Scenario-Based Bayesian Probabilistic Projections of the Sex Ratio at Birth and Missing Female Births

Fengqing Chao
fengqingchao.com

King Abdullah University of Science and Technology (KAUST)

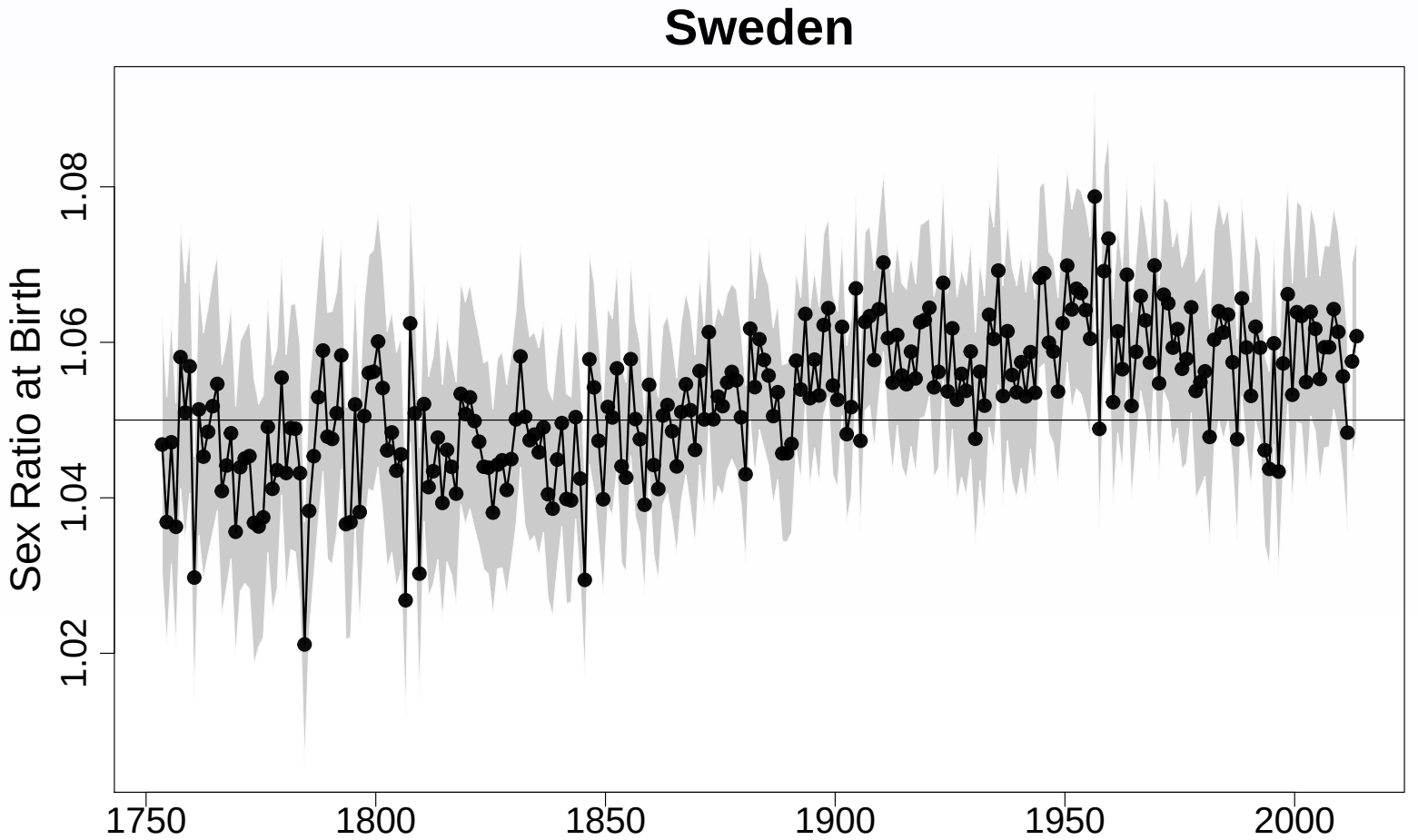
JSM 2020 (virtual)
Statistical Demography, Social Statistics Section
Aug 4th, 2020

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

STATISTICS @KAUST

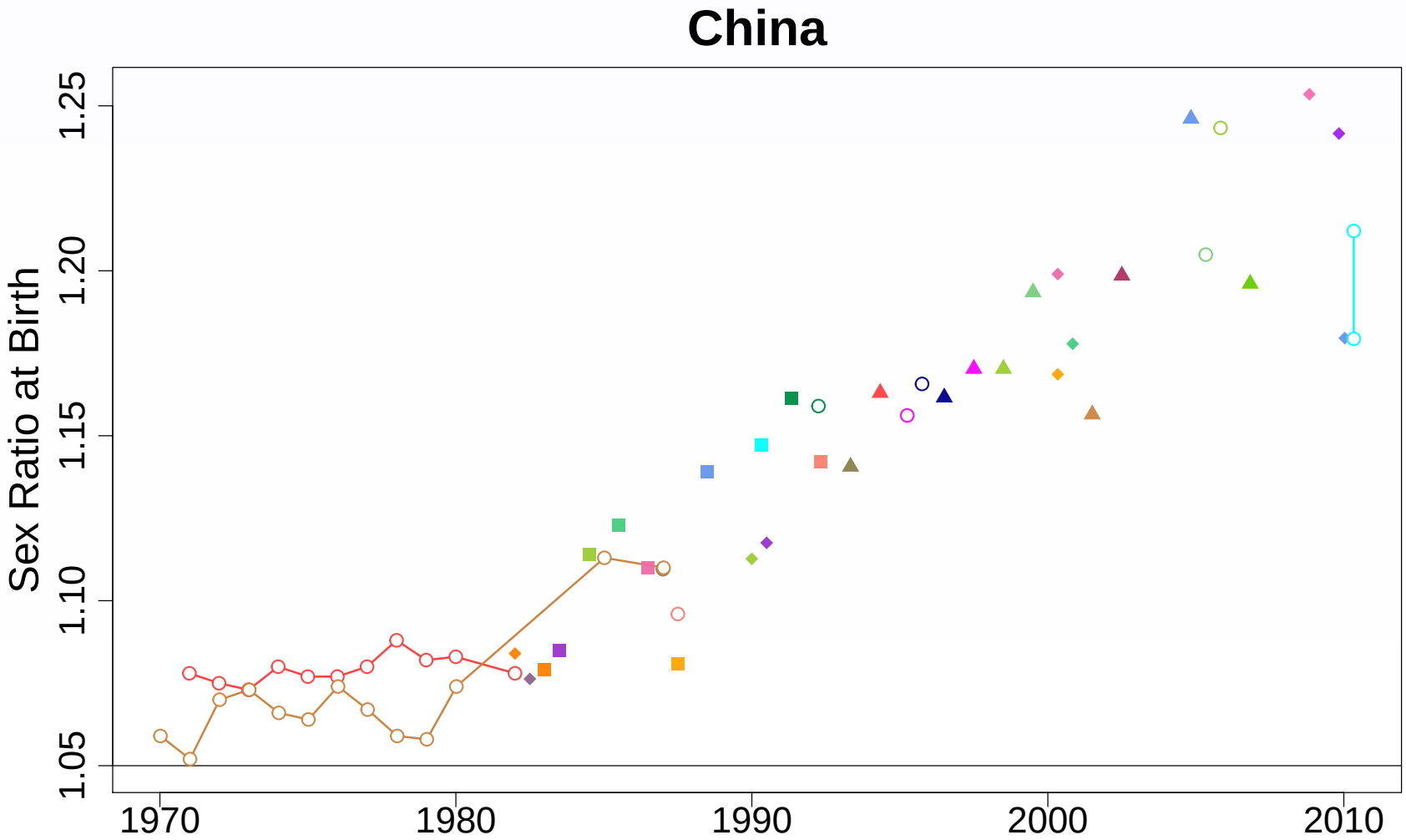
Sex Ratio at Birth (SRB)

Sex Ratio at Birth (SRB): ratio of male to female live births. The SRB biological (baseline) level varies between 1.03 and 1.07 with minor difference across ethnicity groups.



Imbalanced SRB since 1970

Observed SRB from many countries in Asia and Eastern Europe have been well above the baseline level for decades.



Prior work on SRB estimation

In our prior work[1] [Chao et al. (2019)], we developed a model to estimate SRB and imbalances for 212 countries from 1950 to 2017:

- Accounts for the difference in the SRB reference levels across regions and varying uncertainty associated with SRB observations.
- Fitted a model for SRB levels and trends in country-years without risk of sex-selective abortion and obtained estimates for national and regional SRB baseline values.
- Estimated sex imbalances using a sex ratio transition model to capture periods of increasing, constant and decreasing stages.

[1] Chao, F., Gerland, P., Cook, A. R., and 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.

Research gap and objectives

Research gap

- Existing SRB projections are deterministic and based on expert-based opinions.
- No work to date has quantified the possible additional SRB imbalances for countries where sex ratio transitions may start in the future.

Objectives

- Construct a Bayesian model to project SRB till 2100 for all countries[2], based on different scenarios of the occurrence of a sex ratio transition.
- To project missing female births as the effect of SRB imbalance due to sex-selective abortion.

[2]We use the term "country" to refer to populations that are considered as countries or areas in the UN classification.

Data

Data type	# obs.	(% of total)
Census	48	(0.4)
DHS	2,257	(21)
Other DHS[3]	1,392	(9)
Others	142	(1)
CRVS/SRS	6,996	(65)
total	**10,835**	(100)

Table 1: **SRB Database**. DHS: Demographic and Health Surveys. CRVS: civil registration vital statistics. SRS: sampling registration system. Available at here.

- Available for 202 countries.
- 51 observations from each of the 212 countries estimated.

[3]Non-standard DHS, including Special, Interim and National DHS, Malaria Indicator Surveys, AIDS Indicator Surveys, World Fertility Surveys, Reproductive Health Survey, Multiple Indicator Cluster Surveys, Pan Arab Project for Child Development and Pan Arab Project for Family Health.

Data quality model

Data model $\Theta_{c,t}$ is the true SRB for country c in year t. We assume the i-th observed SRB y_i follows a normal distribution on the log-scale:

$$\log(y_i)|\Theta_{c[i],t[i]}, \omega_{s[i]} \sim \mathcal{N}\left(\log(\Theta_{c[i],t[i]}), \omega^2_{s[i]} + v_i^2\right).$$

The variance for the i-th log-scaled SRB observation $\log(y_i)$ is the sum of:

- v_i^2: known stochastic/sampling variance.
- $\omega^2_{s[i]}$: unknown non-sampling error variance for data source type $s[i]$.

$$\omega_s = 0, \text{ for } s = \text{CRVS/SRS},$$
$$\omega_s \overset{\text{iid}}{\sim} \mathcal{U}(0, 0.5), \text{ for } s \in \{\text{Census, DHS, Other DHS, Other}\}.$$

Method main idea

We use two models to estimate SRB $\Theta_{c,t}$ for two groups of countries/areas:

- **Basic model**: for country-years **without** SRB inflation;
- **Extended model**: for selected country-years **with** past/current/potential future SRB inflation.

Method details are available at Chao et al. (2020a):
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. *arXiv preprint arXiv:2006.07101*.

Model setup

Model for country-years without SRB inflation We assume the true value of SRB $\Theta_{c,t}$ to be the product of national biological norm β_c and multiplier $\eta_{c,t}$:

$$\begin{aligned} \Theta_{c,t} &= \beta_c \eta_{c,t}, \\ \log(\beta_c)|\beta_{r[c]}^{(\text{region})}, \sigma_\beta &\sim \mathcal{N}\left(\log(\beta_{r[c]}^{(\text{region})}), \sigma_\beta^2\right). \end{aligned}$$

β_r is estimated by fitting the model with reduced database, with uniform priors: $\beta_r \overset{\text{iid}}{\sim} \mathcal{U}(1, 1.1)$.
$\eta_{c,t}$ is estimated by an AR(1) time series model:

$$\begin{aligned} \log(\eta_{c,t})|\phi &\sim \mathcal{N}(0, (1-\rho^2)/\sigma_\epsilon^2), \text{ for } t = 1950, \\ \log(\eta_{c,t}) &= \rho \log(\eta_{c,t-1}) + \epsilon_{c,t}, \text{ for } t \in \{1951, \cdots, 2100\}, \\ \epsilon_{c,t}|\sigma_\epsilon &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2). \end{aligned}$$

Non-informative priors were assigned to ρ and σ_ϵ.

Select countries with SRB inflation

Model for country-years at risk of SRB inflation

$$\Theta_{c,t} = \beta_c \eta_{c,t} + \delta_c \Omega_{c,t}.$$

- Selection criteria for countries/areas with potential SRB inflation:
  - observed SRB is suspected to be beyond biological norm as supported by literature; OR
  - desired sex ratio at birth > 120 and/or sex ratio of last birth > 130 suggested in [Bongaarts (2013)].
- 29 selected countries/areas at risk of SRB inflation:
  - **Asia (17)** : Afghanistan, Armenia, Azerbaijan, Bangladesh, China, Georgia, Hong Kong (China), India, Jordan, Korea Rep, Nepal, Pakistan, Singapore, Taiwan (China), Tajikistan, Turkey, Vietnam;
  - **SSA (7)**: Gambia, Mali, Mauritania, Nigeria, Senegal, Tanzania, Uganda;
  - **elsewhere (5)**: Albania, Egypt, Montenegro, Morocco, Tunisia.

SRB inflation indicator

Model for country-years at risk of SRB inflation

$$\Theta_{c,t} = \beta_c \eta_{c,t} + \delta_c \Omega_{c,t}.$$

- δ_c: a binary country-specific indicator indicating presence or absence of SRB inflation.
- $\Omega_{c,t}$: a non-negative sex ratio transition process.

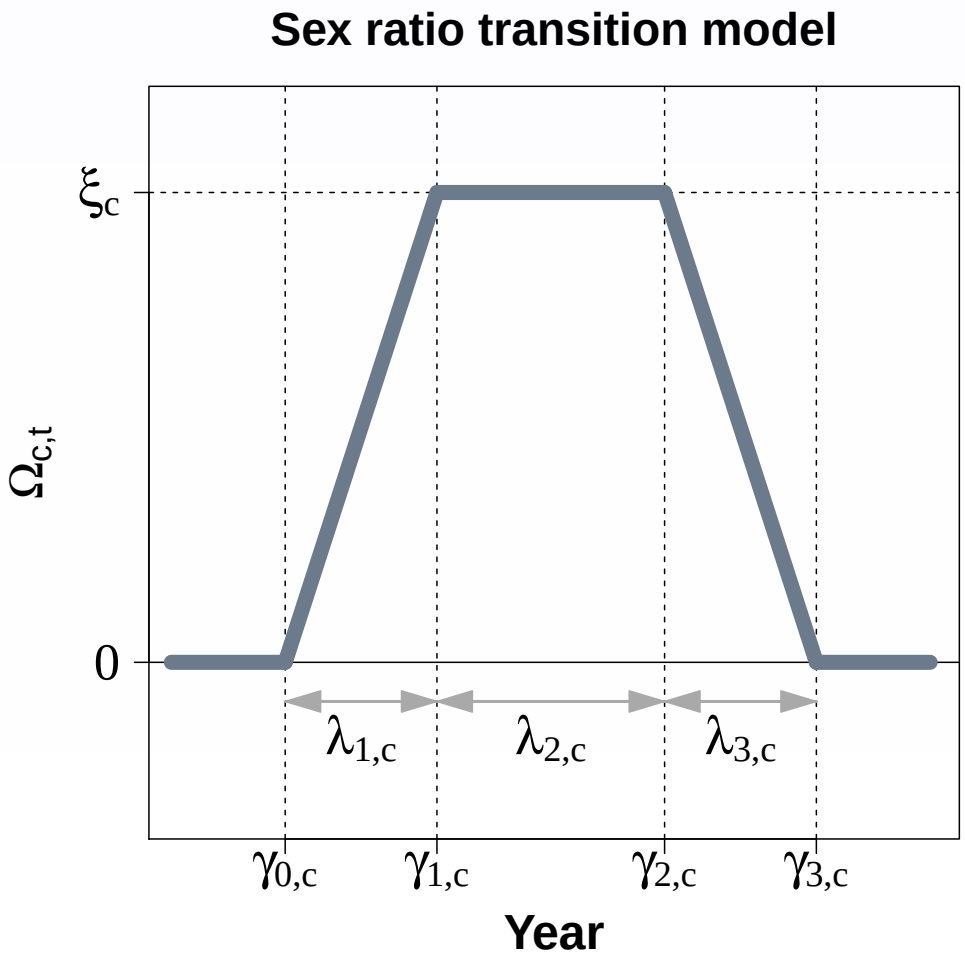
δ_c is modeled with a Bernoulli distribution with value 0 (no inflation) or 1 (with inflation):

$$\begin{aligned} \delta_c|\pi_c &\sim \mathcal{B}(\pi_c), \\ \text{logit}(\pi_c)|\mu_\pi, \sigma_\pi &\sim \mathcal{N}(\mu_\pi, \sigma_\pi^2). \end{aligned}$$

Vague priors are assigned to σ_π and μ_π.

Sex ratio transition process

For an at-risk country c, $\Omega_{c,t}$ is modeled with a trapezoid function with hierarchical structure.



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

Sex ratio transition process

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

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

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

$$\begin{aligned} \gamma_{0,c}|\sigma_\gamma &\sim t_3(x_c, \sigma_\gamma^2)T(z_c,), \\ \xi_c|\mu_\xi, \sigma_\xi &\sim \mathcal{N}(\mu_\xi, \sigma_\xi^2)T(0,), \\ \lambda_{1,c}|\mu_{\lambda 1}, \sigma_{\lambda 1} &\sim \mathcal{N}(\mu_{\lambda 1}, \sigma_{\lambda 1}^2)T(0,), \\ \lambda_{2,c}|\mu_{\lambda 2}, \sigma_{\lambda 2} &\sim \mathcal{N}(\mu_{\lambda 2}, \sigma_{\lambda 2}^2)T(0,), \\ \lambda_{3,c}|\mu_{\lambda 3}, \sigma_{\lambda 3} &\sim \mathcal{N}(\mu_{\lambda 3}, \sigma_{\lambda 3}^2)T(0,). \end{aligned}$$

Non-informative priors are assigned to hyper-parameters.

Scenario-based projections

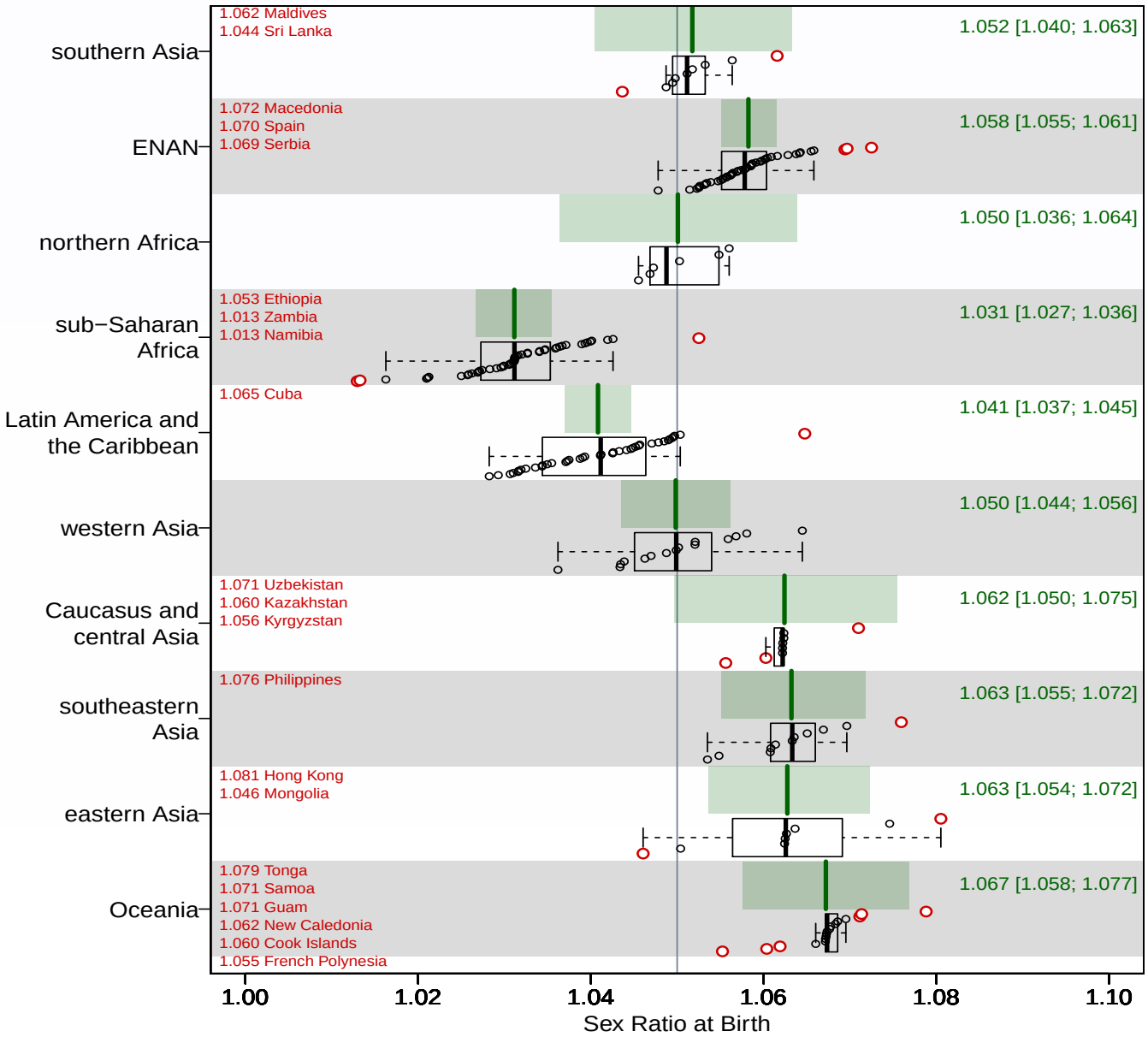
We construct scenario-based SRB projections based on the sex ratio transition model with varying assumptions regarding the occurrence of a sex ratio transition in at-risk countries:

- **Scenario 1** no inflation ($\delta_c = 0$): $\Theta_{c,t} = \beta_c \eta_{c,t}$.
- **Scenario 2** country-specific inflation probability: $\Theta_{c,t} = \beta_c \eta_{c,t} + \delta_c \Omega_{c,t}$.
- **Scenario 3** SRB inflation to occur ($\delta_c = 1$): $\Theta_{c,t} = \beta_c \eta_{c,t} + \Omega_{c,t}$.

In all scenarios, we use $\beta_c \eta_{c,t} + \delta_c \Omega_{c,t}$ to project SRB for countries with strong statistical evidence of ongoing inflation:

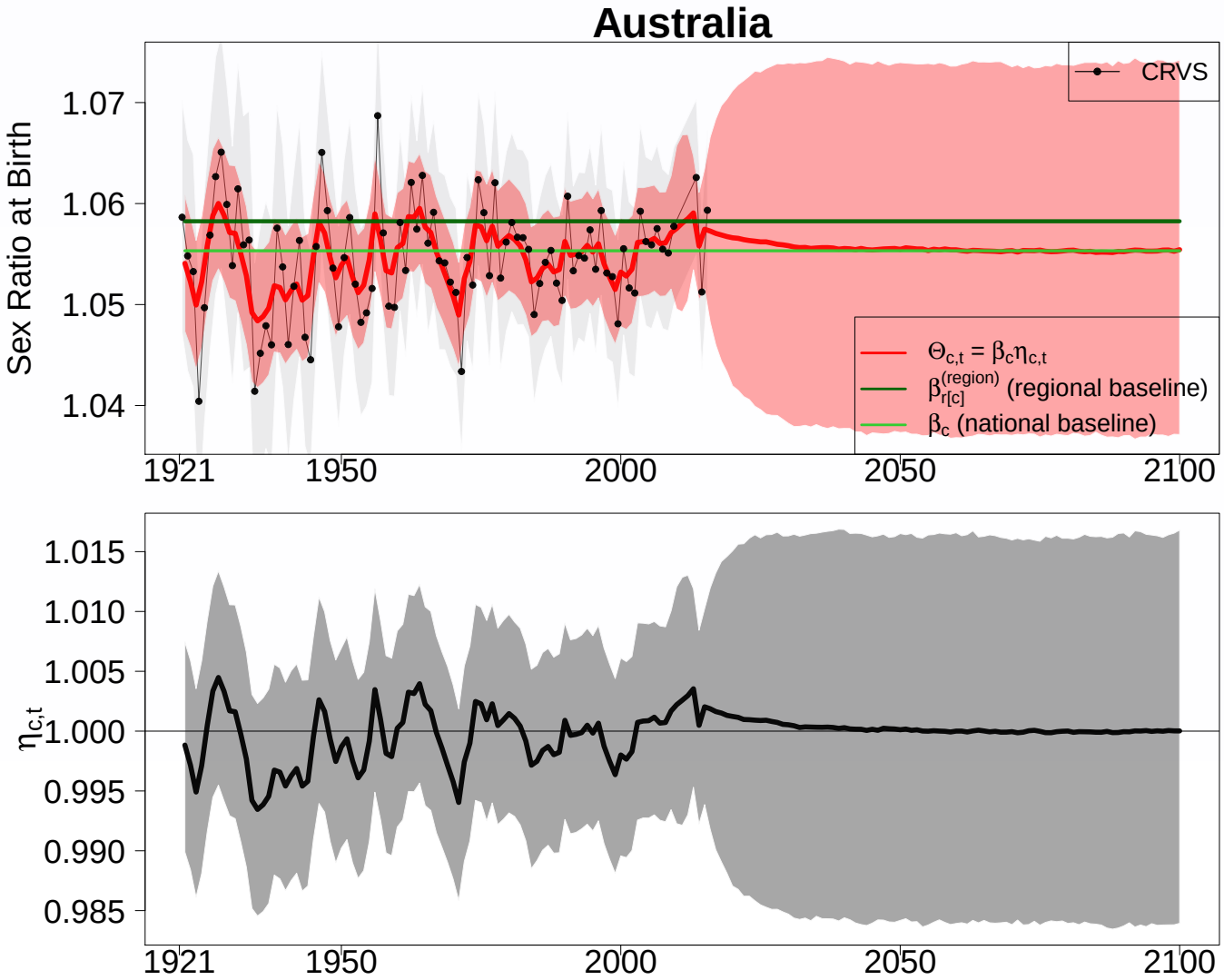
- **12 countries/areas**: Albania, Armenia, Azerbaijan, China, Georgia, Hong Kong (China), India, Republic of Korea, Montenegro, Taiwan (China), Tunisia, and Vietnam.

SRB regional and national baselines



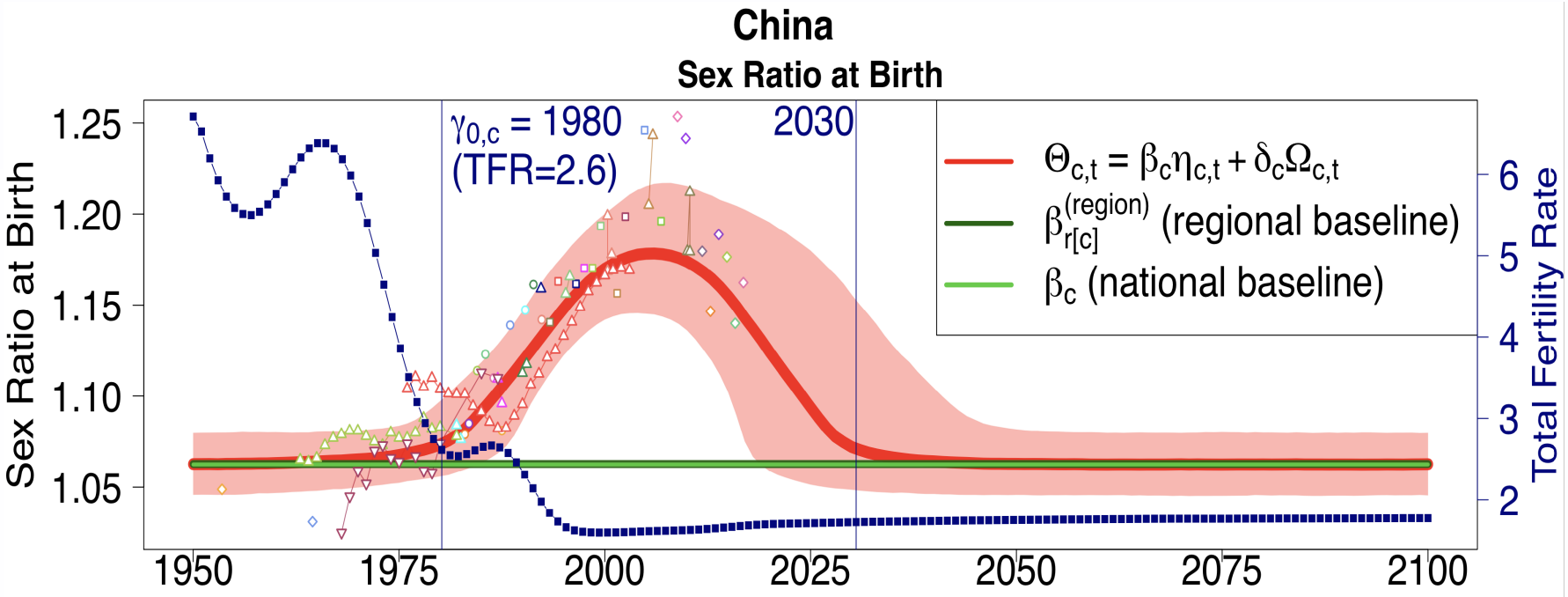
Country results without SRB inflation

For country-year **without** SRB inflation: $\Theta_{c,t} = \beta_c \eta_{c,t}$



Country results with SRB inflation

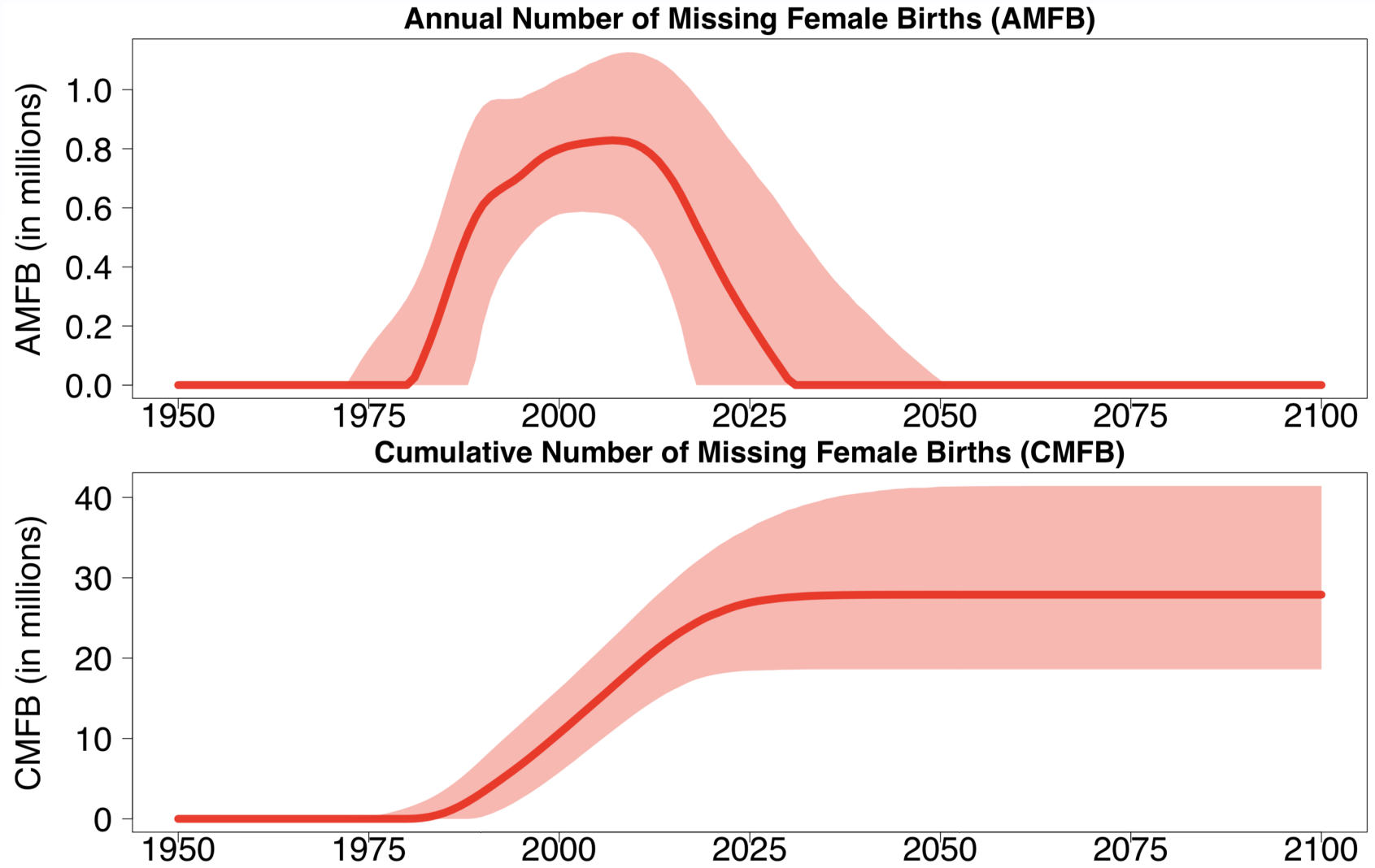
For country-year **with** SRB inflation: $\Theta_{c,t} = \beta_c \eta_{c,t} + \delta_c \Omega_{c,t}$



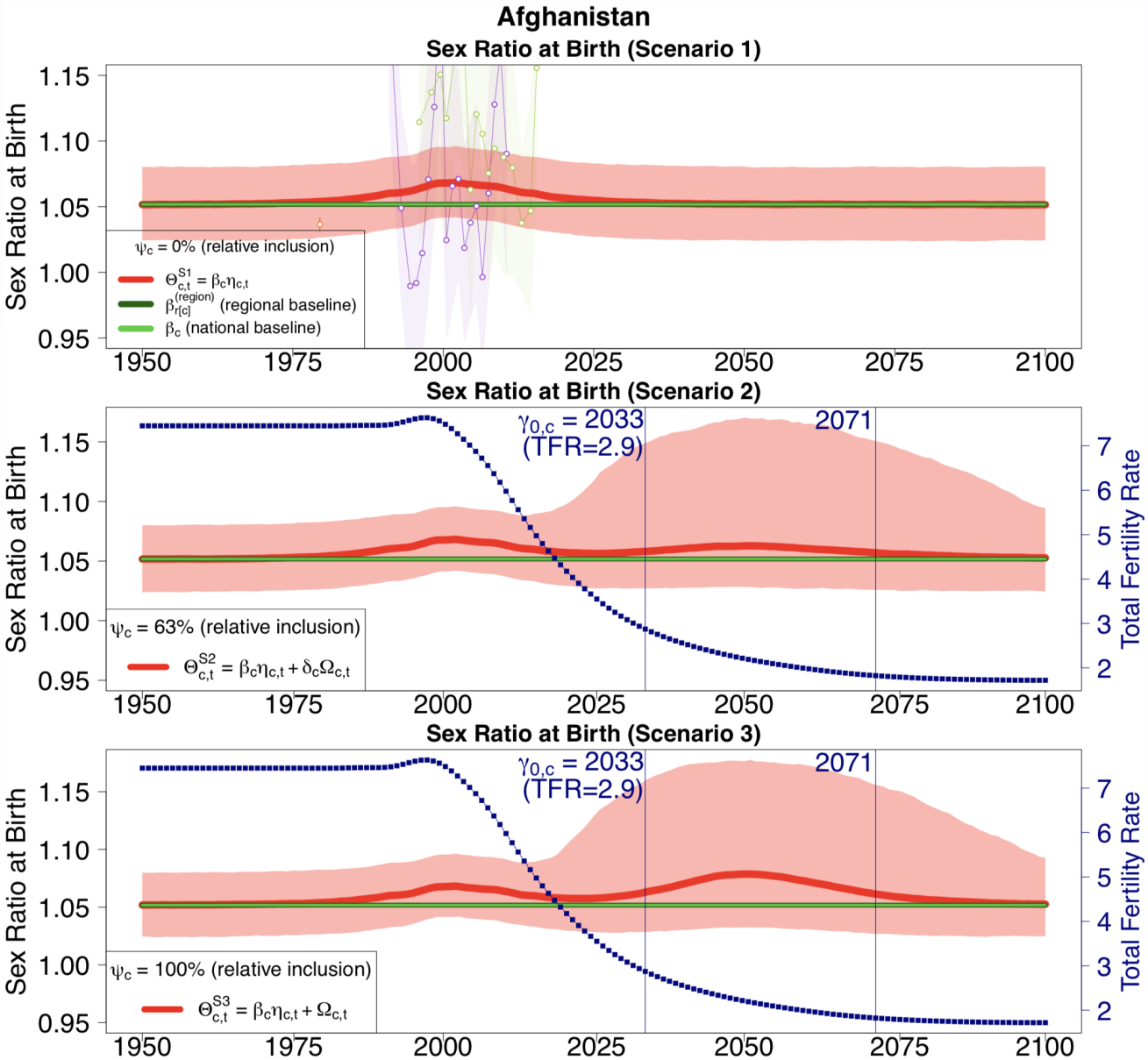
The median estimates and projections of total fertility rate (TFR) are from the UN WPP 2019 (blue squared dots). The median estimates of inflation start year $\gamma_{0,c}$ and end year $\gamma_{3,c}$ are the vertical lines. The TFR value in the year $\gamma_{0,c}$ is shown.

Country results with SRB inflation

Annual and cumulative missing female births in China:



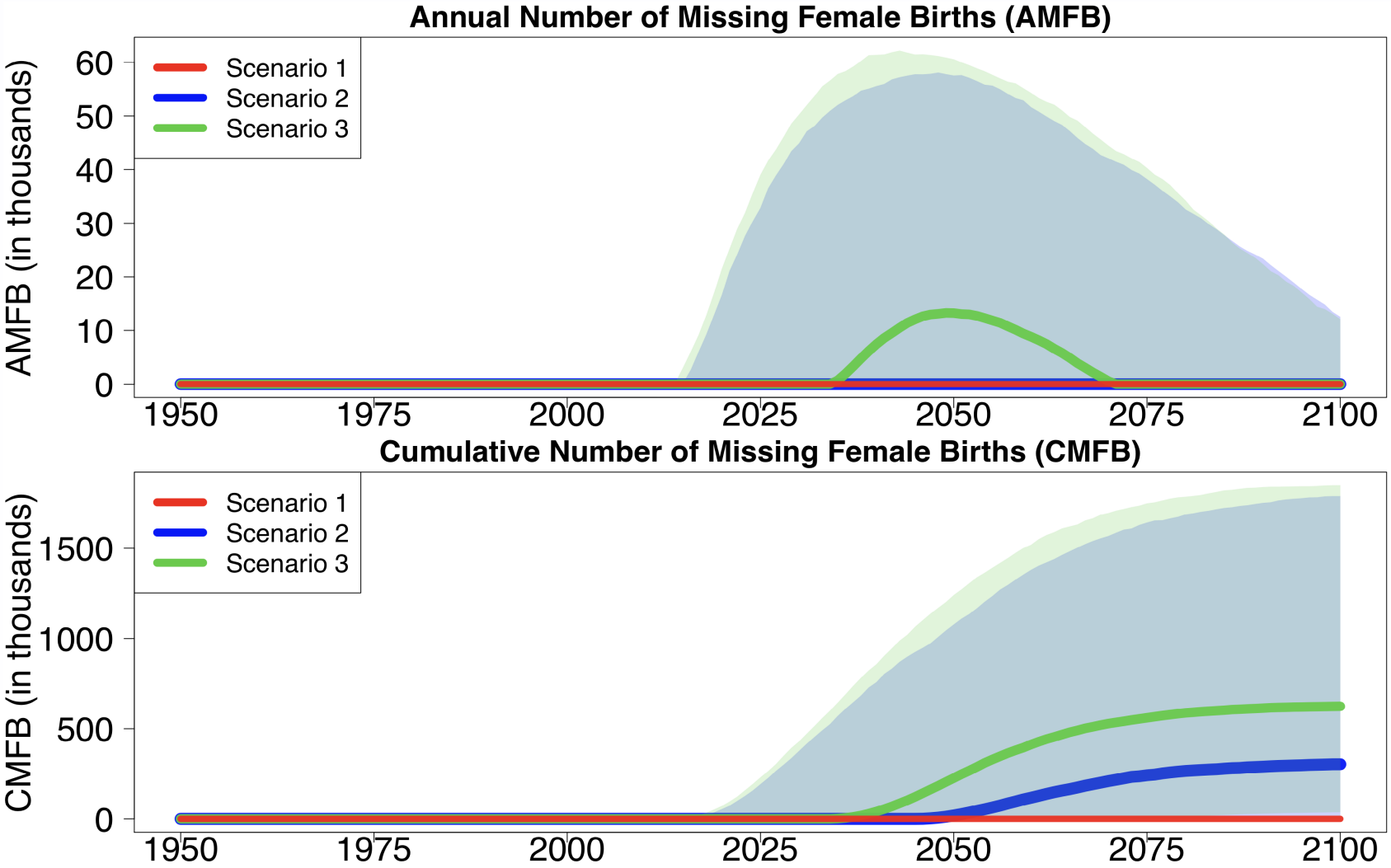
Scenario-based projection: SRB



Scenario-based projection: missing female births

KAUST

Annual and cumulative missing female births in Afghanistan:



Summary

Conclusion

- We described a Bayesian hierarchical time series model for producing SRB estimates and scenario-based projections for all countries from 1950 to 2100.
- We provide scenario-based projections that quantify the SRB inflation:
  - S1: no inflation.
  - S2: country-specific probability of having inflation.
  - S3: inflation will occur with probability 1.

Future research

- Same/similar model can be applied to estimate and project SRB and sex ratio transition on subnational levels.
- E.g. SRB estimation and projection by Nepal province [Chao et al. (2020b)].

Acknowledgment

Collaborators This is a joint work with:

- **Patrick Gerland** from United Nations Population Division, DESA, United Nations.
- **Alex Cook** from Saw Swee Hock School of Public Health, National University of Singapore and National University Health System.
- **Leontine Alkema** from School of Public Health and Health Sciences, University of Massachusetts, Amherst.

Funding This work was supported by a research grant from the National University of Singapore.

References

Bongaarts, J. (2013). The Implementation of Preferences for Male Offspring. *Population and Development Review*, 39(2):185–208.

Chao, F., Gerland, P., Cook, A. R., and 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., Gerland, P., Cook, A. R., and Alkema, L. (2020a). Global estimation and scenario-based projections of sex ratio at birth and missing female births using a bayesian hierarchical time series mixture model. *arXiv preprint arXiv:2006.07101*.

Chao, F., KC, S., and Ombao, H. (2020b). 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:2007.00437*.

For more details, refer to our preprint:

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. *arXiv preprint arXiv:2006.07101*.

Thank you!

Fengqing Chao

fengqingchao.com

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