Levels and trends in sex ratio at birth in provinces of Pakistan from 1980 to 2020 with scenario-based missing female birth projections 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 (KAUST), Thuwal, Saudi Arabia

[2]UNFPA-Pakistan, Pakistan

November 2, 2021

Abstract

BACKGROUND: Pakistan has a strong preference for boys over girls; previous evidence on sex preference is primarily reported at the postnatal stage in which the child mortality rate is higher for females than males. Prenatal sex discrimination in Pakistan, reflected in the inflated sex ratio at birth (SRB; ratio of male to female births) has been barely mentioned before this study.

OBJECTIVE: We estimate the SRB and missing female births in Pakistan provinces from 1980 to 2020 and identify provinces with imbalanced SRB. We provide scenario-based projections of missing female births in provinces without the existing SRB inflation.

METHODS: An extensive SRB database of 832,091 birth records was compiled from all available surveys and censuses. To synthesize different data sources and provide annual estimates and their associated uncertainties of SRBs across provinces, we adopted a Bayesian hierarchical time series model.

RESULTS: As per our model, Balochistan has had SRB imbalance since 1980. The maximum SRB was estimated as 1.121 (95% credible interval [1.066; 1.142]) in 1997. Assuming different start year of SRB inflation process in provinces without existing imbalance, the largest female birth deficit is projected to be 76.2 thousand in Punjab in 2033 when the SRB inflation starts in 2021.

CONTRIBUTION: This is the first study on estimating the SRB from 1980 to 2020 and providing scenario-based projections of missing female births up to 2050 by Pakistan province. We identified the Balochistan province with imbalanced SRB and demonstrated important disparities in the occurrence and quantity of female birth deficits before 2050.

*Corresponding author (FC); Email: fengqing.chao@kaust.edu.sa

1 Introduction

The annual measure of sex ratio at birth (SRB; ratio of male to female births) is not only required for population estimation and projection but also reflects the persistent social, cultural, political, and economic discrimination against women [1, 2]. SRBs are distorted from their natural levels in several countries, primarily clustered in South Asia, Eastern Asia, and Eastern Europe. The imbalanced values reach as high as 1.2 (i.e. 120 male births per 100 female births) in certain regions [3, 4, 5, 6, 7, 8, 9, 10, 11]. SRB is the prime indicator of prenatal sex discrimination and sex imbalance in human populations. Distortion in the SRB has been primarily attributed to three interlinked factors [2, 10]: 1) son preference; (2) technological advances in prenatal diagnosis; and (3) preferences for smaller family size and consequent fertility decline. In countries having 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 in the family. The SRB in such populations is male-biased. SRB imbalance has been reported in 12 countries/areas in the world since 1970 [11].

Pakistan is a country that has a strong preference for sons [12, 13, 14, 15]. Preference for male birth in Pakistan stems from lineage, economic and social conditions, caste, and identity. At least one son in a strongly patrilineal society is considered to be important 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 one son [16]. Son preference is evidenced by the excess mortality of female children over male children under five years old in Pakistan, indicating possible differential treatment between girls and boys in this age group [17, 14]. 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 [18]. Between 2018 and 2019, 36% of girls (age 5–16) were out of school versus 25% of boys [19]. However, little evidence of prenatal sex preference has been reported in Pakistan. Previous studies identified no imbalanced SRB at the national level [20, 11]. Other studies suggested that, among couples in Pakistan, the desire for a large family might dominate preferences for the children of a certain type [21]. High prevalence of sex-selective abortion was identified in two rural districts in Balochistan province [22]. However, the above-mentioned results are based on survey data with small sample sizes.

The national-scale levels and trends in SRB can mask the disparity among sub-regions in a country. Even in countries such as China and India with overall strong preference for sons, the SRB is not imbalanced in every province or state [23, 24, 25, 26]. In Pakistan, a subnational-level assessment of SRB is important 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) 2017–18 revealed a high heterogeneous SRB across provinces: a 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) [18]. To 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; 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. In the absence of a fully developed vital registration system in Pakistan, administrative birth records are lacking and vital events are mostly estimated based on household surveys. Only few sample surveys provided the birth histories over different time periods since the 1990s. Second, the data quality of census counts is typically low because of age heaping [27]. 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 [28]. 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 one 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. Given the lack of reliable administrative birth data in Pakistan, it is crucial to estimate the levels and trends in the SRB by a reproducible statistical model.

Figure 1 illustrates the large uncertainty in SRB observations on natinal and provincial levels. It shows the observed SRB from Pakistan 2021–13 and 2017–18 DHSs on the national and provincial levels. In data series Pakistan DHS 2012–13, the national SRB observations range from 0.405 in year 1976 to 1.195 in year 2003. The average sampling error associated with these observations is 0.098, corresponding to 13% of the observation values. On the provincial level, the SRB observations have a wider range with greater uncertainties because of smaller sample sizes. The smallest provincial SRB observation is 0.290 in Balochistan in year 1979 and the biggest value is 2.134 in Balochistan in year 1984. The average sampling error associated with the provincial observations is 0.131, corresponding to 12% of the observation values. Similarly, for Pakistan DHS 2017–18, the observed SRB on national and provincial levels range from 0.914 to 1.439 and from 0.079 to 2.363, respectively. The average sampling errors are 0.100 and 0.172 for national and provincial data, respectively.

In this study, we estimate SRBs among provinces of Pakistan from 1980 to 2020 and provide scenario-based projections to 2050 using a reproducible Bayesian statistical model. In addition to estimating the levels and trends in provincial SRBs, the model detects the existence and transition process of the sex ratio imbalance. 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. 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 reminder of this paper is organized as follows. Section 2 summarizes the database compiled for statistical modeling, and Section 3 summarizes the Bayesian statistical model used for provincial SRB estimation and post-modeling process (identifying provinces with imbalanced SRB and calculating the number of missing female births). Section 4 shows the SRB results by province, the provincial SRB imbalances, the corresponding missing female births, and the scenario-based missing female birth projections. Section 5 summarizes the primary contributions and limitations and concludes the study.

2 Data

Table 1 summarizes our database of provincial SRBs in Pakistan, with 531 SRB data points available in eight provinces of Pakistan. The reference years of these observations range from 1965 to 2019. The database contains 832,091 birth records.[1] The SRB observations were generated from the individual birth records in data sources with full birth histories (see Appendix for details of the data processing steps).

The DHS 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. Furthermore, the Pakistan Social & Living Standards Measurement Survey (PSLM) is a provincial-level survey with a high coverage of households in Pakistan. The PSLM records births over the 12 months prior to the date of the survey interview. The census is conducted once per decade and collects births in the 12 months preceding the census.[2]

[1] Summation of available number of births records. The number of birth records is unknown in some data series.

[2] Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. `https://doi.org/10.18128/D020.V7.2`. Accessed on Aug 16th, 2020.

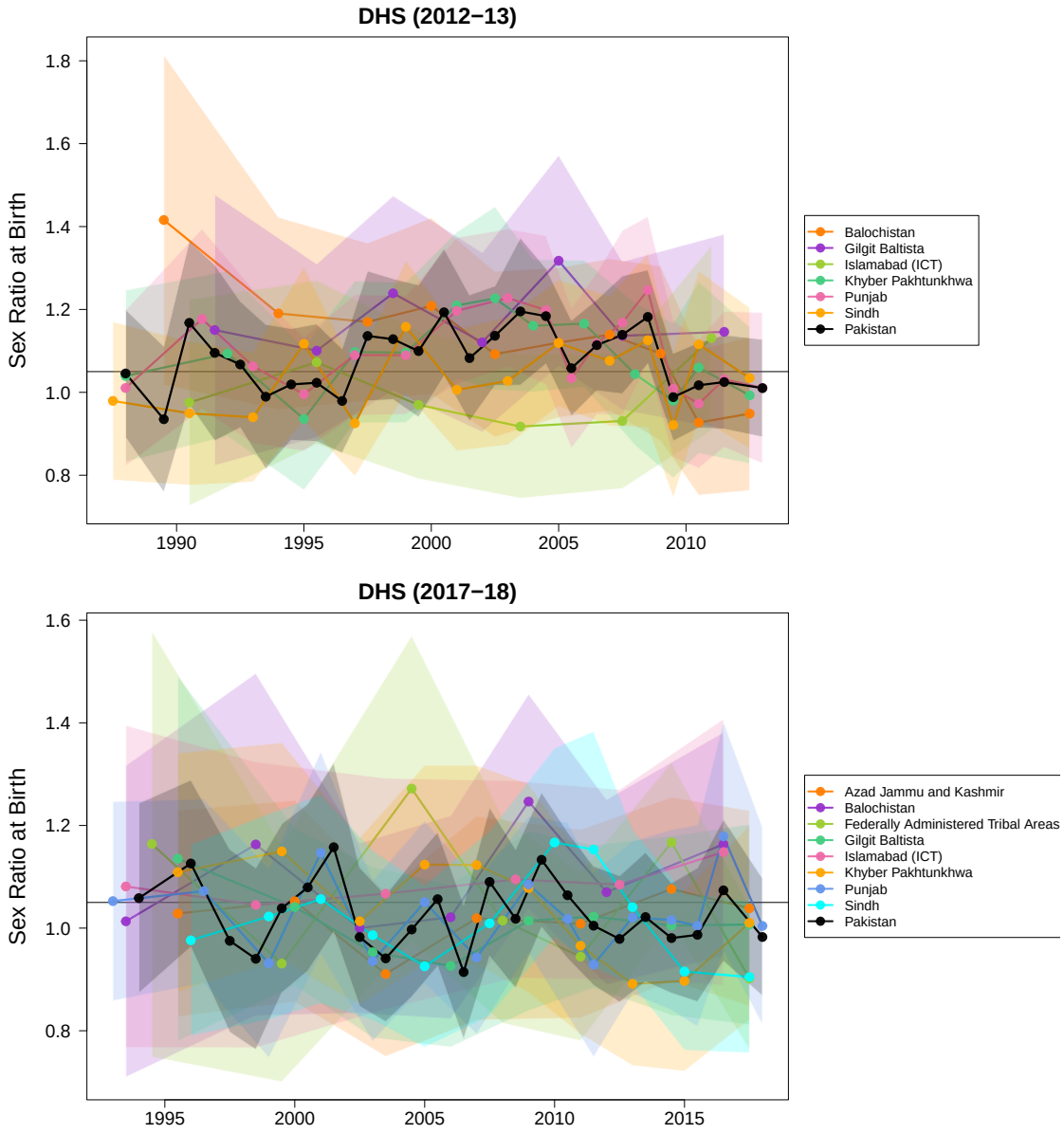

Figure 1: **SRB data in 2012-13 and 2017–18 DHS on national and provincial levels.** The SRB observations (dots) from the same area (province or the whole Pakistan) are connected with line segments of the same color. Shaded areas around the observation series represent the sampling errors in the series (quantified by twice the sampling standard errors).

Survey name	Survey year	# SRB obs.	# Births records
Census	1973, 1981, 1998	15	424,739
DHS	1990–91, 2006–07, 2012–13, 2017–18, 2019	301	253,580
MICS	2010, 2011, 2014, 2003–04, 2007–08, 2016–17, 2017–18	37	153,772†
PSLM	1995–16, 2005–06, 2007–08, 2013–14, 2018–19	228	–
Total		**531**	**832,091***

Table 1: **Pakistan provincial SRB database.** DHS: Demographic and Health Survey. MICS: Multiple Indicator Cluster Survey. PSLM: Pakistan Social & Living Standards Measurement Survey. †: number of birth records available only in MICS 2017–18. –: number of births records is unavailable. *: the total number of births obtained by summing the available number of births records in the 25 years prior to the survey conducted.

3 Methods

The full model specifications, detailed descriptions of model elements, priors for the Bayesian method, and statistical computation are included in the Appendix. The model performance and predictive power were assessed by an out-of-sample validation exercise (leaving out recent observations) and simulation exercises (see 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.

3.1 Bayesian model for provincial SRB estimation and projection

The model for sex ratio at birth by Pakistan province is based on the model described previously [29, 30]. In this study, we made a few modifications in the model to better address the data quality and availability of provincial SRBs in Pakistan. The outcome of interest $\Theta_{p,t}$, the SRB in Pakistan province p in year t is modeled as follows:

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

where $b = 1.056$ 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 [11, 31]. $\Phi_{p,t}$ follows an AR(1) time series model to capture the natural year-by-year fluctuations in SRB within each province.

δ_p is the binary identifier of the sex ratio transition at the provincial level. $\delta_p = 1$ indicates an SRB imbalance in province p, whereas $\delta_p = 0$ indicates no imbalance in province p.

$\alpha_{p,t}$ 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.

We consider both sampling and non-sampling errors when fitting the Bayesian model to the SRB observations. The sampling errors show the uncertainties in the SRB observations introduced by the survey sampling design. These known values are calculated using a jackknife method, as explained in the Appendix. Non-sampling errors include the uncertainties resulting from non-response, recall errors, and data input errors. Usually, they are immeasurable but difficult to eliminate. We treat the non-sampling error as a parameter to be estimated in the Bayesian model.

3.2 Post-modeling process

Identifying provinces of Pakistan with SRB imbalance SRB imbalance in a Pakistan province is detected if $\delta_p = 1$ for more than 95% of the posterior samples (indicating SRB inflation).

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 seven 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.

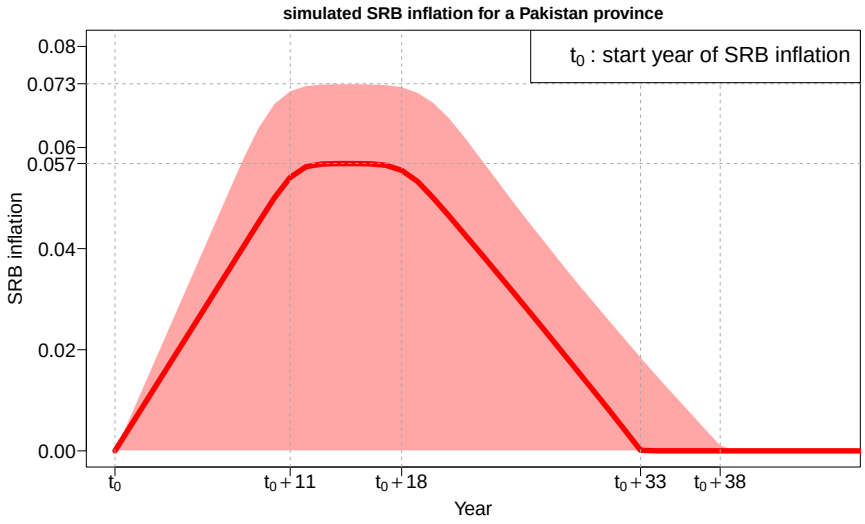

Figure 2: **SRB inflation simulation for a Pakistan province.** Medians are in red curve. 95% credible intervals are in red shades. t_0 is the start year of the SRB inflation process.

Computing the number of missing female births Let $\Psi_{p,t}$ and $\Psi_{p,t}^{(\text{inflation-free})}$ 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 $\Psi_{p,t}^{(\text{inflation-free})} = (B_{p,t} - \Psi_{p,t})/\Theta_{p,t}^{(\text{inflation-free})}$ respectively. The total number of births $B_{p,t}$ in a given province and year, is obtained from [32]. The number of inflation-free female births $\Psi_{p,t}^{(\text{inflation-free})}$ is obtained from the estimated number of male births $(B_{p,t} - \Psi_{p,t})$, and the inflation-free SRB $\Theta_{p,t}^{(\text{inflation-free})} = b\Phi_{p,t}$ in the respective given province-year. The number of missing female births is calculated using a method introduced in [33], which was reviewed and validated in [34].

The annual number of missing female births (AMFB) in province p in year t is defined as: $\Psi_{p,t}^{(\text{missing})} = \Psi_{p,t}^{(\text{inflation-free})} - \Psi_{p,t}$. 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: $\Lambda^{(\text{missing})}_{p,[t_1,t_2]} = \sum_{t=t_1}^{t_2} \Psi^{(\text{missing})}_{p,t}$.

4 Results

4.1 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 (see Section 4.2 for details).

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.

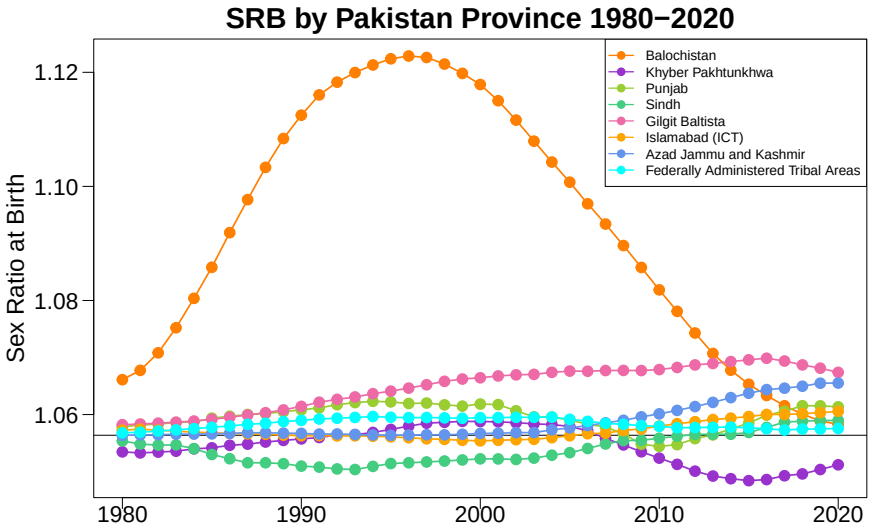

Figure 3: **Median SRB estimates in Pakistan provinces over the 1980–2020 period.**

4.2 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 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 an existing SRB imbalance.

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

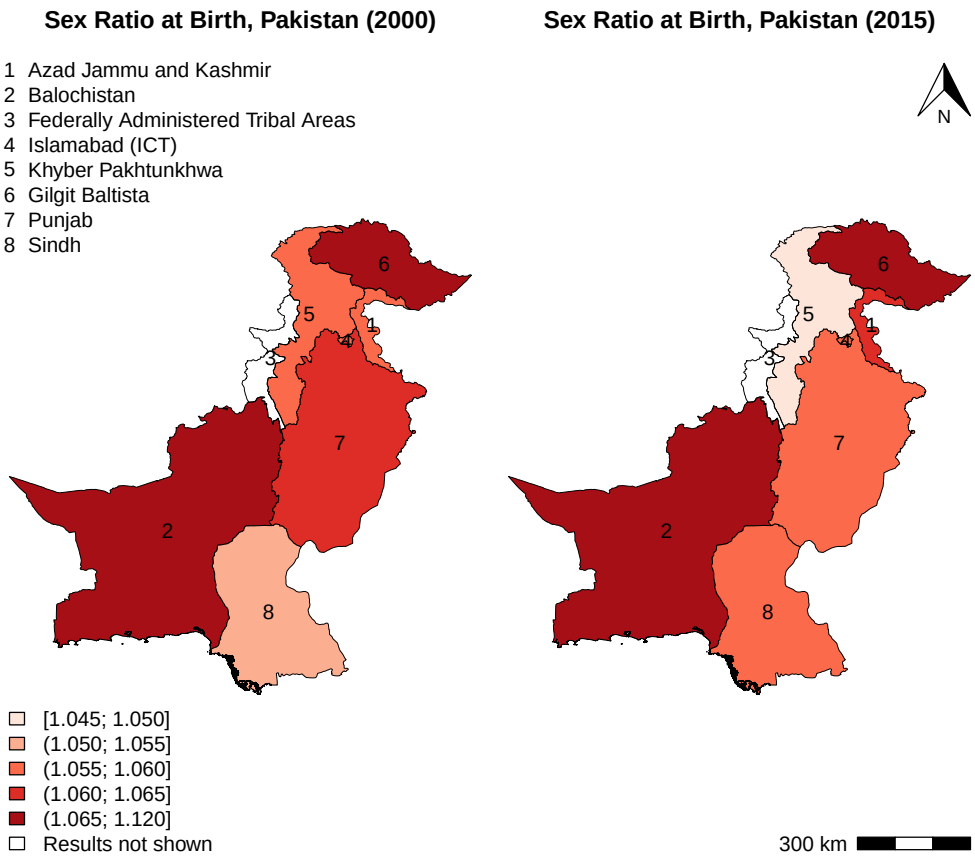

Figure 4: **Geographic disparities of SRB estimates, 2000 and 2015.** Median estimates are presented. Results for Federally Administered Tribal Areas are omitted.

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

Table 2: **SRB inflation probability by Pakistan province.** Provinces are listed in descending order of the inflation probability.

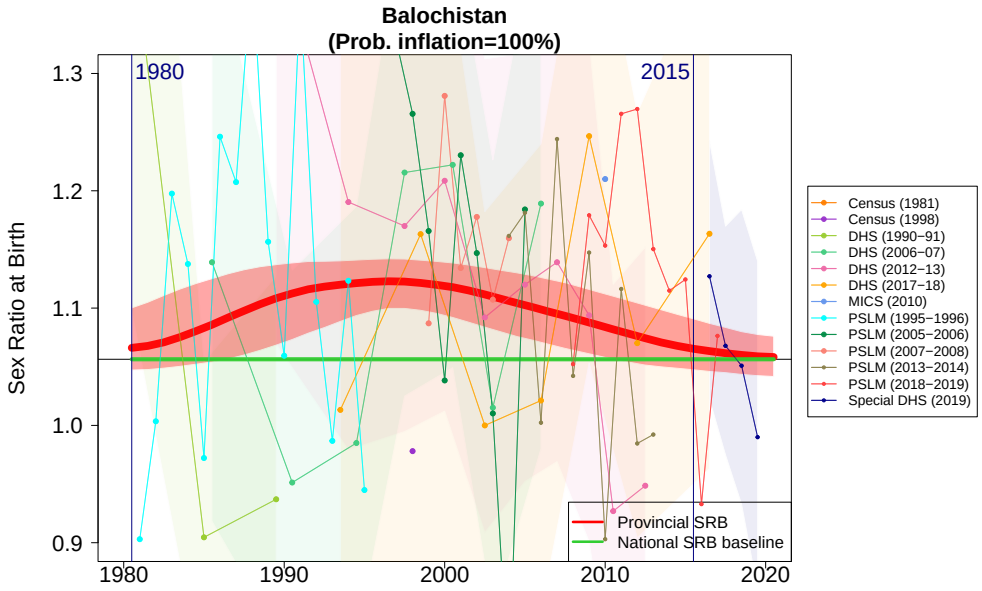

Figure 5: **SRB estimates in Balochistan, 1980–2020.** Red line and shaded areas are the median and 95% credible intervals of the province-specific SRB, respectively. The green horizontal line is the SRB baseline for the whole Pakistan at 1.056 [11]. The SRB observations (dots) from the same source are 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.

4.3 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 at 4.7 [3.4; 6.1] thousand. On a decade-by-decade basis, the average AMFB increase 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.

4.4 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 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 more years in which the projected AMFB are near zero. In the heatmap, these trends manifested as increasing extents of blue areas as the rows are traversed form bottom to top. When the simulated SRB become imbalanced during the projection period, the three stages of 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

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

Table 3: **Missing female births model results in Balochistan.** 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 periods. AMFB: annual number of missing female births. CMFB: cumulative number of missing female births.

but also by the levels and trends in the total numbers of births over time. Accounting for both the SRB inflation and 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.

5 Discussion

This is the first study on estimating SRB in Pakistan from 1980 to 2020 and provide 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.

Our study highlights the importance of Bayesian statistical models in assessing subnational SRB imbalances in Pakistan. Provincial SRB observations are available from censuses and surveys and the data in each province cover most of the estimation period since 1980. However, most survey data of provincial SRB contain large sampling errors introduced by small sample sizes and sampling design. Furthermore, data from administrative records, which usually cover a large extent of the population, are lacking in Pakistan. Our Bayesian model tackles these data challenges and synthesizes observations from different sources using a reproducible approach. The subnational Bayesian model of SRBs is extended from the global SRB model [29] with modifications to handle provincial-level data in Pakistan. Subnational SRB models have been applied to other culturally and demographically heterogeneous countries with son preference such as Nepal [35] and Vietnam [36].

The Bayesian model identifies Balochistan as the only province in Pakistan with an imbalanced SRB during the 1980–2020 period, with an estimated cumulative number of missing female births at 194.6 [140.0; 248.9] thousand. Our results are consistent with previous results [22] of sex-selective abortion in

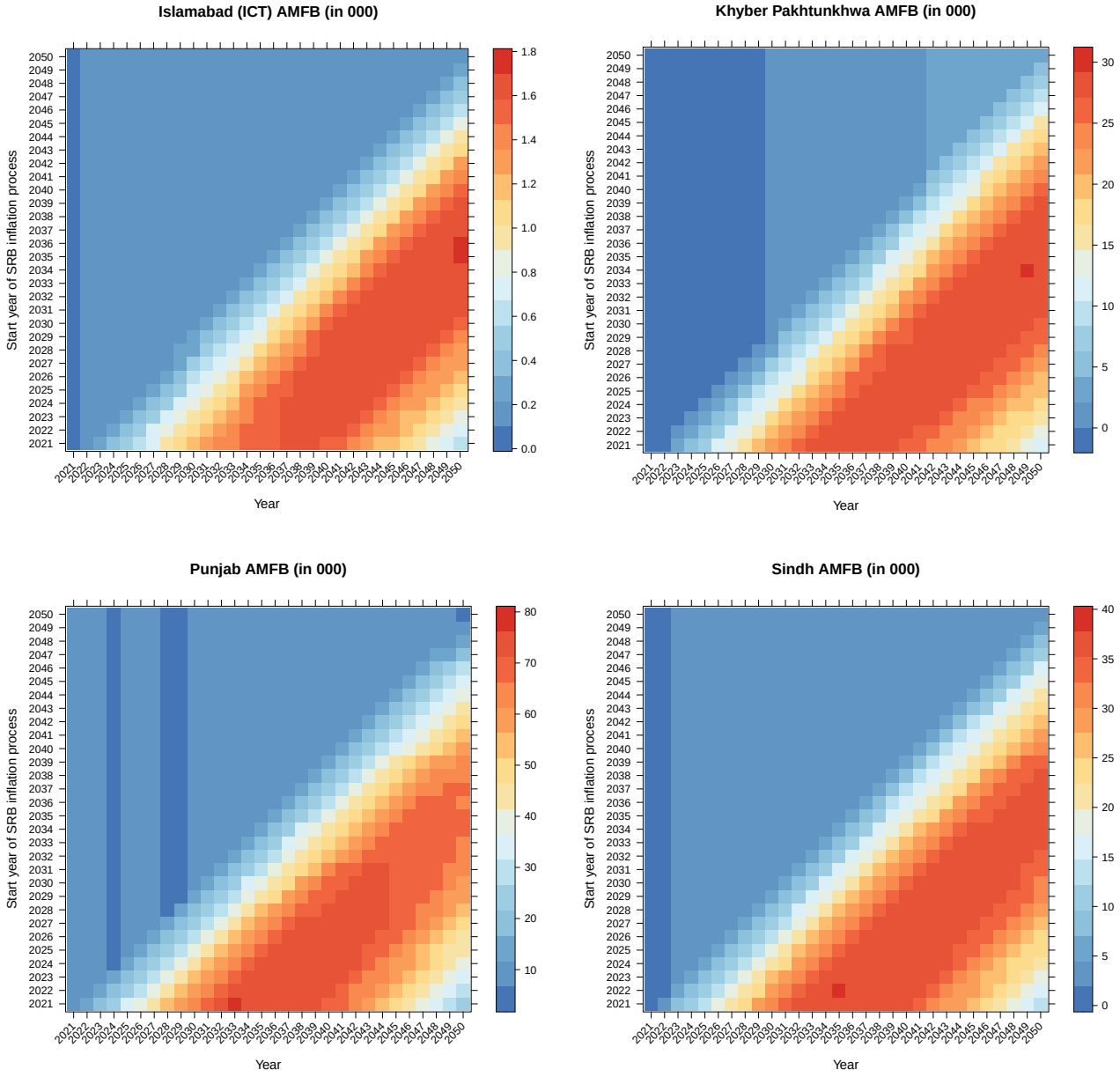

Figure 6: **Post-2020 missing female births simulation in provinces without past imbalanced SRB.** Median projections of AMFB (in thousand) are presented. AMFB: annual number of missing female births. Results are shown for provinces with projections of annual total births. X-axis: projection period 2021–2050. Y-axis: simulation scenarios with different start years of SRB inflation from 2021 to 2050.

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.

Using the scenario-based projections of the number of missing female births in the provinces without ongoing SRB inflation, policy planners can prepare guidelines for preventing prenatal sex discrimination. A recent study provided missing female birth projections at the national level for all countries in the world [30]. 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 2017–18 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 [18].

Several limitations of the study are resulting from the quality and availability of data and the model assumptions. First, as the sampling sizes of birth records are smaller in provincial SRB data than in national-level data, the uncertainties in Pakistan provincial SRB observations are relatively large. Furthermore, due to the small number of births records sampled (which are still representative samples for each province), the levels and trends across different data sources vary greatly. Second, we assume 95% of probability as the cutoff for provincial SRB inflation. The different choices of cutoff value will identify different sets of provinces with SRB imbalance. Gilgit Baltistan has a 68% chance of having SRB imbalance meanwhile it is estimated at 42.5% for Azaz Jammu and Kahmir. To confirm whether the SRBs are truly imbalanced in these provinces, in-depth within-province analyses are needed.

In conclusion, 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 for the monitoring of subnational SRB disparities in Pakistan.

Acknowledgment The author 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).

Supplementary material The Appendix is available at `https://doi.org/10.6084/m9.figshare.16917622`.

References

[1] Das Gupta M, Zhenghua J, Bohua L, Zhenming X, Chung W, Hwa-Ok B. 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. 2003;40(2):153–187.

[2] Guilmoto CZ. The sex ratio transition in Asia. Population and Development Review. 2009;35(3):519–549.

[3] Park CB, Cho NH. Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea. Population and development review. 1995:59–84.

[4] Attané I, Guilmoto CZ. Watering the neighbour's garden: The growing demographic female deficit in Asia. Paris: Committee for International Cooperation in National Research in Demography; 2007.

[5] Guilmoto CZ, Hoang X, Van TN. Recent increase in sex ratio at birth in Viet Nam. PLoS One. 2009;4(2):e4624.

[6] Lin Tc. The decline of son preference and rise of gender indifference in Taiwan since 1990. Demographic research. 2009;20:377.

[7] Goodkind D. Child underreporting, fertility, and sex ratio imbalance in China. Demography. 2011;48(1):291–316.

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

[9] Guilmoto CZ, Ren Q. Socio-economic differentials in birth masculinity in China. Development and change. 2011;42(5):1269–1296.

[10] Guilmoto C. Sex Imbalances at Birth: Current trends, consequences and policy implications. Bangkok, Thailand: UNFPA Asia and Pacific Regional Office; 2012. Available from: http://www.unfpa.org/webdav/site/global/shared/documents/publications/2012/SexImbalancesatBirth.PDFUNFPAAPROpublication2012.pdf.

[11] Chao F, Gerland P, Cook AR, Alkema L. 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. 2019;116(19):9303–9311.

[12] Khan MA, Sirageldin I. Son preference and the demand for additional children in Pakistan. Demography. 1977;14(4):481–495.

[13] Hussain R, Fikree FF, Berendes H. The role of son preference in reproductive behaviour in Pakistan. Bulletin of the World Health Organization. 2000;78:379–388.

[14] Sathar Z, Rashida G, Hussain S, Hassan A. Evidence of son preference and resulting demographic and health outcomes in Pakistan. Islambad: Population Council. 2015:doi:10.31899/pgy9.1085. https://knowledgecommons.popcouncil.org/departments_sbsr-pgy/665/.

[15] Atif K, Ullah MZ, Afsheen A, Naqvi SAH, Raja ZA, Niazi SA. Son preference in Pakistan; a myth or reality. Pakistan journal of medical sciences. 2016;32(4):994.

[16] Wazir MA, Shaheen K. Understanding and Measuring Pre–and Post–Abortion Stigma About Women Who Have Abortions: Results from Explorative Study. In: 2016 Annual Meeting. PAA; 2016. .

[17] Alkema L, Chao F, You D, Pedersen J, Sawyer CC. 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. 2014;2(9):e521–e530.

[18] National Institute of Population Studies (NIPS) [Pakistan], ICF. Pakistan Demographic and Health Survey 2017–18. Islamabad, Pakistan, and Rockville, Maryland, USA: NIPS, ICF. 2019.

[19] Pakistan Bureau of Statistics (PBS). Pakistan Social and Living Standards Measurement Survey: 2013-14 and 2018-19. Statistics Division, Planning Commision, Islamabad Government of Pakistan. 2019. Available at: https://www.pbs.gov.pk/content/pakistan-social-living-standards-measurement-survey-pslm-2018-19-national-provincial.

[20] Zaidi B, Morgan SP. In the pursuit of sons: Additional births or sex-selective abortion in Pakistan? Population and development review. 2016;42(4):693.

[21] De Tray D. Son preference in Pakistan: an analysis of intentions vs. behavior. Research in Population Economics. 1984;5:185–200.

[22] Qayyum K, Rehan N. Sex-selective Abortion in Rural Pakistan. Journal of Advances in Medicine and Medical Research. 2017:1–7.

[23] Ge T, Mei L, Tai X, Jiang Q. Change in China's SRB: A Dynamic Spatial Panel Approach. International Journal of Environmental Research and Public Health. 2020;17(21):8018.

[24] Jiang Q, Zhang C. Recent sex ratio at birth in China. BMJ Global Health. 2021;6(5):e005438.

[25] Chao F, Yadav AK. 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. 2019;1(2):177.

[26] Chao F, Guilmoto CZ, KC S, Ombao H. Probabilistic projection of the sex ratio at birth and missing female births by State and Union Territory in India. PloS one. 2020;15(8):e0236673.

[27] Feeney G, Alam I. Fertility, population growth, and accuracy of census enumeration in Pakistan: 1961-1998. In: Kemal A, Irfan M, Mahmood N, et al., editors. Population of Pakistan: An Analysis of 1998 Population and Housing Census. Islamabad: Pakistan Institue of Development Economics (PIDE) and UNFPA. Islamabad: Pakistan Institue of Development Economics (PIDE) and UNFPA; 1998. .

[28] Ali SM, Hussain J, Chaudhry MA. Fertility Transition in Pakistan: Evidence from Census [with Comments]. The Pakistan Development Review. 2001;40(4):537–550.

[29] Chao F, Gerland P, Cook AR, Alkema L. Global estimation and scenario-based projections of sex ratio at birth and missing female births using a Bayesian hierarchical time series mixture model. Ann Appl Statist. 2021;15(3):1499–1528.

[30] Chao F, Gerland P, Cook AR, Guilmoto CZ, Alkema L. 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. 2021;6(8):e005516.

[31] Chao F, Gerland P, Cook AR, Alkema L. Web Appendix Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. 2019. DOI: 10.6084/m9.figshare.12442373. Available at https://www.pnas.org/content/pnas/suppl/2019/04/10/1812593116.DCSupplemental/pnas.1812593116.sapp.pdf.

[32] Wazir MA, Goujon A. Assessing the 2017 census of Pakistan using demographic analysis: A sub-national perspective. Vienna Institute of Demography Working Papers; 2019. Available at: https://www.oeaw.ac.at/fileadmin/subsites/Institute/VID/PDF/Publications/Working_Papers/WP2019_06.pdf.

[33] Dréze J, Sen A. Hunger and public action. Oxford: Clarendon Press; 1990.

[34] Guilmoto CZ, Chao F, Kulkarni PM. On the estimation of female births missing due to prenatal sex selection. Population studies. 2020;74(2):283–289.

[35] Chao F, KC S, Ombao H. Levels and trends in the sex ratio at birth in seven provinces of Nepal between 1980 and 2016 with probabilistic projections to 2050: a Bayesian modeling approach. arXiv preprint arXiv:200700437. 2020. Available at https://arxiv.org/abs/2007.00437.

[36] Chao F, Guilmoto CZ, Ombao H. 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. 2021;16(7):e0253721.