developing countries and developed countries, respectively. We will develop the new methods, user-friendly software and databases for PHPs by substantially extending the ProFamy cohort-component methods/software, which use commonly available demographic rates as input and project detailed household types, sizes, and living arrangements of all individuals of various ages, as well as detailed population size, age and sex distributions (Zeng, Land, Gu and Wang et al., 2014). As discussed in Zeng et al. (2024), the ProFamy cohort-component methods/software provide an ideal and highly feasible modelling framework for research on PHPs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11113-024-09926-y.

Acknowledgements The research reported in this paper were jointly supported by the grant awarded by the Ministry of Finance and Ministry of Foreign Affairs of China under grant section 2 of the Collaborative Program between China and Association of Southeast Asian Nations (ASEAN), and by National Natural Science Foundation of China (72104004,72204008).

Author Contributions YZ designed the study and drafted the paper. ML and SWZ performed the demographic and statistical data analysis and worked with YZ and KL to draft and revise the paper. YZ, KL, ML, SWZ and all other co-authors discussed and contributed to the theoretical framework, interpretation of the results, and revised and gave final approval of this manuscript.

Funding The study funders provided financial supports for data analyses but had no role in the writing of the paper, interpretation of the results, and submission for consideration of publication. The co-corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. Ministry of Finance and Ministry of Foreign Affairs of China under grant Sect. 2 of the Collaborative Program between China and Association of Southeast Asian Nations (ASEAN), National Natural Science Foundation of China,72104004,Man Li,72204008

Data availability The data is publicly on this website: https://bayespop.csss.washington.edu/download/index.html.

Declarations

Conflict of interests We declare no any potential conflicts of financial or non-financial interests.

References

Alesina, A. F., and Giuliano, P. (2007). Divorce, fertility and the value of marriage. Harvard Institute of Economic Research Discussion Paper, (2136).

Alkema, L., Gerland, P., Raftery, A., & Wilmoth, J. (2015). The United Nations probabilistic population projections: an introduction to demographic forecasting with uncertainty. *Foresight (Colchester, Vt.), 2015*, 19.

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., & Heilig, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography, 48*(3), 815–839.

Aughinbaugh, A., Robles, O., & Sun, H. (2013). Marriage and divorce: Patterns by gender, race, and educational attainment. *Monthly Lab. Rev., 136*, 1.

Awan, A. G., & Abbas, N. (2015). Impact of demographic factors on impulse buying behavior of consumers in Multan-Pakistan. *European Journal of Business and Management, 7*, 96–105.

Bai, R., Liu, Y., Zhang, L., Dong, W., Bai, Z., & Zho, M. (2023). Projections of future life expectancy in China up to 2035: a modelling study. *The Lancet Public Health., 8*, e915.

Bao, L., Raftery, A. E., & Reddy, A. (2015). Estimating the sizes of populations at risk of hiv infection from multiple data sources using a Bayesian hierarchical model. *Statistics and Its Interface, 8*, 125–136.

Barlow, R., & Vissandjée, B. (2011). Determinants of national life expectancy. *Canadian Journal of Development Studies / Revue Canadienne D études Du Développement, 20*(1), 9–29.

Camarota, S. A. (2011). *A record-setting decade of immigration, 2000 to 2010*. Center for Immigration Studies.

De Tray, D. N. (1977). Age of marriage and fertility: A policy review. *The Pakistan Development Review, 16*(1), 89–100.

Delavari, S., Zandian, H., Rezaei, S., Moradinazar, M., Delavari, S., Saber, A., & Fallah, R. (2016). Life expectancy and its socioeconomic determinants in Iran. *Electronic Physician, 8*(10), 3062–3068.

Esteban Ortiz-Ospina and Max Roser (2020) - "Marriages and Divorces". Published online at OurWorldInData.org. Retrieved from: ‘https://ourworldindata.org/marriages-and-divorces [Online Resource]

Gerland, P., Raftery, A. E., Sevcikova, H., Li, N., Gu, D. A., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., & Wilmoth, J. (2014). World population stabilization unlikely this century. *Science, 346*(6206), 234–237.

Girum, T., Muktar, E., & Shegaze, M. (2018). Determinants of life expectancy in low and medium human development index countries. *Medical Studies, 34*(3), 218–225. https://doi.org/10.5114/ms.2018.78685

Husain, A. R. (2002). Life expectancy in developing countries: A cross-section analysis. *The Bangladesh Development Studies, 28*(1/2), 161–178.

Jafari, H., Jaafaripooyan, E., Vedadhir, A. A., Foroushani, A. R., Ahadinejad, B., & Pourreza, A. (2016). Socio-economic factors influencing on total fertility rate in Iran: A panel data analysis for the period of 2002–2012. *Electronic Physician, 8*(6), 2551.

Jafrin, N., Masud, M. M., Saif, A. N. M., Mahi, M., & Khanam, M. (2021). A panel data estimation of the determinants of life expectancy in selected SAARC countries. *Operations Research and Decisions*. https://doi.org/10.37190/ord210404

Jalovaara, M. (2002). Socioeconomic differentials in divorce risk by duration of marriage. *Demographic Research, 7*, 537–564.

Jannat, Z. (2022). The dynamics of female life expectancy and total fertility rate in Bangladesh: A time-series analysis. *International Journal of New Economics and Social Sciences, 15*(1), 9–20.

Kabir, M. (2008). Determinants of life expectancy in developing countries. *The Journal of Developing Areas, 41*(2), 185–204.

Kalita, U. (2018). An econometric investigation into variations in life expectancy at birth across countries. *International Journal of Science and Research, 8*(8), 1832–1837.

Kontis, V., Bennett, J. E., Mathers, C. D., Li, G., Foreman, K., & Ezzati, M. (2017). Future life expectancy in 35 industrialised countries: Projections with a Bayesian model ensemble. *Lancet, 389*(10076), 1323–1335.

Krupp, D. B. (2012). Marital, reproductive, and educational behaviors covary with life expectancy. *Archives of Sexual Behavior, 41*(6), 1409–1414.

Lee, R. (2003). The demographic transition: Three centuries of fundamental change. *Journal of Economic Perspectives, 17*(4), 167–190.

Low, B. S., Parker, N., Hazel, A., & Welch, K. B. (2013). Life expectancy, fertility, and women s lives. *Cross-Cultural Research, 47*(2), 198–225.

Miladinov, G. (2021). Impact of unemployment by sex and marriage rate on fertility decline: Estimates for Turkey and Greece using CCR model. *Population and Economics, 5*(3), 76–89.

Mondal, M. N. I., & Shitan, M. (2013). Impact of socio-health factors on life expectancy in the low and lower middle income countries. *Iranian Journal of Public Health, 42*(12), 1354.

Mondal, M. N., & Shitan, M. (2014). Relative importance of demographic, socioeconomic and health factors on life expectancy in low- and lower-middle-income countries. *Journal of Epidemiology, 24*(2), 117–124.

Mondal, M. N. I., Ullah, M. M. M. N., Islam, M. R., Rahman, M. S., Khan, M. N., & Ahmed, K. M. (2015). Sociodemographic and health determinants of inequalities in life expectancy in least developed countries. *International Journal of Maternal and Child Health and AIDS (IJMA), 4*(1), 57–66.

Nettle, D. (2011). Flexibility in reproductive timing in human females: Integrating ultimate and proximate explanations. *Philosophical Transactions of the Royal Society of London. Series b, Biological Sciences, 366*(1563), 357–365.

OECD, Department of Economic and Social Affairs (2018), "Marital and divorce rates, births outside of marriage, and cohabitation status", OECD Family Database, https://www.oecd.org/els/family/database.htm.

Onwube, O., Chukwu, A. B., Ahamba, K., Emenekwe, C., & Enyoghasim Oguwuike, M. (2021). Determinants of life expectancy in Nigeria: An autoregressive distributed lag approach. *Journal of Sustainability Science and Management, 16*(8), 177–192.

Raftery, A. E., Alkema, L., & Gerland, P. (2014). Bayesian population projections for the United Nations. *Statistical Science, 29*(1), 58–68.

Raftery, A. E., Chunn, J. L., Gerland, P., & Ševčíková, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography, 50*, 777–801.

Requena, M., & Salazar, L. (2014). Education, marriage, and fertility: The Spanish case. *Journal of Family History, 39*(3), 283–302.

Roser, M., Ritchie, H., Ortiz-Ospina, E. and Hasell, J. (2020) Coronavirus Pandemic (COVID-19) Published Online at OurWorldInData.org.

Sarlak, A., & Savari, A. (2016). The impact of socio-economic factors on life expectancy in Iran. *Journal of Healthcare Management, 7*, 7–15.

Ševčíková, H., Raftery, A. E., and Gerland, P. (2013). Bayesian probabilistic population projections using r. measuring uncertainty in population forecasts: A new approach., 347.

Ševčíková, H., & Raftery, A. E. (2016). bayesPop: probabilistic population projections. *Journal of Statistical Software, 75*(5), 1–29.

Singariya, M. (2013). Principal component analysis of socioeconomic factors and their association with life expectancy in India. *Journal of Economics, 113*, 163–172.

Stevenson, B., & Wolfers, J. (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives, 21*(2), 27–52.

Summoogum, J. P., & Fah, B. C. Y. (2016). A comparative study analysing the demographic and economic factors affecting life expectancy among developed and developing countries in Asia. *Asian Development Policy Review, 4*(4), 100–110.

United Nations Population Division (UNPD), Department of Economic and Social Affairs (2011). World Population Prospects 2011, Online Edition.

UNPD, Department of Economic and Social Affairs (2011). Population Facts.

UNPD, Department of Economic and Social Affairs (2013). World Population Prospects 2013, Online Edition.

UNPD, Department of Economic and Social Affairs (2019). World Population Prospects 2019, Online Edition.

UNPD, Department of Economic and Social Affairs (2022). World Population Prospects 2022: Methodology Of the United Nations population estimates and projections (UN DESA/POP/2022/TR/NO. 4).

UNPD, Department of Economic and Social Affairs (2022). World Population Prospects 2022, Online Edition.

Vaupel, J. (2019). Forecasting life expectancy: The scope approach, chapter 6. In T. Bengtsson & N. Keilman (Eds.), *Old and New Perspectives on Mortality Forecasting* (p. 2019). Springer.

Venieris, Y. P., & Stewart, D. B. (1988). Long-run social and economic responses of fertility in the United States. *The Pakistan Development Review, 27*, 137–157.

Yang, G., & Zhang, L. (2023). Relationship between aging population, birth rate and disposable income per capita in the context of COVID-19. *PLoS ONE, 18*(8), e0289781.

Yi Zeng, Junni Zhang, Wei Tang, Yan Mo Minhui Liu, Yi Sun, Zuqi Cai, Man Li, Shanwen Zhu, Emily Grundy, Michael Murphy, Jason Hilton, John Bryant, Fengqing Chao, Lijing Yan, Can Jia, Liquan

Zhou, Michael Lutz, Scott M. Lynch, David Banks1 Kenneth Land, Qiushi Feng, and Zhenglian Wang (2024). "Illustrative Applications of Bayesian Model to Estimate and Probabilistically Project TMR(t) and TDR(t) for Probabilistic Household and Living Arrangement Projections"; manuscript submitted to the peer-review journal for considerations of publication and is now under peer-reviews (available upon request).

Zeng, Y., K.C. Land, D. Gu, and Z. Wang. 2014. Household and Living Arrangement Projections: The Extended Cohort-Component Method and Applications to the U.S. and China. New York: Springer Publisher.

Zeng, Y., Chen, H., Wang, Z., & Land, K. C. (2013b). Implications of changes in households and living arrangements on future home-based care costs for disabled elders in China. *Journal of Aging and Health, 27*(3), 519–550.

Zeng, Y., Land, K. C., Wang, Z., & Gu, D. (2006). U.S. family household momentum and dynamics: An extension and application of the ProFamy method. *Population Research and Policy Review, 25*(1), 1–41.

Zeng, Y., Land, K. C., Wang, Z., & Gu, D. (2013a). Household and living arrangements projections at the sub-national level: An extended cohort-component approach". *Demography, 50*, 827–852.

Zeng, Y., Li, L., Wang, Z., Huang, H., & Norris, J. (2013c). Effects of changes in household structure on future housing demand in Hebei province. *China. GENUS, 69*(2), 85–111.

Zeng, Y., Yang, H., Wang, Z., & Li, L. (2021). Impacts of family household dynamics on residential energy demands in Hebei province of China. *Genus, 77*, 35.

Zhang, Junni, John Bryant, Fengqing Chao, David Banks, and Yi Zeng (2024). Bayesian Models for Estimating and Probabilistically Projecting Total Marriage Rates (TMR(t)) and Total Divorce Rates (TDR(t)); manuscript under preparations.

Authors and Affiliations

Man Li[1] · Shanwen Zhu[2] · Zhenglian Wang[3] · Qiushi Feng[4] · Junni Zhang[5] · Fengqing Chao[6] · Wei Tang[5] · Linda George[7] · Emily Grundy[8] · Michael Murphy[9] · Michael Lutz[10] · Adrian Dobra[11] · Kenneth Land[12] · Yi Zeng[5,13,14]

✉ Kenneth Land
kland@duke.edu

✉ Yi Zeng
zengyi@duke.edu

1 School of Social Science and Soochow University Think Tank, Soochow University, Suzhou, China

2 School of Humanities, Shanghai Institute of Technology, Shanghai, China

3 China Population and Development Research Center, Beijing, China

4 Department of Sociology and Centre for Family and Population Research, National University of Singapore, Singapore, Singapore

5 National School of Development, Peking University, Beijing, China

6 Chinese University of Hongkong, Shenzhen, China

[7] Department of Sociology, Duke University, Durham, USA

[8] Institute for Economic and Social Research, University of Essex, Colchester, UK

[9] London School of Economics and Political Science, London, UK

[10] Department of Neurology, Duke University School of Medicine, Durham, USA

[11] Department of Statistics, University of Washington, Seattle, USA

[12] Department of Sociology and Social Science Research Institute, Duke University, Durham, USA

[13] Center for Study of Aging and Human Development and Geriatrics Division, School of Medicine, Duke University, Durham, USA

[14] National School of Development, Peking University, Beijing, China