Supporting Information

Raftery et al. 10.1073/pnas.1211452109

SI Text

Probabilistic Projection of Female and Male Life Expectancy. We generated probabilistic projections of the period female life expectancy at birth using the Bayesian hierarchical model of Raftery et al. (1). Female life expectancy at birth for country c in period t is assumed to follow a random walk with drift, given by

$$\ell_{c,t+1} = \ell_{c,t} + g(l_{c,t}|\theta^{(c)}) + \varepsilon_{c,t+1}, \quad \textbf{[S1]}$$

where $\varepsilon_{c,t} \overset{\text{ind}}{\sim} N(0, \omega f(\ell_{c,t}))$, with $f(\ell_{c,t})$ a smooth declining function of $\ell_{c,t}$. In Eq. **S1**, the expected 5-y gain in life expectancy is a double-logistic function of the current level of life expectancy, namely

$$g(\ell_{c,t}|\theta^c) = \frac{k^c}{1+\exp(-\frac{A_1}{\Delta_2^c}(\ell_{ct} - \Delta_1^c - A_2\Delta_2^c))} + \frac{z^c - k^c}{1+\exp(-\frac{A_1}{\Delta_4^c}(\ell_{ct} - \sum_{i=1}^3 \Delta_i^c - A_2\Delta_4^c))},$$

where $\theta^c = (\Delta_1^c, \Delta_2^c, \Delta_3^c, \Delta_4^c, k^c, z^c)$ are the six parameters of the double-logistic function for country c, and A_1 and A_2 are constants.

Each of the parameters of the double-logistic function for country c is in turn assumed to be drawn from a world distribution of the parameter

$$\Delta_i^c|\sigma_{\Delta_i} \sim^{\text{iid}} \text{Normal}_{[a_i,100]}(\Delta_i, \sigma_{\Delta_i}^2), \qquad i = 1, \ldots, 4, \quad \textbf{[S2]}$$

$$k^c|\sigma_k \sim^{\text{iid}} \text{Normal}_{[0,10]}(k, \sigma_k^2), \quad \textbf{[S3]}$$

$$z^c|\sigma_z \sim^{\text{iid}} \text{Normal}_{[0,1.15]}(z, \sigma_z^2), \quad \textbf{[S4]}$$

where $\text{Normal}_{[a,b]}(\mu, \sigma^2)$ denotes a normal distribution with mean μ and standard deviation σ, truncated to lie between a and b, $a_1 = a_2 = a_4 = 0$ and $a_3 = -20$.

Prior Distributions for the Model of Female Life Expectancy at Birth. Estimation of the Bayesian hierarchical model for female life expectancy at birth given by Eqs. **S1–S4** requires the specification of prior distributions for 13 world parameters. The parameters $\Delta_i (i = 1, \ldots, 4)$, k and z were given normal prior distributions with prior means equal to the values specifying the medium-pace double-logistic function used by the United Nations in the 2008 *World Population Prospects* (WPP 2008) (2), and prior variances equal to the variances of the parameters among the five double-logistic functions used by the UN in WPP 2008.

For the world variance parameters, $\sigma_{\Delta_i}^2 (i = 1, \ldots, 4)$, σ_k^2 and σ_z^2, we used inverse gamma priors. The inverse gamma distribution with shape parameter α and rate parameter β, denoted by $IG(\alpha, \beta)$, has probability density function at x proportional to $x^{-(\alpha+1)}e^{-\beta/x}$ for $x > 0$. We set the shape parameter equal to 2 for all inverse gamma priors. To set the rate parameters, we first fit the double-logistic model by least squares to the data from each country individually, and then for each parameter we computed the empirical average squared deviations from the values for the UN WPP 2008 medium-pace double-logistic function. We then set the prior means of the reciprocals of the world variance parameters equal to the reciprocals of these values. The resulting prior distributions are much more spread out than the posterior distribution, and the posterior distribution is relatively insensitive to reasonable changes in these priors.

The quantities $\sum_{i=1}^4 \Delta_i$ and $\sum_{i=1}^4 \Delta_i^c$ correspond roughly to the average and country-specific ages at which the expected increase in life expectancy in the next 5-y period becomes approximately constant at its asymptote, z^c. We therefore restricted them to lie within the range of life expectancies that have either been historically observed or are plausible before 2100, taken to be [30,110]. This eliminated a small number of implausible trajectories but did not substantially change the results.

Finally, a diffuse Uniform[0,10] prior was used for ω. The priors are summarized in Table S1.

Estimation of Model for Life Expectancy Gap. The model in Eq. **1** of the main article for the gap in life expectancy between females and males was estimated by maximum likelihood (3, 4), using the `hett` R package. The top equation in Eq. **1** in the main article was estimated using the 12 time periods during 1950–2010 for the 159 countries in our analysis. The bottom equation in Eq. **1** in the main article was estimated using only period/country combinations for which female life expectancy at birth was greater than 80 years, which resulted in 83 data points. The parameter estimates are shown in Table S2.

Probabilistic Population Projections for Brazil, The Netherlands, Madagascar, China, and India. Probabilistic projections of the major population indicators considered for the Netherlands, Madagascar, China, and India are shown in Figs. S1–S4. Probabilistic population projections for all age groups in Brazil are shown in Fig. S5.

1. Raftery AE, Chunn JL, Gerland P, Ševčíková H (2012) Bayesian probabilistic projections of life expectancy for all countries. *Demography* 49:in press.
2. United Nations (2009) *World Population Prospects: The 2008 Revision* (United Nations, New York).
3. Lange KL, Little RJA, Taylor JMG (1989) Robust statistical modeling using the *t* distribution. *J Am Stat Assoc* 84:881–896.
4. Taylor J, Verbyla A (2004) Joint modelling of location and scale parameters of the *t* distribution. *Stat Model* 4:91–112.

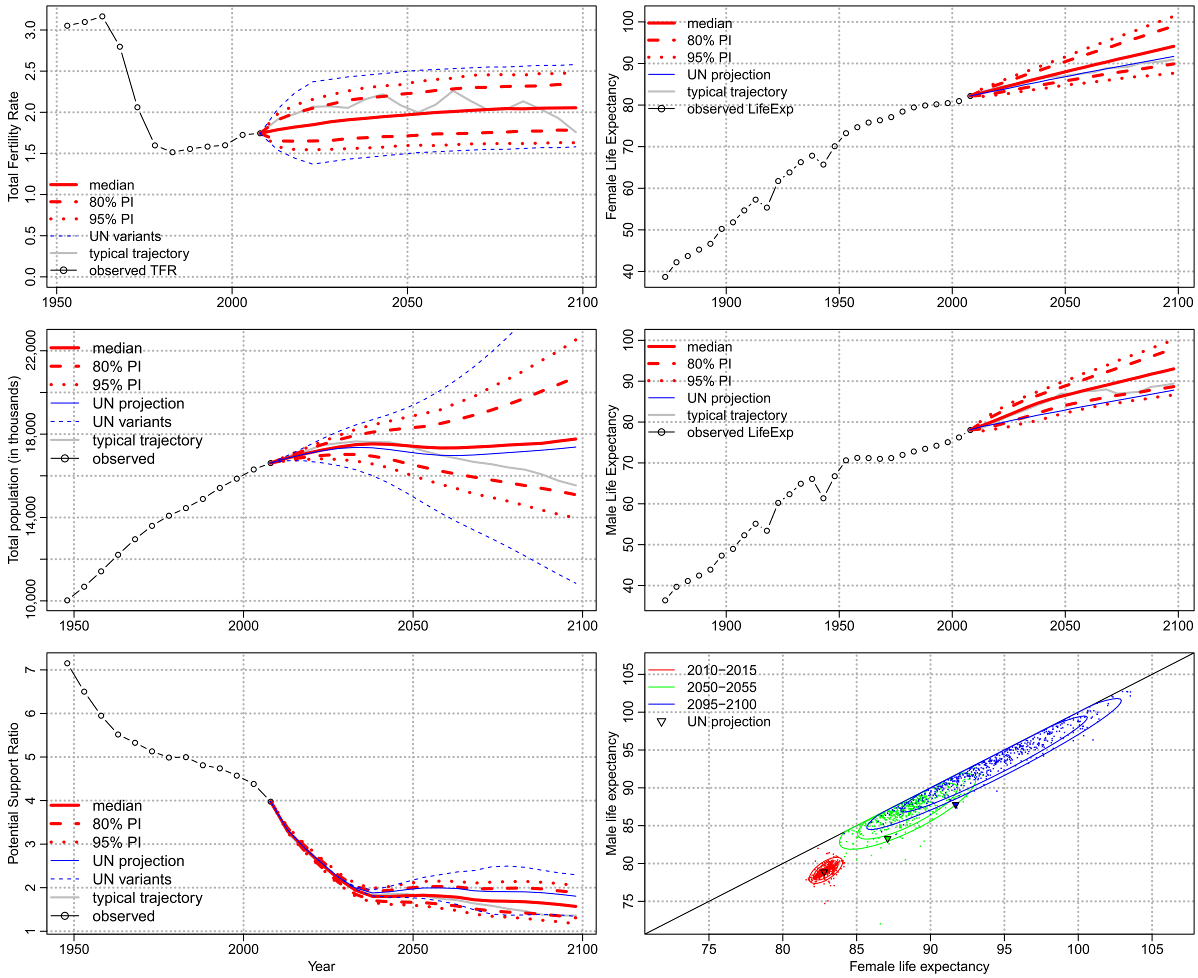

Fig. S1. Bayesian probabilistic population projections for the Netherlands, 2010–2100: major population indicators. *Left*, *Top* to *Bottom*: total fertility rate; total population; potential support ratio (20–64 population/65+ population). *Right*, *Top* to *Bottom*: female life expectancy; male life expectancy; joint predictive distribution of female and male life expectancy for 2010–2015, 2050–2055 and 2095–2100. The Bayesian predictive distributions are shown in red: median—solid; 80% prediction interval—dashed; 95% prediction interval—dotted. The UN WPP 2010 projection is shown as a solid blue line. The typical trajectory is shown as a solid gray line.

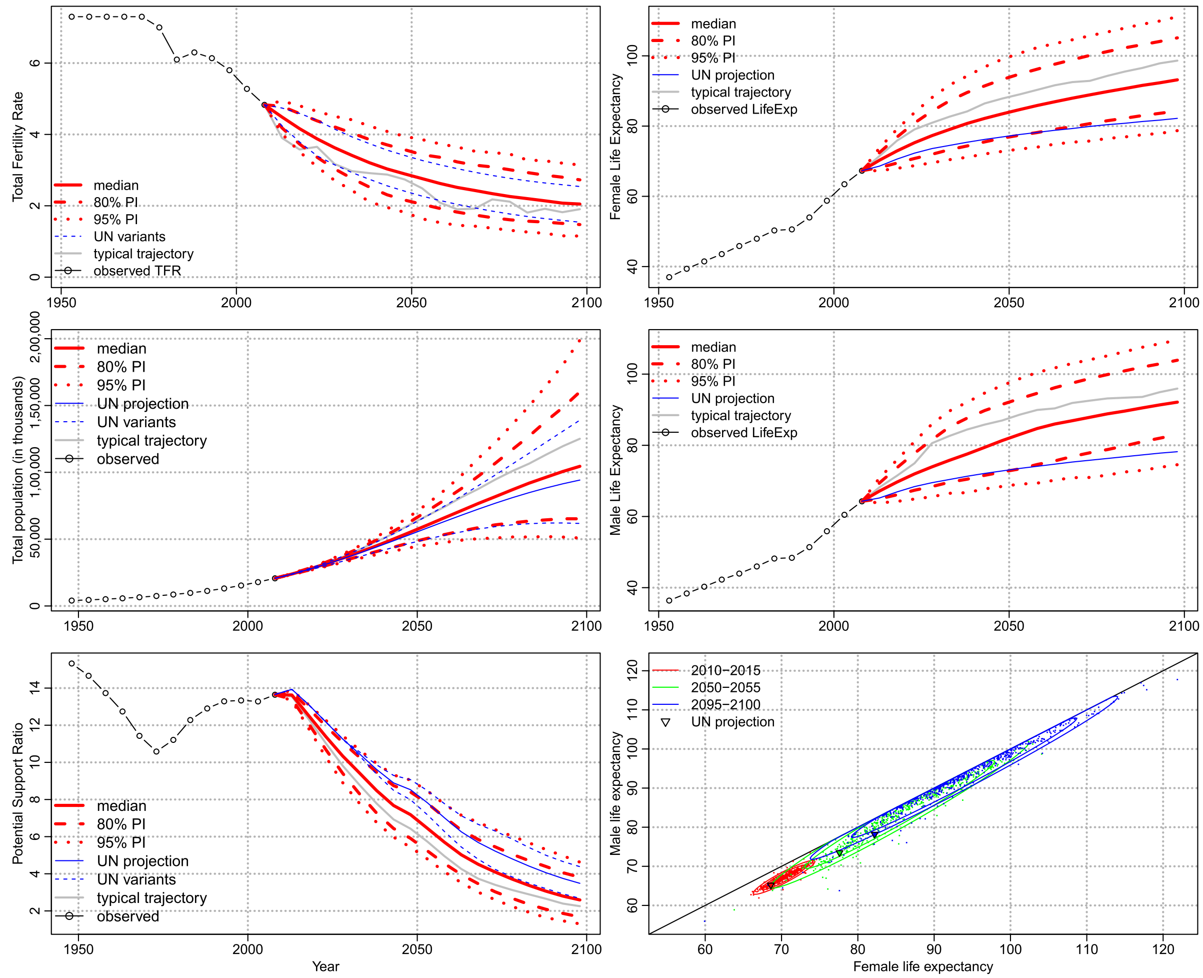

Fig. S2. Bayesian probabilistic population projections for Madagascar, 2010–2100: major population indicators. *Left*, *Top* to *Bottom*: total fertility rate; total population; potential support ratio (20–64 population/65+ population). *Right*, *Top* to *Bottom*: female life expectancy; male life expectancy; joint predictive distribution of female and male life expectancy for 2010–2015, 2050–2055 and 2095–2100. The Bayesian predictive distributions are shown in red: median—solid; 80% prediction interval—dashed; 95% prediction interval—dotted. The UN WPP 2010 projection is shown as a solid blue line. The typical trajectory is shown as a solid gray line.

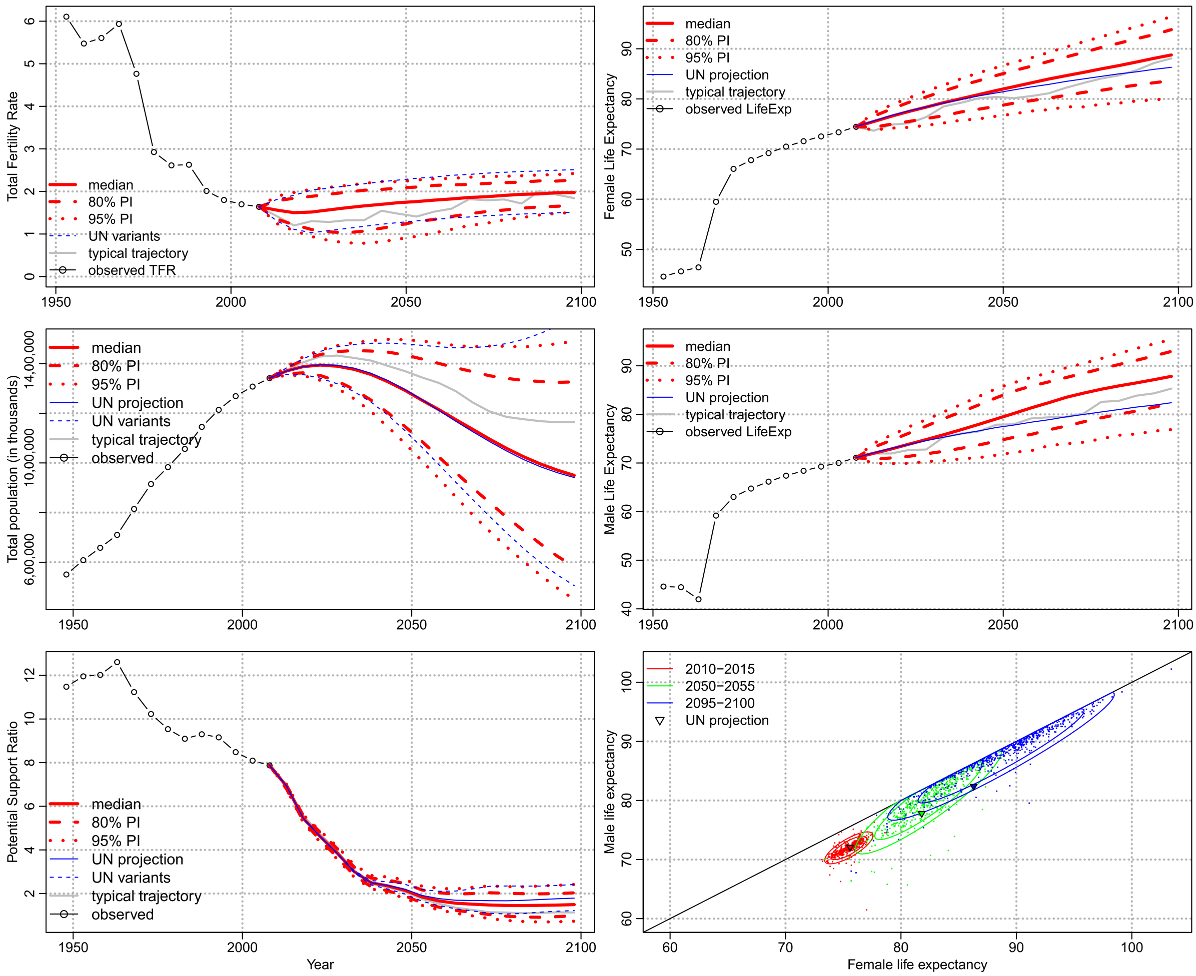

Fig. S3. Bayesian probabilistic population projections for China, 2010–2100: major population indicators. *Left*, *Top* to *Bottom*: total fertility rate; total population; potential support ratio (20–64 population/65+ population). *Right*, *Top* to *Bottom*: female life expectancy; male life expectancy; joint predictive distribution of female and male life expectancy for 2010–2015, 2050–2055 and 2095–2100. The Bayesian predictive distributions are shown in red: median—solid; 80% prediction interval—dashed; 95% prediction interval—dotted. The UN WPP 2010 projection is shown as a solid blue line. The typical trajectory is shown as a solid gray line.

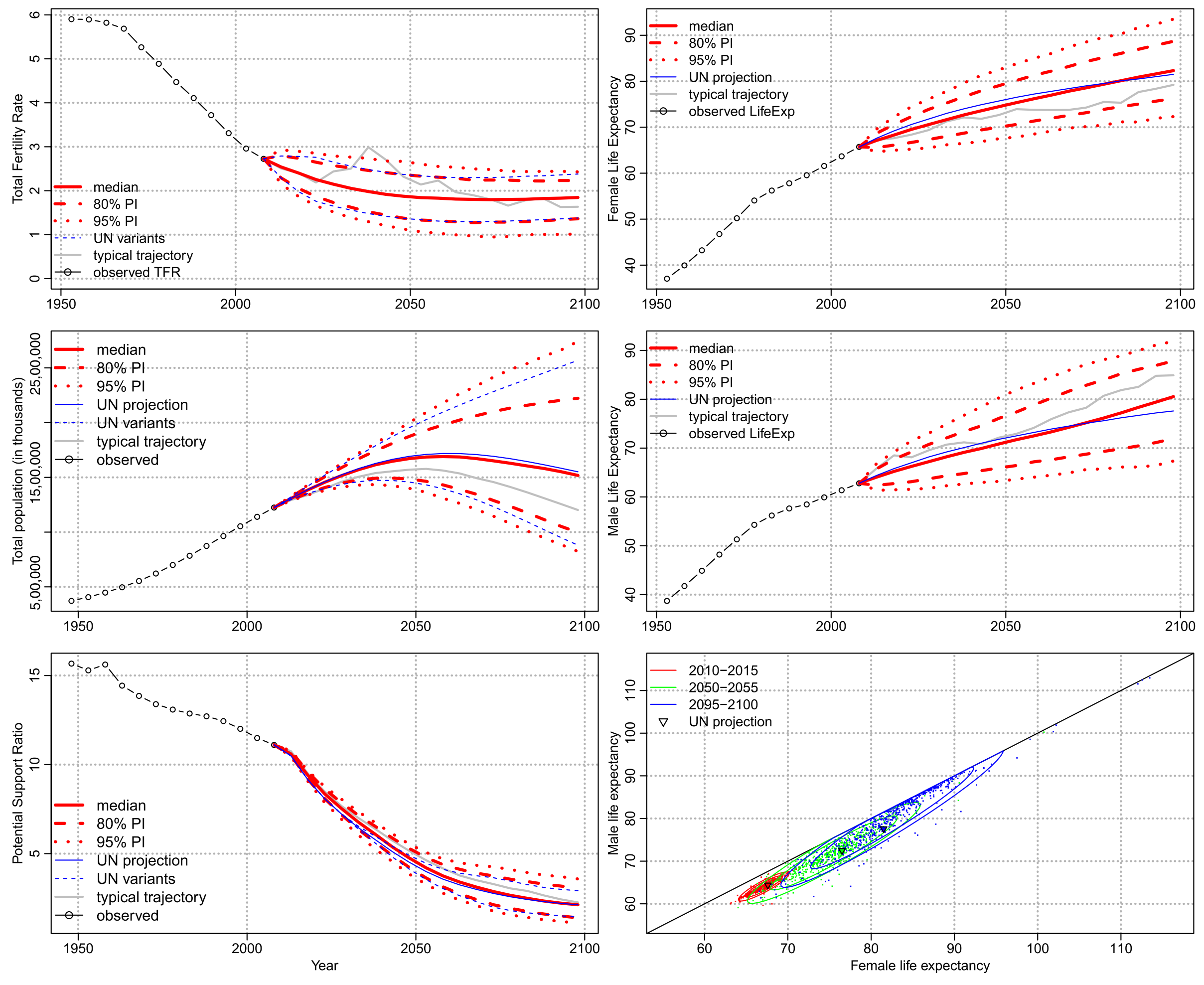

Fig. S4. Bayesian probabilistic population projections for India, 2010–2100: major population indicators. *Left*, *Top* to *Bottom*: total fertility rate; total population; potential support ratio (20–64 population/65+ population). *Right*, *Top* to *Bottom*: female life expectancy; male life expectancy; joint predictive distribution of female and male life expectancy for 2010–2015, 2050–2055 and 2095–2100. The Bayesian predictive distributions are shown in red: median—solid; 80% prediction interval—dashed; 95% prediction interval—dotted. The UN WPP 2010 projection is shown as a solid blue line. The typical trajectory is shown as a solid gray line.

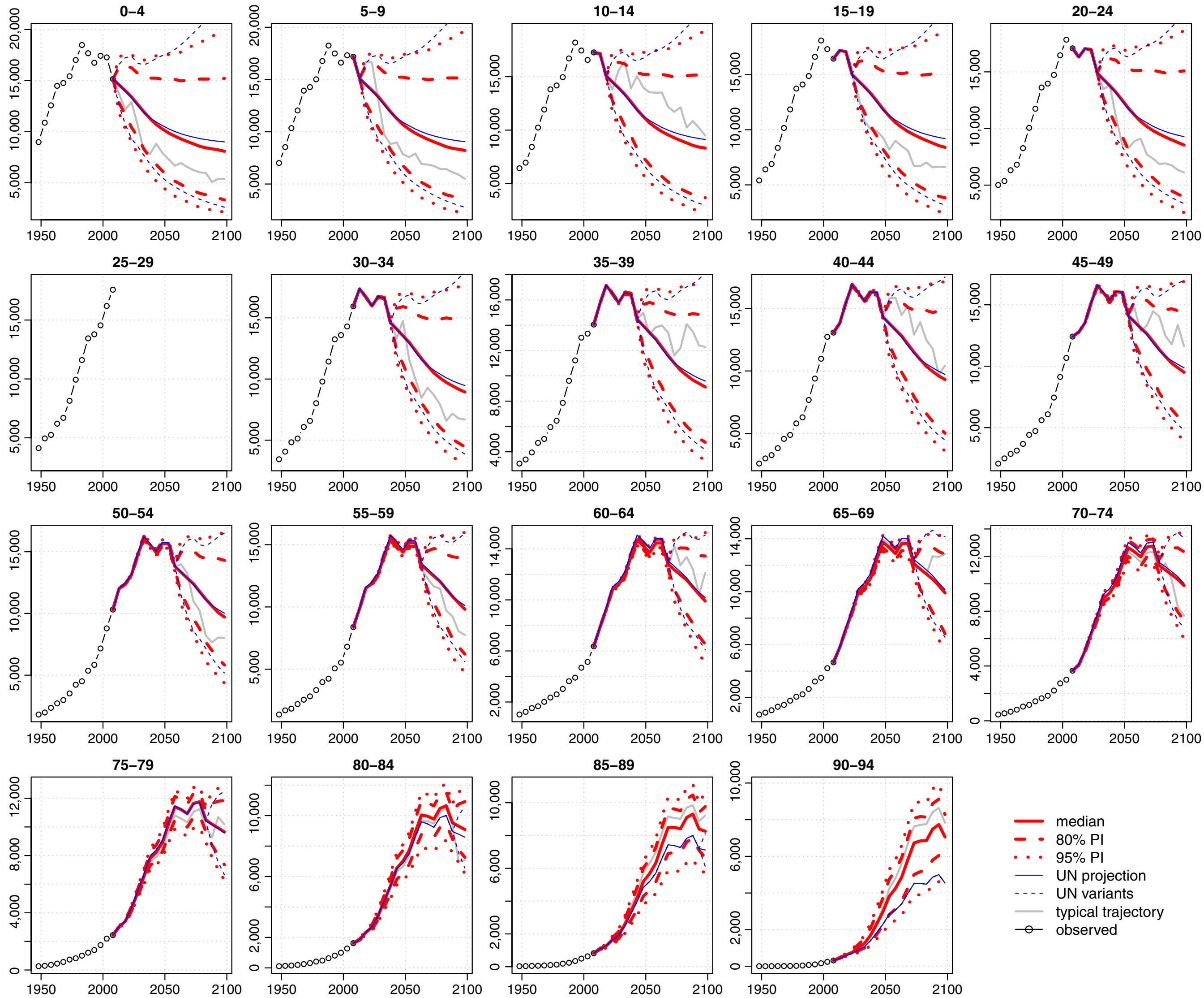

Fig. S5. Bayesian probabilistic population projections for all age groups for Brazil, 2010–2100.

Table S1. Prior distributions for the world parameters of the Bayesian hierarchical model for female life expectancy at birth in Eqs. S1–S4 *

Parameter	Prior distribution
Δ_1	$N_{[0,100]}(13.22, 3.85^2)$
Δ_2	$N_{[0,100]}(41.07, 4.03^2)$
Δ_3	$N_{[-20,100]}(9.24, 11.54^2)$
Δ_4	$N_{[0,100]}(17.60, 5.64^2)$
k	$N_{[0,10]}(2.84, 0.9^2)$
z	$N_{[0,1.15]}(0.38, 0.4^2)$
$\sigma^2_{\Delta_1}$	IG $(2, 15.6^2)$
$\sigma^2_{\Delta_2}$	IG $(2, 23.5^2)$
$\sigma^2_{\Delta_3}$	IG $(2, 14.5^2)$
$\sigma^2_{\Delta_4}$	IG $(2, 14.7^2)$
σ^2_k	IG $(2, 3.5^2)$
σ^2_z	IG $(2, 0.6^2)$
ω	$U[0, 10]$

*$N_{[a,b]}(\mu, \sigma^2)$ denotes a truncated normal distribution with mean μ and standard deviation σ, truncated to lie between a and b. IG(α, β) denotes an inverse gamma distribution with shape parameter α and rate parameter β. $U[a, b]$ denotes a uniform distribution between a and b.

Table S2. Parameter estimates for the gap regression model in Eq. 1 in the main article

Coefficient	Variable	Estimate	Std. Error	*t* value
β_0	Intercept	−0.217	0.058	−3.725
β_1	$\ell_{c,1953}$	0.008	0.001	6.808
β_2	$G_{c,t}$	0.963	0.004	233.374
β_3	$\ell_{c,t}$	0.002	0.001	1.498
β_4	$(\ell_{c,t} - 75)_+$	−0.093	0.006	−15.513
γ_1	$G_{c,t}$	0.950	0.006	167.200
σ_1		0.267		
ν_1	dof *	1.963	0.118	
σ_2		0.299		
ν_2	dof	20.107	38.066	

*ν_1 and ν_2 are the numbers of degrees of freedom of the *t* distributions of the error term in Eq. **1** in the main article.