Sex ratio at birth estimation and projection for Pakistan provinces, a Bayesian modeling approach

Fengqing Chao[1] Muhammad Asif Wazir[2] Hernando Ombao[1]

[1]Statistics Program, CEMSE, KAUST [2]UNFPA–Pakistan

Introduction

Sex Ratio at Birth (SRB)

- SRB: ratio of male to female births;
- An important indicator:
  - For population estimation and projection;
  - To assess the prenatal gender equality.

SRB imbalance

- The natural SRB fluctuates between 103 to 107 male births per 100 female births;
- Since 1970, observed SRB in some Asian and Eastern European countries are much higher than the natural level.
- The imbalanced SRB is due to the coexistence of 3 main factors:
  1. Strong son preference at population level;
  2. Sex determine and abortion technology is accessible and affordable;
  3. Family size is getting smaller over time.
- On national level, my previous study (Chao, F. et al 2019) reports SRB inflation in 12 countries/areas.

Pakistan demography and SRB imbalance

- Pakistan is a country with strong preference of boys to girls.
- Prior evidence on the sex preference is mainly during postnatal stage: higher female child mortality than that among males (Alkema, L. et al 2014).
- Great heterogeneity of Pakistan demography.
- Gender discrimination in Pakistan at the prenatal stage, reflected in the inflated SRB has been barely mentioned before.

Objective

- Develop a Bayesian model for subnational SRB estimation in Pakistan.
- Estimate SRB across eight provinces in Pakistan during 1980–2018:
  - Balochistan, Khyber Pakhtunkhwa, Punjab, Sindh, Gilgit Baltista, Islamabad (ICT), Azad Jammu and Kashmir.

Contact

Twitter

- Webpage: https://www.fengqingchao.com/
- Feel free to reach out if you have any question regarding this study.

Results

Provincial SRB estimates We identify Balochistan with imbalanced SRB. Meanwhile SRBs in the other provinces (e.g. Khyber Pakhtunkhwa) are around normal level.

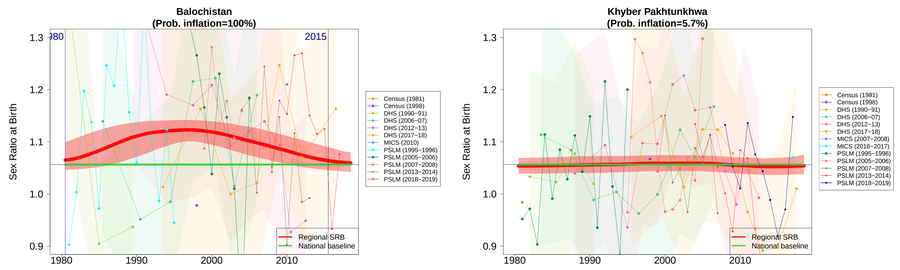

Figure 1: **SRB estimates by Pakistan province, 1980–2018.** The red line and shades are the median and 95% credible intervals of the region-specific SRB. The green horizontal line refers to the SRB baseline for the whole Pakistan at 1.056 (Chao, F. et al 2019). SRB observations are displayed with dots and observations are connected with lines when obtained from the same source. Shaded areas around observation series represent the sampling variability in the series (quantified by two times the sampling standard errors). The start and end years of sex ratio transition are shown in blue vertical lines.

SRB median estimates in 2000 and 2015, by Pakistan province

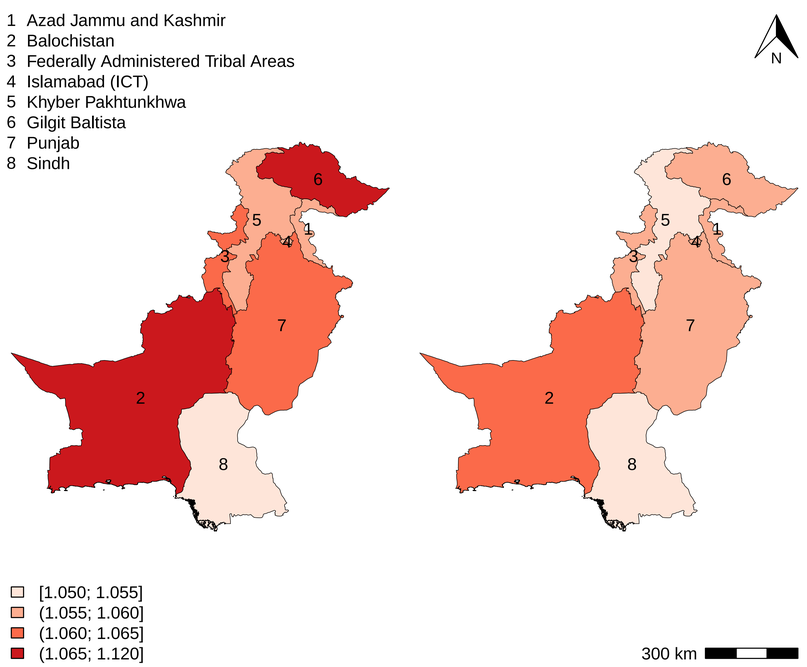

Figure 2: **Geographic disparity in SRB estimates, 2000 and 2015.** Median estimates are shown.

Method

Model for provincial SRB

Built upon and is based on the global SRB model (Chao, F. et al 2021).

$\Theta_{p,t}$, the SRB in Pakistan province p in year t is modeled as:

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

$b = 1.056$ is the SRB baseline level for the entire Pakistan (Chao, F. 2019).

$\Phi_{p,t}$ follows an AR(1) times series model on the log scale to capture the natural fluctuations of SRB within each province over time.

$$\begin{aligned} \log(\Phi_{p,t}) &\sim \mathcal{N}(0, (1-\rho^2)/\sigma_\epsilon^2), \text{ if } t = 1980, \\ \log(\Phi_{p,t}) &= \rho \log(\Phi_{p,t-1}) + \epsilon_{p,t}, \text{ if } t \in \{1981, \cdots, 2018\}, \\ \epsilon_{p,t} &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2). \end{aligned}$$

$\rho = 0.9$ and $\sigma_\epsilon = 0.004$ based on previous study (Chao, F. 2019).

δ_p is the binary identifier of the sex ratio transition:

$$\begin{aligned} \delta_p|\pi_p &\sim \mathcal{B}(\pi_p), \text{ for } p \in \{1, \cdots, 8\}, \\ \text{logit}(\pi_p)|\mu_\pi, \sigma_\pi &\sim \mathcal{N}(\mu_\pi, \sigma_\pi^2), \text{ for } p \in \{1, \cdots, 8\}. \end{aligned}$$

$\alpha_{p,t}$ refers to the province-specific SRB imbalance process to represent the increase, stagnation, and decrease of the transition stages:

$$\alpha_{p,t} = \begin{cases} (\xi_p/\lambda_{1p})(t-\gamma_p), & t_{0p} < t < t_{1p} \\ \xi_p, & t_{1p} < t < t_{2p} \\ \xi_p - (\xi_p/\lambda_{3p})(t-t_{2p}), & t_{2p} < t < t_{3p} \\ 0, & t < t_{0p} \text{ or } t > t_{3p} \end{cases}$$

$$t_{1p} = t_{0p} + \lambda_{1p}, t_{2p} = t_{1p} + \lambda_{2p}, t_{3p} = t_{2p} + \lambda_{3p},$$

$$t_{0p} \sim \mathcal{U}(1970, 2050), \text{ for } p \in \{1, \cdots, 8\}.$$

Data Model For the i-th observation r_i:

$$\log(r_i)|\Theta_{p[i],t[i]} \sim \mathcal{N}(\log(\Theta_{p[i],t[i]}), \sigma_i^2), \text{ for } i \in \{1, \cdots, 507\},$$

σ_i^2 is the known sampling error variance for $\log(r_i)$.

References

- Chao, F., Gerland, P., Cook, A. R., Alkema, L. (2019). Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *PNAS, 116*(19), 9303-9311.
- Alkema, L., Chao, F., You, D., Pedersen, J. and 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), pp.e521-e530.
- Chao, F., Gerland, P., Cook, A.R. and Alkema, L., 2020. Global estimation and scenario-based projections of sex ratio at birth and missing female births using a Bayesian hierarchical time series mixture model. (forthcoming in AOAS) *arXiv preprint arXiv:2006.07101*.

2021 PAA virtual meeting. Session P7 – Gender and Sexuality; and Health and Health Behavior