ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Levels and trends estimate of sex ratio at birth for seven provinces of Pakistan from 1980 to 2020 with scenario-based probabilistic projections of missing female birth to 2050: A Bayesian modeling approach

Fengqing Chao[1]*, Muhammad Asif Wazir[2], and Hernando Ombao[1]

[1]Statistics Program, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[2]Population and Development Advisor (Freelance), Islamabad (ICT), Pakistan

Abstract

Most evidence on son preference in Pakistan is reflected in the higher child mortality among females than males. The sex discrimination before birth is rarely reported in Pakistan. This is the first study to quantify prenatal sex discrimination in Pakistan on a subnational level. We provide annual estimates of the sex ratio at birth (SRB) from 1980 to 2020 and scenario-based projections of the number of missing female births up to 2050 by Pakistan province. The results are based on a comprehensive database consisting of 832,091 birth records from all available surveys and censuses. We adopted a Bayesian hierarchical time series model to synthesize different data sources. We identified Balochistan with an existing imbalanced SRB since 1980. For the rest provinces without past or ongoing SRB inflation, we projected the largest female birth deficit to occur in Punjab in 2033 under the scenario that the SRB transition process starts in 2021. We demonstrated important disparities in the occurrence and quantification of missing female births up to 2050.

Keywords: Bayesian hierarchical model; Pakistan; Scenario-based projection; Sex ratio at birth; Son preference; Sex-selective abortion; Subnational modeling; Time series models

***Corresponding author:**
Fengqing Chao
(fengqing.chao@kaust.edu.sa)

Citation: Chao, F., Wazir, M.A., & Ombao, H. (2022). Levels and trends estimate of sex ratio at birth for seven provinces of Pakistan from 1980 to 2020 with scenario-based probabilistic projections of missing female birth to 2050: A Bayesian modeling approach. *International Journal of Population Studies*, 8(2):51-70. https://doi.org/10.36922/ijps.v8i2.332

Received: August 27, 2022

Accepted: November 14, 2022

Published Online: December 14, 2022

1. Introduction

The ratio of the number of male live births to the number of female live birth, namely, the sex ratio at birth (SRB), is an essential element in estimating and projecting population size and its dynamics. Furthermore, an imbalanced SRB in a population reflects discrimination and disadvantages that females face in social, political, economic, and cultural contexts (Gupta *et al.*, 2003; Guilmoto, 2009). SRBs are distorted from their natural levels in several countries, primarily clustered in South Asia, East Asia, and East Europe. The imbalanced values reach as high as 1.2 (i.e., 120 male births per 100 female births) in certain regions (Attané & Guilmoto, 2007; Chao *et al.*, 2019a; Duthé *et al.*, 2012; Goodkind, 2011; Guilmoto *et al.*, 2009; Guilmoto, 2012; Guilmoto & Ren, 2011; Lin, 2009; Park & Cho, 1995). SRB is the prime indicator of prenatal sex discrimination and sex imbalance in human populations.

The objectives of this study are (i) to provide annual estimates of SRBs among seven provinces of Pakistan from 1980 to 2020, (ii) to provide scenario-based projections to 2050 using a reproducible Bayesian statistical model, and (iii) to identify provinces with SRB imbalance. Our study has several contributions as a result of achieving the research objectives. First, to the best of our knowledge, this is the first study on Pakistan SRB that has produced provincial estimates and projections during 1980 – 2050. Second, it is the 1st time Balochistan is identified with the existence and transition process of the sex ratio imbalance using a Bayesian model. Third, based on the SRB imbalance results of the Bayesian hierarchical time series model, we compute the number of missing female births over time in provinces with imbalanced SRB and quantify the female birth deficits in each Pakistan province. Our study included seven provinces of Pakistan: Balochistan, Khyber Pakhtunkhwa, Punjab, Sindh, Gilgit Baltistan, Islamabad (ICT), and Azad Jammu and Kashmir. The results for Federally Administered Tribal Areas are omitted because of the unavailability of the longer time series data on SRB.

The remainder of this paper is organized as follows. Section 1.1 provides the theoretical background of the study. Section 2.1 summarizes the database compiled for statistical modeling and Section 2.2 summarizes the Bayesian statistical model used for provincial SRB estimation and the post-modeling process (identifying provinces with imbalanced SRB and calculating the number of missing female births). Section 3 presents the SRB results by province, the provincial SRB imbalances, the corresponding missing female births, and the scenario-based missing female birth projections. Sections 4 and 5 summarize the primary contributions and limitations and conclude the study.

1.1. Theoretical background

Distortion in the SRB has been primarily attributed to three interlinked factors (Guilmoto, 2009; 2012): (1) Son preference, (2) technological advances in prenatal diagnosis, and (3) preferences for smaller family size and consequent fertility decline. In countries with a patrilineal culture and shrinking family size, when prenatal sex determination and abortion technology are available, couples practice sex-selective abortion to secure at least one son. The SRB in such populations is male biased. SRB imbalance has been reported in 12 countries/areas since 1970 (Chao *et al.*, 2019a).

Pakistan is a country that has a strong preference for sons (Atif *et al.*, 2016; Hussain *et al.*, 2000; Khan & Sirageldin, 1977; Sathar *et al.*, 2015). Preference for male births in Pakistan stems from lineage, economic and social conditions, caste, and identity. At least one son in a strongly patrilineal society is essential for living arrangements in old age. One study suggested that the ideal family size in Pakistan (four children) has remained constant since the 1970s; moreover, the ideal sex composition of the children is more than 1 son (Wazir & Shaheen, 2016). Son preference is evidenced by the excess mortality of female children over male children under five in Pakistan, indicating possible differential treatment between girls and boys in this age group (Alkema *et al.*, 2014; Sathar *et al.*, 2015). The education attainment gap between females and males is large in Pakistan. Between 2017 and 2018, 30% of young women (age 15 – 24) completed middle or higher education compared to 50% of young men (National Institute of Population Studies (NIPS) [Pakistan] & ICF, 2019). Between 2018 and 2019, 36% of girls (ages 5 – 16) were out of school versus 25% of boys (Pakistan Bureau of Statistics (PBS) 2019). However, little evidence of prenatal sex preference has been reported in Pakistan. The previous studies identified no imbalanced SRB at the national level (Zaidi & Morgan, 2016; Chao *et al.*, 2019a). Other studies suggested that, among couples in Pakistan, the desire for a large family might dominate preferences for children of a particular type (De Tray, 1984). A high prevalence of sex-selective abortion was identified in two rural districts in Balochistan province (Qayyum & Rehan, 2017). However, the results mentioned above are based on survey data with small sample sizes.

The national scale levels and trends in SRB can mask the disparity among subregions in a country. Even in countries such as China and India, with an overall strong preference for sons, the SRB is not imbalanced in every province or state (Chao & Yadav, 2019; Chao *et al.*, 2020; Ge *et al.*, 2020; Jiang & Zhang, 2021). In Pakistan, a subnational level assessment of SRB is essential because the demography, socioeconomic status, and cultures (i.e., caste and ethnicity) are considerably heterogeneous. The latest estimates from the Pakistan Demographic and Health Survey (DHS) 2017 – 2018 revealed a high heterogeneous SRB across provinces: High inflation at 1.16 in Balochistan, a roughly normal SRB in Punjab at 1.05, and a female bias in Sindh and Khyber Pakhtunkhwa (SRB is 0.91 and 0.95, respectively) (National Institute of Population Studies (NIPS) [Pakistan] & ICF, 2019). To the best of our knowledge, no study has provided the annual estimates of the provincial SRB in Pakistan using all available data since 1980. To accurately determine whether the SRB is imbalanced in Pakistan and if so, where the imbalance occurs, it is essential to estimate the SRB on the subnational level.

Estimating the SRB in Pakistan is challenging for two reasons. First, limited data are available on birth histories in the past. Without a fully developed vital registration system in Pakistan, administrative birth records are lacking, and vital events are mostly estimated based on household surveys. Only a few sample surveys provided birth histories over different periods since the 1990s. Second, the data quality of census counts is typically low because of age heaping (Feeney & Alam, 1998). In historical census data, the number of children ever born in Pakistan is either unavailable or is unreliably reported. For example, birth histories were not collected in the 1981 Pakistan census (Ali *et al.*, 2001). Accordingly, the individual-level data of the three most recent censuses in Pakistan (conducted in 1981, 1998, and 2017) contain only the populations of boys and girls under 1 year old. The SRB data from sample surveys such as Pakistan DHS are suffering from large uncertainties because of the small sample sizes and misreporting of female births.

When estimating and projecting the provincial SRB in Pakistan, it is crucial to assess the levels and trends in the SRB by a reproducible statistical model. Using a Bayesian modeling approach for estimation and projection, observations from different data sources with varying levels of uncertainties can be synthesized and pooled together in a systematic and reproducible fashion. The Bayesian method can take into account both provincial SRB observations and external information on the SRB imbalance process to assist in model estimation and projection.

2. Data and methods

2.1. Data sources

Table 1 summarizes our database of provincial SRBs in Pakistan, with 531 SRB observations available in eight provinces of Pakistan. The reference years of these observations range from 1965 to 2019. The database contains 832,091 birth records by summing up the number of birth records where available. The number of birth records is unknown in some data series. Hence, the actual number of birth records involved in the study is more than what we reported here. The SRB observations were generated from the individual birth records in data sources with full birth histories (appendix for details of the data processing steps). The database is available as Supplementary File 1 (https://doi.org/10.6084/m9.figshare.21548082).

The DHS (ICF International, 2022) and Multiple Indicator Cluster Survey (MICS) provide the birth histories (either the full birth histories or the birth histories during the past 24 months before the survey interview) of women interviewed in retrospective survey questionnaires. Birth records are excluded if they were born more than 20 years prior to the year in which surveys were conducted to minimize recall bias from older women. Furthermore, the Pakistan Social and Living Standards Measurement Survey (PSLM) is a provincial-level survey with high coverage of households in Pakistan (Pakistan Bureau of Statistics [PBS], 2019). The PSLM records births over the 12 months before the date of the survey interview. The census is conducted once per decade and collects births in the 12 months preceding the census (Minnesota Population Center, 2019).

Given Pakistan's lack of reliable administrative birth data, it is essential to include all available data from different surveys to produce more reliable estimates and projections. The practice of making use of data from multiple data sources in estimation and projection has been widely used by international agencies, including the UNICEF, UN Population Division, and the Global Burden of Disease, and researchers in global and public health to reduce systematic bias from a single data source, to increase the length of the period that is covered by data (Alkema *et al.*, 2016; Bearak *et al.*, 2018; Gerland *et al.*, 2014; Liu & Raftery, 2020; Masquelier *et al.*, 2018; Wang *et al.*, 2020; You *et al.*, 2015). The data sources, we used as listed in Table 1, are based on provincial representative samples. If any future in-depth survey-specific consistency checks provide concrete evidence of bias in the examples of certain data sources, that particular data source should not be included

Table 1. Pakistan provincial SRB database

Survey name	Sample design	Survey year	# SRB observations	# Births records
Census	Census enumeration	1973, 1981, 1998	15	424,739
DHS	Two-stage stratified sample design	1990 – 1991, 2006 – 2007, 2012 – 2013, 2017 – 2018, 2019	301	253,580
MICS	Two-stage stratified sample design	2010, 2011, 2014, 2003 – 2004, 2007 – 2008, 2016 – 2017, 2017 – 2018	37	153,772†
PSLM	Two-stage stratified sample design	1995 – 2016, 2005 – 2006, 2007 – 2008, 2013 – 2014, 2018 – 2019	228	–
Total			531	832,091*

Note: DHS: Demographic and Health Survey; MICS: Multiple Indicator Cluster Survey; PSLM: Pakistan Social and Living Standards Measurement Survey. †: Number of birth records available only in MICS 2017 – 2018; –: The number of birth records is unavailable; *: The total number of births obtained by summing the available number of birth records in the 20 years before the survey conducted.

in the modeling. As of now, no evidence of biased sampling has been reported for any of these data sources. With their own objectives of the survey, these data sources provide a wealth of information, and there is no supporting evidence to choose one at the cost of the others. Hence, we use all these sources in our model estimation and projection.

Figure 1 illustrates the large uncertainty in SRB observations on national and provincial levels. It shows the observed SRB from Pakistan 2012 to 2013 and 2017 to 2018 DHSs on the national and provincial levels. In the data series Pakistan DHS 2012 – 2013, the national SRB observations range from 0.935 in the year 1989 to 1.195 in the year 2003. The average sampling error associated with these observations is 0.073, equivalent to 7% of the coefficient of variation (CV; refers to the ratio of sampling error to the value of observations, indicating the extent of variability given the observation value (Brown, 1998). On the provincial level, the SRB observations have a wider range with greater uncertainties because of smaller sample sizes. The smallest provincial SRB observation for DHS 2012 – 2013 is 0.917 in Islamabad (ICT) in year 2002 and the biggest value is 1.416 in Balochistan in year 1989. The average sampling error of all the provincial observations for DHS 2012 – 2013 is 0.087, corresponding to 8% of CV. Similarly, for Pakistan DHS 2017 – 2018, the observed SRB on national and provincial levels ranges from 0.914 to 1.178 and from 0.891 to 1.272, respectively. The average sampling errors are 0.087 and 0.139 for national and provincial data, respectively.

2.2. Methods

The model performance and predictive power were assessed by an out-of-sample validation exercise (leaving out recent

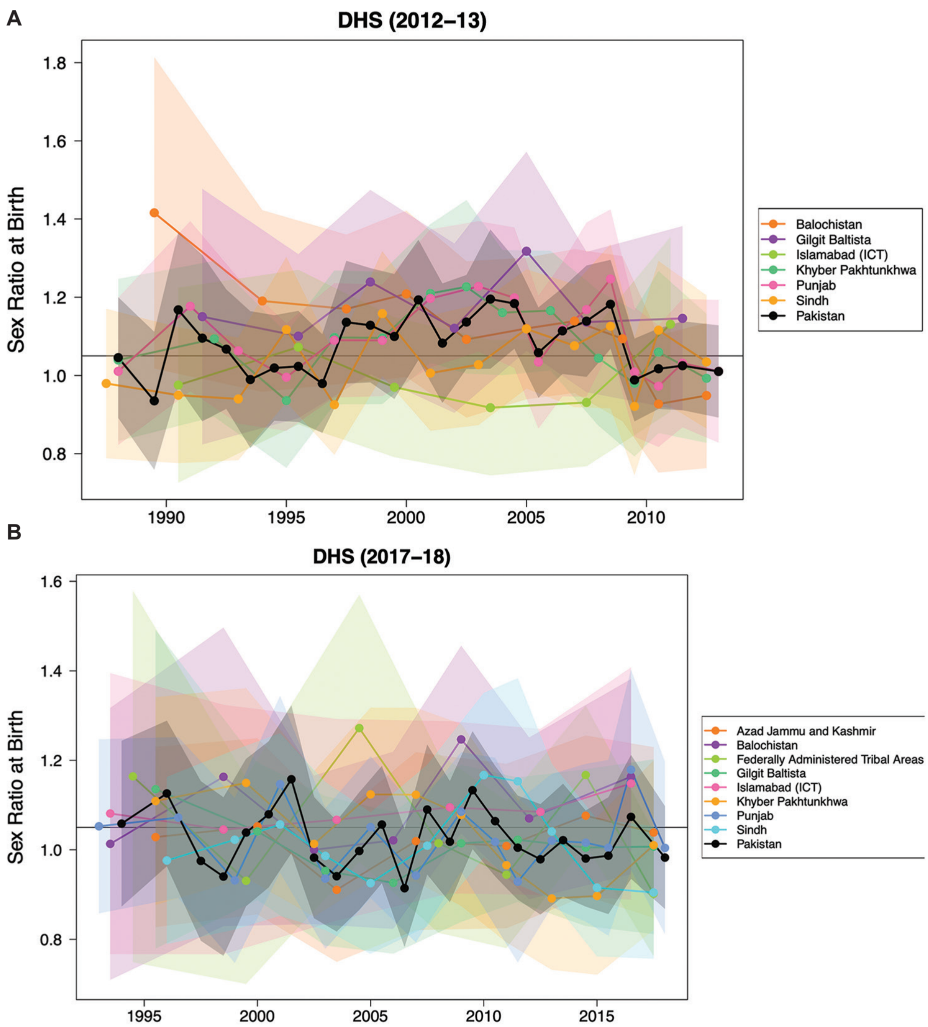

Figure 1. SRB data in (A) 2012 – 2013 and (B) 2017 – 2018 DHS on national and provincial levels

Note: The SRB observations (dots) from the same area (province or the whole of Pakistan) are connected with line segments of the same color. Shaded areas around the observation series represent the sampling errors in the series (quantified by twice the sampling standard errors).

observations) and simulation exercises (appendix for details). The validation and simulation results suggest good calibration and predictive power of the model.

The remainder of this section overviews the SRB Bayesian model.

2.2.1. Bayesian model for provincial SRB estimation and projection

The model for the SRB in Pakistan province is based on the model described previously by Chao *et al.* (2021a, b). In this study, we made a few modifications to the model to better address the data quality and availability of provincial SRBs in Pakistan. Subnational SRB models have been applied to other culturally and demographically heterogeneous countries with son preference such as Nepal (Chao, *et al.*, 2022) and Vietnam (Chao *et al.*, 2021c).

The outcome of interest $\Theta_{p,t}$, the SRB in Pakistan province p in year t is modeled as follows:

$$\Theta_{p,t} = b\Phi_{p,t} + \delta_p \alpha_{p,t},$$

b = 1.056 is the SRB baseline level for the entire Pakistan. The Pakistan SRB baseline b is estimated based on national SRB observations in Pakistan before the reference year 1970 (Chao *et al.*, 2019a, b). $p \in \{1,\ldots,k\}$ is the province index where $k = 7$. $t \in \{0,\ldots,h\}$ is the time index where $t = 0$ refers to the year 1980 and $t = h$ refers to the year 2050.

$\Phi_{p,t}$ follows an AR (1) time series model on the log scale, which captures the natural fluctuations of SRB in each province over time. The values of ρ and σ_ϵ ($\rho = 0.9$) and $\sigma_\epsilon = 0.004$) were not estimated but were borrowed from a previous study (Chao *et al.*, 2019a,b), which robustly estimated the parameters from an extensive national SRB database. We assume that log ($\Phi_{p,t}$) is causal and weakly stationary AR (1) process with parameters ρ and σ_ϵ, and hence, log ($\Phi_{p,t-s}$) is uncorrelated with $\epsilon_{p,t}$ for all s>0. The unconditional variance at $t = 0$ is expressed as $\sigma_\epsilon^2/\left(1-\rho^2\right)$ for log ($\Phi_{p,t}$). $\Phi_{p,t}$ is modeled as follows:

$$\log\left(\Phi_{p,t}\right) \sim \mathcal{N}\left(0, \sigma_\epsilon^2/\left(1-\rho^2\right)\right), if\ t = 0,$$

$$\log\left(\Phi_{p,t}\right) = \rho \log\left(\Phi_{p,t-1}\right) + \epsilon_{p,t}, if\ t \in \{1,\cdots,h\},$$

$$\epsilon_{p,t} \sim_{i.i.d.} \mathcal{N}\left(0, \sigma_\epsilon^2\right).$$

δ_p is the binary identifier of the sex ratio transition at the provincial level with only two possible values 1 and 0. $\delta_p = 1$ indicates an SRB imbalance in province p, whereas $\delta_p = 0$ indicates no imbalance in province p. The provincial sex ratio transition identifier parameter δ_p is meant to detect whether the transition process exists based on the levels and trends in the SRB observations. To ensure that the probability parameter π_p lies in the interval [0, 1], we use the logit transformed π_p follows a hierarchical normal distribution with a global mean and variance μ_π and σ_π^2, respectively. δ_p follows a Bernoulli distribution:

$$\delta_p \mid \pi_p \sim \mathfrak{B}(\pi_p), for\ \ p \in \{1,\cdots,k\},$$

$$logit\ (\pi_p) \mid \mu_\pi, \sigma_\pi \sim \mathcal{N}\left(\mu_\pi, \sigma_\pi^2\right), for\ p \in \{1,\cdots,k\}.$$

Vague priors are assigned to the parameters related to the indicator that detects sex ratio transitions:

$$inverse-logit\left(\mu_\pi\right) \sim \mathcal{U}\left(0,1\right),$$

$$\sigma_\pi \sim \mathcal{U}\left(0,2\right).$$

$\alpha_{p,t}$ refers to the province-specific SRB imbalance process. The process is assumed to be non-negative and is modeled by a trapezoidal function representing the three consecutive stages (increase, stagnation, and decrease) of the sex ratio transition. The trapezoidal function specification of $\alpha_{p,t}$ is motivated by the patterns of national-level SRB observations in countries with strong statistical son preference to capture the three-stage transition process (Chao *et al.*, 2021a). The trapezoidal functional form for $\alpha_{p,t}$ can capture the shape of the observed SRB transition process according to Chao *et al.* (2021c) for those countries. $\alpha_{p,t}$ is modeled as:

$$\alpha_{p,t} = \left(\xi_p / \lambda_{1p}\right)\left(t - t_{0p}\right), if\ t_{0p} < t < t_{1p}$$

$$\alpha_{p,t} = \xi_p, if\ t_{1p} < t < t_{2p}$$

$$\alpha_{p,t} = \xi_p - \left(\xi_p / \lambda_{3p}\right)\left(t - t_{2p}\right), if\ t_{2p} < t < t_{3p}$$

$$\alpha_{p,t} = 0, if\ t < t_{0p}\ or\ t > t_{3p}$$

Where,

$$t_{1p} = t_{0p} + \lambda_{1p},$$

$$t_{2p} = t_{1p} + \lambda_{2p},$$

$$t_{3p} = t_{2p} + \lambda_{3p}.$$

The start year of the SRB inflation t_{0p} is modeled by a continuous uniform prior distribution with a lower bound at the year 1980 and an upper bound at the year 2050, respectively. For $p \in \{1,\ldots,k\}$, we have:

$$t_{0p} \sim \mathcal{U}\left(0,h\right).$$

The province-specific period lengths of the three stages of the SRB inflation (λ_{1p}, and λ_{3p}) are assigned with informative priors. The means of prior distributions are taken from a systematic study (Chao *et al.*, 2021a) which modeled the sex ratio transition of multiple countries, including Pakistan. The standard deviations of prior distribution are set such that the CV is 0.1. The informative priors assist the provincial-level modeling of the sex ratio transition in Pakistan by exploiting the corresponding information at the national level. For $p \in \{1,\ldots,k\}$, we have:

$$\xi_p \sim \mathcal{N}_{(0,\infty)}\left(0.06, 0.006^2\right),$$

$$\lambda_{1p} \sim \mathcal{N}_{(0,\infty)}\left(11.0, 1.1^2\right),$$

$$\lambda_{2p} \sim \mathcal{N}_{(0,\infty)}\left(7.6, 0.8^2\right),$$

$$\lambda_{3p} \sim \mathcal{N}_{(0,\infty)}\left(16.1, 1.6^2\right).$$

2.2.2. Data quality model

r_i is the i^{th} observed SRB in province p[i] in year t[i], where i indexes all SRB observations across the provinces over time. r_i is assumed to follow a normal distribution on the log scale with mean of $\log\left(\Theta_{p[i],t[i]}\right)$ and variance of σ_i^2:

$$\log\left(r_i\right)|\Theta_{p[i],t[i]} \sim \mathcal{N}\left(\log\left(\Theta_{p[i],t[i]}\right), \sigma_i^2 + \omega^2\right),$$
$$for\, i \in \{1,\cdots,n\},$$

Where, n = 531 is the total number of observations. σ_i^2 is the sampling error variance of log (r_i), which reflects the uncertainty in log-scaled SRB observations because of the survey sampling design. σ_i^2 is calculated using a jackknife method (Appendix A.1). ω^2 is the non-sampling error variance representing the uncertainty contributed by non-responses, recall errors, and data input errors. We assume that ω^2 is immeasurable and is estimated using the model by assigning a vague prior:

$$\omega \sim \mathcal{U}\left(0, 0.5\right).$$

2.2.3. Posterior distribution

Likelihood

For the i^{th} observation r_i, let $v_i = \log(r_i)$ and $V_{p,t} = \log\left(\Theta_{p,t}\right)$.

The likelihood on log-scale up to proportion is:

$$p\left(v_i|V_{p[i],t[i]},\omega\right) \propto \frac{1}{\sqrt{\sigma_i^2+\omega^2}} \exp\left\{-\frac{\left(v_i - V_{p[i],t[i]}\right)^2}{2\left(\sigma_i^2+\omega^2\right)}\right\}.$$

For $\in\{1,\ldots n\}$, the likelihood can be written as:

$$p\left(v_i|\Phi_{p[i],t[i]},\omega,\delta_{p[i]},\alpha_{p[i],t[i]}\right) \propto \frac{1}{\sqrt{\sigma_i^2+\omega^2}} \exp\left\{-\frac{\left(v_i - b\Phi_{p[i],t[i]} - \delta_{p[i]}\alpha_{p[i],t[i]}\right)^2}{2\left(\sigma_i^2+\omega^2\right)}\right\}.$$

Posterior Density

The posterior density for $V_{p,t}$, the true SRB on the log scale for province p at time t, up to proportion is:

$$p\left(\begin{matrix}V_{1:k,0:h},\omega,\Phi_{1:k,0:h},\delta_{1:k},t_{0,1:k},\xi_{1:k},\lambda_{1,1:k},\lambda_{2,1:k},\lambda_{3,1:k}\\ ,\rho,\sigma_\epsilon,\mu_\pi,\sigma_\pi|v_{1:n}\end{matrix}\right) \propto$$

$$\frac{\left(1-\rho^2\right)^{k(h-1)/2}\prod_{\delta_p\in\{0,1\}}\pi_p^{\sum_{p=1}^{k}\delta_p}\left(1-\pi_p\right)^{k-\sum_{p=1}^{k}\delta_p}}{\sigma_\delta^{k(h-1)}\sigma_\pi^k\left[\prod_{t=1}^{h}\left(1-\rho^{2t}\right)^{k/2}\right]\prod_{i=1}^{n}\left(\sigma_i^2+\omega^2\right)^{1/2}}\times$$

$$\exp\left\{\sum_{p=1}^{k}\frac{2\mu_\pi\pi_p-\pi_p^2}{2\sigma_\pi^2}-\sum_{i=1}^{n}\frac{\left(v_i-b\Phi_{p[i],t[i]}-\delta_{p[i]}\alpha_{p[i],t[i]}\right)^2}{2\left(\sigma_i^2+\omega^2\right)}\right\}\times$$

$$\exp\left\{-\sum_{p=1}^{k}\sum_{t=1}^{h}\frac{\left(1-\rho^2\right)\left(\Phi_{p,t}-\rho^t\Phi_{p,0}\right)^2}{2\sigma_\epsilon^2\left(1-\rho^{2t}\right)}-\sum_{p=1}^{k}\frac{\sigma_\epsilon^2\Phi_{p,0}^2}{2\left(1-\rho^2\right)}\right\}.$$

2.2.4. Statistical computing and Bayesian Inference

We obtained posterior samples of all the model parameters and hyperparameters using a Markov chain Monte Carlo (MCMC) algorithm, implemented in the open source software R 4.2.1 (R Core Team, 2022) and JAGS 4.3.0 (Plummer, 2003), using R-packages R2jags (Su & Yajima, 2015) and rjags (Plummer, 2018). Results were obtained from 10 chains with a total of 5000 iterations in each chain, while the first 1000 iterations were discarded as burn-in. After discarding burn-in iterations and proper thinning, the final posterior sample size for each parameter by combining all chains is 25,000. The convergence of the MCMC algorithm and the sufficiency of the number of samples obtained were checked through visual inspection of trace plots and convergence diagnostics of Gelman & Rubin (1992), implemented in the coda R-package (Plummer *et al.*, 2006).

2.2.5. Post-modeling process

2.2.5.1. Identifying provinces of Pakistan with SRB imbalance

SRB imbalance in a Pakistan province is detected if δ_p = 1 for more than 95% of the posterior samples (indicating SRB inflation).

2.2.5.2. Simulating SRB imbalance after 2020

In provinces of Pakistan without past/ongoing SRB inflation (assessed in the model), we simulate the SRB imbalance process after 2020 for different starting years of SRB inflation.

The simulated province-specific SRB imbalance process $\delta_p\alpha_{p,t}$ is based on posterior samples in the model. The simulated $\delta_p\alpha_{p,t}$ is added to the projected $\Theta_{p,t}$ for different starting years of SRB inflation in each province. The simulation process is detailed in the appendix.

Figure 2 shows the simulated SRB imbalance process $\delta_p\alpha_{p,t}$ in a Pakistan province, with a given start year of the inflation process t_0. The SRB inflation process spans 38 years. After approximately one decade, the imbalance reaches its maximum level and remains around that level for approximately 7 years. The SRB imbalance then deflated toward the normal/reference level of SRB (i.e., the SRB inflation becomes zero) over the next 15 years.

2.2.5.3. Computing the number of missing female births

Let $\psi_{p,t}$ and $\Psi_{p,t}^{inflation-free}$ denote the estimated and expected inflation-free numbers of female live births respectively, in province p in year t. The estimated and expected numbers of female birth are calculated as $\psi_{p,t} = B_{p,t}/(1+\Theta_{p,t})$ and $\Psi_{p,t}^{(inflation-free)} = \left(B_{p,t} - \Psi_{p,t}\right) / \Theta_{p,t}^{(inflation-free)}$, respectively.

The total number of births $B_{p,t}$ in a given province and year is obtained from Wazir and Goujon (2019). The number of inflation-free female births $\Psi_{\mathrm{p,t}}^{(inflation-free)}$ is obtained from the estimated number of male births $(B_{p,t} - \psi_{p,t})$ and the inflation-free SRB $\Theta_{p,t}^{(inflation-free)} = b\Phi_{p,t}$ in the respective given province year. The number of missing female births is calculated using a method introduced in Dréze and Sen (1990), which was reviewed and validated in Guilmoto *et al.* (2020).

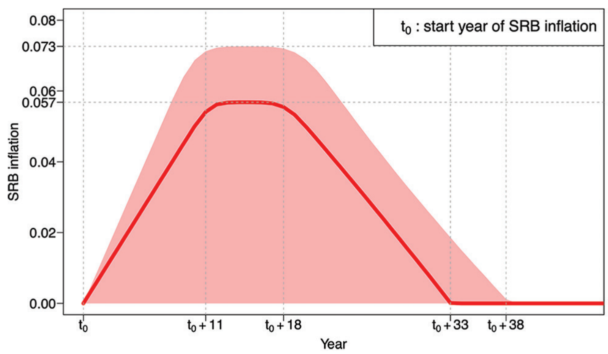

Figure 2. SRB inflation simulation of a Pakistan province
Note: Medians are in red curve. About 95% credible intervals are in red shades. t0 is the start year of the SRB inflation process.

The annual number of missing female births (AMFBs) in province p in year t is defined as: $\Psi_{p,t}^{(missing)} = \Psi_{p,t}^{(inflation-free)} - \Psi_{p,t}$. The cumulative number of missing female births (CMFB) from t_1 to t_2 in province p is obtained by adding the AMFB from year t_1 to year t_2: $\Lambda_{p,[t_1,t_2]}^{(missing)} = \sum_{t=t_1}^{t_2} \Psi_{p,t}^{(missing)}$.

3. Results

3.1. Levels, trends, and geographic disparities in provincial SRB estimates

Figure 3 shows an overview of the levels and trend in provincial SRBs in Pakistan from 1980 to 2020. The median estimates fluctuate around the national SRB reference level (1.056) except in Balochistan and Gilgit Baltistan. In Gilgit Baltistan, the SRB gradually increased from 1.058 (95% credible interval [1.041; 1.102]) in 1980 to 1.070 [1.047; 1.125] in 2016. After reaching its provincial maximum, the SRB continuously declines. The SRB in Gilgit Baltistan is not statistically significantly different from the national baseline level because the 95% credible intervals overlap with the national baseline throughout the whole period. However, the SRB in Balochistan is an outlier from the SRBs in all other provinces (Section 2.2.2 for details). The SRB estimates from 1980 to 2020, including uncertainty for the seven Pakistan provinces, are presented in Supplementary File 2 (https://doi.org/10.6084/m9.figshare.21548103).

Figure 4 illustrates the disparity in SRB across geographic locations. The SRB is most inflated in the southwest and northeast regions, including the Balochistan and Gilgit Baltistan provinces.

3.2. SRB imbalance at the provincial level

Table 2 lists the modeled SRB inflation probability in each Pakistan province. The probability of having a past or

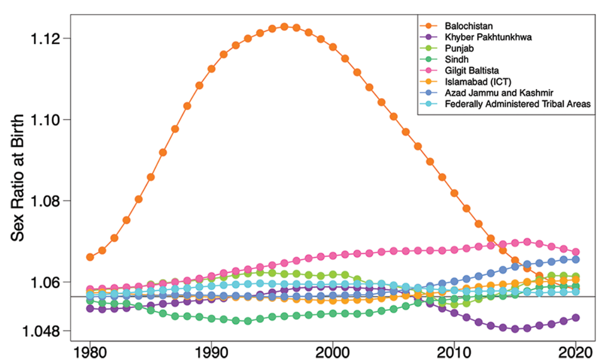

Figure 3. Median SRB estimates in Pakistan provinces over the 1980 – 2020 period

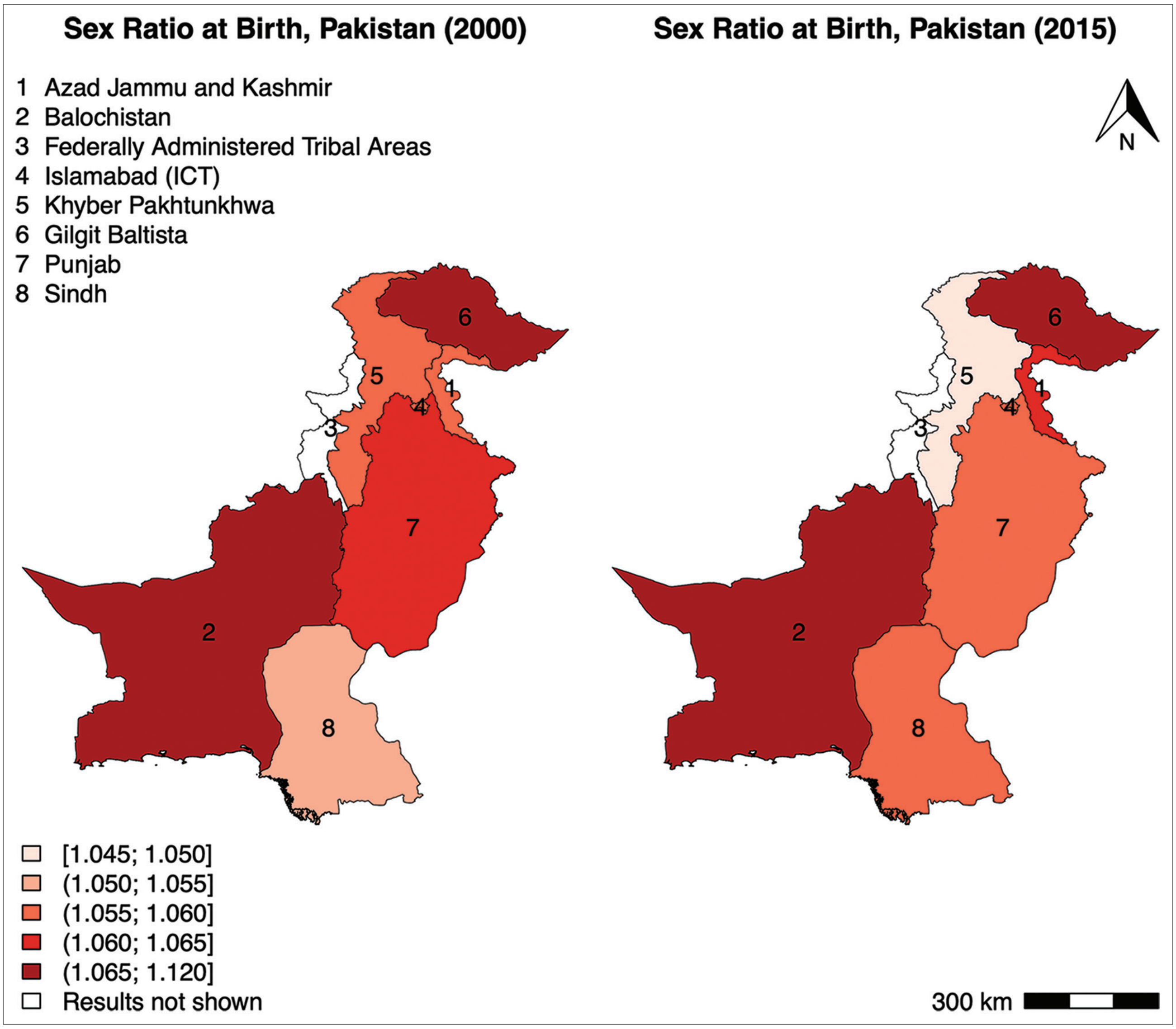

Figure 4. Geographic disparities of SRB estimates, 2000 and 2015
Note: Median estimates are presented. Results for Federally Administered Tribal Areas are omitted.

Table 2. SRB inflation probability by Pakistan province

Pakistan province	SRB inflation probability
Balochistan	100%
Gilgit Baltistan	65.4%
Azad Jammu and Kashmir	39.4%
Islamabad (ICT)	25.5%
Punjab	17.2%
Sindh	13.7%
Khyber Pakhtunkhwa	5.6%

Note: Provinces are listed in descending order of the inflation probability.
SRB: Sex ratio at birth.

ongoing SRB inflation is the highest in Balochistan at 100%. As the probability in Balochistan is the only probability above the cutoff value (95%), we identify Balochistan as the only province in Pakistan with SRB imbalance that happened before 2020.

Figure 5 illustrates the SRB model results in Balochistan. The SRB imbalance process in this province started in 1980 and ended around 2015. The maximum SRB in this province is estimated to occur in 1996, with SRB median estimate at 1.123 and the 95% credible interval at [1.100; 1.142]. The SRB in Balochistan is significantly above the national baseline (1.056) from 1986 to 2010. Over this

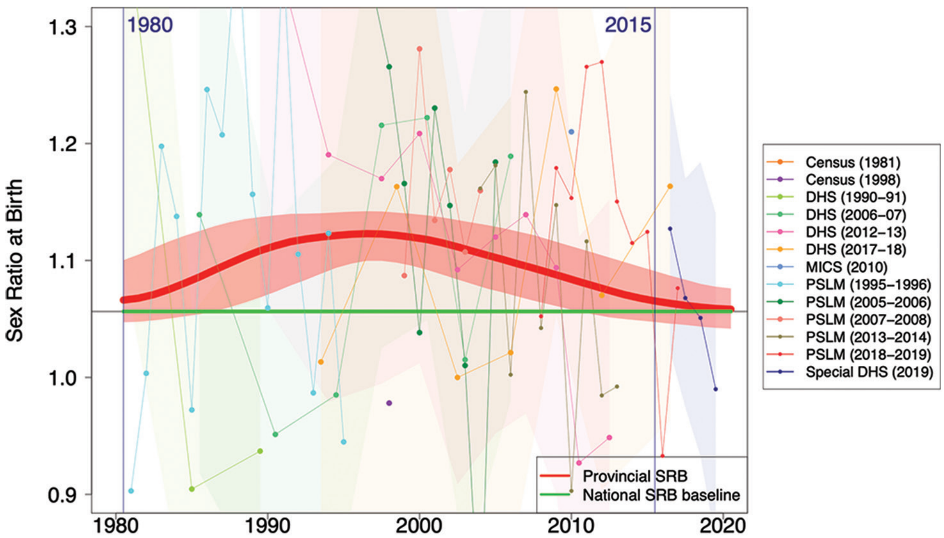

Figure 5. SRB estimates in Balochistan, 1980 – 2020

Note: Red line and shaded areas are the medians and 95% credible intervals of the province-specific SRB, respectively. The green horizontal line is the SRB baseline for the whole of Pakistan at 1.056 (Chao *et al.*, 2019a). The SRB observations (dots) from the same source are connected with line segments of the same color. Shaded areas around the observation series represent the sampling errors in the series (quantified by twice the sampling standard errors). Blue vertical lines denote the start and end years of the sex ratio transition.

Table 3. Missing female births model results in Balochistan

Values are in thousands	Time period				
	1980 – 1990	1991 – 2000	2001 – 2010	2011 – 2020	1980 – 2020
Estimated female births	124.2	124.2	153.2	174.9	143.6
	[122.3; 125.8]	[123.3; 125.5]	[151.5; 155.1]	[173.1; 176.2]	[142.9; 144.4]
Expected female births	127.8	131.7	159.2	176.4	148.3
	[126.3; 129.6]	[130.5; 132.7]	[157.4; 160.9]	[175.2; 178.2]	[147.6; 149.0]
AMFB	3.7	7.6	6.0	1.8	4.7
	[0.7; 7.4]	[5.0; 9.4]	[2.3; 9.4]	[0.1; 5.1]	[3.4; 6.1]
CMFB	40.2	76.1	60.1	18.2	194.6
	[7.4; 81.2]	[50.0; 93.9]	[23.5; 94.1]	[0.5; 51.0]	[140.0; 248.9]

Note: All values are in thousands. Numbers in front of the brackets are the posterior medians, and those inside brackets are the 95% credible intervals. The estimated female births, expected female births, and AMFBs are averaged over each period. The CMFBs are the cumulative values over each period. AMFB: Annual number of missing female births. CMFB: Cumulative number of missing female births.

period, the lower bound of the 95% credible intervals of SRB exceeded the national baseline.

3.3. Missing female births before 2020 in Balochistan

Table 3 lists the missing female births in Balochistan over different time periods. From the estimated and expected female births, we demonstrate the baseline magnitude of female births in Balochistan. The estimated average AMFB over the four decades from 1980 to 2020 is 4.7 [3.4; 6.1] thousand. On a decade-by-decade basis, the average AMFB increased from 3.7 [0.7; 7.4] in 1980 – 1990 to 7.6 [5.0; 9.4] in 1991 – 2000, then gradually declined from 2000, reaching 1.8 [0.1; 5.1] in the 2011 – 2020 period. Consequently, as the cumulative result of the AMFB, the CMFB between 1980 and 2020 is estimated at 194.6 [140.0; 248.9] in Balochistan province. Our results are consistent with previous results (Qayyum & Rehan, 2017) of sex-selective abortion in Balochistan. Although the data of that study were collected only from the rural areas (and hence may not be provincially representative), Balochistan had the highest rate of sex-selective abortion during the 2011 – 2014 period.

3.4. Scenario-based missing female births simulation after 2020

Although we identify Balochistan as the only province with past/ongoing SRB inflation, we do not rule out the possibility that the imbalanced SRB will emerge in other

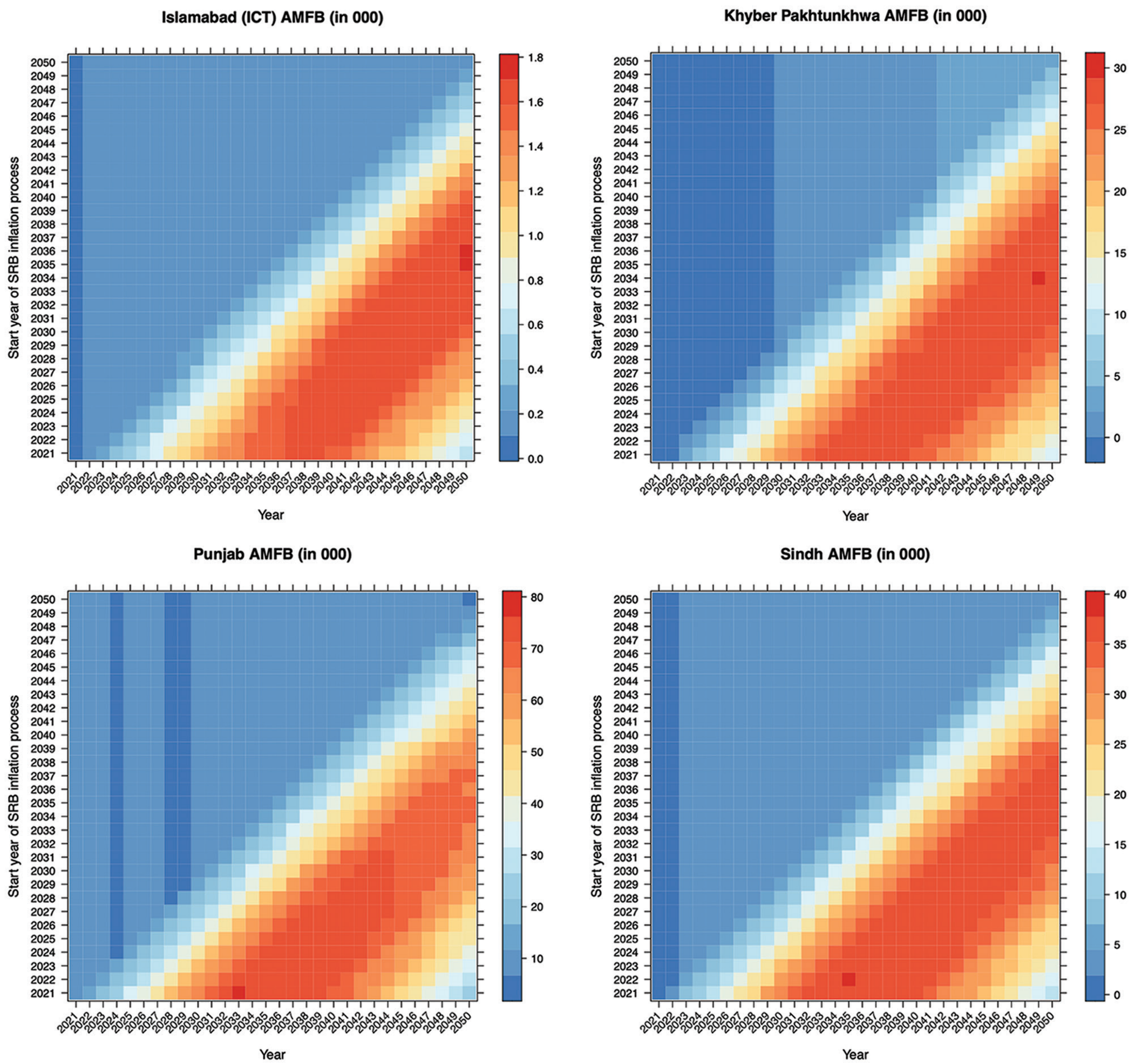

Figure 6. Post-2020 missing female births simulation in provinces without past imbalanced SRB

Note: Median projections of AMFB (in thousand) are presented. AMFB: Annual number of missing female births. Results are shown for provinces with projections of annual total births. X-axis: Projection period 2021 – 2050. Y-axis: Simulation scenarios with different start years of SRB inflation from 2021 to 2050.

provinces. Hence, we present the results of scenario-based projections of missing female births in provinces without current SRB inflation.

Figure 6 shows the AMFB results in four provinces: Islamabad (ICT), Khyber Pakhtunkhwa, Punjab, and Sindh, in which the projected annual total numbers of births are available and are not identified with past SRB inflation. Each row in each heatmap shows the simulated AMFBs over the projection period 2021 – 2050 under the scenario that the SRB inflation process starts in a certain year.

In general, the later the assumed start year of the SRB imbalance process, the AMFBs in more years are projected near zero. In Figure 6, the heatmap, these trends manifested as increasing extents of blue areas as the rows are traversed from bottom to top. In the heatmap, darker blue means a smaller number of missing female births while redder means a higher number of missing female births. This implies that while the start year of the SRB inflation process delays year by year (i.e., moving up on the y-axis), there are delayed effects on the occurrence of the number of missing female births. When the simulated SRB becomes imbalanced during the projection period, the three stages of the sex ratio transition (increase, stagnation, and return to normal) are visible in the resulting AMFBs. However, the AMFB is influenced not only by the SRB imbalance process but also by the levels and trends in the total number of births over time. Accounting for both the SRB inflation and the total number of births, we project the maximum AMFB for different combinations of start-year scenarios and the year in the projection period. In Islamabad (ICT), the maximum AMFB is projected to occur in 2050 when the SRB imbalance process starts in 2035, with the AMFB at 1.7 thousand. In Khyber Pakhtunkhwa, the maximum AMFB is projected at 29.2 thousand in 2049 when the SRB inflation starts in 2034. The largest AMFB in Punjab is projected to occur in 2033 when the SRB inflation stats in 2021 and its projected value is 76.2 thousand. In Sindh, when the SRB imbalance process starts in 2022, the maximum AMFB across all scenarios is projected to occur in 2035 at 37.8 thousand.

4. Discussion

Policy planners can prepare guidelines for preventing prenatal sex discrimination using scenario-based projections of the number of missing female births in the provinces without ongoing SRB inflation. A recent study provided missing female birth projections at the national level for all countries in the world (Chao *et al.*, 2021b). According to that study, if sex-selective abortion were to happen in Pakistan national wide, the missing female births may contribute as high as 14% of the global numbers during the 2021 – 2100 period. Our study reveals the potential future missing female births in Pakistan in four provinces, namely, Islamabad (ICT), Khyber Pakhtunkhwa, Punjab, and Sindh. For different start years of the SRB imbalance in each province, we identify the years in which the number of missing female births will possibly deficit the greatest. The projections results reflect the fact that the number of missing female births is a combined effect of the SRB inflation process and fertility transition. Given the speed of the fertility transition in Pakistan, and the estimates from the 2017 – 2018 DHS, we revealed that the decline in fertility rates in Pakistan has slowed at both the national and subnational levels since the 1990s, and the total fertility rate at 3.6 remains higher than in neighboring countries (National Institute of Population Studies (NIPS) [Pakistan] & ICF, 2019).

The findings of this study reinforce the persistence of strong son preferences in Pakistan and identified high-level distortion in Balochistan. The three conditions of Guilmoto (2009) hypothesis for the distortion of SRB in Pakistan are currently well aligned: Inherited gender discrimination, preference for large families, and technological advancement. Pakistan fares poorly on gender equality, ranking 135 of 191 countries on the Gender Inequality Index (United Nations Development Programme, 2022). The deep-rooted social and cultural norms and practices continue to be the underlying cause of gender inequalities.

Furthermore, inflated SRB influences the demand of a larger number of children, currently manifesting as higher fertility in Pakistan for the sake of more sons. The fertility stagnation in Pakistan since the onset of the 21st century is evident primarily attributed to the higher ideal family size with at least two sons (Wazir, 2018). The imbalanced SRB leads to prolonged consequences in both demographic and social aspects. Imbalanced SRB is one of the main factors that lead to the phenomenon of “missing women” first endeavored by Amartya Sen, referring to the females that should have survived or been born in the absence of sex discrimination and excess mortality among females (Sen, 1990). A large number of “missing women” results in a marriage squeeze, increased levels of antisocial behavior and violence, and may eventually have a long-term impact on stability and social development.

Despite the strong son preference persists in Pakistan, one study implied that son preference has not resulted in nationwide sex-selective abortions but occurred in subpopulations such as in urban clinics (Sathar *et a*l., 2015). Although abortion is illegal in Pakistan, the abortion rate significantly increased from 27 to 50/1000 for women aged 15 – 49 over the period 2000 – 2012. Meanwhile, Balochistan experienced the highest abortion rate of 60/1000 for women aged 15 – 49 among the provinces (Sathar *et al.*, 2014). In addition, our study also reinforces the distortion of SRB in Balochistan is highest among the provinces. These numbers are associated with well-documented demographic phenomena of “missing women” at the national and provincial levels in Pakistan. Although, the compelling evidence of sex selection and excess mortality are not prevailing in Pakistan on the national level, the absence of evidence is not the evidence of absence. Our subnational study pinpoints the disparity in SRB and sex-selective abortion on the provincial level that is masked by national-level results. There is an urgent need to generate high-quality data in Pakistan, particularly through census at the subnational level followed by in-depth research on the persistence of discriminatory practices of sex selection and excess mortality for females.

This is the first study on estimating SRB in Pakistan from 1980 to 2020 and provides scenario-based projections of missing female births up to 2050 by province based on a Bayesian hierarchical time series model. Our results revealed important SRB disparity across geographic locations in Pakistan. Among the seven provinces included in the study, Balochistan presents a decisively imbalanced SRB. In the other provinces without the existing SRB inflation, we demonstrate important disparities in the occurrences and quantities of female birth deficits before 2050.

5. Conclusions

Our study provides model-based and data-driven SRB estimates and projections for provinces in Pakistan from 1980 to 2050. Our model results demonstrate important disparities in SRB levels and trends across provinces over time. Balochistan is identified as the only province in Pakistan with an existing SRB imbalance and, consequently missing female births. In future work, in-depth provincial studies and the collection of high-quality birth data are required to monitor subnational SRB disparities in Pakistan.

Effective program and policy solutions to curb sex discrimination remain elusive in Pakistan because the practices, leading to excess mortality among females and sex selection, are often poorly understood. Therefore, the institutional response is primarily focused on the improvement of the provision of health care. The last two decades have witnessed the adoption of several pro-women laws such as prevention of sexual violence and harassment, protection, domestic violence, and early marriages. However, the implementation remains challenging, primarily because of federal and provincial autonomy to deliver basic social services. Advancing gender equality and discriminatory practices require accountability mechanisms for policy implementation and enforcement of laws, adequate financing at the provincial level, and community engagement to address discriminatory gender and social norms.

Acknowledgments

The authors acknowledge the baseline research fund support from the King Abdullah University of Science and Technology. The authors also wishes to acknowledge the statistical offices that provided the underlying data making this research possible: Statistics Division, Pakistan. The views expressed in this article are those of the authors and do not necessarily reflect the views of the United Nations Population Fund (UNFPA).

Funding

None.

Conflict of interest

No conflicts of interest were reported by all authors.

Author contributions

Conceptualization: Fengqing Chao
Formal analysis: Fengqing Chao and Hernando Ombao
Supervision: Muhammad Asif Wazir
Validation: Fengqing Chao and Hernando Ombao
Writing – original draft: Fengqing Chao
Writing – review & editing: Fengqing Chao, Muhammad Asif Wazir, and Hernando Ombao

Ethics approval and consent to participate

The human data used in our study are secondary and publicly available datasets from surveys and censuses. The survey data are available at https://dhsprogram.com/ for the DHS Program and at https://mics.unicef.org/surveys for MICS. The census data are available from IPUMS International at https://international.ipums.org/international/.

Consent for publication

Not applicable.

Availability of data

Supplementary File 1: Pakistan provincial SRB database. DOI: https://doi.org/10.6084/m9.figshare.21548082

Supplementary File 2: SRB estimates by Pakistan province from 1980 to 2020. DOI: https://doi.org/10.6084/m9.figshare.21548103

References

Ali, S.M., Hussain, J., & Chaudhry MA. (2001). Fertility transition in Pakistan: Evidence from census [with Comments]. *The Pakistan Development Review*, 40(4):537-550.

Alkema, L., Chao, F., You, D., Pedersen, J., & Sawyer, C.C. (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*, 2(9):e521-e530.

https://doi.org/10.1016/s2214-109x(14)70280-3

Alkema, L., Chou, D., Hogan, D., Zhang, S., Moller, A.B., Gemmill, A., *et al.* (2016). Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: A systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *The Lancet*, 387(10017):462-474.

https://doi.org/10.1016/S0140-6736(15)00838-7

Alkema, L., Wong, M.B., & Seah, P.R. (2012). Monitoring progress towards Millennium Development Goal 4: A call for improved validation of under-five mortality rate estimates. *Statistics, Politics and Policy*, 3(2):1-19.

https://doi.org/10.1515/2151-7509.1043

Atif, K., Ullah, M.Z., Afsheen, A., Naqvi, S.A., Raja, Z.A., & Niazi, S.A. (2016). Son preference in Pakistan; A myth or reality. *Pakistan Journal of Medical Sciences*, 32(4):994-998.

https://doi.org/10.12669/pjms.324.9987

Attané, I., & Guilmoto, C.Z. (2007). Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia. Paris: Committee for International Cooperation in National Research in Demography.

Bearak, J., Popinchalk, A., Alkema, L., & Sedgh, G. (2018). Global, regional, and subregional trends in unintended pregnancy and its outcomes from 1990 to 2014: Estimates from a Bayesian hierarchical model. *The Lancet Global Health*, 6(4):e380-e389.

https://doi.org/10.1016/S2214-109X(18)30029-9

Brown, C.E. (1998). Coefficient of variation. In: Applied Multivariate Statistics in Geohydrology and Related Sciences. Berlin, Heidelberg: Springer, p. 155-157.

Chao F, Guilmoto CZ & Ombao H. (2021c). Sex ratio at birth in Vietnam among six subnational regions during 1980-2050, estimation and probabilistic projection using a Bayesian hierarchical time series model with 2.9 million birth records. *PLoS One*, 16(7):e0253721.

https://doi.org/10.1371/journal.pone.0253721

Chao, F., & Yadav, AK. (2019). Levels and trends in the sex ratio at birth and missing female births for 29 states and union territories in India 1990-2016: A Bayesian modeling study. *Foundations of Data Science*, 1(2):177-196.

https://doi.org/10.3934/fods.2019008

Chao, F., Gerland, P., Cook, A.R., & Alkema, L. (2019a). Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *Proceedings of the National Academy of Sciences,* 116(19):9303-9311.

https://doi.org/10.1073/pnas.1812593116

Chao, F., Gerland, P., Cook, A.R., & Alkema, L. (2019b). "Web appendix systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *Proceedings of the National Academy of Sciences*, 116:9303-9311.

https://doi.org/10.6084/m9.figshare.12442373

Chao, F., Gerland, P., Cook, A.R., & Alkema, L. (2021a). Global estimation and scenario-based projections of sex ratio at birth and missing female births using a Bayesian hierarchical time series mixture model. *Annals of Applied Statistics*, 15(3):1499-1528.

https://doi.org/10.1214/20-AOAS1436

Chao, F., Gerland, P., Cook, A.R., Guilmoto, C.Z., Alkema, L. (2021b). Projecting sex imbalances at birth at global, regional and national levels from 2021 to 2100: Scenario-based Bayesian probabilistic projections of the sex ratio at birth and missing female births based on 3.26 billion birth records. *BMJ Global Health*, 6(8):e005516.

https://doi.org/10.1136/bmjgh-2021-005516

Chao, F., Guilmoto, C.Z., Samir, K.C., & Hernando, O. (2020). Probabilistic projection of the sex ratio at birth and missing female births by State and Union Territory in India. *PLoS One*, 15(8):e0236673.

https://doi.org/10.1371/journal.pone.0236673

Chao, F., Samir, K.C., & Ombao, H. (2022). Estimation and probabilistic projection of levels and trends in the sex ratio at birth in seven provinces of Nepal from 1980 to 2050: A Bayesian modeling approach. *BMC Public Health*, 22(1): 1-5.

https://doi.org/10.1186/s12889-022-12693-0

Chao, F., You, D., Pedersen, J., Hug, L., & Alkema, L. (2018a). National and regional under-5 mortality rate by economic status for low-income and middle-income countries: A systematic assessment. *The Lancet Global Health*, 6(5):e535-e547.

https://doi.org/10.1016/S2214-109X(18)30059-7

Chao, F., You, D., Pedersen, J., Hug, L., & Alkema, L. (2018b). National and regional under-5 mortality rate by economic status for low-income and middle-income countries: A systematic assessment. *The Lancet Global Health*, 6:e535-e547.

https://doi.org/10.6084/m9.figshare.12442244

De Tray, D. (1984). Son preference in Pakistan: An analysis of intentions vs. behavior. *Research in Population Economics*, 5:185-200.

Dréze, J., & Sen, A. (1990). Hunger and Public Action. Oxford: Oxford University Press.

Duthé, G., Meslé, F., Vallin, J., Badurashvili, I., & Kuyumjyan, K. (2012). High sex ratios at birth in the Caucasus: Modern technology to satisfy old desires. *Population and Development Review*, 38(3):487-501.

https://doi.org/10.1111/j.1728-4457.2012.00513.x

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1):36-48.

https://doi.org/10.2307/2685844

Efron, B., & Tibshirani, RJ. (1944). An Introduction to the Bootstrap. Abingdon: Chapman and Hall/CRC.

Feeney, G., & Alam, I. (1998). Fertility, population growth, and accuracy of census enumeration in Pakistan: 1961-1998. In: A Kemal, M Irfan and N. Mahmood (eds.). Population of Pakistan: An Analysis of 1998 Population and Housing Census. Islamabad: Pakistan Institute of Development Economics (PIDE) and UNFPA.

Ge, T., Mei, L., Tai, X., & Jiang, Q. (2020). Change in China's SRB: A dynamic spatial panel approach. *International Journal of Environmental Research and Public Health*, 17(21):8018.

https://doi.org/10.3390/ijerph17218018

Gelman, A., & Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457-472.

https://doi.org/10.1214/ss/1177011136

Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., *et al.* (2014). World population stabilization unlikely this century. *Science*, 346(6206):234-237.

https://doi.org/10.1126/science.1257469

Goodkind, D. (2011). Child underreporting, fertility, and sex ratio imbalance in China. *Demography*, 48(1):291-316.

https://doi.org/10.1007/s13524-010-0007-y

Guilmoto, C. (2012). Sex Imbalances at Birth: Current Trends, Consequences and Policy Implications. Bangkok, Thailand: UNFPA Asia and Pacific Regional Office. Available from:https://www.unfpa.org/publications/sex-imbalances-birth [Last accessed on 2022 Nov 07].

Guilmoto, C.Z. (2009). The sex ratio transition in Asia. *Population and Development Review*, 35(3):519-549.

https://doi.org/10.1111/j.1728-4457.2009.00295.x

Guilmoto, C.Z., & Ren, Q. (2011). Socio-economic differentials in birth masculinity in China. *Development and Change*, 42(5):1269-1296.

https://doi.org/10.1111/j.1467-7660.2011.01733.x

Guilmoto, C.Z., Chao, F., & Kulkarni, P.M. (2020). On the estimation of female births missing due to prenatal sex selection. *Population Studies*, 74(2):283-289.

https://doi.org/10.1080/00324728.2020.1762912

Guilmoto, C.Z., Hoang, X., & Van, T.N. (2009). Recent increase in sex ratio at birth in Viet Nam. *PLoS One*, 4(2):e4624.

https://doi.org/10.1371/journal.pone.0004624

Gupta, M.D., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., & Hwa-Ok, B. (2003). Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *The Journal of Development Studies*, 40(2):153-187.

https://doi.org/10.1080/00220380412331293807

Hussain, R., Fikree, F.F., & Berendes, H. (2000). The role of son preference in reproductive behaviour in Pakistan. *Bulletin of the World Health Organization*, 78:379-388.

ICF International. (2012). Demographic and Health Survey Sampling and Household Listing Manual. Calverton, Maryland, U.S.A.: Measure DHS, p. 78-79. Available from: https://dhsprogram.com/pubs/pdf/DHSM4/DHS6_Sampling_Manual_Sept2012_DHSM4.pdf [Last accessed on 2022 Nov 07].

ICF International. (2022). The DHS Program. Available from: https://dhsprogram.com [Last accessed on 2022 Nov 07].

Jiang, Q., & Zhan, C. (2021). Recent sex ratio at birth in China. *BMJ Global Health*, 6(5):e005438.

https://doi.org/10.1136/bmjgh-2021-005438

Khan, M.A., & Sirageldin, I. (1977). Son preference and the demand for additional children in Pakistan. *Demography*, 14(4):481-495.

https://doi.org/10.2307/2060591

Lin, T. (2009). The decline of son preference and rise of gender indifference in Taiwan since 1990. *Demographic Research*, 20:377.

https://doi.org/10.4054/DemRes.2009.20.16

Liu, P., & Raftery, A.E. (2020). Accounting for uncertainty about past values in probabilistic projections of the total fertility rate for most countries. *The Annals of Applied Statistics*, 14(2):685.

https://doi.org/10.1214/19-aoas1294

Masquelier, B., Hug, L., Sharrow, D., You, D., Mathers, C., Gerland, P., *et al.* (2018). Global, regional, and national mortality trends in older children and young adolescents (5-14 years) from 1990 to 2016: An analysis of empirical data. *The Lancet Global Health*, 6(10):e1087-e1099.

https://doi.org/10.1016/S2214-109X(18)30353-X

Minnesota Population Center. (2019). Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS.

https://doi.org/10.18128/D020.V7.2

National Institute of Population Studies (NIPS) [Pakistan], & ICF. (2019). Pakistan Demographic and Health Survey 2017-18. Islamabad, Pakistan, and Rockville, Maryland, USA: NIPS, ICF. Available from: https://dhsprogram.com/pubs/pdf/FR354/FR354.pdf [Last accessed on 2022 Nov 07].

Pakistan Bureau of Statistics (PBS). (2019). Pakistan Social and Living Status Measurement Survey: 2013-14 and 2018-19. Statistics Division, Planning Commission, Islamabad. Government of Pakistan: Pakistan Bureau of Statistics (PBS). Available from: https://www.pbs.gov.pk/content/pakistan-social-and-living-standards-measurement [Last accessed on 2022 Nov 07].

Park, C.B., & Cho, N.H. (1995). Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea. *Population and Development Review*, 21(1):59-84.

https://doi.org/10.2307/2137413

Pedersen, J., & Liu, J. (2012). Child mortality estimation: Appropriate time periods for child mortality estimates from full birth histories. *PLoS Medicine*, 9(8):e1001289.

https://doi.org/10.1371/journal.pmed.1001289

Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Vol. 124. Vienna, Austria: Proceedings of the 3rd International Workshop on Distributed Statistical Computing. pp. 1-10.

Plummer, M. (2018). RJAGS: Bayesian Graphical Models using MCMC. R Package Version 4-8. Available from: https://CRAN.R-project.org/package=rjags [Last accessed on 2022 Nov 07].

Plummer, M., Best, N., Cowles, K., & Karen, V. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1):7-11.

Qayyum, K., & Rehan, N. (2017). Sex-selective abortion in rural Pakistan. *Journal of Advances in Medicine and Medical Research*, 22(12):1-7.

R Core Team. (2022). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Core Team. Available from: https://www.R-project.org [Last accessed on 2022 Nov 07].

Sathar, Z., Rashida, G., Hussain, S., & Hassan, A. (2015). Evidence of son preference and resulting demographic and health outcomes in Pakistan. Islambad: Population Council. Available from: https://knowledgecommons.popcouncil.org/departments_sbsr-pgy/665 [Last accessed on 2022 Nov 07].

Sathar, Z., Singh, S., Rashida, G., Shah, Z., Niazi, R. (2014). Induced abortions and unintended pregnancies in Pakistan. *Studies in Family Planning*, 45(4): 471-491.

https://doi.org/10.1111/j.1728-4465.2014.00004.x

Sen, A. (1990). More than 100 Million Women are Missing. Vol. 37. New York City: New York Review of Books. pp. 61-66. Available from: https://www.nybooks.com/articles/1990/12/20/more-than-100-million-women-are-missing [Last accessed on 2022 Nov 07].

Su, Y.S., & Yajima, M. (2015). R2jags: Using R to Run 'JAGS'. R Package Version 0.5-7. Available from: https://CRAN.R-project.org/package=R2jags [Last accessed on 2022 Nov 07].

United Nations Development Programme. (2022). Gender Inequality Index (GII). Available from: https://hdr.undp.org/data-center/thematic-composite-indices/gender-inequality-index/indicies/GII [Last accessed on 2022 Nov 07].

Verma, V., & Le, T. (1996). An analysis of sampling errors for the demographic and health surveys. *International Statistical Review*, 64(3):265-294.

https://doi.org/10.2307/1403786

Wang, H., Abbas, K.M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., *et al.* (2020). Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: A comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258):1160-1203.

https://doi.org/10.1016/S0140-6736(20)30977-6

Wazir, M.A. (2018). Fertility preferences in Pakistan. In: S Gietel-Basten, J Casterline, and M Choe, (eds.). Family Demography in Asia: A Comparative Analysis of Fertility Preferences. Cheltenham: Edward Elgar, pp. 247-259.

Wazir, M.A., & Goujon, A. (2019). Assessing the 2017 Census of Pakistan Using Demographic Analysis: A Sub-National Perspective. Austria: Vienna Institute of Demography Working Papers. Available from: https://www.oeaw.ac.at/fileadmin/subsites/Institute/VID/PDF/Publications/Working_Papers/WP2019_06.pdf [Last accessed on 2022 Nov 07].

Wazir, M.A., & Shaheen, K. (2016). Understanding and Measuring Pre-and Post-Abortion Stigma about Women Who Have Abortions: Results from Explorative Study. In: 2016 Annual Meeting. Washington, DC: Population Association of America (PAA). Available from: https://paa.confex.com/paa/2016/mediafile/ExtendedAbstract/ Paper3246/Extended%20abstract%20PAA%202016.pdf [Last accessed on 2022 Nov 07].

You, D., Hug, L., Ejdemyr, S., Idele, P., Hogan, D., Mathers, C., *et al.* (2015). Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: A systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *The Lancet*, 386(10010):2275-2286.

https://doi.org/10.1016/S0140-6736(15)00120-8

Zaidi, B., & Morgan, S.P. (2016). In the pursuit of sons: Additional births or sex-selective abortion in Pakistan? *Population and Development Review*, 42(4):693-710.

https://doi.org/10.1111/padr.12002

Appendix

Data processing

Appendix A. Sampling errors in the DHS and MICS data

Both Demographic and Health Survey (DHS) and Multiple Indicator Cluster Survey (MICS) provide individual-level data with the full birth history of each woman of reproductive age interviewed during the survey fieldwork period. We calculated the sampling error in the log-transformed sex ratio at birth (SRB) obtained from the DHS and MICS data series using the jackknife method (Efron & Gong, 1983; Efron & Tibshirani, 1994; ICF International, 2012). For a certain DHS or MICS data series, let U denote the total number of clusters (based on the cluster/primary sampling unit numbers in the survey data (Verma & Le 1996)). The u-th partial prediction of SRB is determined by the following equation:

$$r_{-u} = \frac{\sum_{n=1}^{N} I_n\left(x_n = male; d_n \neq u\right) w_n}{\sum_{n=1}^{N} I_n\left(x_n = female; d_n \neq u\right) w_n}, for\, u \in \{1,\ldots,U\},$$

Where, n indexes the live births in each state-survey-year; N is the total number of live births; and x_n, d_n, and w_n are the sex, cluster number, and sampling weight for the n-th live birth, respectively. The sampling weight of each birth w_n is extractable from the survey data and reflects the survey sampling design (Verma & Le, 1996). We define I_n (.) = 1 if the condition inside the brackets is true and I_n (.) = 0 otherwise. The u-th pseudo-value estimate of the SRB on the log-scale is:

$$\log(r)_u^* = U\log\left(r^{'}\right) - (U-1)\log\left(r_{-u}\right), where$$

$$r^{'} = \frac{\sum_{n=1}^{N} I_n\left(x_n = male\right) w_n}{\sum_{n=1}^{N} I_n\left(x_n = female\right) w_n}.$$

The sampling variance is:

$$\sigma^2 = \frac{\sum_{u=1}^{U}\left(\log(r)_u^* - \overline{\log(r)_u^*}\right)^2}{U(U-1)}, where$$

$$\overline{\log(r)_u^*} = \frac{1}{U}\sum_{u=1}^{U}\log(r)_u^*.$$

In the DHS or MICS data, the annual log-transformed SRB observations are merged such that the coefficient of variation (CV) for log-transformed SRB is below 0.1 or the merged period reaches 5 years (Pedersen & Liu, 2012). For a certain DHS/MICS data series, let $\{t_n, t_{n-1},\ldots,t_1\}$ be years with recorded births from recent to past. The merge starts from the most recent year t_n and goes back year by year to $t_{n-1},\ldots,1$. The process is performed by the following algorithm for each DHS and MICS data series:

Step #	Merging process of DHS and MICS data
1	for $t \in \{t_n, t_n-1,\ldots, t_1\}$ do
2	if $t=t_n$ then
3	Compute σ as explained above. Compute
4	if $CV<0.1$ or $t_n-t_{n-1} > 1$ then
5	stop and move to the previous time point
6	else
7	Set $t=t_{n-1}$, merge births from tn and, t_{n-1} by summing them up
8	Repeat step 3
9	if $CV<0.1$ or $t_{n-1} - t_{n2} >1$ or $t_n-t_{n-1} = 5$ then
10	stop and move to the previous time point
11	else
12	Set $t=t_{n-2}$, merge births from t_n, t_{n-1}, t_{n-2} by summing them up
13	Repeat steps 8 – 12 for $t \in \{tn-2,\ldots, t1\}$

A. 1. Further explanations of the motivation and assumptions of merging DHS/MICS data

The above algorithm is to merging observations from single calendar years into observations over short time periods from DHS and MICS surveys where full birth histories are available. The merge refers to summing up the number of sex-specific births across multiple years before computing the SRB for that period. The purpose of the merging process is meant to reduce uncertainties associated with observations to a reasonable level. Without merging the births from each calendar year, the sampling errors in these population indicators become unacceptably large due to the small sample sizes (Pedersen & Liu, 2012). The underlying assumptions for the following expressions are:

- Step 1 for $t \in \{t_n, t_{n-1},\ldots,t_1\}$ do: this means that we are merging the 1-year observations from the most recent year t_n to the earliest year with data t_1. The reason we merge the observations backward in time rather than merge forward is that: Usually in countries where DHS and MICS surveys are conducted, more births were sampled in recent years than in older years. This is largely due to the improvement of surveying technology, more mothers were still alive to recall recent births than births born decades ago, and less recall bias happened to births born in recent years than in earlier periods. Hence, the 1-year observations in recent years are usually less likely to be merged compared to those in earlier periods because the sample sizes are usually larger in recent years. By merging backwards, we are able to preserve more observations in recent years without merging them. If we merge forward, it is likely that some of the observations in recent periods would be merged with observations in early periods.

- Step 4 if $CV < 0.1$ or $t_n - t_{n-1} > 1$: When time t is at the most recent year t_n, the condition to stop merging the birth to t_{n-1}, the previous year where data is available is (i) the $CV < 0.1$; or (ii) t_{n-1} is not the adjacent calendar year of t_n, i.e. $t_n - t_{n-1} > 1$.
- Step 9 if $CV < 0.1$ or $t_{n-1} - t_{n-2} > 1$ or $t_n - t_{n-1} = 5$: When time t is not at the most recent year t_n, the condition to stop merging the birth to t_{n-2} is (i) the $CV < 0.1$; or (ii) t_{n-2} is not the adjacent calendar year of t_{n-1}, that is, $t_{n-2} - t_{n-2} > 1$; or (iii) the merged period is already 5 years, that is, $t_n - t_{n-1} = 5$. We set the upper limit of merging period to 5 years because we want to still generate a time series of observations even after merging. Without setting an upper limit, too few numbers of observations may be generated just to satisfy the CV condition. The 5-year upper bound follows the conventional period length used in other population indicators as suggested by Pedersen & Liu, 2012.

Appendix B. Bayesian model for provincial SRB estimation and projection

B. 1. Notations

The notations are listed in the table below.

B. 2. Scenario-based simulated projections for SRB inflation

The province-specific SRB imbalance process $\delta_p\alpha_{p,t}$ is simulated using posterior samples from the model. The simulated $\delta_p\alpha_{p,t}$ is added to the projected $b\Phi_{p,t}$ for different starting years of the SRB inflation in each province.

For $g \in \{1,\ldots,G\}$, the g[th] simulated SRB inflation is denoted as $\alpha_{p,t}^{'(g)}\delta_p^{(g)}$. $\delta_p^{(g)}$ is the g[th] posterior sample of parameter δ_p. $\alpha_{p,t}^{'(g)}$ is the g[th] simulated SRB imbalance process, with the start year of inflation fixed at $t_0 \in \{2021,\ldots,2050\}$. $\alpha_{p,t}^{'(g)}$ is simulated as below for $g \in \{1,\ldots,G\}$:

$$\alpha_{p,t}^{'(g)} = \left(\xi_p^{(g)} / \lambda_{1p}^{(g)}\right)\left(t - t_0\right), if\ t_0 < t < t_{1p}^{(g)}$$

$$\alpha_{p,t}^{'(g)} = \xi_p^{(g)}, if\ t_{1p}^{(g)} < t < t_{2p}^{(g)}$$

$$\alpha_{p,t}^{'(g)} = \xi_p^{(g)} - \left(\xi_p^{(g)} / \lambda_{3p}^{(g)}\right)\left(t - t_{2p}^{(g)}\right), if\ t_{2p}^{(g)} < t < t_{3p}^{(g)}$$

Symbol	Description
	Index
i	Indicator of the i^{th} SRB observation across all province-years, $i \in \{1,\ldots,531\}$
t	Indicator of year, $t \in \{1980,\ldots,2050\}$
p	Indicator of provinces of Pakistan, $p \in \{1,7\}$
Unknown parameters	
$\Theta_{p,t}$	Model fitting to the true SRB in Pakistan province p in year t
$\Phi_{p,t}$	Province year-specific multiplier for capturing the natural fluctuation in SRBs around the national baseline b in Pakistan province p in year t
$\alpha_{p,t}$	SRB imbalance in Pakistan province p in year t
t_{0p}	Start year of SRB inflation in Pakistan province p
δ_p	Indicator of the presence ($\delta_p = 1$) or absence ($\delta_p = 0$) of SRB inflation in Pakistan province p
ξ_p	Maximum level of SRB inflation in Pakistan province p
λ_{1p}	Period length of the increase stage of the sex ratio transition in Pakistan province p
λ_{2p}	Period length of the stagnation stage of the sex ratio transition in Pakistan province p
λ_{3p}	Period length of the decrease stage of the sex ratio transition in Pakistan province p, which returns the SRB to the national SRB baseline
ω	Non-sampling error
r_i	The i^{th} SRB observation
σ_i	Sampling error for the ith SRB observation (computed in Appendix A.1)
b	Baseline level of SRB over the whole of Pakistan (Chao *et al.*, 2019a), where $b = 1.063$
ρ	Autoregressive Indicator of $\Phi_{p,t}$, where $\rho = 0.9$ (Chao et al., 2019a, b)
σ_ϵ	Standard deviation of distortion parameter for $\Phi_{p,t}$, where $\sigma_\epsilon = 0.004$ (Chao et al., 2019a, b)

$$\alpha_{p,t}^{'(g)} = 0, if\ t < t_0\ or\ t > t_{3p}^{(g)}$$

Where,

$t_{1p}^{(g)} = t_0 + \lambda_{1p}^{(g)}$, $t_{2p}^{(g)} = t_{1p}^{(g)} + \lambda_{2p}^{(g)}$, $t_{3p}^{(g)} = t_{2p}^{(g)} + \lambda_{3p}^{(g)}$.

$\xi_p^{(g)^{(g)}}$, $\lambda_{1p}^{(g)}$, $\lambda_{2p}^{(g)}$, and $\lambda_{3p}^{(g)}$ are the g^th^ posterior samples of parameters ξ_p, λ_{1p}, λ_{2p}, and λ_{3p}.

Appendix C. Model validation

The performance of the inflation model was evaluated by two approaches: (1) Out-of-sample validation and (2) one-province simulation.

C. 1. Out-of-sample validation

We leave out 12% of the observations since the data collection year 2018 instead of reference year, which has been used for assessing model performance of demographic indicators largely based on survey data (Alkema *et al.*, 2012; Alkema *et al.*, 2014; Chao *et al.*, 2018a,b). There are 64 left-out observations from six Pakistan provinces. After leaving out the data, we fit the model to the training dataset and obtain point estimates and credible intervals that would have been constructed from the available dataset in the selected survey year. Based on the training dataset, we also generate the prediction distribution for each left-out observation.

We calculate the median errors and median absolute errors in the left-out observations. The errors are defined as $e_j = y_j - \widetilde{y}_j$, where $\widetilde{y}_j$ refers to the posterior median of the predictive distribution based on the training dataset for the j^{th} left-out observation y_j. The coverage is given by $1/J \sum oI\,[y_j \geq l_j]\ I\,[y_j \leq u_j]$, where J refers to the number of left-out observations, and l_j and u_j correspond to the lower and upper bounds, respectively, of the 95% prediction interval of the j^{th} left-out observation y_j.

The validation measures are calculated for 1000 sets of left-out observations where each set contains one randomly selected left-out observation from each Pakistan province. The reported validation results are based on the mean outcomes of the 1000 sets of left-out observations. This technique of validation exercise is used to reduce the correlation of validation results within each province and has been used in validation exercises in the previous studies (Alkema *et al.*, 2014; Chao *et al.*, 2018a; You *et al.*, 2015).

Specifically, the final validation results regarding the left-out observations are calculated as follows:

1. For $k \in \{1,\ldots,1000\}$, we select a set of left-out observations $\{y_{k,1},\ldots y_{k,p},\ldots, y_{k,6}\}$. $y_{k,p}$ is the only selected left-out observation from province p and six provinces have left-out observations. Hence, we have $y_{k,1}$ to $y_{k,6}$.
2. For the k^{th} set of left-out observations $\{y_{k,1},\ldots y_{k,p},\ldots, y_{k,6}\}$, we can get the following results:
   - Corresponding errors $\{e_{k,1},\ldots e_{k,p},\ldots, e_{k,6}\}$ for these selected left-out observations.
   - Median of this set of error: medium $(e)_k$.
   - Coverage for this set: Coverage $e_k = \frac{1}{6}\sum_{p=1}^{6} I\left[y_{k,p} \geq l_{k,p}\right] I\left[y_{k,p} \leq u_{k,p}\right]$. Here $l_{k,p}$ and $u_{k,p}$ correspond to the lower and upper bounds of the 95% prediction interval of the left-out observation $y_{k,p}$.
3. Compute the mean of these results for the 1000 set of observations:
   - Corresponding errors $\{e_{k,1},\ldots e_{k,p},\ldots,e_{k,6}\}$ for these selected left-out observations.
   - Final median of error: $\frac{1}{1000}\sum_{k=1}^{1000} median(e)_k$.
   - Final coverage: $\frac{1}{1000}\sum_{k=1}^{1000} coverage_k$.

For the point estimates obtained from the full and training datasets, we define the errors in the true SRB as $e(\Theta)_{p,t} = \widehat{\Theta}_{p,t} - \tilde{\Theta}_{p,t}$, where $\widehat{\Theta}_{p,t}$ is the posterior median in province p in year t obtained from the full dataset, and $\tilde{\Theta}_{p,t}$ is the posterior median in the same province-year obtained from the training dataset.

Similarly, the error in the sex ratio transition process with probability is defined as $e(\alpha\delta)_{p,t} = \widehat{\alpha}_{p,t}\widehat{\delta}_p - \tilde{\alpha}_{p,t}\tilde{\delta}_p$. The coverage is computed similarly to the left-out observations and is based on the lower and upper bounds of the 95% credible interval of $\tilde{\Theta}_{p,t}$ from the training dataset.

C.2. One-province simulation

We assess the inflation model performance in a one-province simulation setting. We simulate SRB for a province prior observing data. In this simulation exercise, we consider all observations as the test data and simulate the SRB using the posterior samples of only the global parameters (instead of province-specific parameters) obtained from the sex ratio transition model using the full dataset. Hence, we simulate the SRB for a province without data and check how well the simulated results can align and cover the SRB observations in each province.

The g^{th} simulated SRB for a "new" province $\Theta(new)_t^{(g)}$ in year t are obtained as follows for $g \in \{1,\ldots,G\}$:

$$\Theta\left(new\right)_t^{(g)} = b\Phi\left(new\right)_t^{(g)} + \alpha\left(new\right)_t^{(g)}\delta\left(new\right)^{(g)}$$

Where the simulated $\Phi\left(new\right)_t^{(g)}$, $\alpha\left(new\right)_t^{(g)}$ and $\delta\left(new\right)^{(g)}$ refer to a "new" province without data. This simulation follows the model specifications of these parameters without considering any province-specific data. In particular, $\Phi\left(new\right)_t^{(g)}$ is simulated as:

$$\log\left(\Phi\left(new\right)_t^{(g)}\right) \sim \mathcal{N}\left(0, \frac{\left(\sigma_\epsilon^{(g)}\right)^2}{1-\left(\rho^{(g)}\right)^2}\right), if\ t = 1980,$$

$$\log\left(\Phi\left(new\right)_t^{(g)}\right) = \rho^{(g)}\log\left(\Phi\left(new\right)_{t-1}^{(g)}\right)$$
$$+\epsilon_t^{(g)}, if\ t \in \{1981,\cdots,2020\},$$

$$\epsilon_t^{(g)} \sim_{i.i.d.} \mathcal{N}\left(0, \left(\sigma_\epsilon^{(g)}\right)^2\right).$$

$\delta\ (new)^{(g)}$ is simulated as:

$$logit\left(\pi^{(g)}\right) \sim \mathcal{N}\left(\mu_\pi^{(g)}, \left(\sigma_\pi^{(g)}\right)^2\right),\ \delta\left(new\right)^{(g)} \sim \mathcal{B}\left(\pi^{(g)}\right),$$

$\alpha\left(new\right)_t^{(g)}$ is simulated as:

$$\xi^{(g)} \sim \mathcal{N}_{(0,\infty)}\left(0.06,\ 0.006^2\right),\ \lambda_1^{(g)} \sim \mathcal{N}_{(0,\infty)}\left(11.0, 1.1^2\right),$$
$$\lambda_2^{(g)} \sim \mathcal{N}_{(0,\infty)}\left(7.6, 0.8^2\right),\ \lambda_3^{(g)} \sim \mathcal{N}_{(0,\inf)}\left(16.1,\ 1.6^2\right),$$
$$t_0^{(g)} \sim \mathcal{U}\left(1970,\ 2050\right),$$

$$\alpha\left(new\right)_t^{(g)} = \left(\xi^{(g)} / \lambda_1^{(g)}\right)\left(t - t_0^{(g)}\right), if\ t_0^{(g)} < t < t_1^{(g)},$$
$$\alpha\left(new\right)_t^{(g)} = \xi^{(g)}, if\ t_1^{(g)} < t < t_2^{(g)}$$
$$\alpha\left(new\right)_t^{(g)} = \xi^{(g)} - \left(\xi_p^{(g)} / \lambda_3^{(g)}\right)\left(t - t_2^{(g)}\right), if\ t_2^{(g)} < t < t_3^{(g)}$$
$$\alpha\left(new\right)_t^{(g)} = 0, if\ t < t_0^{(g)}\ or\ t > t_3^{(g)}$$

Where,

$$t_1^{(g)} = t_0^{(g)} + \lambda_1^{(g)},\ t_2^{(g)} = t_1^{(g)} + \lambda_2^{(g)},\ t_3^{(g)} = t_2^{(g)} + \lambda_3^{(g)}.$$

After generating the simulated values, we calculate results as described for the out-of-sample validation (Appendix C.1.1).

Appendix D. Validation and simulation results

Table D.1 summarizes the results of the left-out SRB observations in the out-of-sample validation exercise and one-province simulation. The median errors are nearly zero in the left-out observations. Although the median absolute errors are slightly higher than the median errors, the average coefficient of variance of the absolute errors for left-out observations (calculated as absolute errors divided by the left-out observation values) is only 5.6%. The coverage of the 95% and 80% prediction intervals is more conservative than expected. The wider-than-expected prediction interval in the left-out observations can be primarily attributed to larger uncertainty in more recent observations.

Table D.2 compares the model estimates obtained from the full dataset and the training set in the out-of-sample validation exercise. Here, we examined the model estimates of the true SRB $\Theta_{p,t}$ and the inflation process with province-specific probability $\delta_p\alpha_{p,t}$. The median errors and the median absolute errors are close to zero.

In summary, the validation results indicate reasonably good calibrations and prediction power of the inflation model with conservative credible intervals.

Table D.1. Validation and simulation results for left-out SRB observations

	Validation out of sample	Simulation
# Province in test dataset	6	8
Median error	0.020	0.003
Median absolute error	0.047	0.071
Below 95% prediction interval (%)	0.0	0.2
Above 95% prediction interval (%)	0.0	3.2
Expected (%)	2.5	2.5
Below 80% prediction interval (%)	0.0	7.6
Above 80% prediction interval (%)	8.0	9.2
Expected (%)	10	10

Note: Error is defined as the difference between a left-out SRB observation and the posterior median of its predictive distribution. SRB observations with data collection years since 2018 are left out. Numbers in the parentheses after the proportions indicate the average number of left-out observations that fall below or above their respective 95% and 80% prediction intervals.

Table D.2. Validation results of estimates based on the training set

Model validation (Out-of-sample)	$\Theta_{p,t}$ 1995	$\Theta_{p,t}$ 2005	$\Theta_{p,t}$ 2015	$\delta_p\alpha_{p,t}$ 1995	$\delta_p\alpha_{p,t}$ 2005	$\delta_p\alpha_{p,t}$ 2015
Median error	0.001	0.001	0.002	0.000	0.000	0.000
Median absolute error	0.001	0.001	0.004	0.000	0.000	0.000
Below 95% credible interval (%)	0.0	0.0	0.0	0.0	0.0	0.0
Above 95% credible interval (%)	0.0	0.0	0.0	0.0	0.0	0.0
Expected (%)	≤2.5	≤2.5	≤2.5	≤2.5	≤2.5	≤2.5
Below 80% credible interval (%)	0.0	0.0	0.0	0.0	0.0	0.0
Above 80% credible interval (%)	0.0	0.0	0.0	0.0	0.0	0.0
Expected (%)	≤10	≤10	≤10	≤10	≤10	≤10

Note: Error defines the differences between the model estimates (i.e., $\Theta_{p,t}$ or $\delta_p\alpha_{p,t}$) obtained from the full and training datasets, and proportions refer to the proportions (%) of countries in which the median estimates from the full dataset fall below or above their respective 95% and 80% credible intervals, respectively, in the training set.