# Introduction

The sex ratio at birth (SRB) is a key component of population dynamics, as it has major implications for both future population structure and overall reproductive capacity. The natural equilibrium for SRB generally hovers around 1.05 males per female with some country-specific deviations [1]. Prior to the advent of ultrasound technologies and feasibility of sex-selective abortion, SRB remained stable over time. Since this period, however, remarkable shifts in SRB have manifested through systematic sex preferences for children, especially in the Caucasus, South Asia, and East Asia. To date, standardized methods have not been used to produce global estimates of SRB. To capture both historic equilibria and more recent deviations, we developed a model to estimate SRB in all national GBD locations, Hong Kong, and Macau from 1950 to 2017.

# Methods

## Data Sources

Registered live births by sex are the gold standard source for deriving SRB; however, their quality in some settings may suffer from sex-differentials in registration or reporting. We extracted total live births by sex from all United Nations Demographic Yearbook Reports back to 1950 and the most recent update to the WHO mortality database. Where VR data are poor or non-existent, precedent exists to use other source types from which SRB can be calculated, including surveys, censuses, and population registries [2]. To supplement registered births, we extracted sex ratios from all complete birth histories identified for the GBD 2017 analysis of age-specific fertility using a 25-year recall and averaging over five-year windows. We also extracted under-1 counts by sex from censuses and population registries, and where the former were unavailable, under-5 counts by sex from censuses and population registries. The GBD 2017 analysis of mortality did not show large or systematic enough sex-specific differences in infant or under-5 mortality to bias these sources. For this analysis, we could not, however, assess the effects of sex-specific differences in migration under-5. For India and China, respectively, where there have been documented periods of male preference but poor registry data, we utilized information available from the Sample Registration System and the 1% Population Sample Survey. In total, we extracted 4690 unique location years of registered births by sex, 1457 of under-1 counts by sex and 299 of under-5 counts by sex from censuses and population registries, and 2490 unique location -years form 446 complete birth histories.

## Modelling Approach

We estimated time trends in SRB using the GBD spatiotemporal Gaussian-process regression (ST-GPR) framework [3]. ST-GPR is three-step modelling approach whereby a linear prior is refined by weighting and adding residuals proximal in space and time and then finally fed into a Gaussian-process regression that takes into account both data and model variance to produce a robust measure of uncertainty. Data variance for all source types was calculated using a binomial distributional assumption. We then specified a linear prior for the mean as follows:

$$
\begin{aligned}
&logit\left(m_{ly}\right) = \alpha_r + \epsilon_{ly} \\
&\alpha_r \sim N(\gamma_{sr}, \sigma_\alpha^2) \\
&\gamma_{sr} \sim N(0, \sigma_\gamma^2)
\end{aligned}
$$

Where $m_{ly}$ is the proportion of live births that are male in a given location *l* and year *y* and $\gamma_{sr}$ and $\alpha_r$ are nested random intercepts by GBD super-region and region. We chose to model the proportion of male live births in lieu of the mathematically equivalent SRB to take advantage of the stability of estimation in transformed logit space. Additionally, we used a flat prior, as we would not expect SRB to change over time except in clear cases of sex-preference. We refined this prior with spatiotemporal smoothing and subsequently passed it to the GPR to generate final estimates of mean SRB. Hyperparameters for spatiotemporal smoothing and GPR for a location *l* were chosen on the basis of a data density score derived as follows:

$$dd_l = .5 * sourceYears_{l,CBH} + sourceYears_{l,census} + sourceYears_{lVR}$$

We down-weighted CBH relative to other source types due to the comparative magnitude of noise in the data apparent as a result of smaller sample sizes. Score thresholds for hyperparameter selection were chosen from their empirical distribution. Expert judgement was used to override automated hyperparameter selection, mostly in countries with small populations and medium data density. In these cases we took a more conservative estimation approach since small denominators introduced additional volatility in the data.

## Results

Averaging over trends prior to 1980, we observed that the natural equilibrium of SRB is markedly different by region and different from the common assumption of 1.05 males per female. East Asia had the highest natural equilibrium at 1.063 while Southern Sub-saharan Africa had the lowest natural equilibrium at 1.023, compared to a global equilibrium of 1.050. After this period, as countries started to adopt specific policies and practices related to sex-specific abortion, we started to see more individual departures in SRB compared to the regional baselines. The largest positive deviations from regional baselines were observed in the Caucasus, East Asia, and South Asia.

Both highest absolute SRB and the highest SRB relative to the regional baseline over the entire estimation period was observed in China in 2006 at 1.214, followed by Azerbaijan in 2008 at 1.162 and Armenia in 2003 at 1.158. A total of 7 countries have had a SRB over 1.100 at some point in the last forty years of their population's history, all of which have previously had SRBs over 1.100 for more than 5 consecutive years, implying sustained patterns of sex-selection. In 2017, the SRB of every country that had a peak SRB above 1.100 has moved back towards baseline, though SRBs in countries like India and China have remained high. Examining the relationship between SRB and the sociodemographic index, we notice that most large deviations occurred at middle levels of development, perhaps at a balance between access to ultrasound technology and the existence of legal mechanisms to limit sex-specific abortion. This also potentially foreshadows paths that could be taken countries that are currently graduating into these middle-levels of development. Last, we examined the relationship between the 20-year lagged sex ratio and net reproductive rate. As expected, we found that high SRBs in the past were correlated with lower net reproductive rates in the future (Pearson correlation -.2866). In particular, countries that had SRBs above 1.100 in the past almost all experienced net reproductive rates below 1 twenty years later. More time is required to assess whether instability introduced by high SRBs.

## Forthcoming Analyses

This paper is a work in progress. Over the coming months, we mainly intend to validate our measure of uncertainty. We are also eager to explore a number of other avenues. These include 1) the relationships between cohort SRBs and the sex ratios of those cohorts over time, 2) the correlation between SRB and the sex ratio of the population, and 3) the relationships between lags of SRB other than 20 years and the net reproductive rate. All these analyses are intended to try and address the important issue of whether massive shifts in SRB could have long-term consequences of population decline or collapse.

## References

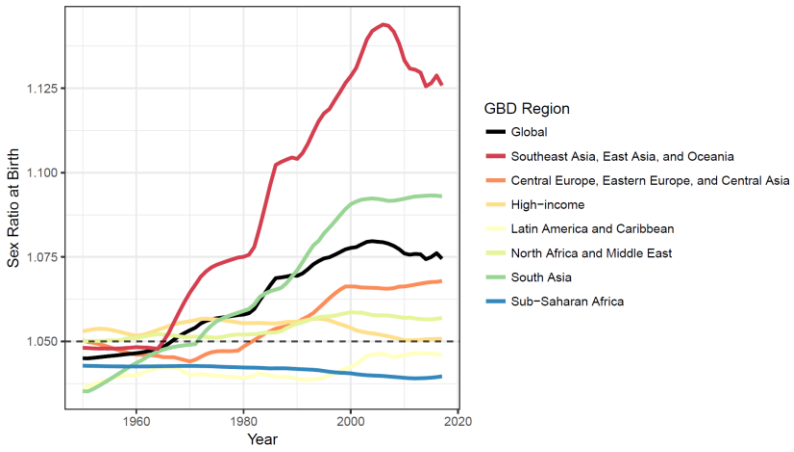

[1] Preston, S. H., Heuveline, P. and Guillot, M. Demography: Measuring and Modeling Population Processes. 2001 Malden: Blackwell.

[2] Wheldon MC, Raftery AE, Clark SJ, Gerland P. Bayesian reconstruction of two-sex populations by age: estimating sex ratios at birth and sex ratios of mortality. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2015 Oct 1;178(4):977-1007.

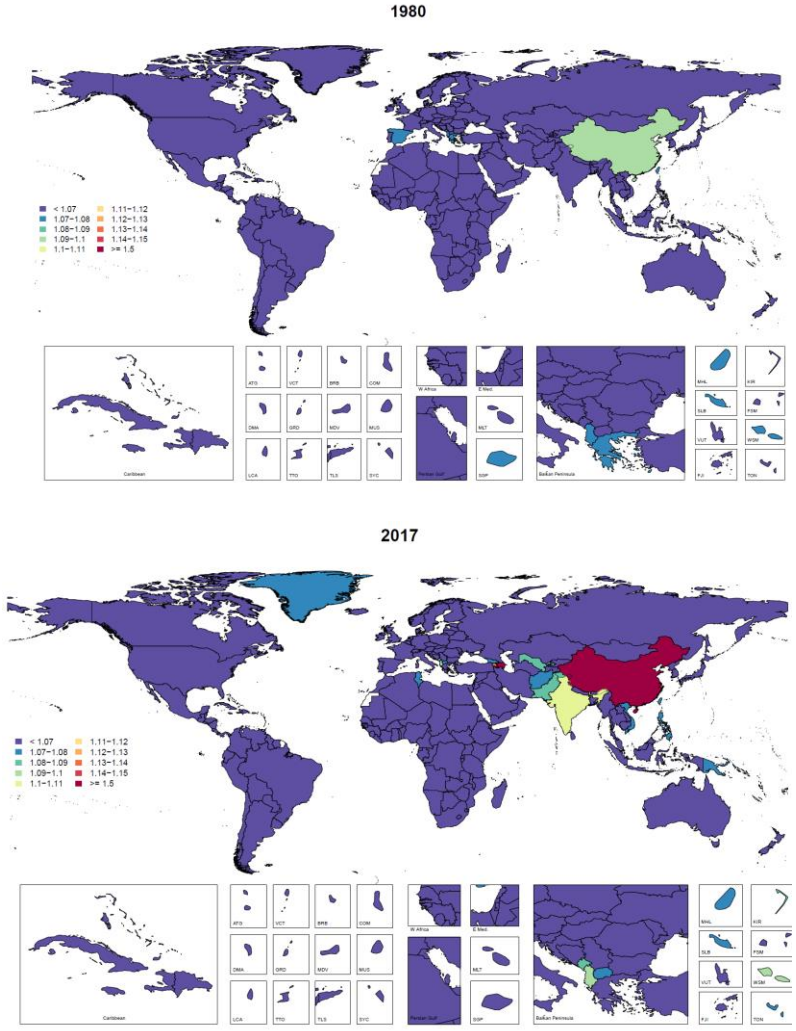
[3] Wang H, Abajobir AA, Abate KH, *et al.* Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* 2017; 390: 1084–150


## Figures

### Figure 1: Global and Regional Sex Ratios at Birth, 1950-2017



## Figure 2: Global Sex Ratio at Birth, 1980 (A) and 2017 (B)




## Figure 3: 20-year lagged SRB against Net Reproductive Rate