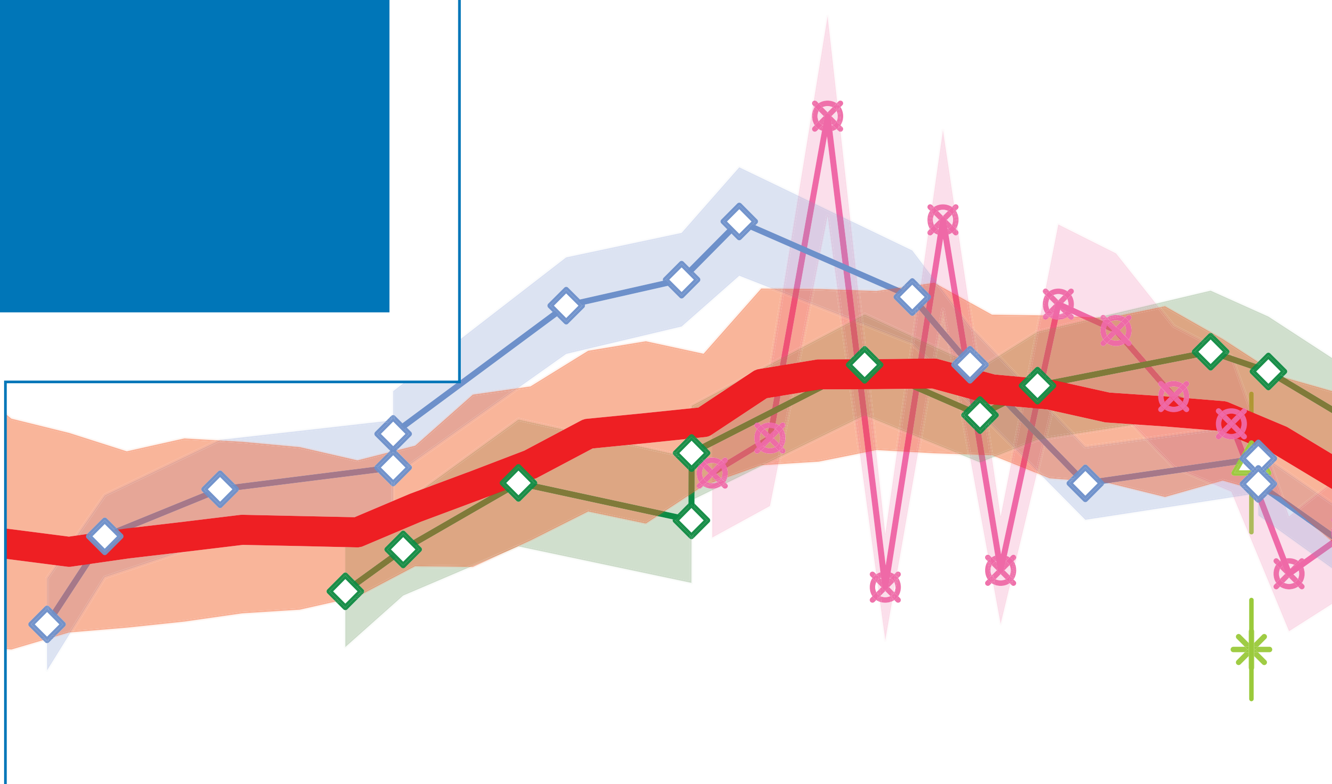
United Nations | Department of Economic and Social Affairs

Population Division
UN DESA/POP/2023/TP/No. 6
October 2023

Estimating age-specific fertility rate in the World Population Prospects: A Bayesian modelling approach

Technical Paper

Fengqing Chao, Vladimíra Kantorová, Giulia Gonnella, Lina Bassarsky, Lubov Zeifman and Patrick Gerland



Population Division
Technical Paper
UN DESA/POP/2023/TP/No.6
October 2023

The Population Division of the Department of Economic and Social Affairs provides the international community with timely and accessible population data and analysis of population trends and development outcomes. The Division undertakes studies of population size and characteristics and of the three components of population change (fertility, mortality and migration).

The purpose of the **Technical Paper series** is to publish substantive and methodological research on population issues carried out by experts both within and outside the United Nations system. The series promotes a scientific understanding of population issues among Governments, national and international organizations, research institutions and individuals engaged in social and economic planning, research and training.

Suggested citation: Chao, Fengqing and others (2023). *Estimating age-specific fertility rate in the World Population Prospects: a Bayesian modelling approach*. United Nations, Department of Economics and Social Affairs, Population Division, Technical Paper No. UN DESA/POP/2023/TP/No.6.

This technical paper is available in electronic format on the Division's website at www.unpopulation.org. For further information, please contact the Population Division, Department of Economic and Social Affairs, Two United Nations Plaza, DC2-1950, New York, 10017, USA; phone: +1 212-963-3209; e-mail: population@un.org.

Estimating age-specific fertility rate in the World Population Prospects: a Bayesian modelling approach[*]

Fengqing Chao,[**] Vladimíra Kantorová,[***] Giulia Gonnella,[***] Lina Bassarsky,[***] Lubov Zeifman[***] and Patrick Gerland[***]

Abstract

As part of its work in revising population estimates and projections for the biennial publication of *the World Population Prospects (WPP)*, the United Nations Population Division produces age-specific fertility estimates for all countries and areas of the world, starting from 1950 up to today. These estimates are based on data from several reference data sources, such as civil registration and vital statistics systems, sample registration systems, surveys, national estimates and population censuses, and calculated using standard demographic techniques and approaches. Available estimates are often affected by biases and inconsistencies that need to be examined and considered while producing the annual series of age-specific fertility estimates.

This technical paper details the Bayesian hierarchical model (BHM) that the Population Division developed to estimate the levels and trends in age-specific fertility rates (ASFR) for all countries and areas since 1950. The model uses an extensive database of fertility data from various data sources maintained by the Population Division. The BHM allows sharing of information across countries and periods to inform annual estimates for the countries and periods with sparse, biased or non-available data.

The information included in *World Population Prospects* is used widely by the United Nations system, academia and civil society, among others, including for monitoring several indicators of the Sustainable Development Goals. The age-specific fertility estimates from the *World Population Prospects* are used to monitor the global and regional trends of the Sustainable Development Goal 3.7.2 Adolescent birth rate (aged 10–14 years; aged 15–19 years) per 1,000 women in that age group.

Keywords: Fertility, estimation, Bayesian hierarchical model

Sustainable Development Goals: 3

[*] The authors gratefully acknowledge Helena Cruz Castanheira from United Nations, Economic Commission for Latin America and the Caribbean, Santiago de Chile, Chile, as well as Danan Gu, Sara Hertog, Thomas Spoorenberg, Ivan Williams and Mark Wheldon from the United Nations Population Division for their insightful comments and useful inputs on the data, method development and results. The authors would also like to thank Bruno Schoumaker and David Sánchez-Páez from Université catholique de Louvain, Belgium for their assistance with data preparation. Additionnally, the authors are thankful to Stephen Kisambira for reviewing this report. This work was supported, in part, by a grant from the Bill and Melinda Gates Foundation, Making Family Planning Count 3.0 (Grant No. INV-033016).

[**] Statistics Program, Computer, Electrical and Mathematical Sciences and Engineering Division, 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia.

[***] United Nations Department of Economic and Social Affairs, Population Division.

Contents

EXPLANATORY NOTES

The following symbols have been used in the tables throughout this report:

A full stop (.) is used to indicate decimals.

References to countries, territories and areas:

The designations employed in this publication and the material presented in it do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The term "country" as used in this publication also refers, as appropriate, to territories or areas.

The following abbreviations have been used:

AR1	first-order autoregression
AR3	third-order autoregression
ASFR	age-specific fertility rate
BHM	Bayesian Hierarchical Model
BRC	birth-reporting completeness
CCM	cohort component method
CI	Credible Interval
CPS	Contraceptive Prevalence Surveys
CRVS	civil registration and vital statistics
CS	Calibrated Spline
DHS	Demographic and Health Surveys
DYB	Demographic Yearbook
EDU	female educational attainment
GBD	Global Burden of Disease
HFD	Human Fertility Database
IHME	Institute for Health Metrics and Evaluation, University of Washington
INLA	Integrated Nested Laplace Approximation
IPUMS	Integrated Public Use Microdata Series
M49	Standard Country or Area Codes for Statistical Use (Series M, No. 49)
MIS	Malaria Indicators Surveys
MICS	Multiple Indicator Cluster Surveys
NA	Not available
OECD	Organization for Economic Co-operation and Development
PAPCHILD	Pan Arab Project for Child Development
PAPFAM	Pan Arab Project for Family Health
PC	Penalized Complex
PIs	Prediction Intervals
PMA	Performance Monitoring and Accountability
RHS	Reproductive Health Surveys
RW1	first-order random walk

RW2	second-order random walk
SDG	Sustainable Development Goals
SQL	Structured Query Language
SRS	Sample Registration System
TFR	Total fertility rate
UNAIDS	Joint United Nations Programme on HIV/AIDS
UN DESA	United Nations Department of Economic and Social Affairs
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
UNSD	United Nations Statistics Division
WFS	World Fertility Surveys
WHO	World Health Organization
WPP	World Population Prospects

I. INTRODUCTION

The Population Division of the United Nations Department of Economic and Social Affairs (UN DESA) releases a set of population estimates and projections every other year, known as the *World Population Prospects* (WPP). It forms a comprehensive set of demographic data to assess population trends at the global, regional, and national levels. The WPP consists of a prospective population reconstruction from 1950 to the present (i.e., population estimates) and various scenarios of future population development (i.e., population projections) (United Nations, 2022a).

In the WPP, the cohort component method (CCM) is used to estimate and project populations by age and sex. The CCM offers a consistent framework for reconciling historical population estimates with estimated levels and trends in fertility, mortality, and net international migration. This method relies on the population balancing equation (Equation 1), whereby the national population can only increase or decrease between two points in time (e.g., t and $t + n$ where t is the initial date and n the time interval) as the result of births, deaths, and movements of the population across national boundaries (i.e., emigration and immigration).

$$\mathrm{Pop(t+n) = Pop(t) + Births(t, t+n) - Deaths(t, t+n) + NetMigrants(t, t+n)} \tag{1}$$

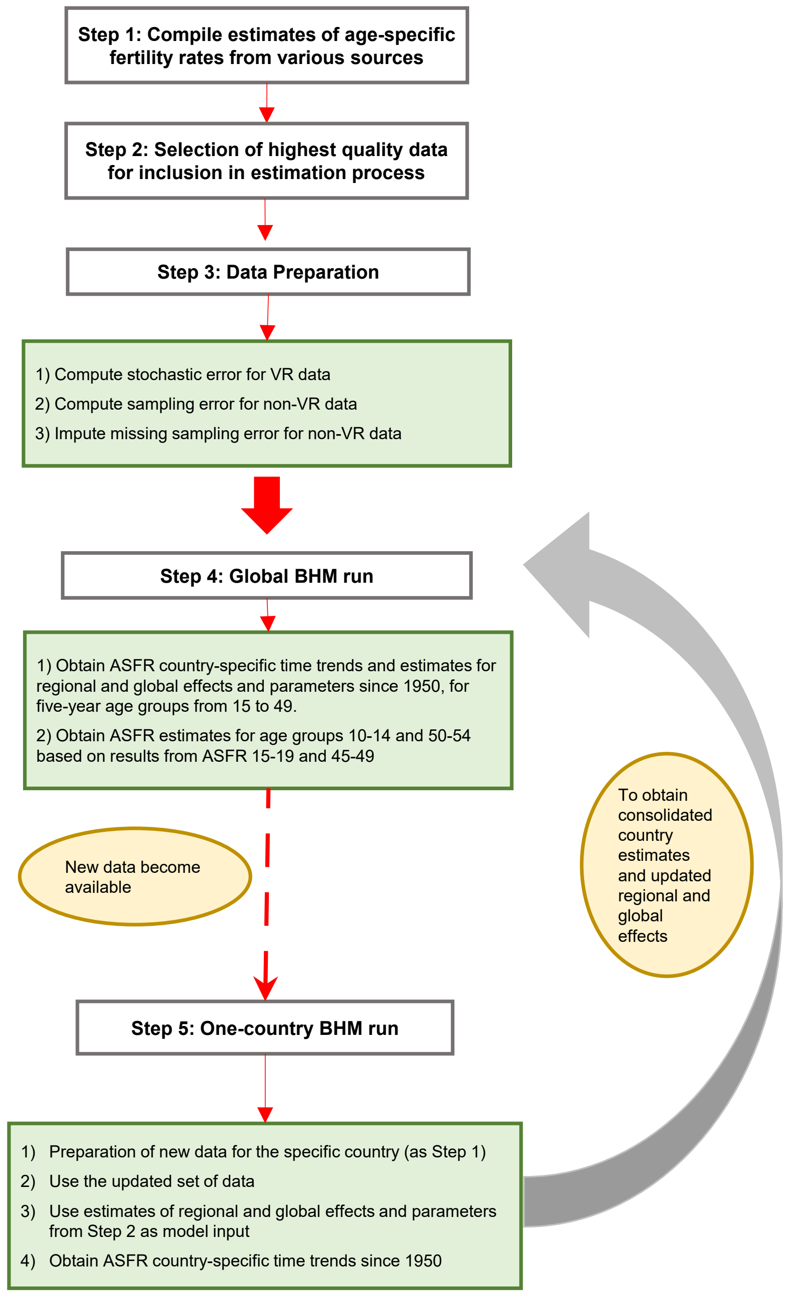
As input for CCM, births within the time interval are compiled as the product of the estimates of age-specific fertility rates (ASFR) and the estimated number of women in a given age group. Since the 2022 revision of the *World Population Prospects*, the time interval and age group on which CCM is applied are one-year periods and single years of age (1x1).

This technical report describes the Bayesian hierarchical model (BHM) used to produce annual estimates of the levels and trends in age-specific fertility rates (ASFR) by five-year age groups from 10 to 54 years for all countries and areas since 1950. The model uses an extensive database of age-specific fertility estimates from reference data sources maintained by the Population Division. The BHM allows the sharing of information across countries and periods to inform annual estimates for countries and periods with sparse, biased, or missing data. The annual estimates of ASFR for all five-year age groups produced with the BHM are eventually rescaled such that the total fertility rate obtained by appropriately aggregating the age-specific fertility rates is equal to that produced by bayesTFR, a model for total fertility only (Liu and Raftery, 2022; United Nations, 2022b), which is the model used in the WPP 2022 to estimate the total fertility rate based on a larger set of input fertility data (i.e., including data sources and estimation methods for which only the total fertility rate is available, but not the age-specific rates). The overall process is illustrated in figure 1.

In the following steps of the WPP process (United Nations, 2022b), the full annual time series of estimated fertility rates by five-year age groups was graduated to a single year of age using the Calibrated Spline (CS) method (Schmertmann, 2014), informed by a large set of empirical single-year fertility rates representing a diverse range of fertility age patterns (United Nations, 2022b). Lastly, preparing the final estimates of ASFR to be used in CCM entailed adjusting the graduated rates, as needed, for consistency with the total fertility rate each year (United Nations, 2022b).

The resulting age-specific fertility estimates from the *World Population Prospects* are used to monitor the global and regional trends of the Sustainable Development Goal (SDG) indicator 3.7.2 - Adolescent birth rate (aged 10–14 years; aged 15–19 years) per 1,000 women in that age group. SDG indicator 3.7.2 is one of two indicators used for the global monitoring of the progress made towards SDG target 3.7, which aims to ensure universal access to sexual and reproductive health care services, including for family planning, information and education, and the integration of reproductive health into national strategies and programs by 2030. The Population Division is the custodian agency for SDG indicator 3.7.2.

Figure 1. Workflow to estimate age-specific fertility rates



II. Data compilation and data preparation

A. Data compilation

Analysts from the Population Division collected available data from various reference data sources, such as population censuses, surveys, vital and population registers, analytical reports and other sources for a given country. [1] The preferred data source for fertility is counts of live births by the age of the mother from a system of civil registration and vital statistics (CRVS) with national coverage and a high level of completeness (United Nations, 2017a). In cases where birth registration is deficient or lacking, fertility estimates are typically obtained through household sample surveys. Demographic sample surveys may provide estimates of fertility by asking women detailed questions to obtain their complete childbearing status. Current global survey programmes collecting detailed birth histories include the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS). [2] Separate from the global programmes, some countries field their national demographic surveys, and a few have established sample vital registration systems. Population censuses serve as additional sources of information on fertility through questions about the number of children ever born and the births in the last 12 (or 24) months before the census date. Moreover, census or survey household population counts can be used to estimate age-specific fertility rates through "own-children estimates" methods (Cho, Retherford and Choe, 1986; United Nations, 2004).

From all compiled data in the empirical demographic database of the Population Division[3], when multiple sources of information were available, only the series that met the highest authoritative standards were selected for inclusion in the estimation model. For example, birth registration data were used only for combinations of country and year for which birth registration completeness was higher than 60 per cent (Preston, 1984). Additionally, to avoid duplicates, when multiple sources of information existed, only one series of estimates and one series of vital registration were selected for each country and year combination. However, for the same time period, if results from multiple surveys, censuses or vital registration were available, each of these different data sources (or estimation methods) were included.

Overall, 41.3 per cent of data selected for the estimation of fertility rates by the age of mother were from vital registration covering 168 countries or areas (table 1). Additional common data sources were surveys (31.1 per cent of all observations, covering 141 countries) and estimates (24.7 per cent covering 93 countries). Other data sources providing a smaller percentage of selected data were censuses (2.7 per cent of all observations covering 128 countries) and sample registration systems (only 0.1 per cent of all observations, but a particularly important source of data for two countries - Bangladesh and India).

---

[1] Traditionally, the data on number of births and fertility rates are obtained from the United Nations Statistics Division (Demographic Yearbook), national statistical offices and regional ones (e.g., Eurostat, OECD), United Nations Regional Commissions, other United Nations entities (e.g., UNAIDS, UNFPA, UNESCO, UNICEF, WHO, World Bank), and complemented using international databases (the Human Fertility Database (Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria), 2021) and Human Fertility Collection (Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria), 2020), the International Data Base (U.S. Bureau of the Census, 2020), the Global Burden of Disease project (Institute for Health Metrics and Evaluation, 2020), and public use microdata archives (e.g., DHS, MICS, IPUMS-International).

[2] Fertility estimates from some other international survey programs were also considered, for example, the Performance Monitoring and Accountability (PMA) surveys. Other international survey programs that provided fertility estimates in decades prior to 2010 included the World Fertility Survey (WFS), the Contraceptive Prevalence Surveys (CPS), the Reproductive Health Surveys (RHS), and the Pan-Arab Project for Family Health (PAPFAM).

[3] DemoData SQL database available at https://population.un.org/DemoData/web/.

TABLE 1. DATA AVAILABILITY BY TYPE OF DATA SOURCE

Source	Number of countries	Number of observations	Proportion of observations (percentage)
Census	128	3683	2.7
Survey	141	42648	31.1
Estimate	93	33866	24.7
CRVS	168	56568	41.3
SRS	2	154	0.1
Total	237	136919	100

Note: Observations were selected from the DemoData SQL database (https://population.un.org/DemoData/web/) for analysis of age-specific fertility rates as of March 2022. CRVS: Civil Registration and Vital Statistics. SRS: Sample Registration System.

More than half of the observations from the surveys were obtained through DHS (more than 25,000 observations, 59.1 per cent of all survey observations) and MICS (more than 6,000 observations, 14.5 per cent) (table 2). Most of the observations from the demographic surveys were calculated using the full birth (or pregnancy) histories (representing 93.8 per cent of all observations from surveys), which reconstructed the list of births a woman had, including information on the date of birth (table 3). This type of data allows for the calculation of the ASFR for the periods preceding the survey. In this context, the period between the interview and the event – birth, in this case – can be classified into different intervals (5-year periods, 3-year periods or irregular intervals). As shown in table 4 (for the observations from birth histories classified by five-year periods), the availability of such data was critical, especially for the youngest age groups (10-14 and 15-19), providing estimates of age-specific fertility for long periods preceding the survey.

TABLE 2. DATA AVAILABILITY BY TYPE OF SURVEY

Survey	Countries	Observations	Proportion of observations (percentage)
DHS	93	25184	59.1
MICS	64	6201	14.5
WFS	40	2624	6.2
RHS	14	1386	3.2
PAPFAM/PAPCHILD	14	1292	3
MIS	18	875	2.1
Panel	1	86	0.2
Other	93	5000	11.7
Total	141	42648	100

Note: DHS: Demographic and Health Surveys. MICS: Multiple Indicator Cluster Surveys. WFS: World Fertility Surveys. RHS: Reproductive Health Surveys. PAPFAM: Pan Arab Project for Family Health. PAPCHILD: Pan Arab Project for Child Development. MIS: Malaria Indicators Surveys.

TABLE 3. ESTIMATION METHODS USED TO CALCULATE AGE-SPECIFIC FERTILITY RATES FOR CENSUS AND SURVEY

Source	*Method*	*Observations*	*Proportion of observations (percentage)*
Census	Recent births	2866	77.8
	Population methods	584	15.9
	Other methods	233	6.3
Survey	Birth histories	40017	93.8
	Recent births	2187	5.1
	Other methods	444	1.0

TABLE **4**. NUMBER OF OBSERVATIONS FOR AGE-SPECIFIC FERTILITY RATES BY 5-YEAR PERIODS PRECEDING THE SURVEY, ACCORDING TO AGE GROUP

	Number of years preceding survey						
Age group	*0-4*	*5-9*	*10-14*	*15-19*	*20-24*	*25-29*	*30-34*
10-14	458	430	429	426	426	404	390
15-19	462	432	431	429	428	397	
20-24	466	433	432	430	417		
25-29	466	433	432	423			
30-34	466	433	425				
35-39	466	426					
40-44	465						
45-49	453						

Note: Only observations from birth histories using 5-year periods are presented in this table. Other observations used 3-year or irregular periods, and are not presented in this table.

The years of the most recent observations of age-specific fertility rates from all available data sources and estimation methods vary greatly among countries. Among the 236 countries or areas with 1,000 inhabitants or more in 2021, all but 38 had available fertility data collected in 2015 or later (table 5). For 20 countries, the most recent data were collected between 2011-2014, for 16 countries between 2007 and 2010, and for only 2 countries, the most recent national data were from 2004 (Western Sahara) and 2003 (Lebanon). In terms of the differences in recent data availability across regions (figure 2), the majority of countries in Europe and Northern America and Australia and New Zealand had the latest data available in 2019 or 2020 (figure 1), while 52.4 per cent of countries in sub-Saharan Africa and 40 per cent in Oceania (excluding Australia and New Zealand) had the latest data available in 2016 or earlier. In the remaining regions, some countries had recent data from 2019 or 2020, whereas other countries had no recent data available.

The metadata associated with the 2022 revision of WPP, available online, provide further details about the age-specific fertility data used for each country[4] and for the development of BHM methods presented in this technical report. In many cases, estimates derived from different data sources or methods vary significantly.

[4] https://population.un.org/wpp/Download/Metadata/Documentation/ and https://population.un.org/wpp/DataSources/.

TABLE 5. NUMBER OF COUNTRIES BY THE LATEST YEAR WITH DATA AVAILABLE

Year	Countries
2014 and earlier	38
2015	12
2016	18
2017	19
2018	30
2019	76
2020	39
2021	5

Figure 2. Proportion of countries by latest year of available data by SDG regions

B. Data preparation

This section describes the preparation of the input dataset used in BHM, including the calculation of the stochastic error for CRVS data, and the calculation of sampling error for non-CRVS data.

1. Stochastic error of CRVS data

The first step in data preparation entails the computation of stochastic errors for the CRVS data. As previously stated, CRVS data were used only for combinations of country and year in which the birth reporting completeness was above 60 per cent. The number of births by women in a specific age group is computed as the product of the observed fertility rates for the women in that age group from CRVS and the number of women in that age group (from WPP estimates of the female population by age). Generally, the number of births computed is smaller than the actual number, as births are subject to under-reporting. In the calculation of the stochastic errors of the CRVS data, the uncertainties from both under-reported and reported births were included.

First, the reported birth-reporting completeness[5] (BRC) for country c in year t, denoted as $z_{c,t}$, accounts for the uncertainty of under-reported births. The reported BRC $z_{c,t}$ was assumed to be uniformly distributed. The g-th simulated BRC $z_{c,t}^{g}$ is obtained by:

$$z_{c,t}^{g} \sim U\left(z_{c,t} - \delta(z)_{c,t}, z_{c,t} + \delta(z)_{c,t}\right) \tag{2}$$

where $\delta(z)_{c,t}$ is the standard error of the reported BRC for country c in year t. T was assumed to decrease linearly from 0.25 to 0.05 when the reported BRC $z_{c,t}$ was within the interval [60%, 95%]. When $z_{c,t}$ further increased to 100%, $\delta(z)_{c,t}$ was assumed to further decline linearly to zero. $\delta(z)_{c,t}$ was imputed as follows:

$$\begin{aligned} \delta(z)_{c,t} &= 0.25 - \frac{0.25 - 0.05}{0.95 - 0.6}\left(z_{c,t} - 0.6\right) && if\ 60\% \leq z_{c,t} < 95\% \\ \delta(z)_{c,t} &= 0.05 - \frac{0.05}{1 - 0.95}\left(z_{c,t} - 0.95\right) && if\ z_{c,t} \geq 95\% \end{aligned} \tag{3}$$

It is worth noting that the assumptions made in simulating BRC are largely based on expert opinions. When additional information becomes available about the distribution of $z_{c,t}$, the simulation steps can be updated accordingly.

The g-th simulated number of under-reported births $B_{c,t}^{\text{under}(g)}$ was calculated as the product of the number of births reported $B_{c,t}^{\text{report}}$ and the difference between 1 and the g-th simulated BRC $z_{c,t}^{(g)}$:

[5] The completeness of birth registration corresponds to the proportion of all births that occurred in a given year and were reported to civil registration authorities. The degree of completeness of birth registration can be evaluated through various analytical methods, including aggregated analysis (comparing observed vital events with reference figures from an alternative source believed to represent the true potential value of expected events), individual-level analysis (comparing and linking individual records of vital events from multiple data sources to identify matched records and those present in one data source but not another), indirect demographic analysis (comparing reported births with those expected from the reverse survival of enumerated children in censuses or from health or education statistics with universal coverage), or census or survey assessments (asking questions about whether vital events reported in the survey or census have been registered with local authorities) (Rao and others, 2020).

$$B_{c,t}^{\text{under}(g)} = B_{c,t}^{\text{report}}\left(1 - z_{c,t}^{(g)}\right). \tag{4}$$

The g-th simulated total number of births $B_{c,t}^{(g)}$ was obtained as the sum of reported births and g-th simulated under-reported births. The uncertainty in the estimates was included by assuming that the simulated number of births had a Poisson distribution:

$$B_{c,t}^{(g)} \sim Poisson\left(B_{c,t}^{\text{report}} + B_{c,t}^{\text{under}(g)}\right). \tag{5}$$

At this point, the g-th simulated number of total births $B_{c,t}^{(g)}$ was divided by the number of women in the specific age group $N_{c,t}^{\text{female}}$ to obtain the g-th simulated $\text{ASFR}_{c,t}^{(g)}$:

$$\text{ASFR}_{c,t}^{(g)} = B_{c,t}^{(g)} / N_{c,t}^{\text{female}} \tag{6}$$

The stochastic error is the standard deviation of the simulated $\text{ASFR}_{c,t}^{(g)}$:

$$\sigma_{c,t} = \sqrt{\frac{\sum_{g=1}^{G}\left(\text{ASFR}_{c,t}^{(g)} - \overline{\text{ASFR}}_{c,t}\right)^2}{G - 1}}, \tag{7}$$

where:

$$\overline{\text{ASFR}}_{c,t} = \frac{\sum_{g=1}^{G} \text{ASFR}_{c,t}^{(g)}}{G} \tag{8}$$

2. Sampling errors for non-CRVS data

Whenever available, sampling errors were calculated from the micro-datasets. If the sampling errors were missing, they were imputed as the median of the sampling errors within each combination of age groups and the period between the interview and the event. The sampling errors were then calculated using a set of simulated normally distributed ASFR, with a mean equal to the observed ASFR and a standard deviation equal to the sampling error computed from the microdata. The computed standard deviation of the simulated ASFR on the logit scale provided the sampling error of the non-CRVS data.

3. Data inclusion criteria

Additionally, observations with implausible extreme values were removed based on age-specific inclusion criteria: (1) exclude observations with zero values in some age groups and (2) exclude observations above age-specific upper cut-off values based on the estimated maximum natural fertility in human populations (Henri, 1961). Table 6 summarizes the age-specific inclusion criteria for the ASFR observations.

Table 6. Age-specific inclusion criteria for ASFR observations (per 1000 women)

Age group	Henri (1961) natural fertility for married women* (1)	Proportion married or in a union (maximum observed)** (2)	ASFR for natural fertility for all women (3) = (1)*(2)	Upper cut-off values to include observations (4)	Exclude true zeros (5)
10-14				30	YES
15-19		0.60		250	YES
20-24	435	0.90	392	450	YES
25-29	407	0.98	399	475	YES
30-34	371	0.98	364	425	YES
35-39	298	0.98	292	350	YES
40-44	152	0.95	144	200	YES
45-49	22	0.95	21	50	NO
50-54		0.90		2	NO

*Estimates of the average number of live births among women of a given age group based on 13 historical populations with no deliberate use of contraception or other fertility control methods. Henry (1961) defined natural fertility as "the maximum fertility of a population, which would be achieved if all women were married, lived in stable unions, and had no recourse to voluntary contraception".
**United Nations (2019e).

III. Bayesian hierarchical model for estimating age-specific fertility

Bayesian hierarchical models (BHM) in the context of demographic analysis allow the use of data from different sources, while accounting for uncertainty and potential biases (Bijak, 2016). This section provides the technical details of the BHM used to estimate the age-specific fertility rates by five-year age groups from 10 to 54; first, explaining the process models (section III A-B), which are theoretical models used to describe the levels and trends in the true underlying ASFR, and second, the data models (section III C-D), which are models that describe patterns and uncertainties in the ASFR observations, given the process models. The process and data models are presented separately for age groups 10-14, 45-49, and 50-54 because different parameters are used.

A. Process model for age group 15 to 44

The BHM estimates the logit ASFR for all the 5-year age groups between 15-19 and 40-44 in a specific country and a specific year. The logit scale constrains the ASFR to fall between 0 and 1. The true logit-scaled underlying ASFR from country c in year t for all five-year age groups, $\text{logit}\left(ASFR_{c,t}\right)$, is modelled as the combination of (i) the country-specific effect of female educational attainment, $\alpha_c \log\left(\text{EDU}_{c,t}\right)$, (ii) regional effect[6] from the total fertility rate (TFR), $W_{r,d[c,t]}$, (iii) country-specific temporal effect, $P_{c,t}$, and (iv) country-specific offset, η_c. Specifically,

$$\text{logit}\left(ASFR_{c,t}\right) = \alpha_c \log\left(EDU_{c,t}\right) + W_{r,d[c,t]} + P_{c,t} + \eta_c \tag{9}$$

$W_{r,d[c,t]}$ models the regional non-linear relationship between ASFR and TFR for country c year t, where the index r refers to the SDG region r to which country c belongs. Specifically, let $V_{c,t}$ denote the log of TFR multiplied by 1000 for country c year t,[7] taken from the WPP estimates (in this case from the 2022 revision of WPP (United Nations, 2022a)). A grid of values κ_d was defined for $d \in \{1, \cdots, x\}$, x is the number of locations where $W_{r,d}$ was evaluated, where $x = 151, \kappa_1 = \log(745/1000)$ and κ_x as the 99.5th percentile of $V_{c,t}$ across all country-years with data. Each $V_{c,t}$ was matched to the κ_d with the smallest absolute difference from $V_{c,t}$, denoting the dth index for country-year c, t as $d[c,t]$. To model the relationship between $W_{r,d}$ and κ_d. $W_{r,d}$ was assumed to be constant outside the range of κ_1 and κ_x across all c and t. In particular:

$$\Delta^2\left(W_{r,d}\right) = W_{r,d} - 2W_{r,d+1} + W_{r,d+2}, \tag{10}$$

$$\Delta^2\left(W_{r,d}\right) \sim \mathcal{N}\left(0, \frac{1}{\tau_r^W}\right), \quad \text{for } r \in \{1, \ldots, 7\},\ d \in \{1, \cdots, x-2\} \tag{11}$$

[6] The regional classification used are SDG regions. Further details about the classification used: https://population.un.org/wpp/DefinitionOfRegions/.

[7] Multiplying by 1000 is done for computational purposes only, such that a finer grid of TFR values can be evaluated in the process model.

$P_{c,t}$ accounts for the within-country temporal fluctuations. We used a first-order random walk (RW1) to model $P_{c,t}$ as follows:

$$\Delta P_{c,t} = P_{c,t} - P_{c,t-1}, \tag{12}$$

$$\Delta P_{c,t} \sim \mathcal{N}\left(0, \frac{1}{\tau_c^p}\right), \qquad \text{for } t \in \{1950, \dots, \omega\}, c \in \{1, \dots, 237\} \tag{13}$$

where ω is the latest year in the estimation period.

The country-specific regression coefficients for female educational attainment (the proportion of females with any education, GBD 2019[8]) follow hierarchical normal distributions:

$$\alpha_c \sim \mathcal{N}(0, 1/\tau_\alpha), \qquad \text{for } c \in \{1, \dots, 237\} \tag{14}$$

Penalized Complex (PC) priors to the regional precision parameter τ_r^w for $r \in \{1, \dots, 7\}$, and country-specific precision parameter τ_c^p for $c \in \{1, \dots, 237\}$ are assigned as follows:

$$\tau_r^w \sim \mathrm{PC}(z, 0.01), \qquad \text{for } r \in \{1, \dots, 7\}, \tag{15}$$

$$\tau_c^p \sim \mathrm{PC}(1{,}0.01), \text{for } c \in \{1, \cdots, 237\} \tag{16}$$

where z is the standard deviation of all observations. The PC prior is a vague prior. Simpson and others (2017) documented the PC prior specification in detail.

Non-informative priors for the global precision parameters $\tau_\alpha = 1/\sigma_\alpha^2$ are assigned as follows:

$$\tau_\alpha \sim \mathrm{Gamma}(1{,}0.00005) \tag{17}$$

Figure 3.1 presents an overview of the BHM used for ASFR for an example country (Afghanistan). Part (i) consists of estimates of the proportion of the female population receiving any level of education in each country over time. This effect is modelled for each country and can be shared across all countries. In part (ii), the main assumption is that the levels and trends in TFR have similar effects on ASFR across countries within the same region. Since TFR patterns are usually non-linear, their effect on ASFR patterns is assumed to be non-linear across regions. In addition, using the three-phase Bayesian hierarchical model of fertility transition (Alkema and others, 2011), the effect of TFR in phase I (pre-fertility transition period) is assumed to be constant. For example, in Afghanistan, the TFR phase I transition ended in 1995. The TFR effect is constant during phase I. After 1995, the BHM assumed that for a given level of TFR, the TFR effect on ASFR is the same as in other countries in the same region. As the TFR in Afghanistan declined after 1995, the model estimated a decreasing effect on ASFR. Part (iii) models the correlation over time for each country, while part (iv) offsets the discrepancy between the logit-scaled true level of the ASFR and all other effects.

[8] Institute for Health Metrics and Evaluation (2020).

B. Process model for age groups 10–14 and 50–54

The main assumption in this model is that the ASFR for age groups 10–14 and 50–54 are correlated with their neighbouring age groups (ASFR 15–19 and 45–49, respectively). Figures 3 and 4 present an overview of the BHM used for ASFR for the age groups of 15-19 and 10–14 years for an example country (Bangladesh). Instead of modelling the ASFR directly for these two age groups, the model was fitted to the ratio of the logit-scaled targeted ASFR to the logit of the neighbouring ASFR. Hence, the ASFR estimates are informed by empirical observations, and in periods when data are unavailable, they follow the trend and level of neighbouring ASFR. Throughout this section, superscript a refers to the estimated age group, that is, age groups 10–14 or 50–54, and superscript $a*$ refers to the corresponding neighbouring age group, that is, age group 15–19 as the neighbouring age group for 10–14 and 45–49 as the neighbouring age group for 50–54. The process model is expressed in Equation 18 for $a \in \{10-14, 50-54\}$:

$$\frac{\operatorname{logit}(ASFR_{c,t}^{a})}{\operatorname{logit}(\widehat{ASFR}_{c,t}^{a*}(MA_{10}))} = \alpha_c^a \log\left(\mathrm{EDU}_{c,t}\right) + \beta_c^a \log\left(\mathrm{TFR}_{c,t}\right) + W_t^a + P_{c,t}^a, \tag{18}$$

The process model estimates the ratio between two logits: (1) logit of the outcome of interest ASFR in age groups 10-14 or 50-54 $\operatorname{logit}(ASFR_{c,t}^{a})$, and (2) logit of the 10-year moving average of ASFR median estimates from the neighbouring age group, denoted as $\operatorname{logit}(\widehat{ASFR}_{c,t}^{a*}(MA_{10}))$, where the median estimates of the ASFR from the neighbouring age group are obtained based on the model described in section A. Regarding the denominator $\widehat{ASFR}_{c,t}^{a*}(MA_{10})$, the 10-year moving average of ASFR of the neighbouring age groups (15-19 and 45-49, respectively) preserved the general country-specific pattern of the ASFR while eliminating the within-country year-by-year fluctuation. The moving average was kept constant during phase I of the fertility transition (as explained above).

The process model for ages 10-14 and 50-54 includes three country-specific effects: (1) the female educational attainment linear effect, $\alpha_c^a \log\left(\mathrm{EDU}_{c,t}\right)$, where $\mathrm{EDU}_{c,t}$ across all the countries and years was obtained from GBD 2019; (2) the TFR linear effect, $\beta_c^a \log\left(\mathrm{TFR}_{c,t}\right)$, where $\mathrm{TFR}_{c,t}$ across all the countries and years was obtained from WPP (based on the latest set of empirical data modelled using bayesTFR); (3) the flexible temporal effect $P_{c,t}^a$; and (4) one global temporal effect W_t^a that models the non-linear relationship between ASFR and time across all countries. W_t^a was modelled using a second-order random walk (RW2) structure for $a \in \{10-14, 50-54\}$:

$$\Delta^2(W_t^a) = W_t^a - 2W_{t+1}^a + W_{t+2}^a, \tag{19}$$

$$\Delta^2(W_t^a) \sim \mathcal{N}(0, \tau(a)_w^{-1}), \quad \text{for } t \in \{1950, \ldots, 2019\},$$

$P_{c,t}^a$ accounts for the within-country temporal fluctuations. We used a first-order autoregression (AR1) time series model to model $P_{c,t}^a$ for $a \in \{10-14, 50-54\}$:

$$P_{c,t}^a \sim N\left(0, \frac{1}{\tau(a)_c^p(1-\rho^2)}\right), \text{for } t = 1950, c \in \{1, \ldots, 237\} \tag{20}$$

$$P_{c,t}^a = \rho P_{c,t-1}^a + \varepsilon_{c,t}^a, \text{for } t \in \{1951, \dots, 2019\}, c \in \{1, \dots, 237\} \tag{21}$$

$$\varepsilon_{c,t}^a \sim \mathcal{N}\left(0, \frac{1}{\tau(a)_c^p}\right), \qquad \text{for } t \in \{1951, \dots, 2019\}, c \in \{1, \dots, 237\}, \tag{22}$$

We assumed that the time series had a mean of zero and temporal correlation ρ with the previous year ($|\rho| < 1$ so that the process was stationary). $\varepsilon_{c,t}^a$ was the country-year-specific white noise with a mean of zero and country-specific precision parameter $\tau(a)_c^p$.

The country-specific regression coefficients for female educational attainment (the proportion of females with any education) and TFR followed hierarchical normal distributions for $a \in \{10-14, 50-54\}$:

$$\alpha_c^a \sim \mathcal{N}(0, 1/\tau(a)_\alpha), \qquad \text{for } c \in \{1, \dots, 237\} \tag{23}$$

$$\beta_c^a \sim \mathcal{N}\left(0, 1/\tau(a)_\beta\right), \qquad \text{for } c \in \{1, \dots, 237\} \tag{24}$$

We assigned non-informative priors to the global temporal correlation parameter ρ, the country-specific precision parameter $\tau(a)_c^p$ for $a \in \{10-14, 50-54\}$ and $c \in \{1, \dots, 237\}$, and the global precision parameters $\tau(a)_w$, $\tau(a)_\alpha$, and $\tau(a)_\beta$ for $a \in \{10-14, 50-54\}$:

$$\rho \sim PC(0.8, 0.5) \tag{25}$$

$$\tau(a)_c^p \sim \mathrm{PC}(1, 0.01), \text{for } c \in \{1, \cdots, 237\} \tag{26}$$

$$\tau(a)_w \sim \mathrm{PC}(z, 0.01) \tag{27}$$

$$\tau(a)_\alpha \sim \mathrm{Gamma}(1, 0.00005) \tag{28}$$

$$\tau(a)_\beta \sim \mathrm{Gamma}(1, 0.00005) \tag{29}$$

C. Process model for age groups 45–49 and 50-54

Similar to the main model assumption for 10-14 and 50-54, the ASFR for age groups 45-49 and 50-54 was assumed to correlate with its neighbouring age groups (40-44 and 45-49, respectively). Throughout this section, superscript a refers to the estimated age groups, that is, age groups 45–49 or 50-54, and superscript $a*$ refers to its corresponding neighbouring age group, that is, age groups 40–44 or 45-49, respectively.

The process model is expressed in Equation 30 for $a = 45\text{–}49$ and $a* = 40\text{–}44$ as follows:

$$\frac{\text{logit}(ASFR^{a}_{c,t})}{\text{logit}(\widehat{ASFR}^{a*}_{c,t}(MA_{10}))} = \alpha^{a}_{c}\log\left(\text{EDU}_{c,t}\right) + W^{a}_{r,d[c,t]} + P^{a}_{c,t}, \tag{30}$$

where $\widehat{ASFR}^{a*}_{c,t}(MA_{10})$ denotes the 10-year moving average of ASFR median estimates from the neighbouring age group 40-44, and the median estimates of the ASFR from the neighbouring age group were obtained based on the model described in section A for 40-44 and in this section for 45-49.

The process model for these age 45-49 groups includes three parts: (1) female educational attainment linear effect, $\alpha^{a}_{c}\log\left(\text{EDU}_{c,t}\right)$, where $\text{EDU}_{c,t}$ across all the countries and years are obtained from GBD 2019 ; (2) regional effect from the TFR (obtained from the WPP 2022 revision), with the same model specification as described in section A the process model for ages 15-44, Equations 10-11 ; (3) flexible temporal effect $P^{a}_{c,t}$, where we used a third-order autoregression (AR3) time series model to estimate:

$$P^{a}_{c,t} = \phi_1 P^{a}_{c,t-1} + \phi_2 P^{a}_{c,t-2} + \phi_3 P^{a}_{c,t-3} + \varepsilon^{a}_{c,t} \tag{31}$$

$$\varepsilon^{a}_{c,t} \sim \mathcal{N}\left(0, \frac{1}{\tau(a)^{p}_{c}}\right), \qquad \text{for } t \in \{1951, \dots, 2019\}, c \in \{1, \dots, 237\}, \tag{32}$$

The time series $P^{a}_{c,t}$ was assumed to have a mean of zero and a temporal correlation with the previous three years $P^{a}_{c,t-1}$, $P^{a}_{c,t-2}$, and $P^{a}_{c,t-3}$ with correlation parameters ϕ_1, ϕ_2, and ϕ_3 respectively (where ϕ_1, ϕ_2, and ϕ_3 are reparametrized by the partial correlation functions to maintain the stationarity of the time process). The prior for the reparametrized ϕ_1, ϕ_2, and ϕ_3 followed a multivariate normal distribution.[9]

D. Data model for age group 15 to 44

The i-th observed ASFR, $asfr_i$, was modelled on the logit scale to ensure that the ASFR fell within the bounds of 0 and 1, as shown in Equation 33. The logit of the observed fertility rate was assumed to be the sum of (1) the true underlying rate on the logit scale $\text{logit}\left(ASFR_{c[i],t[i]}\right)$ for country $c[i]$ in year $t[i]$, and (2) the measurement error δ_i for the i-th observation $asfr_i$. The indices $c[i]$ and $t[i]$ are used to distinguish multiple observations from the same country-year c and t. The i indexes observations across all country-years.

$$\text{logit}(asfr_i) = \text{logit}\left(ASFR_{c[i],t[i]}\right) + \delta_i \tag{33}$$

As shown in Equation 34, the measurement error δ_i was modelled as the sum of (i) the sampling/stochastic error σ^2_i and (ii) the non-sampling error $\omega^2_{s[i]}$ for the data source type s to which the i-th observation $asfr_i$ belongs:

[9] The reparameterization and the multivariate normal prior for ϕ_1, ϕ_2, and ϕ_3 are achieved by the INLA R-package in the background. Refer to https://inla.r-inla-download.org/r-inla.org/doc/latent/ar.pdf for details on the reparameterization and prior.

$$\delta_i \sim \mathcal{N}\left(0, \sigma_i^2 + \omega_{s[i]}^2\right) \tag{34}$$

The sampling/stochastic errors σ_i^2 were pre-calculated for each observation as described in the previous section. They reflect the uncertainty resulting from the survey sampling design for data from surveys and censuses and stochastic uncertainty from administrative records for vital registration data. The non-sampling errors ω_s^2, for $s \in \{1, \dots, 5\}$ are usually unknown, but inevitable during data collection and processing. They represent uncertainty from non-responses, recall bias, and data input errors, among others. For this reason, we modelled non-sampling errors as data source-specific parameters by assigning vague priors:

$$1/\omega_s^2 \sim \mathrm{PC}(z, 0.01), \text{for } s \in \{1, \dots, 5\}. \tag{35}$$

E. Data model for age groups 10–14, 45-49 and 50–54

In this section, the superscript a refers to the estimated age group, that is, age groups 10–14, 45-49 or 50–54. The superscript $a*$ refers to the corresponding neighbouring age group, that is, age group 15–19 as the neighbouring age group for 10–14, age group 40–44 as the neighbouring age group for 45–49, and 45–49 as the neighbouring age group for 50–54.

The data model for age groups 10-14, 45–49, and 50-54 is presented in Equation 36, where:

$asfr_i^a$ represents the i-th observation of ASFR in the age group 10-14, 45-49, or 50-54, and

$ASFR_{c[i],t[i]}^{a*}$ is the corresponding ASFR estimate from the neighboring age group 15-19 (for 10-14), age group 40-44 (for 45-49), or age group 45-49 (for 50-54).

$$\frac{\mathrm{logit}(asfr_i^a)}{\mathrm{logit}\left(ASFR_{c[i],t[i]}^{a*}\right)} = \frac{\mathrm{logit}(ASFR_{c[i],t[i]}^{a})}{\mathrm{logit}(ASFR_{c[i],t[i]}^{a*})} + \delta_i \tag{36}$$

The measurement error δ_i was modelled in the same manner as shown in Equation 34.

F. Statistical computing

The method used for this study is the Integrated Nested Laplace Approximation (INLA) for Bayesian inference (Rue, Marino and Chopin, 2009), implemented through the R package R-INLA (R Core Team, 2022; Rue, Marino, Lindgren et al., 2013).

G. One-country model

While the BHMs described in sections III A-D are global models that use observations from all countries, running a "*one-country model*" is also possible. During the process of country updates by an analyst, different versions of the model are run with new observations added, adjusted, or removed. When such minor data updates occur, global and regional effects are assumed to be unaffected. Hence, the one-country model uses regional and global effects from the global model as model inputs, and only estimates country-level effects, assuming no change in regional and global effects. This is especially useful for obtaining results using updated data and can provide immediate feedback on the impact of data inclusion, adjustment, or exclusion on the estimates. The advantage of the one-country model is that it runs faster than the global model and can immediately produce updated results.[10] Once the data updates were finalized for all countries, the global model was run to produce the final results for the specific revision of the WPP estimates.

Using the results of the global and regional average effects from the global run, the "*one-country model*" allows to obtain updated estimates for a specific country when new data become available, without having to re-run the full global model. This enables analysts to update their estimates as needed, which will then be consolidated when the next global run is performed. The process model for the age group 15-44 now becomes:

$$\text{logit}\left(ASFR_{c,t}\right) = \alpha_c \log\left(\text{EDU}_{c,t}\right) + \widehat{W}_{r[c],t} + P_{c,t} + \eta_c, \tag{37}$$

where $\widehat{W}_{r[c],t}$ is the median estimate from the global run for region r in country c in year t.

The process model for the age groups 45-49 and 50-54 was as follows.

For $a = 45 - 49$ and $a *= 40 - 44$, we have:

$$\frac{\text{logit}(ASFR_{c,t}^{a})}{\text{logit}(\widehat{ASFR}_{c,t}^{a*}(MA_{10}))} = \alpha_c^a \log\left(\text{EDU}_{c,t}\right) + \widehat{W}_{r,d[c,t]}^{a} + P_{c,t}^{a}, \tag{38}$$

where $\widehat{W}_{r,d[c,t]}^{a}$ is the median estimate from the global run for region r in country c in year t.

The process model for age groups 10-14 and 50-54 was as follows:

For $a \in \{10 - 14, 50 - 54\}$:

$$\frac{\text{logit}(ASFR_{c,t}^{a})}{\text{logit}(\widehat{ASFR}_{c,t}^{a*}(MA_{10}))} = \alpha_c^a \log\left(\text{EDU}_{c,t}\right) + \beta_c^a \log\left(\text{TFR}_{c,t}\right) + \widehat{W}_t^a + P_{c,t}^{a}, \tag{39}$$

where $\widehat{W}_t$ is the median estimate from the global run in year t.

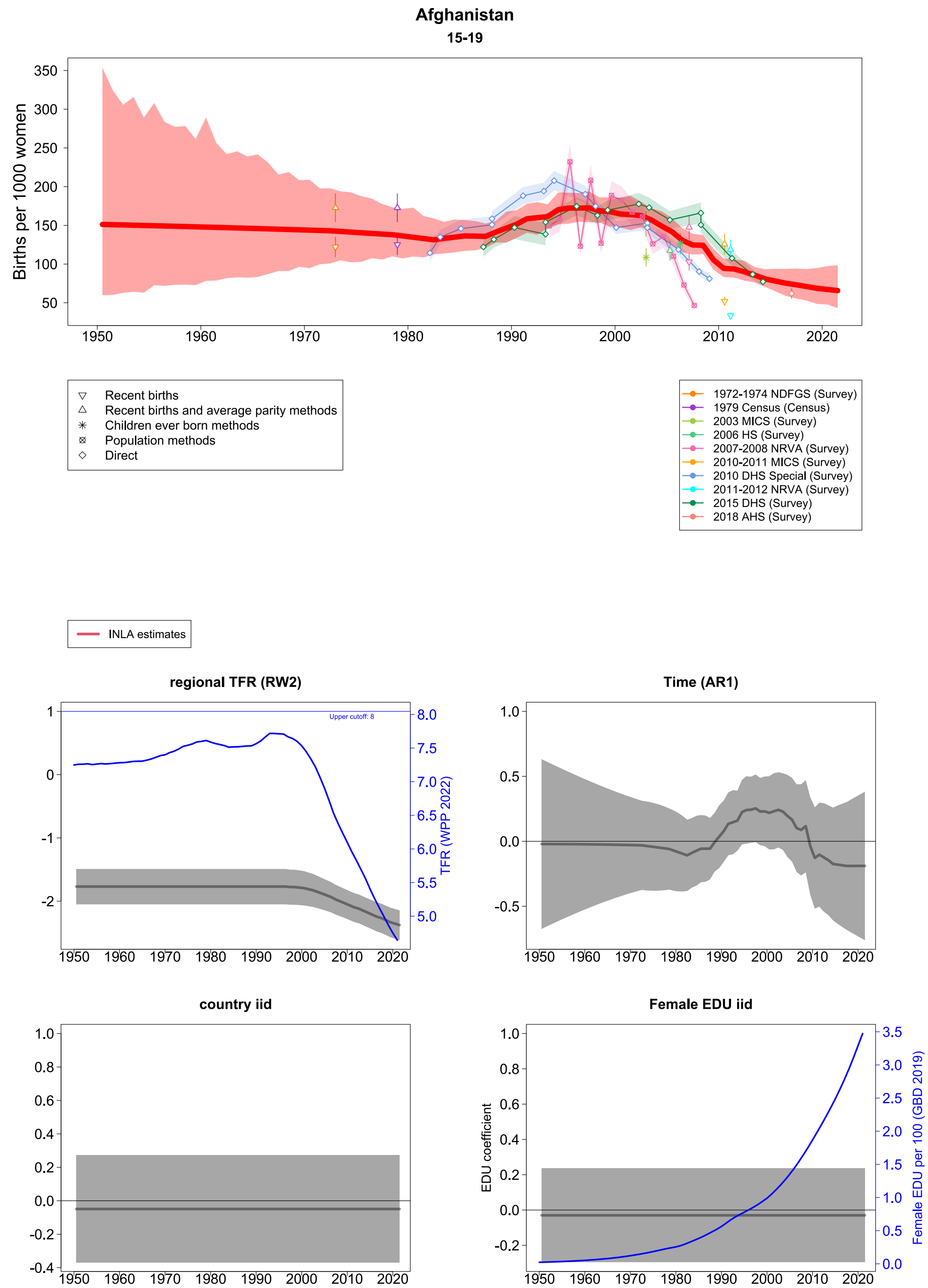
The only change in the data model for all age groups was the distribution of the measurement error δ_i:

$$\delta_i \sim \mathcal{N}\left(0, \sigma_i^2 + \widehat{\omega}_{s[i]}^2\right) \tag{40}$$

where $\widehat{\omega}_{s[i]}^2$ is the median estimate for the non-sampling error variance from the global model for the data source type s to which the i-th observation belongs.

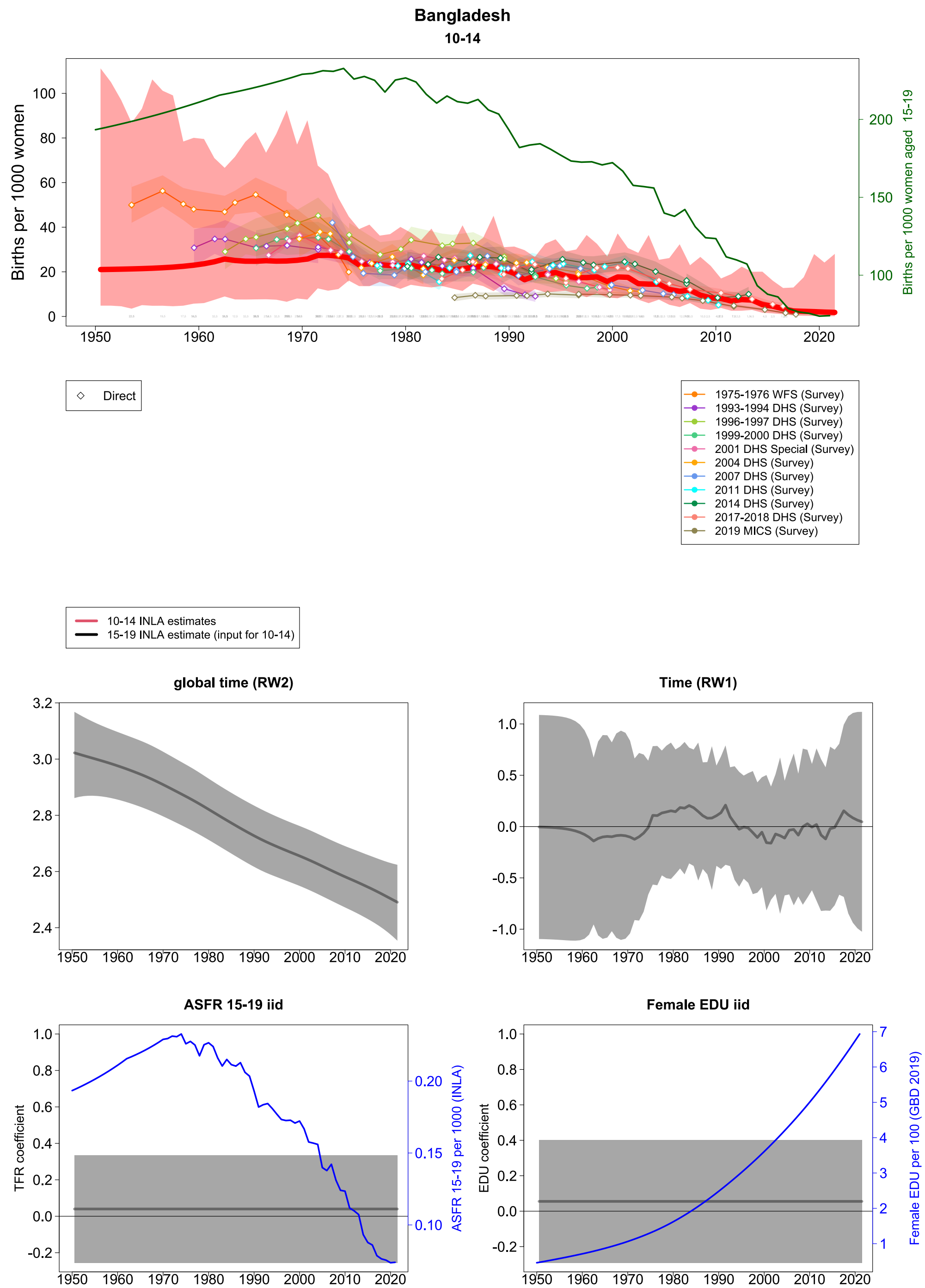
[10] For a typical desktop computer, a global run for an age group takes roughly 18 minutes for 10,000 posterior samples and a one-country run takes 10-15 seconds.

Figure 3. ASFR model illustration for a five-year age group between 15 and 44 years
Example of Afghanistan and age group 15-19



Source: United Nations (2022), GBD (2019) and own calculations.
Note: The curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available). The scale of the y-axis in the 2nd and 3rd rows is an inverse-logit numerical value for each effect.

Figure 4. ASFR model illustration for age group 10-14 using Bangladesh



Source: United Nations (2022), GBD (2019) and own calculations.

Note: The curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The green line and right-hand y-axis represent the results of the model for neighbouring age group (15-19 in this case). The shading and vertical lines around the dots represent sampling errors (if available). The scale of the y-axis in the 2nd and 3rd rows is the inverse-logit numerical value of for each effect.

IV. Results for selected countries

This section presents the BHM estimates of age-specific fertility rates by five-year age groups for the selected countries. The aim is to illustrate the results for three different cases:

1. Countries with high quality and high coverage of data, such as vital registration data from administrative records. Countries in this category are usually high-income or upper-middle-income countries with complete registration systems for recording vital events.
2. Countries without high-quality vital registration data from administrative records but with some data from surveys, censuses, and estimates, and with proper data coverage over time. Countries in this category usually do not have complete vital registration systems, but sampling surveys and censuses are regularly conducted. They are typically middle- or low-income countries.
3. Countries with very limited data.

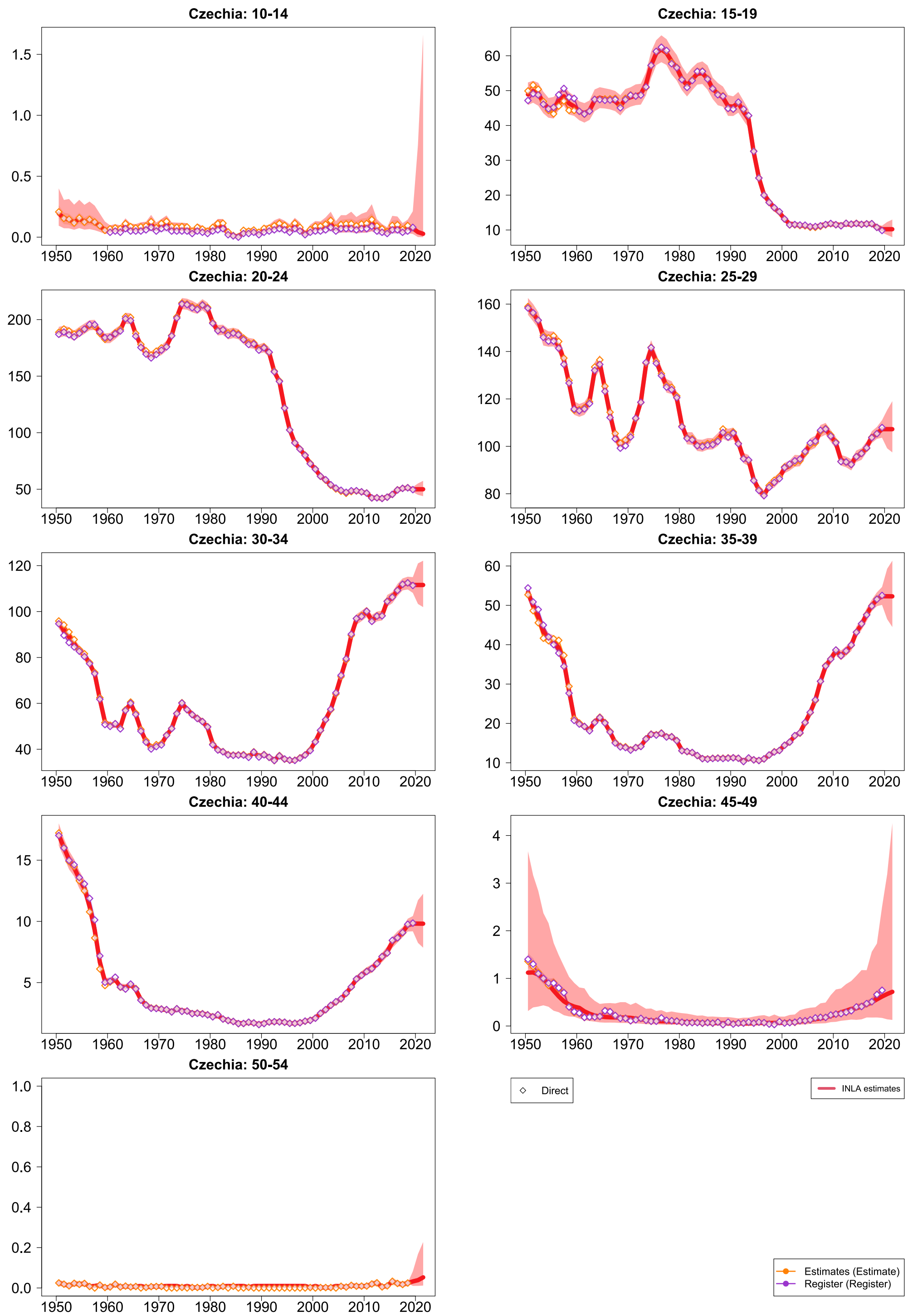
Figure 5 presents the ASFR estimates for each five-year age group from 10 to 54 years from Czechia and the United States of America based on the BHM method. These two countries have high-quality vital registration data for every year since 1950 for all the five-year age groups. The BHM median estimates closely follow the observation levels and trends within a country over time. The uncertainty bounds are narrow for all age groups and reflect the fact that the observations are of high quality and only stochastic errors are accounted for in the model. The wider uncertainty bounds for BHM results in the age groups 10-14, 45-49 and 50-54 in recent years are due to the very small number of births in these age groups. This results in larger stochastic errors in the vital registration data.

Figure 6 shows the ASFR estimates for Egypt and Peru. Egypt had some vital registration observations included in the dataset in the most recent period. Both countries have many observations from other data sources such as international surveys (e.g., DHS and MICS), country-specific surveys, and censuses. The data coverage was generally good in recent decades but low in earlier ones. There are no data for the age group of 50-54. The BHM model estimates follow the data trends for the country-periods in which the data are available. If there are multiple observations for a certain country-year, the BHM model can use all of them. In such cases, the model estimates are closer to the higher-quality observations (corresponding to smaller sampling errors) and less towards the data points with lower quality (equivalent to larger sampling errors). For country-periods with no data, the BHM estimates are mainly driven by the model assumptions: the country-level effect of female education attainment, the regional effect of TFR for all age groups except for 10-14 and the global TFR effect for age group 10-14, country-level temporal effect and offset (i.e., exposure time). For the age group 50-54, the BHM estimates reflect the average experience from all country-periods with data, and also reflect the additional model assumption that the trend in ASFR 50-54 is similar to that in ASFR 45-49.

Figure 7 presents the ASFR estimates for Eritrea and Saudi Arabia. The selected countries have a very limited number of available observations and do not have any data points before the 1980s. The BHM estimates are driven almost entirely by model assumptions for the entire estimation period 1950-1980. Hence, as a result of the hierarchical model structure, the effect of female education attainment is based on the average effect from all countries in the world; the effect of TFR is the average effect from all countries in the same region for age groups between 15 and 49 years or from all countries for age groups 10-14, and the temporal effect, which is the average of all countries.

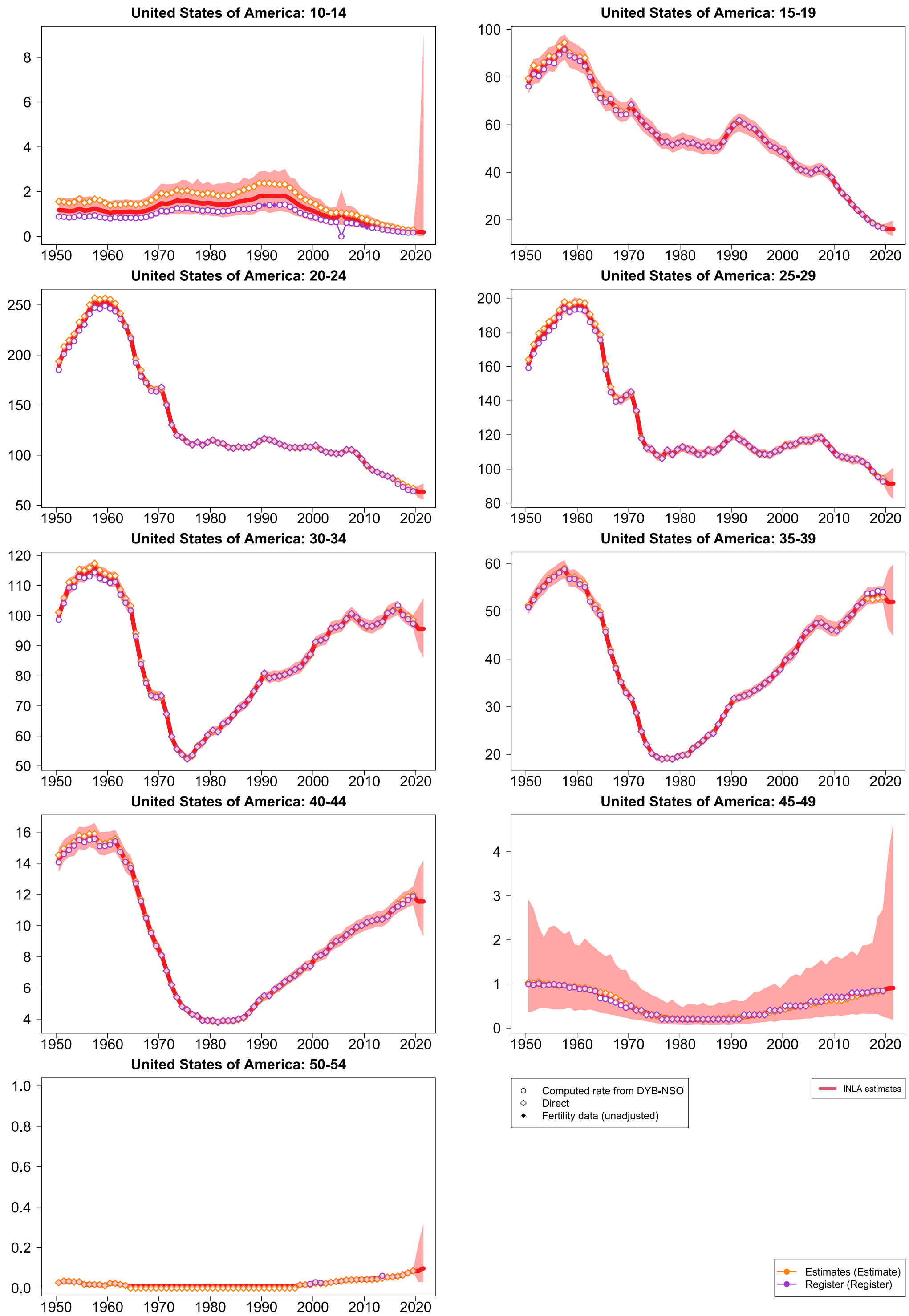
Figure 5. ASFR model estimates for five-year age groups from 10 to 54 years from selected countries with high-quality and high completeness of vital registration data

Examples of Czechia and United States of America



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

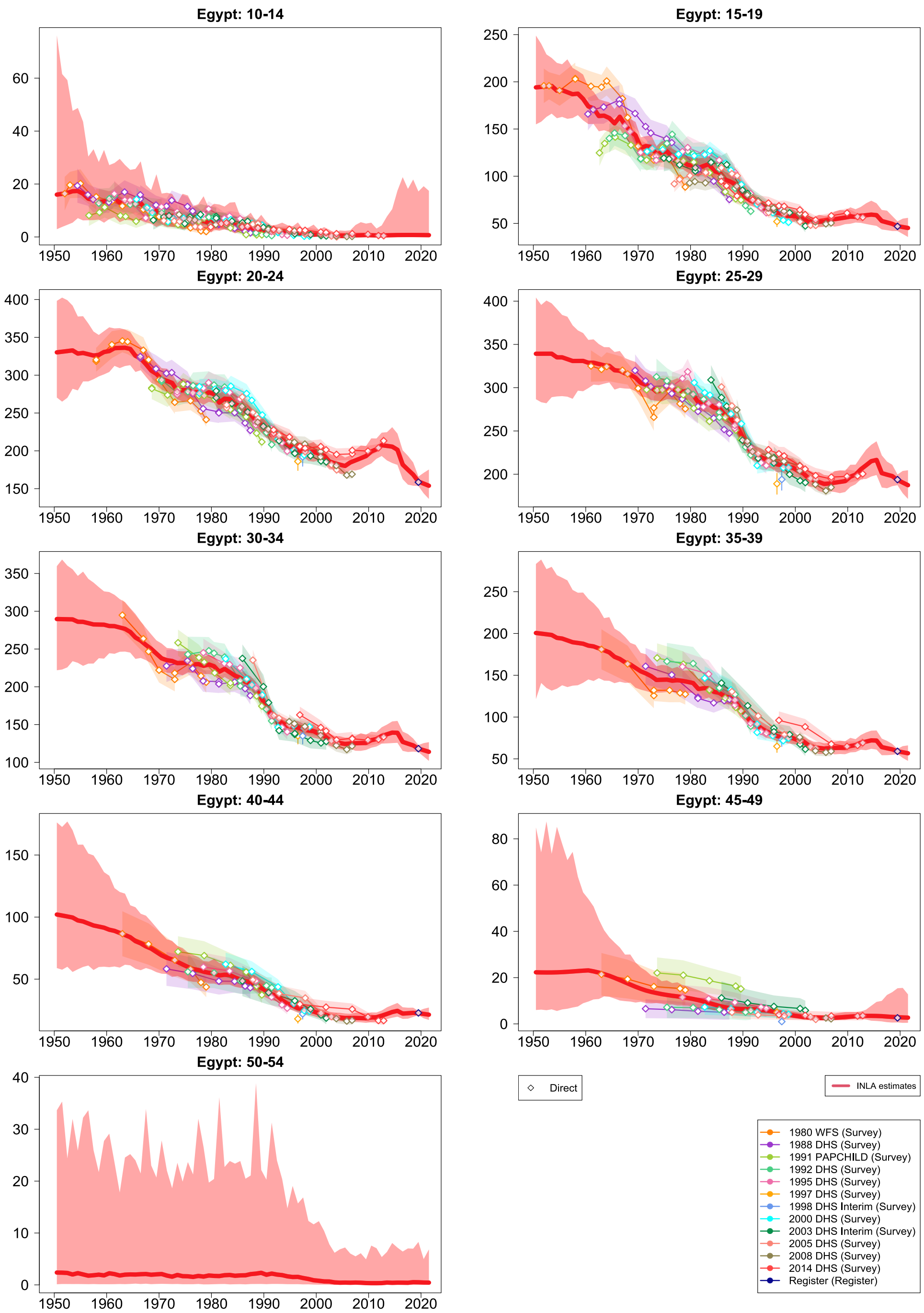
Figure 5. Continued



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

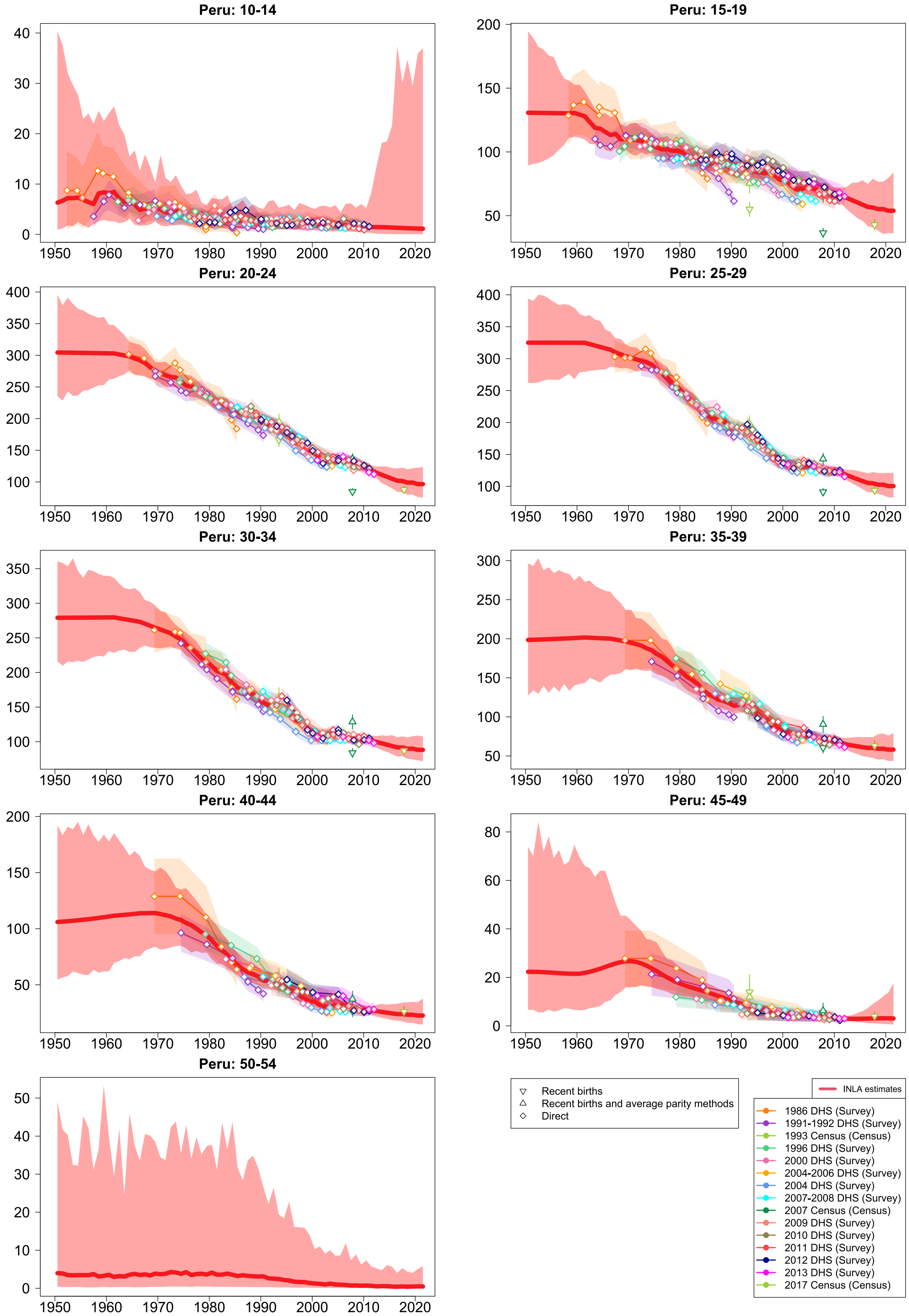
Figure 6. ASFR model estimates for five-year age groups from 10 to 54 from selected countries with data from surveys, censuses and reports and with reasonable data coverage

Examples of Egypt and Peru



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

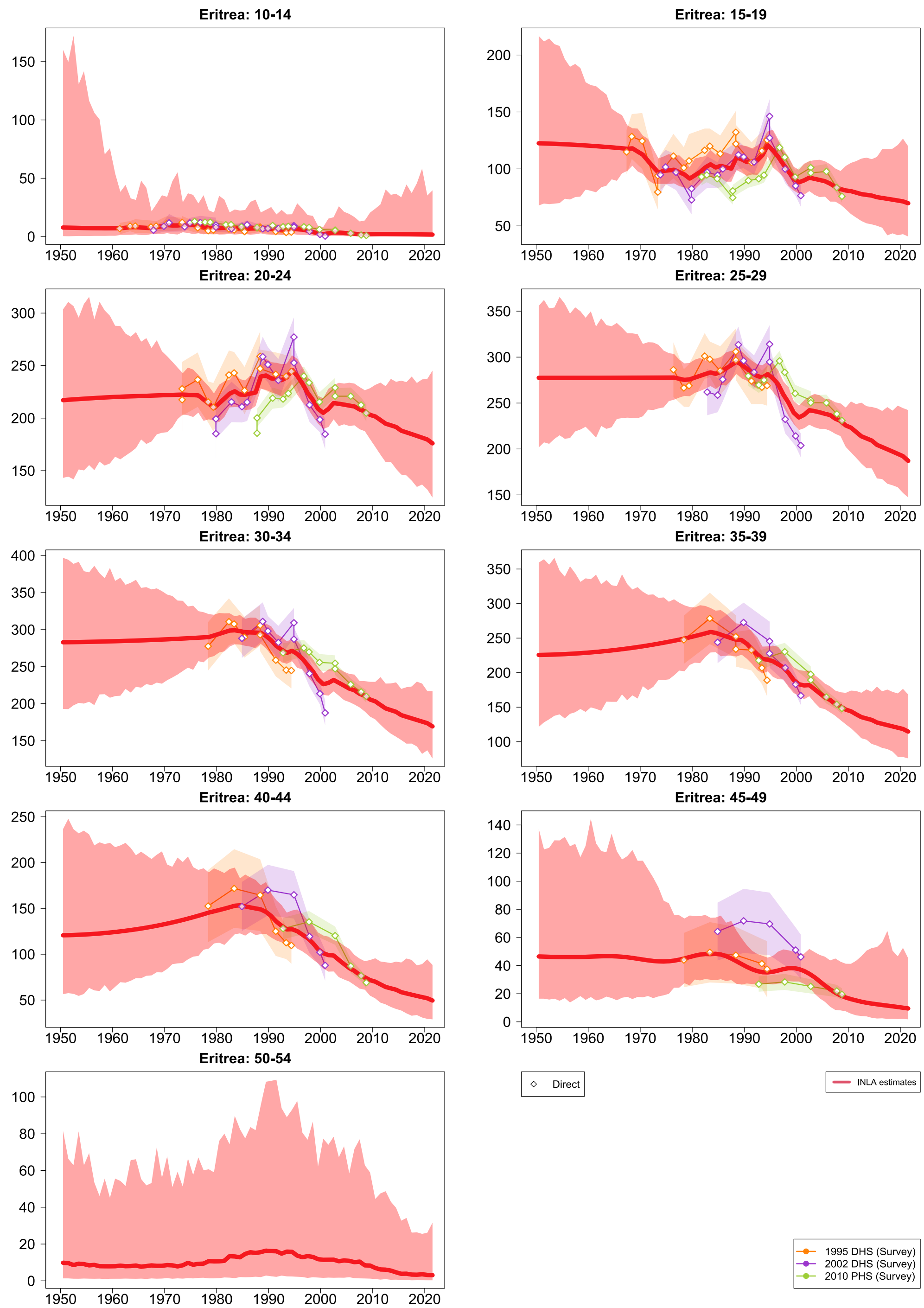
Figure 6. Continued



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

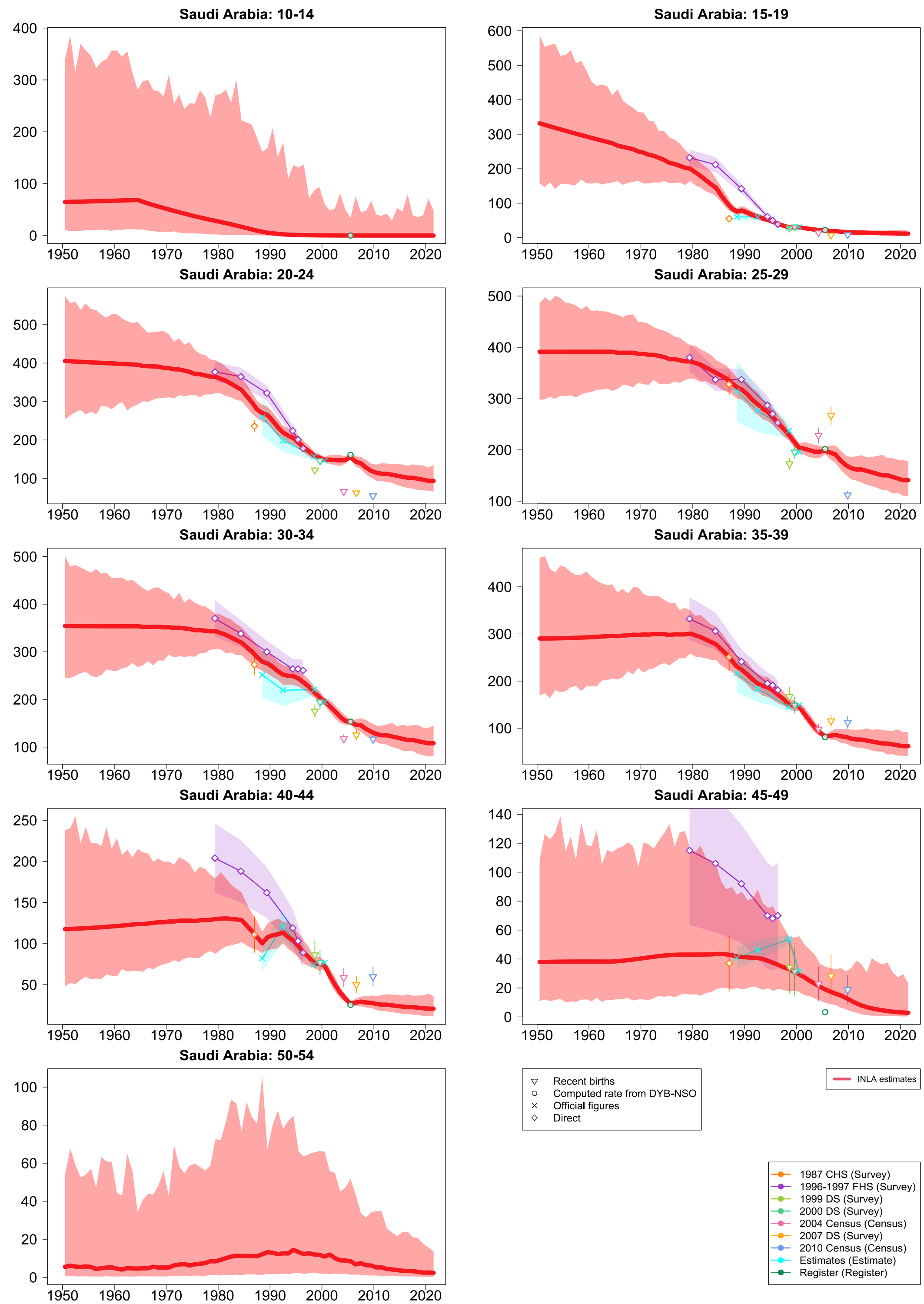
Figure 7. ASFR model estimates for five-year age groups from 10 to 54 years from selected countries with limited data

Examples of Eritrea and Saudi Arabia



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

Figure 7. Continued



Note: The red curves show the posterior medians. The shades show 95 per cent uncertainty bounds. Dots are the observations used for modelling. The shading and vertical lines around the dots represent sampling errors (if available).

V. VALIDATION

The model performance was validated using the validation approach of Alkema and others (2014). ASFR observations within each five-year age group from ages 10 to 54 collected after each age-specific year (referred to as "data source reference year," which is not the reference year of an observation but rather the reference year in which the observation was collected) were excluded from a model run for that particular five-year age group. The left-out observations comprise approximately 20 per cent of the total observations (owing to the varying number of observations collected in each survey reference year, the left-out observations can be slightly above or below 20 per cent). The left-out observations were the *testing dataset* and the remaining observations were the *training dataset*. For each age group, table 7 shows the years for data sources after which the ASFR observations are left out and the corresponding percentage and number of observations left out by age group.

TABLE 7. AGE-SPECIFIC SURVEY REFERENCE YEAR USED FOR TESTING DATASET, PERCENTAGE, AND NUMBER OF LEAVING LEFT-OUT ASFR OBSERVATIONS

Age group	*Cut-off survey year to leave out observations as testing data*	*Percentage of left-out observations*	*Number of left-out observations*
10-14	2013	15.4	3385
15-19	2012	17.7	3795
20-24	2011	20.1	3792
25-29	2011	20.1	3482
30-34	2011	20.1	3158
35-39	2010	22.4	3168
40-44	2010	22.4	2967
45-49	2009	24.1	2985
50-54	2007	28.1	561

For each left-out observation in the *testing dataset*, a posterior predictive distribution was generated based on BHM fittings for the *training* dataset. The error was computed as the difference between the left-out observation and the median of the posterior predictive distribution:

$$e_i = y_i - \widehat{Y}_i, \tag{41}$$

where y_i is the i-th left-out observation and $\widehat{Y}_i$ is the corresponding median of the posterior predictive distribution. The median errors of all the left-out observations and the median of the absolute errors are presented in table 8, together with the coverage of 90 per cent prediction intervals (PIs). [11] The lower and upper bounds of the 90 per cent PIs for each left-out observation are the 5th and 95th percentiles of the posterior predictive distribution, respectively.

[11] Prediction interval is the uncertainty interval based on the posterior predictive distribution that would have been constructed based on left-out observations.

To summarize the coverage, we computed the proportion of left-out observations that fell outside the 90 per cent PIs. Table 8 summarizes the results related to the proportion of left-out observations for the validation exercise. The median errors and median absolute errors were close to zero for all age groups, indicating that the differences were small. The coverage of 90 per cent prediction intervals was higher than expected for all age groups except for the age group 40–44 years. This means that the model predictions are conservative, with wider prediction intervals than expected.

TABLE 8. VALIDATION RESULTS FOR LEFT-OUT OBSERVATIONS BY AGE GROUP

Age group (observation year)	*Median error*	*Median absolute error*	*Proportion of left-out observations below 5th percentile of prediction interval*	*Proportion of left-out observations above 95th percentile prediction interval*
10-14 (2013-2021)	0.000	0.001	4.4%	1.4%
15-19 (2012-2021)	-0.001	0.013	6.0%	2.7%
20-24 (2011-2021)	-0.002	0.016	5.1%	0.9%
25-29 (2011-2021)	0.000	0.011	0.8%	1.1%
30-34 (2011-2021)	0.008	0.015	0.6%	3.1%
35-39 (2010--2021)	0.007	0.012	1.6%	6.5%
40-44 (2010--2021)	0.002	0.005	2.1%	9.0%
45-49 (2009-2021)	0.000	0.001	0.8%	3.1%
50-54 (2007-2021)	0.000	0.000	0.0%	1.4%

The model performance was validated by comparing the model estimates based on the full and training datasets. For each country-year, the error was computed as the difference between the median estimates based on the full dataset and those based on the training dataset. For all country-years, the median of these errors and absolute errors was calculated, and the percentage of country-years in which the median estimates based on the full dataset fell outside the credible bounds of the estimates based on the training dataset was calculated. Table 9 shows the results of the comparison between the estimates obtained based on the full dataset and those based on the training set for the year 2000. Median errors and the median absolute errors were close to zero. The proportion of updated estimates that fell outside the uncertainty intervals constructed based on the training set was generally within the expected range (less than 5 per cent).

The exceptions are for each five-year age group from 15 to 39 years, where the median prediction based on the full run tends to be higher than the 90 per cent credible interval (CI) [12] based on the full training database. This implies that the recent data show that the ASFR from age 15 to 39 declined faster than the historical experience based on the model assumptions.

[12] Credible interval is the uncertainty interval based on the posterior samples from the model fittings.

TABLE 9. SUMMARY OF DIFFERENCES IN ASFR ESTIMATES IN OBSERVATION YEAR 2000, BASED ON THE TRAINING AND FULL DATA SETS

Age group (observation year)	*Median error*	*Median absolute Error*	*Full run prediction below 5th percentile of prediction interval of validation run (%)*	*Full run prediction above 95th percentile prediction interval of validation run (%)*
10-14 (2000)	0.000	0.000	3.0	3.0
15-19 (2000)	0.000	0.002	2.1	8.0
20-24 (2000)	0.000	0.003	3.8	8.4
25-29 (2000)	0.000	0.002	5.5	8.4
30-34 (2000)	0.000	0.001	3.0	8.0
35-39 (2000)	0.000	0.001	1.7	8.4
40-44 (2000)	0.000	0.001	3.4	3.8
45-49 (2000)	0.000	0.000	3.8	3.4
50-54 (2000)	0.000	0.000	0.4	3.8

VI. Conclusions

The model presented in this report produces estimates of age-specific fertility rates for all countries and territories since 1950 in a manner that is reproducible, consistent, and comparable across countries and times. The model uses an extensive database of age-specific fertility estimates to produce annual estimates for all five-year age groups from 10 to 54 years for all countries and areas since 1950. Estimates in countries with limited or no data or low-quality or biased data are supported by borrowing information from other countries in the region and by using covariates in a reproducible fashion. The estimates are based on extensive databases from all available data sources. Estimates in countries with limited or no data or with data of low-quality or biased data are supported by borrowing information from other countries in the region and by using covariates in a reproducible fashion. All underlying data and estimates are publicly available through https://population.un.org/dataportal/home link to data portal for the most recent revision of *World Population Prospects*.

VII. References

Alkema, L., and others (2011). Probabilistic projections of the total fertility rate for all countries. *Demography*, vol. 48, No. 3, pp. 815-39. doi: 10.1007/s13524-011-0040-5.

Alkema L, and others (2014). National, regional, and global sex ratios of infant, child, and under-5 mortality and identification of countries with outlying ratios: A systematic assessment. *The Lancet Global Health*, vol. 2, No. 9, e521-30.

Andreev, K., and others (2013). Demographic Components of Future Population Growth. Technical Paper No. 2013/3, United Nations, New York.

Bijak, J. and J. Bryant (2016). Bayesian demography 250 years after Bayes. Population Studies, vol. 70, No. 1, pp. 1–19.

Cho, L.J., R.D. Retherford, and M.K. Choe (1986). The Own-Children Method of fertility estimation. Honolulu: The East-West Center.

Henry, L. (1961). La fécondité naturelle. Observation-théorie-résultats. *Population* (french edition), pp. 625-636.

Institut Centrafricain des Statistiques et des Études Économiques et Sociales (ICASEES) (2009). Enquête par Grappe à Indicateurs Multiples, couplée avec la Sérologie VIH et Anémie 2006 (MICS-3), Bangui, République Centrafricaine.

Institute for Health Metrics and Evaluation (2020). Global Burden of Disease Study 2019 (GBD 2019). University of Washington. Available from www.healthdata.org/gbd. Accessed 10 October 2020.

Liu, P., and A.E Raftery (2020). Accounting for uncertainty about past values in probabilistic projections of the total fertility rate for most countries. *The annals of applied statistics*, vol. 14 N0. 2, pp. 685.

Preston, S. (1984). The Use of Direct and Indirect Techniques for Estimating the Completeness of Death Registration Systems. In *Data bases for mortality measurements, edited by United Nations*, pp. 66-76. New York.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (www.R-project.org/).

Raftery, A., and others (2009). Probabilistic Projections of the Total Fertility Rate for All Countries for the 2010 World Population Prospects. Paper presented to United Nations Expert Group Meeting on Recent and Future Trends in Fertility, held at the United Nations Headquarters in New York from 2-4 December 2009 by the Population Division Department of Economic and Social Affairs: United Nations.

Rao, C., R. Mswia, M.Brastchi, and P. Setel (2020). *Estimating Completeness of Birth and Death Registration: Methods and Options for Estimating Completeness of Civil Registration*. Vital Strategies, NY.

Rutstein, S. (2018). Data Quality Evaluation of the Niger 2017 Demographic and Health Survey. The Demographic and Health Surveys Program (https://dhsprogram.com/pubs/pdf/OD73/OD73.pdf).

Simpson D, and others (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical science*, vol. 32, No. 1, pp. 1-28.

Susmann, H., M. Alexander, and L. Alkema (2022). Temporal Models for Demographic and Global Health Outcomes in Multiple Populations: Introducing a New Framework to Review and Standardise

Documentation of Model Assumptions and Facilitate Model Comparison. *International Statistical Review*, vol. 90, pp. 437– 467.

United Nations (2004). Handbook on the Collection of Fertility and Mortality Data. Statistical papers. Series F, Studies in methods. ST/ESA/STAT/SER.F/92. Statistics Division Department of Economic and Social Affairs. New York, NY, United Nations. Available from https://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf.

----------------- (2006). *World Population Prospects: The 2004 Revision, Volume III: Analytical Report.* ST/ESA/SER.A/246.

------------------ (2017a). *Principles and Recommendations for Population and Housing Censuses, Revision 3.* ST/ESA/STAT/SER.M/67/Rev.3. New York, NY, United Nations. Available from https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf.

------------------ (2017b). *World Population Prospects 2017*, Online Edition.

------------------ (2017c). *World Fertility Data 2017*. POP/DB/Fert/Rev2017.

------------------ (2019b). *World Population Prospects 2019, Methodology of the United Nations Population Estimates and Projections,* ST/ESA/SER.A/425.

------------------ (2019c). *World Population Prospects 2019, Online Edition. Rev. 1. (metadata).*

------------------ (2019d). *World Fertility Data 2019*. POP/DB/Fert/Rev2019.

------------------ (2019e). *World Marriage Data 2019*. POP/DB/Marr/Rev2019.

------------------ (2022a). *World Population Prospects 2022*, Online Edition.

------------------ (2022b). *World Population Prospects 2022: Methodology of the United Nations population estimates and projections.* UN DESA/POP/2022/TR/NO. 4.

United Nations Economic Commission for Europe (UNECE) (2018). Recommendations on Communicating Population Projections, Prepared by the Task Force on Population Projections, United Nations, New York and Geneva.

Wilson, C. (2011). Understanding Global Demographic Convergence since 1950. *Population and Development Review,* vol. 37, No. 2, pp. 375-388. Retrieved on May 13, 2020, from www.jstor.org/stable/23043287.