Sex Ratio at Birth by Vietnamese Region Estimation and Projection, a Bayesian modeling approach

Fengqing Chao

Statistics Program, King Abdullah University of Science and Technology (KAUST)

Session 56 Time Series Modeling
ENAR Spring Meeting (virtual)

Mar 16th, 2021

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Background

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 biological/natural SRB fluctuates within a narrow band between 1.03–1.07 (i.e. 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.

Background

SRB imbalance (cont.)

- The imbalanced (usually upward inflated) 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, Chao et al. *PNAS* 2019 reports 12 places with SRB inflation:
  - Albania; Armenia; Azerbaijan; China; Georgia; Hong Kong; India; Republic of Korea; Montenegro; Taiwan; Tunisia; **Vietnam**.

Background

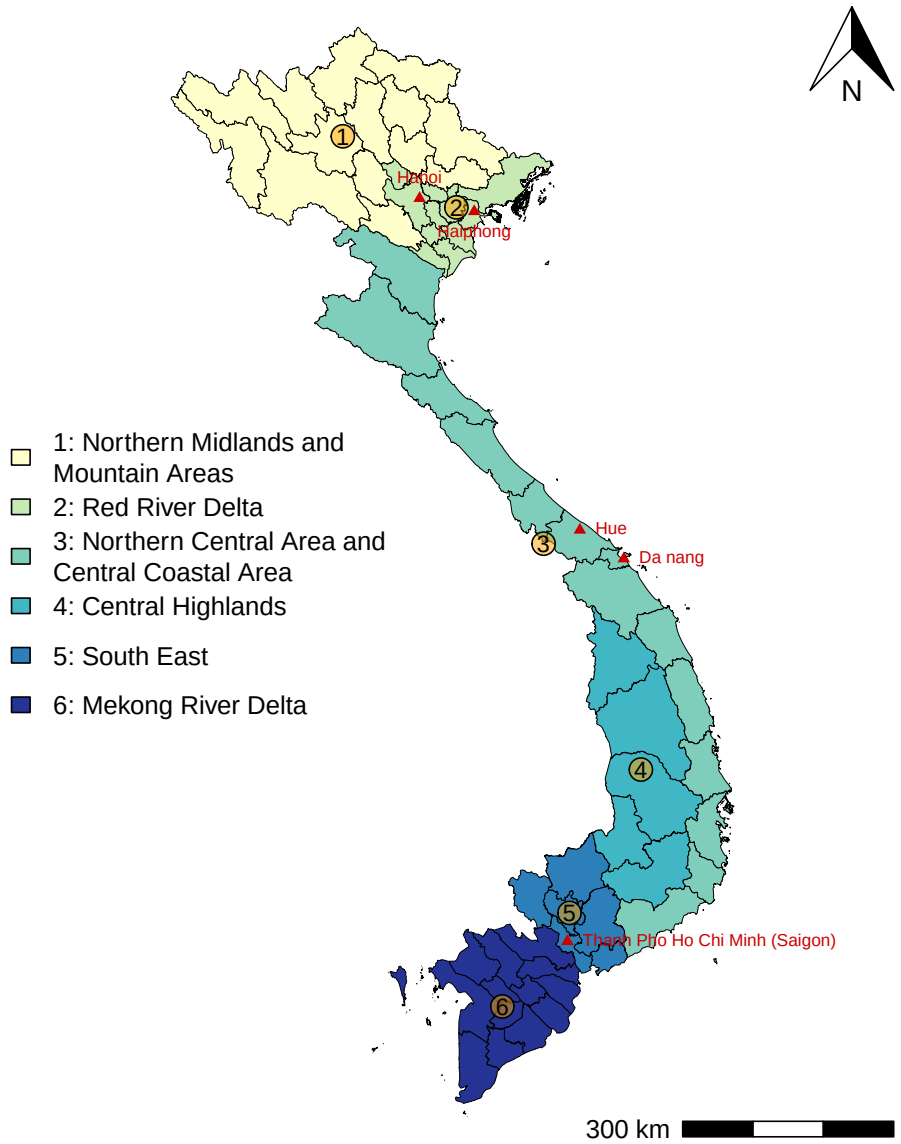
Vietnam demography and SRB imbalance

- Great heterogeneity of Vietnamese culture across regions:
  - North: strict patrilineal family system, influenced by China.
  - South: less acute need for a male child felt by couples, due to the *Khmer* culture and bilateral kinship sytem.
- SRB imbalance in Vietnam began in 2001 (later than most Asian countries).
- Important to estimate and project SRB and assess the SRB imbalance by Vietnam sub-national region.

Objectives

- Estimate and project SRB for six Vietnam regions during 1980–2050.
- Identify Vietnam regions with imbalanced SRB.
- Preprint: Chao F, Guilmoto CZ, Ombao H. *SocArXiv*, 2021. doi:10.31235/osf.io/9xrbk.

Vietnamese regions and provinces matching



Data

- We compiled an extensive database for Vietnam regional SRB from surveys and censuses.
- 526 SRB observations.
- Reference year: 1972–2019.
- More than 2,933,093 births records are included in the database.[1]

[1] Number of births in some data sources are unknown. The total number is the sum of known numbers.

Data quality model

We assume the ith observed SRB y_i follows a normal distribution on the log-scale. For $i \in \{1, \cdots, 526\}$,

$$\log(y_i)|\Theta_{r[i],t[i]}, \omega \sim \mathcal{N}(\log(\Theta_{r[i],t[i]}), v_i^2 + \omega^2),$$

where

- i indexes all the SRB observations across regions over time.
- y_i: the ith observed SRB from region $r[i]$ in year $t[i]$.
- $\Theta_{r,t}$: the true SRB for Vietnamese region r in year t.
- v_i^2: known stochastic/sampling variance.
- ω^2: unknown non-sampling error variance (assign a vague prior to ω).

Model for $\Theta_{r,t}$

$\Theta_{r,t}$, the true SRB, is modeled as:

$$\Theta_{r,t} = b\Phi_{r,t} + \delta_r \alpha_{r,t},$$

where

- b: the SRB baseline level for the entire Vietnam fixed at 1.063, based on Chao et al. *PNAS* 2019.
- $\Phi_{r,t}$: capture the natural fluctuations of SRB within each region over time.
- δ_r: region-specific binary identifier of the sex ratio transition.
- $\alpha_{r,t}$: region-specific SRB imbalance process.

Model for $\Theta_{r,t}$

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

$$\begin{aligned}
\log(\Phi_{r,t}) &\sim \mathcal{N}(0, (1-\rho^2)/\sigma_\epsilon^2), \text{ if } t = 1980, \\
\log(\Phi_{r,t}) &= \rho \log(\Phi_{r,t-1}) + \epsilon_{r,t}, \text{ if } t \in \{1981, \cdots, 2050\}, \\
\epsilon_{r,t} &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2),
\end{aligned}$$

where $\rho = 0.9$ and $\sigma_\epsilon = 0.004$ based on Chao et al. *PNAS* 2019.

Model for $\Theta_{r,t}$

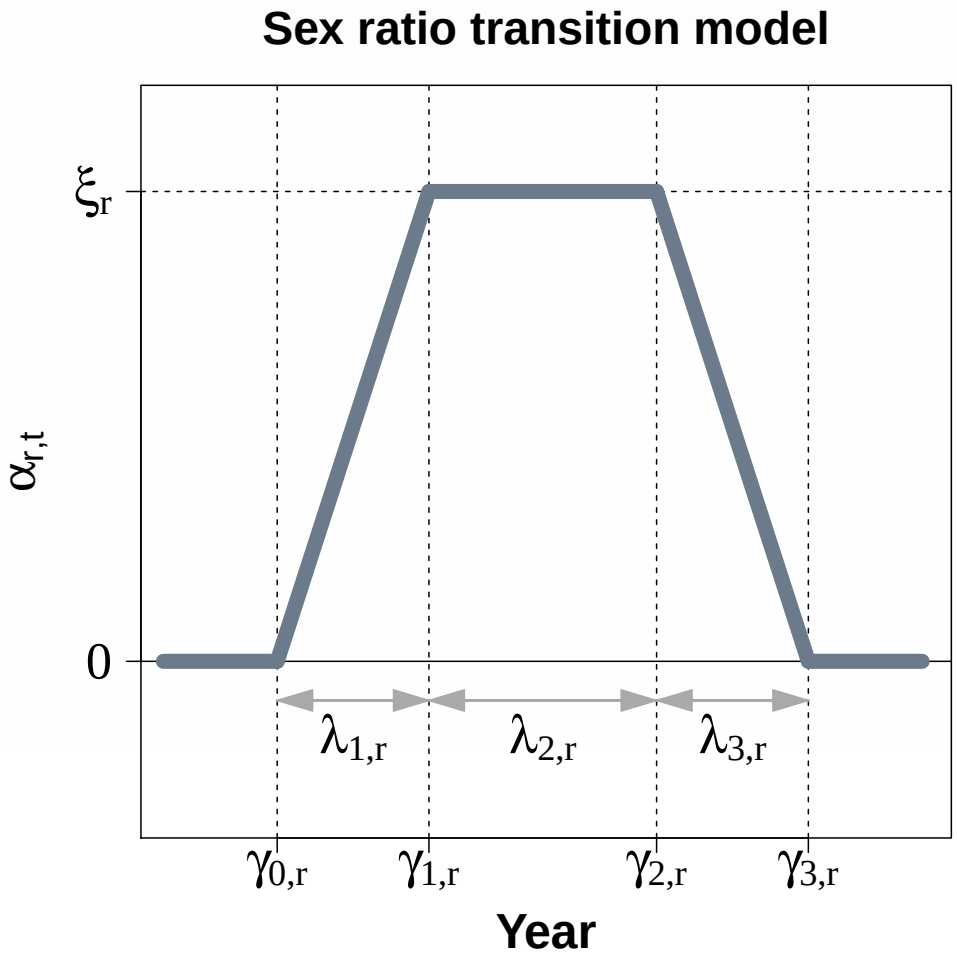
δ_r is the binary identifier of the sex ratio transition, following a Bernoulli distribution:

$$\begin{aligned} \delta_r|\pi_r &\sim \mathcal{B}(\pi_r), \text{ for } r \in \{1, \cdots, 6\}, \\ \text{logit}(\pi_r)|\mu_\pi, \sigma_\pi &\sim \mathcal{N}(\mu_\pi, \sigma_\pi^2), \text{ for } r \in \{1, \cdots, 6\}. \end{aligned}$$

Vague priors are assigned to μ_π and σ_π.

Model for $\Theta_{r,t}$

SRB imbalance process $\alpha_{r,t}$ is modeled with a trapezoid function with hierarchical structure.



- $\gamma_{0,r}$: starting year of inflation period.
- $\lambda_{1,r}, \lambda_{2,r}, \lambda_{3,r}$: period lengths of increase, stagnation and decrease of inflation.
- ξ_r: the maximum value that the adjustment factor could reach.

Model for $\Theta_{r,t}$

The sex ratio transition process $\Omega_{c,t}$ for country c year t is modeled as:

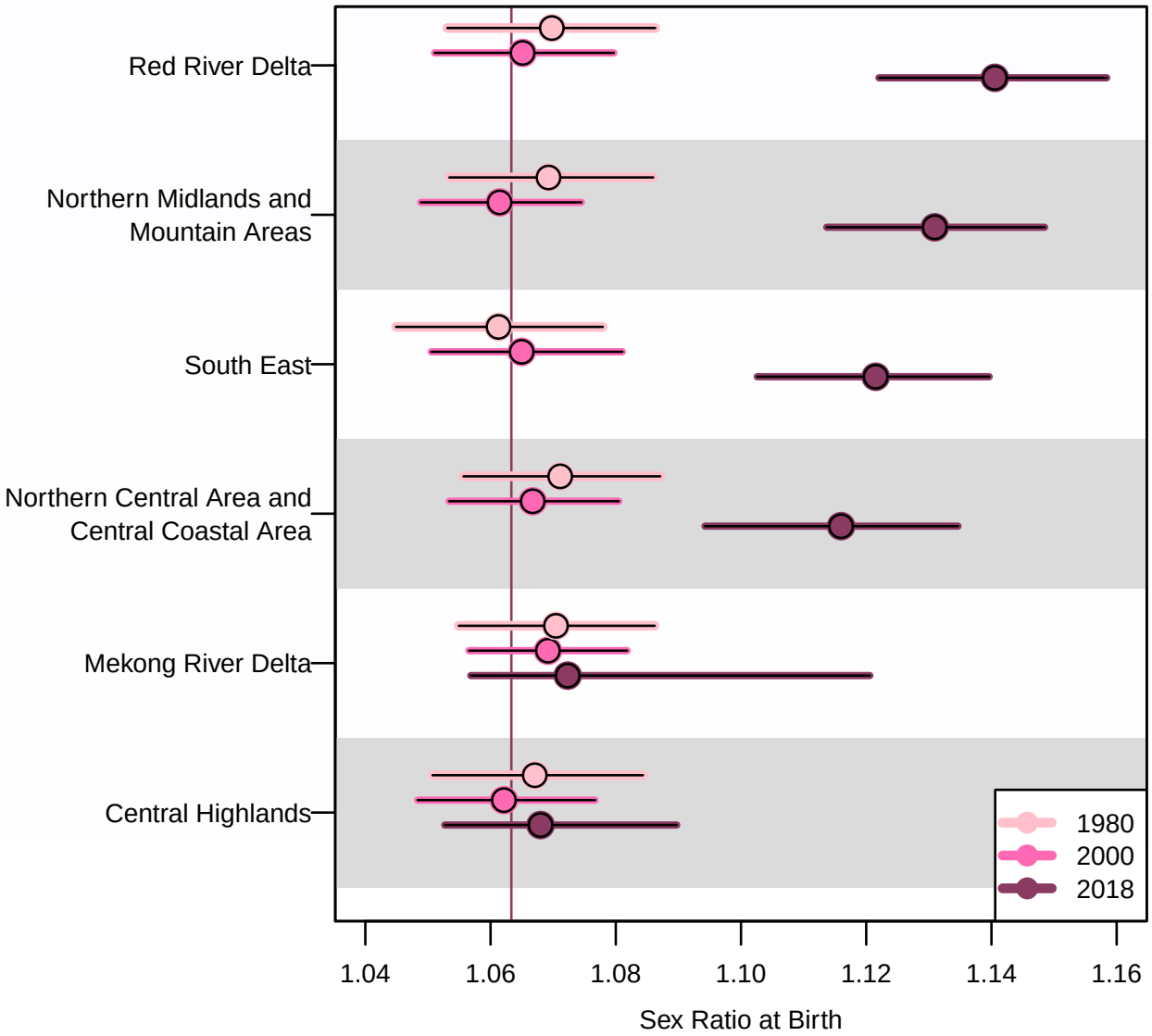
$$\alpha_{r,t} = \begin{cases} (\xi_r/\lambda_{1,r})(t-\gamma_r), & \gamma_{0,r} < t < \gamma_{1,r} \\ \xi_r, & \gamma_{1,r} < t < \gamma_{2,r} \\ \xi_c - (\xi_c/\lambda_{3,r})(t-\gamma_{2,r}), & \gamma_{2,r} < t < \gamma_{3,r} \\ 0, & t < \gamma_{0,r} \text{ or } t > \gamma_{3,r} \end{cases},$$

where $\gamma_{1,r}, \gamma_{2,r}, \gamma_{3,r}$ can be expressed by $\gamma_{0,r}, \lambda_{1,r}, \lambda_{2,r}, \lambda_{3,r}$. Hierarchical distribution and informative priors are assigned to the following regional-level parameters:

$$\begin{aligned} \gamma_{0,r}|\sigma_\gamma &\sim t_3(2001, \sigma_\gamma^2), \\ \xi_r &\sim \mathcal{N}(\mu_\xi, \sigma_\xi^2)T(0,), \\ \lambda_{m,r} &\sim \mathcal{N}(\mu_{\lambda m}, \sigma_{\lambda m}^2)T(0,), \text{for } m \in \{1,2,3\}. \end{aligned}$$

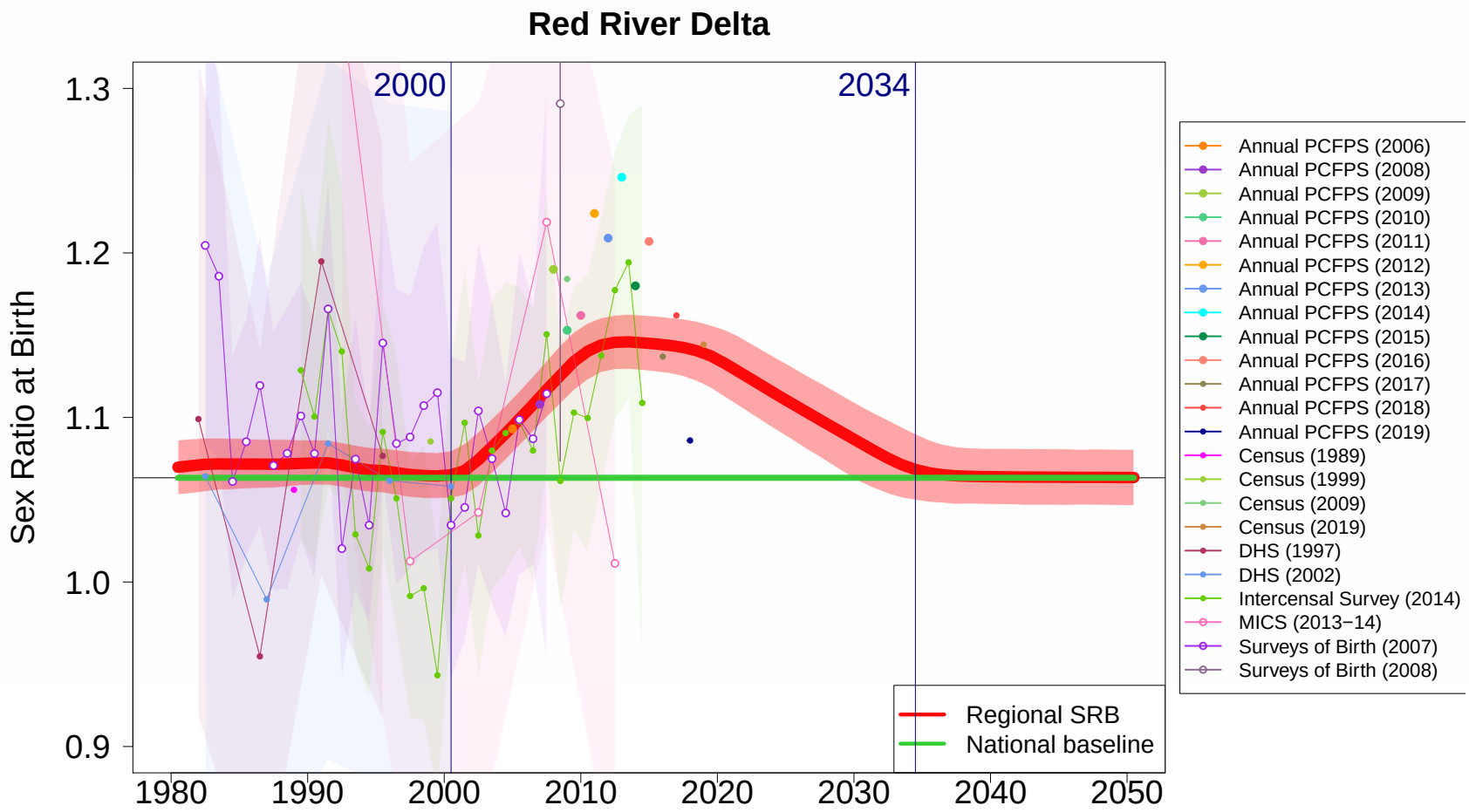
Means and variances of the ξ_r and λ's follows the national level of the Vietnam SRB inflation process, based on Chao et al. *AOAS* 2021 (forthcoming).

SRB estimates by Vietnam regions



Vietnamese regional estimates and projections

Region with imbalanced SRB



Summary

Conclusion

- We estimate and project SRB by Vietnamese region during 1980–2050.
- We identify regions with imbalanced SRB.

Co-authors

- Christophe Z. Guilmoto, CEPED/IRD, Université de Paris.
- Hernando Ombao, Statistics Program, KAUST.

For more details, refer to our preprint:

Chao, F., Guilmoto, C. Z., Ombao, H. (2021). Sex ratio at birth in Vietnam among six subnational regions during 1990–2050, estimation and probabilistic projection using a Bayesian hierarchical time series model. *SocArXiv*, doi:10.31235/osf.io/9xrbk.

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. *Proceedings of the National Academy of Sciences*, 116(19), 9303-9311.

Chao, F., Guilmoto, C. Z., Ombao, H. (2021). Sex ratio at birth in Vietnam among six subnational regions during 1990–2050, estimation and probabilistic projection using a Bayesian hierarchical time series model. *SocArXiv*, doi:10.31235/osf.io/9xrbk.

Chao, F., Gerland, P., Cook, A. R., Alkema, L. (2021, forthcoming). 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*, arXiv:2006.07101.