

Estimating and Projecting Prenatal Sex Discrimination around the World and on Subnational Level in Asia

Fengqing Chao, PhD

Research Scientist, King Abdullah University of Science and Technology (KAUST)

18 Aug 2022

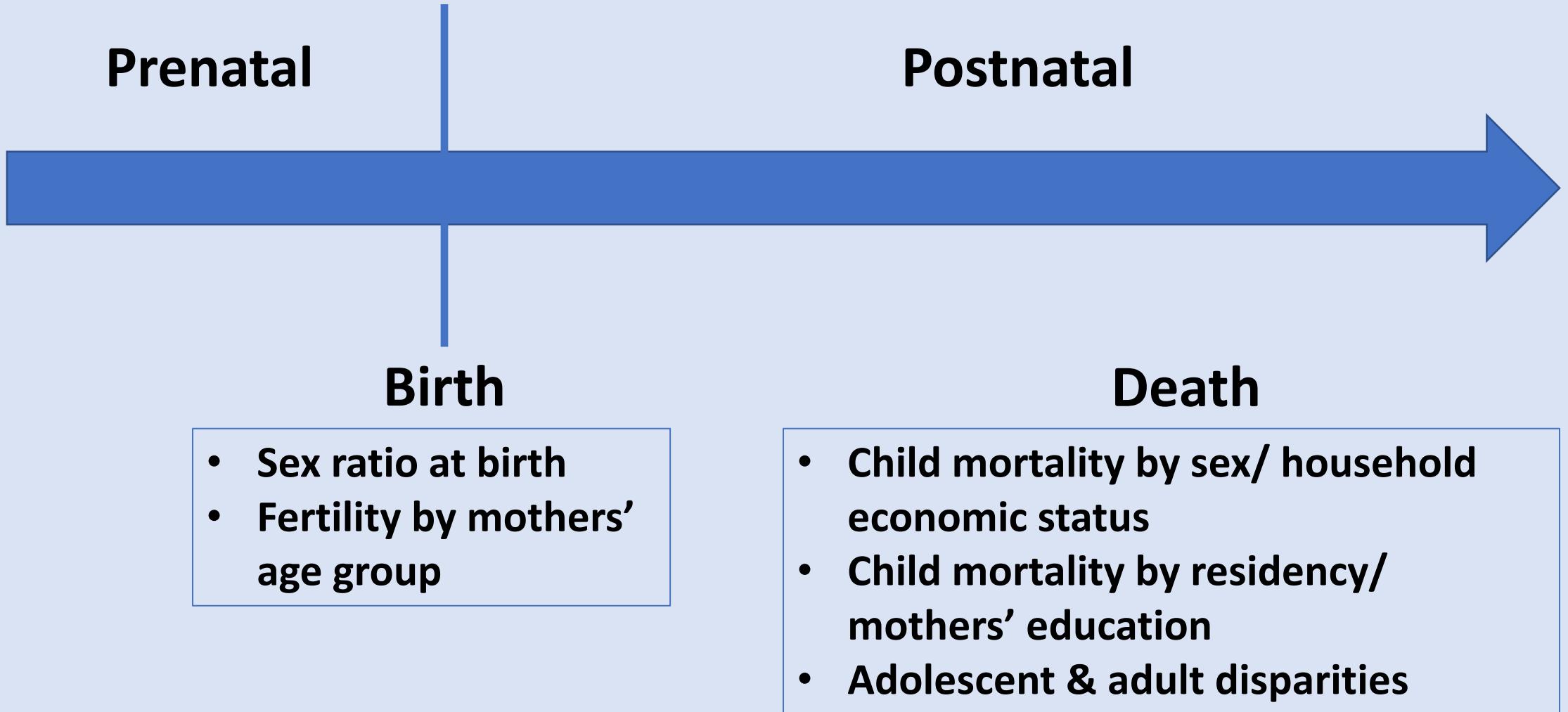
Research talk, Population Association of Singapore



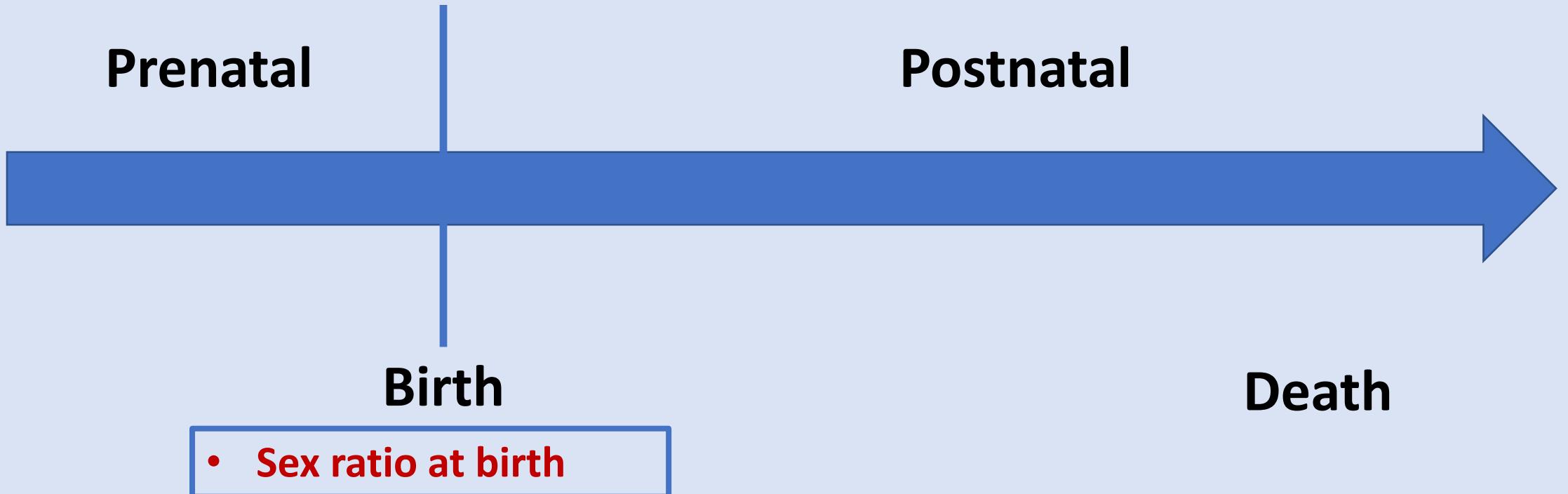
Research Areas

- Statistical Demography
- Global health

Current Research: Disparity in Prenatal and Postnatal Survival



A Boy or a Girl? Sex Ratio at Birth and Prenatal Sex Discrimination

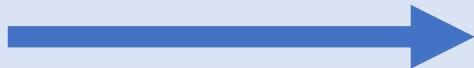


Sex Ratio at Birth (SRB) - It is Not 50/50

Naturally

100
Female
Births

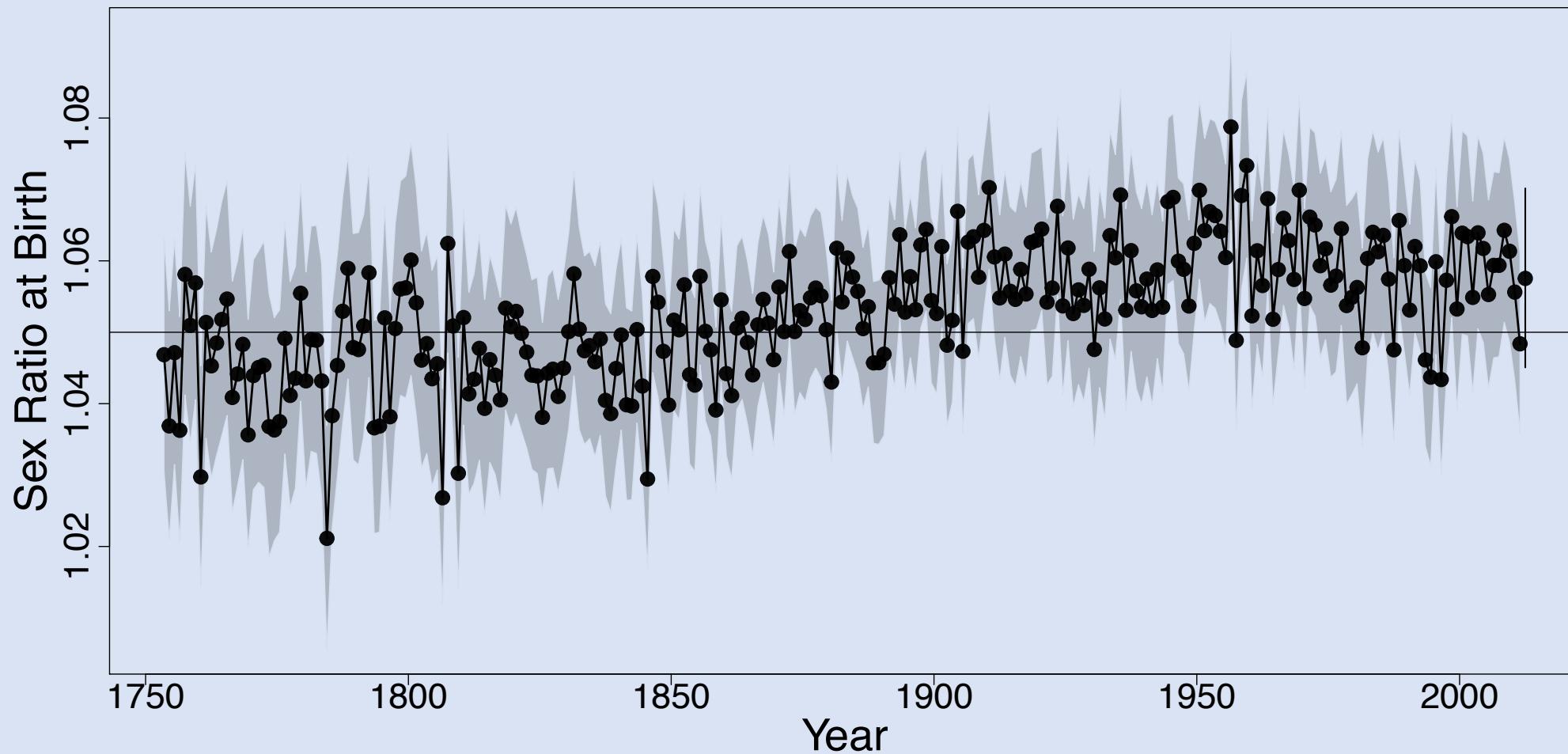
103~107
Male
Births



SRB = Boys/Girls
1.03~1.07

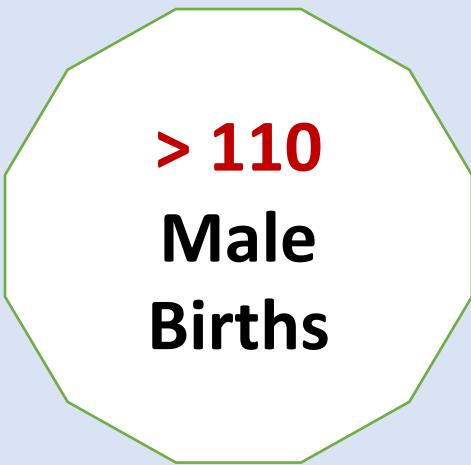
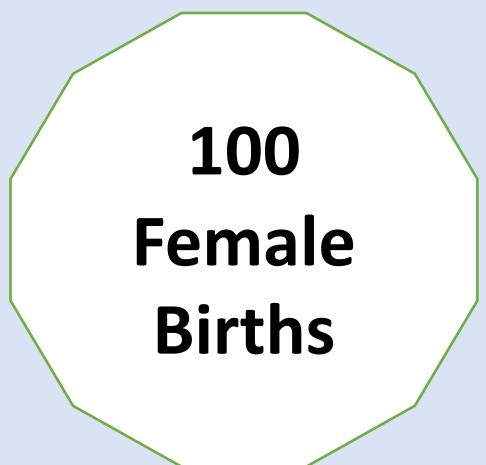
Natural SRB 1.03~1.07

Sweden



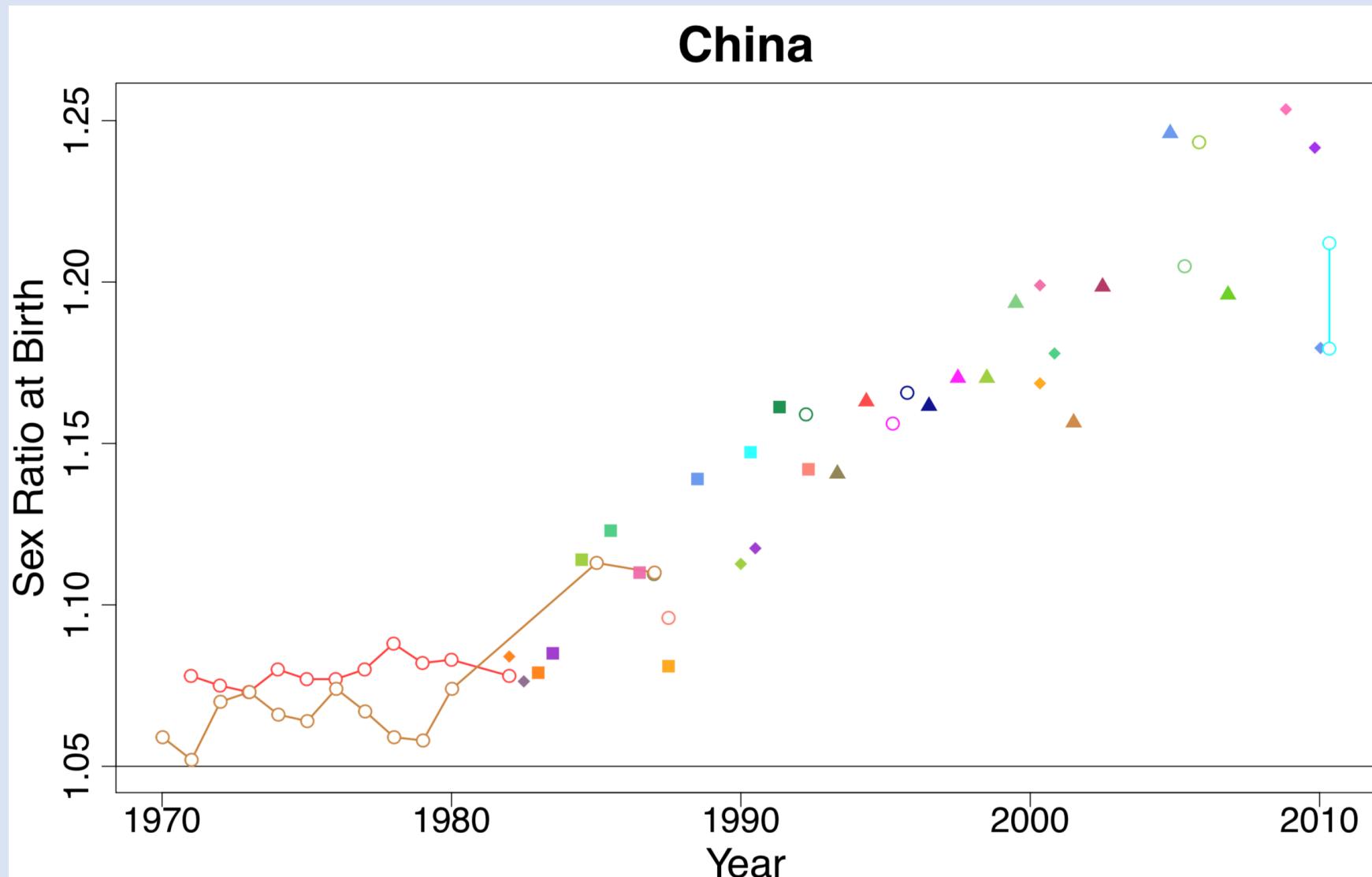
Inflated SRB in Some Countries

In reality,
in some countries



SRB:
>1.10

Inflated SRB in Some Countries



Sex Ratio at Birth (SRB) – Why the Inflation?

Naturally



100
Female
Births

**SRB = boys/girls
1.03~1.07**

Willingness
Son preference

Necessary
Fertility decline

Means
Abortion + Sex
detection

**Sex-selective
abortion**

In reality,
in some countries



<94
Female
Births

**SRB:
>1.10**

Sex Ratio at Birth (SRB) – A Distorted Reality

- Serious social consequences with

prolonged distorted SRB:

- Human trafficking
- Marriage squeeze
- Violation of human right

• Breaks population sex balance at the beginning of the life course:

- Missing female births due to sex selection



40,800 female births doomed in Vietnam every year

By Minh Nga July 19, 2020 | 05:24 pm GMT+7



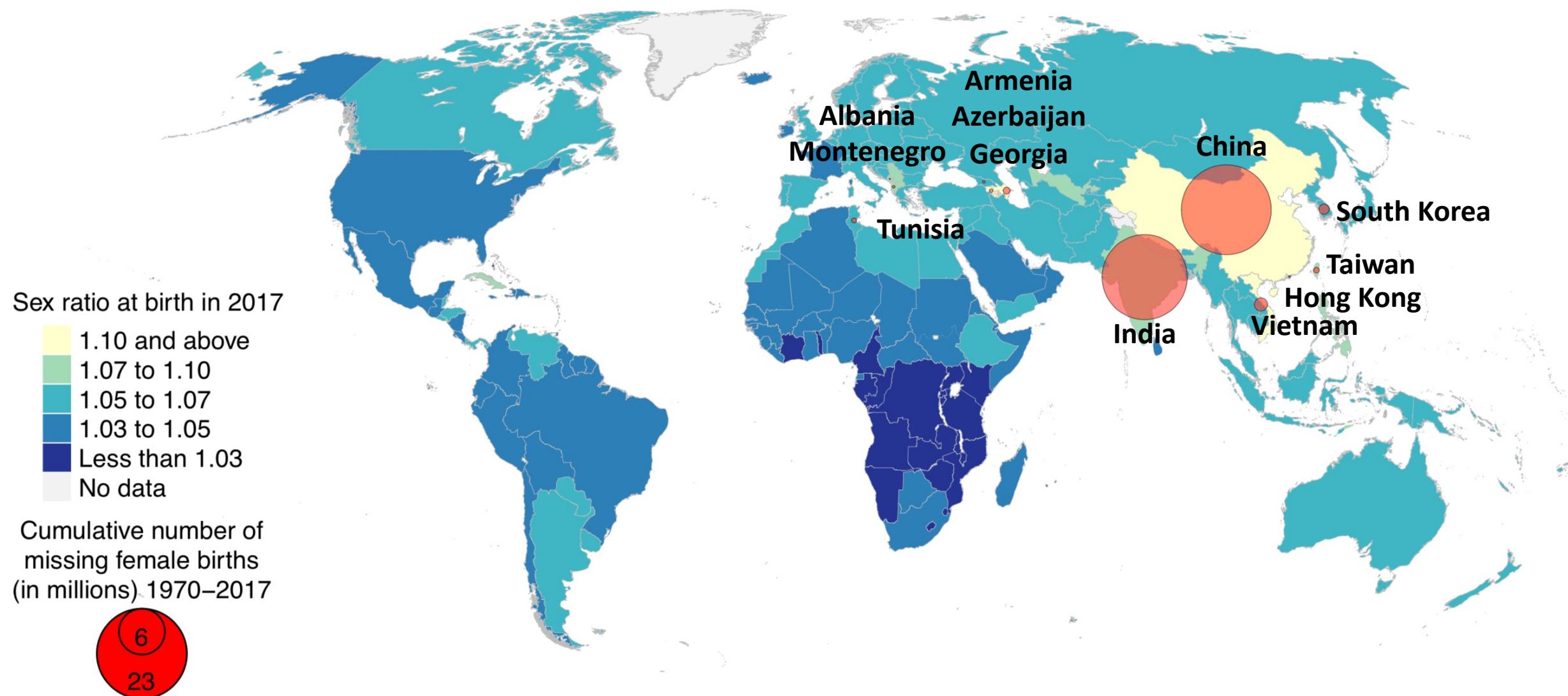
Why many Indian and Chinese men may need to delay marriage or remain bachelors

Radheshyam Jadhav | Pune | Updated on July 01, 2020 | Published on July 01, 2020

Sources: Three women (2020), from [The Conversation](#). 40,800 female births (2020), from [VNExpress](#).

Why many Indian and Chinese men (2020), from [The Hindu Business Line](#).

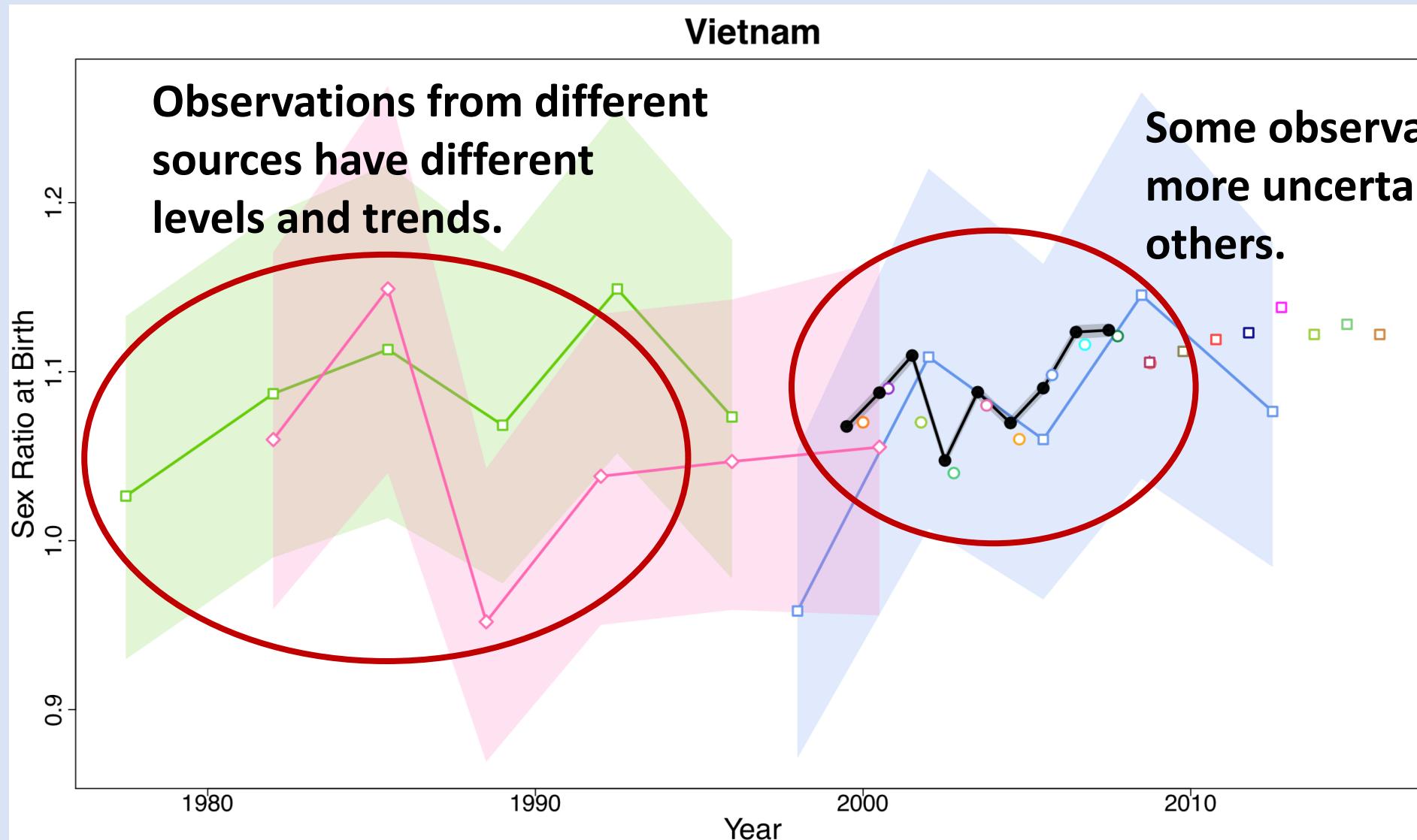
45 Million Missing Female Births during 1970-2017



Database for National SRB (as of June 2021)

- Number of SRB observations: 12,341
- Number of birth records: 3.26 billion
- Number of countries/areas with data: 204
- Reference year covered: 1753 (Sweden) – 2020
- Main data source types:
 - Civil registration and vital statistics
 - Surveys: e.g. Demographic and Health Survey (DHS), Multiple Indicator Cluster Survey (MICS).
 - Census
 - National reports
- Available at:
https://figshare.com/articles/dataset/SRB_database_for_all_countries/14838396

Data Model for SRB – Motivations



Data Model for SRB

Accounts for uncertainty associated with observations

$$\log(y_{c,t}) \sim N(\log(\Theta_{c,t}), \omega_s^2 + v_{c,t}^2)$$

Index c, t : country, year

SRB
observation

True SRB

- ω_s^2 : Non-sampling error variance, e.g. non-response, data input error
- Index s : data source types, e.g. administrative records, survey, census

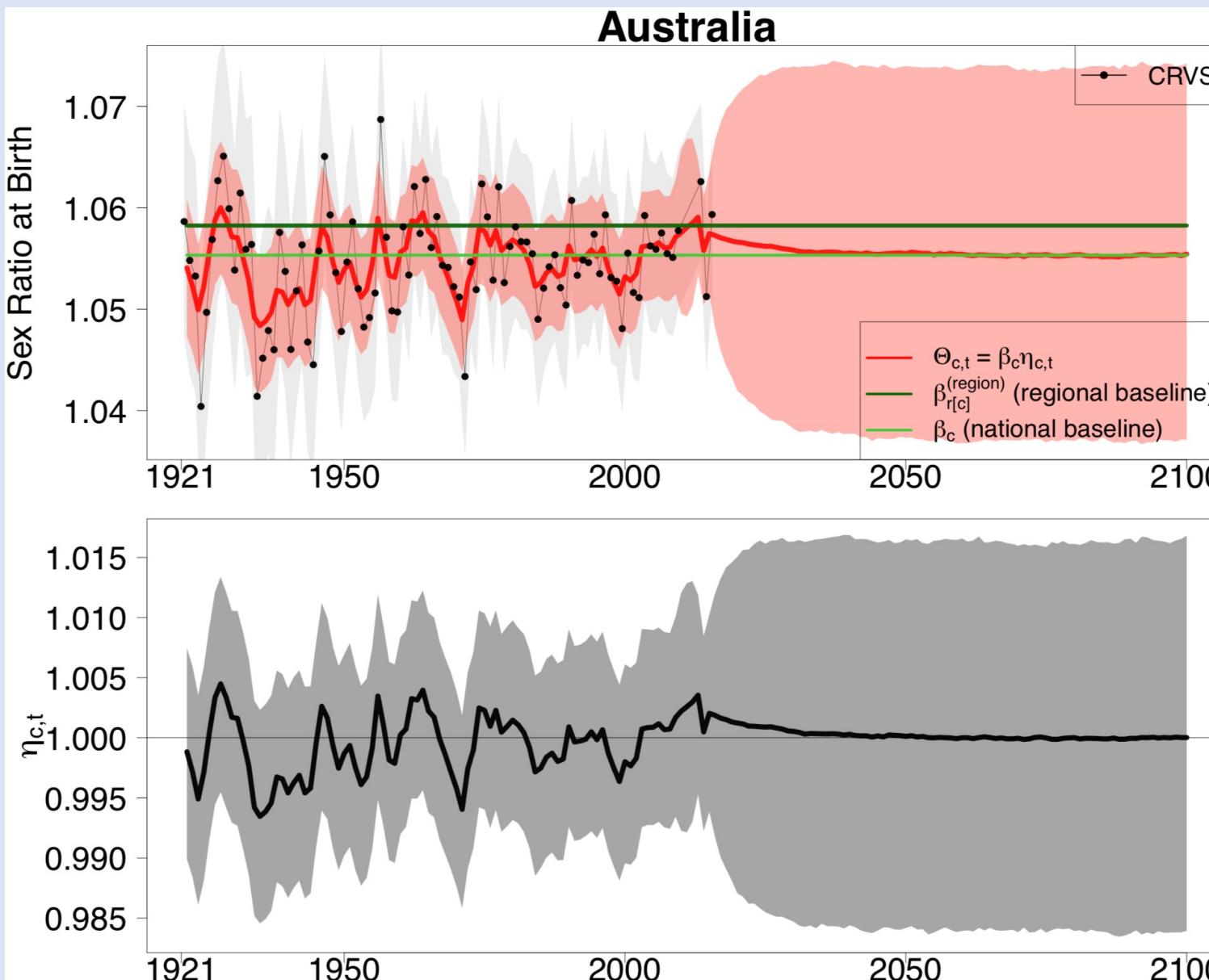
Sampling error variance due to sampling design, pre-computed

Main Idea of the Method

We use two models to estimate SRB for two groups of countries/areas:

- **Baseline model: for countries/areas without SRB inflation;**
- **Inflation model: for selected countries/areas with past/current/potential future SRB inflation.**

Baseline Model Overview



$$\Theta_{c,t} = \beta_c \eta_{c,t}$$

Index c : country
Index t : time, year

- β_c : **country-specific baseline:**
 - Constant within country
 - Differ across countries within a region
- $\eta_{c,t}$: **year-by-year natural fluctuation:**
 - An autoregressive AR(1) time series process

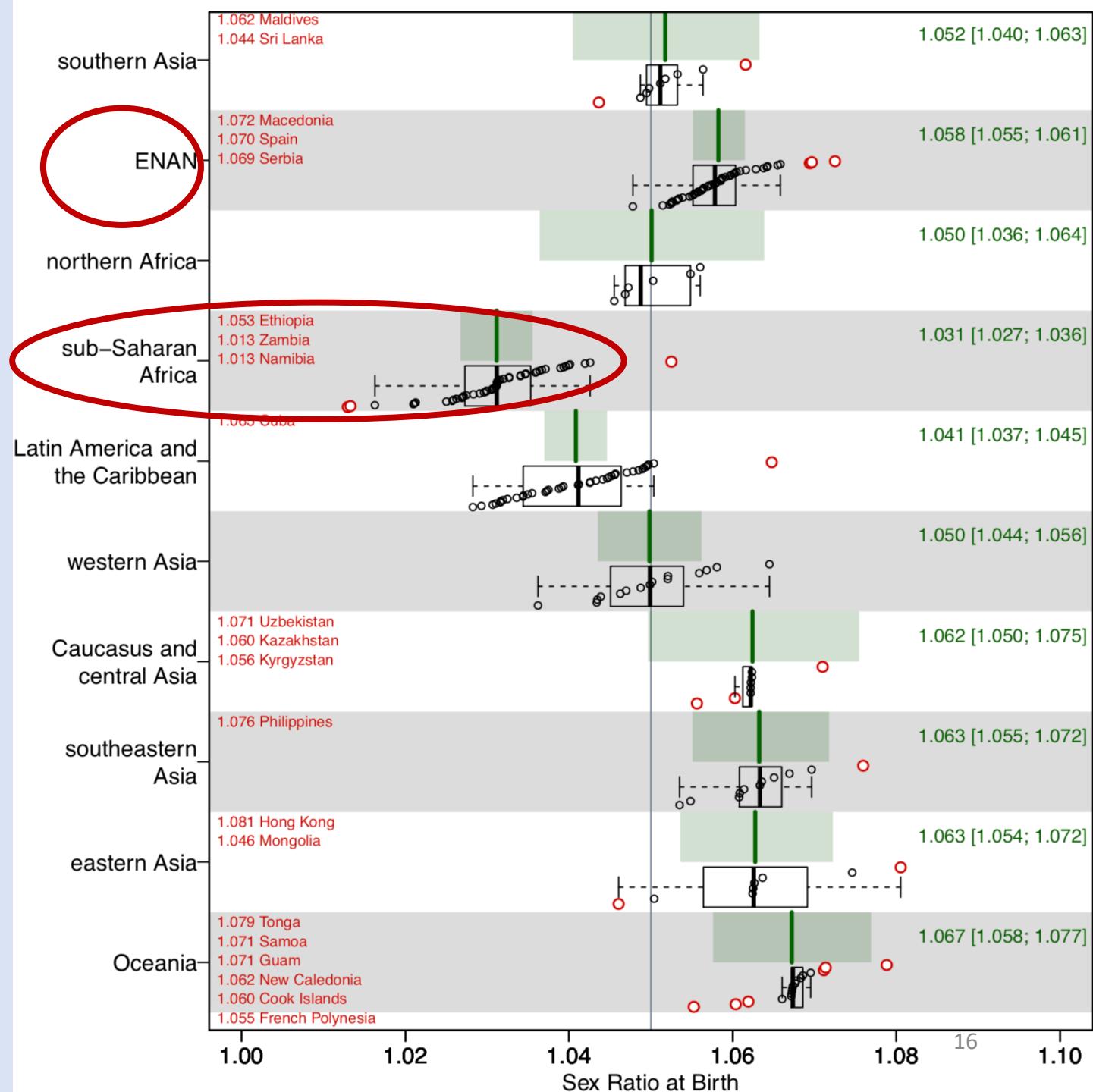
Baseline Model

$$\Theta_{c,t} = \beta_c \eta_{c,t}$$

Country-specific SRB baseline:

$$\beta_c \sim N(\beta_{r[c]}^{(region)}, \sigma_\beta^2)$$

- Mean at $\beta_{r[c]}^{(region)}$, regional SRB baseline
 - Group countries into regions based on their majority ethnicity
 - To account for the heterogeneity in baseline SRB across ethnicity groups

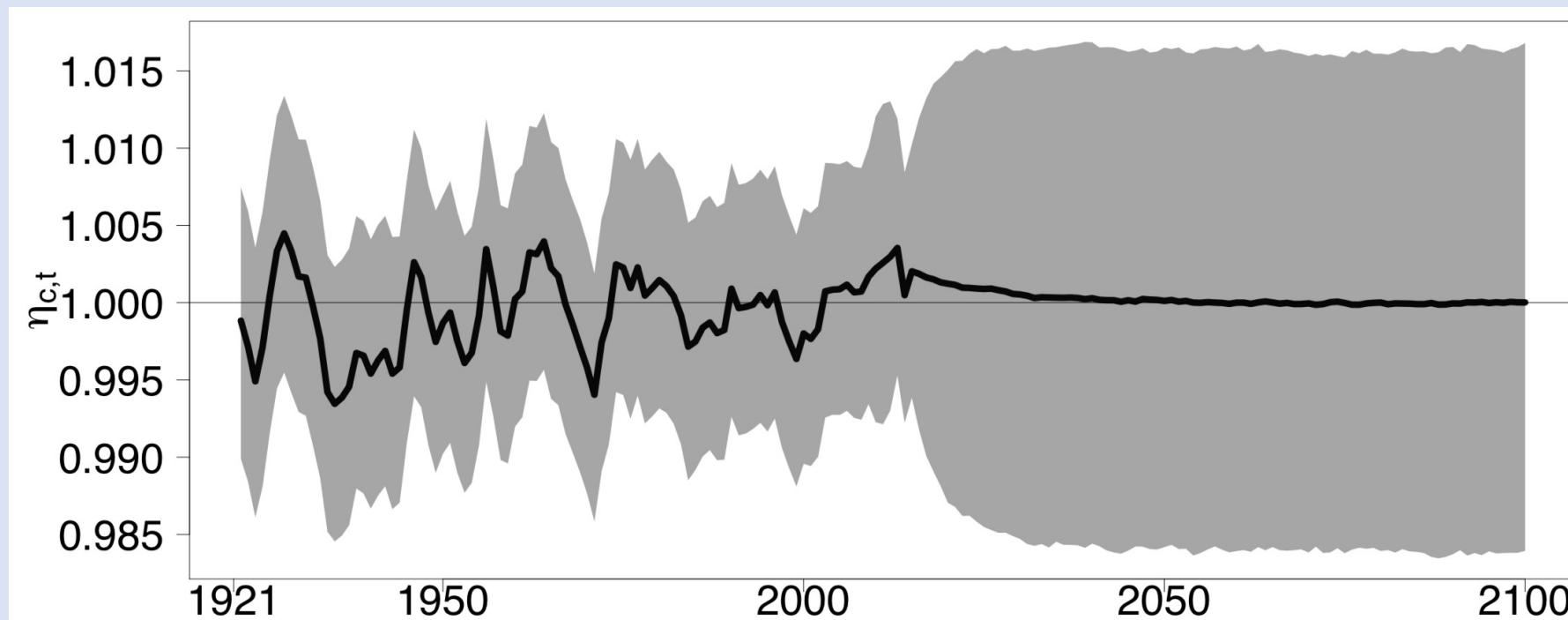


Baseline Model

$$\Theta_{c,t} = \beta_c \eta_{c,t}$$

Within country year-by-year natural fluctuation:

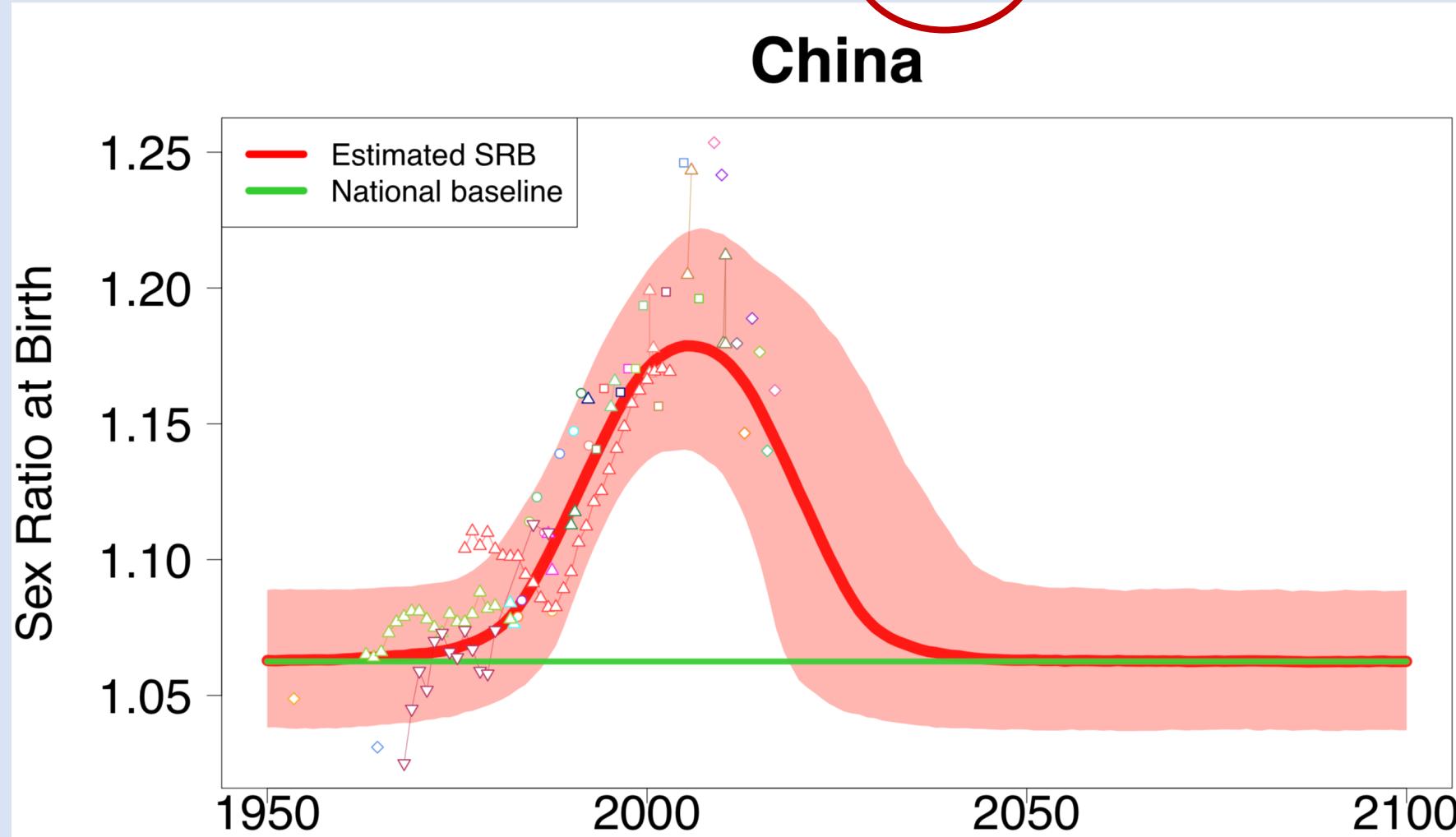
- Fluctuate around 1.
- Capture the natural deviation away from β_c when data suggest.



Inflation Model – Motivation

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

China



Country/Area Selection for Inflation Model

- Selection criteria for countries/areas to model with SRB inflation:
 - Observed SRB is suspected to be beyond biological norm as supported by literature; OR
 - Desired sex ratio at birth > 120 boys per 100 girls, or sex ratio of last birth > 130 boys per 100 girls; OR
 - Outlying female under-5 mortality rate.
- 29 selected countries/areas (12 with existing SRB inflation):
 - Asia (17-9): Afghanistan, Armenia, Azerbaijan, Bangladesh, China, Georgia, Hong Kong, India, Jordan, Korea Rep, Nepal, Pakistan, Singapore, Taiwan, Tajikistan, Turkey, Vietnam;
 - Sub-Saharan Africa (7-0): Gambia, Mali, Mauritania, Nigeria, Senegal, Tanzania, Uganda;
 - Elsewhere (5-3): Albania, Egypt, Montenegro, Morocco, Tunisia.

Inflation Model

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

For 29 countries at risk of SRB inflation: strong son preference.

Country-specific SRB inflation binary detector δ_c :

- 0: no inflation
- 1: with inflation

$$\delta_c \sim Bernoulli(\pi_c)$$

Country-specific SRB inflation probability π_c :

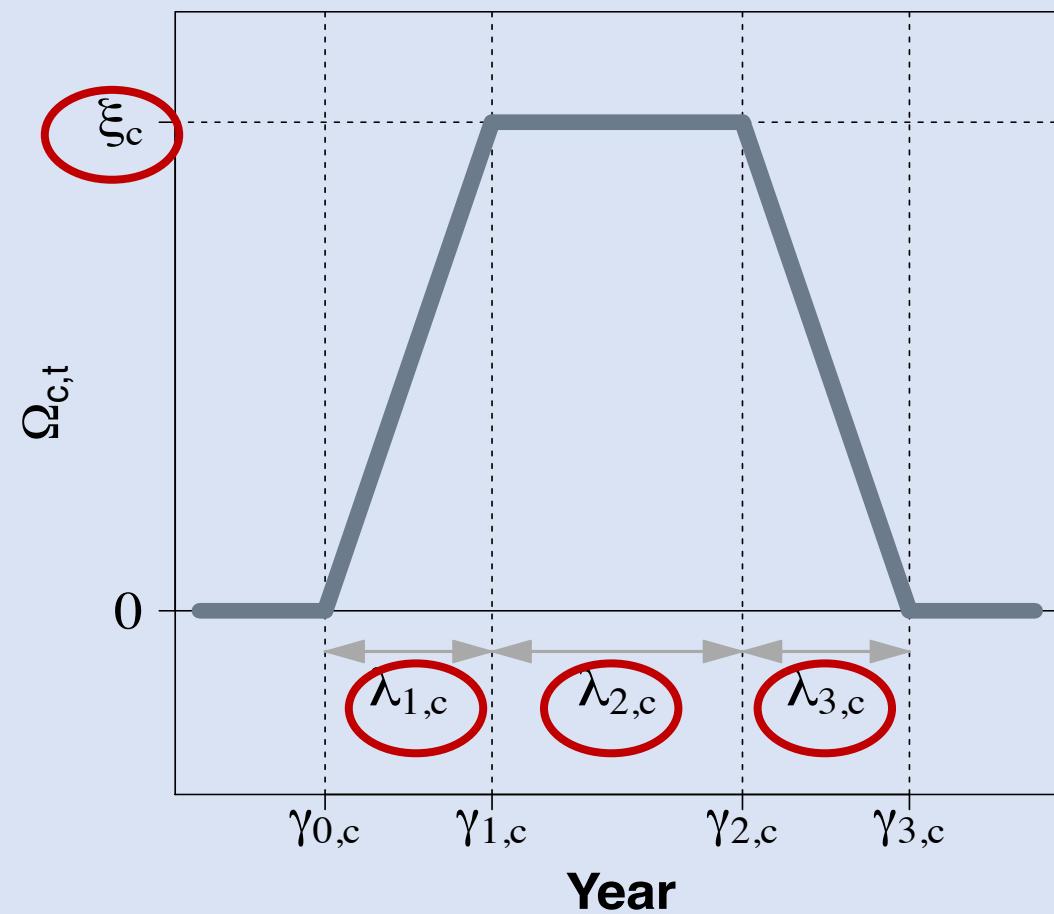
- Above 99.5%: Albania, Armenia, Azerbaijan, China, Georgia, Hong Kong, Korea Rep, Tunisia, Vietnam, Montenegro, Taiwan.
- Below 60%: Jordan, Singapore, Morocco, Bangladesh, Turkey.
- The rest are 62%-64%.

Inflation Model

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

Upward SRB inflation factor: trapezoid function

Sex ratio transition model



- Country-specific increase, stagnation, decrease, max inflation

$$\lambda_{1,c} \sim N(\mu_1, \sigma_1^2) T(0,)$$

$$\lambda_{2,c} \sim N(\mu_2, \sigma_2^2) T(0,)$$

$$\lambda_{3,c} \sim N(\mu_3, \sigma_3^2) T(0,)$$

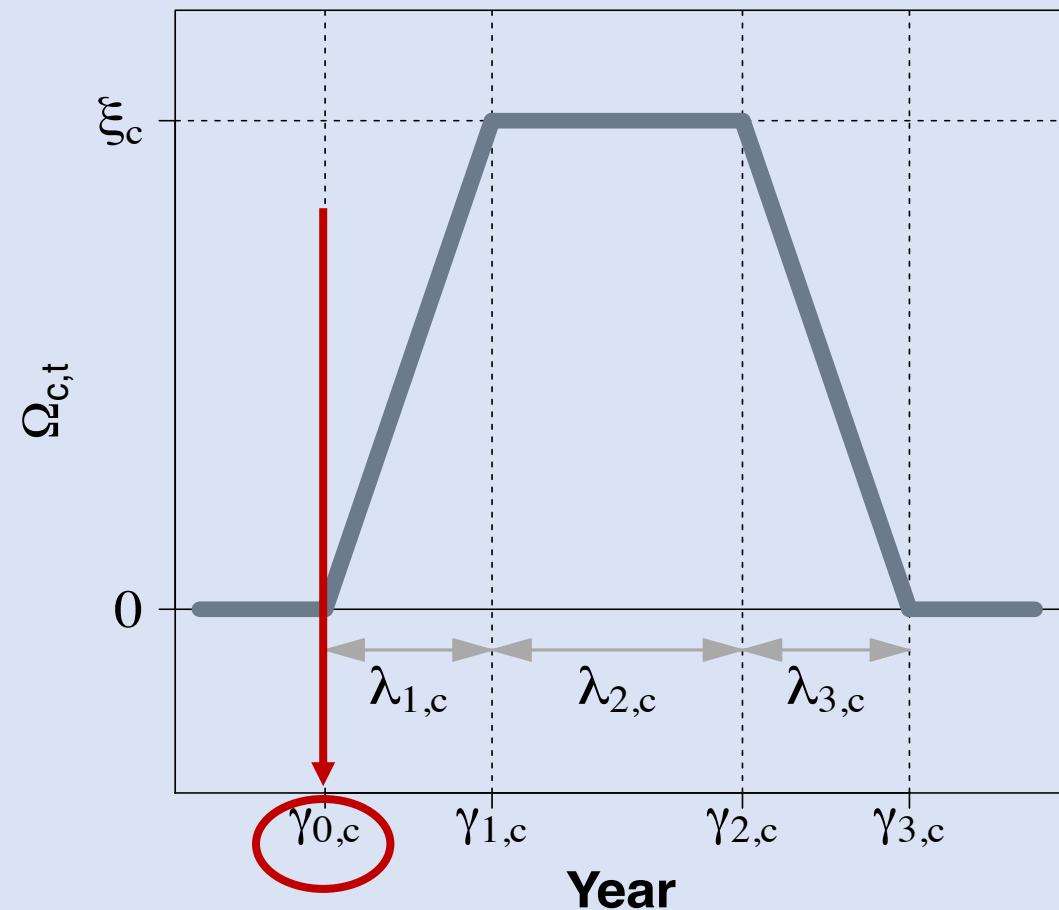
$$\xi_c \sim N(\mu_\xi, \sigma_\xi^2) T(0,)$$

Inflation Model

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

Upward SRB inflation factor: trapezoid function

Sex ratio transition model



- Country-specific start year includes fertility decline effect:

$$\gamma_{0,c} \sim t_3(f_{c,2.9}, \sigma_\gamma^2) T(f_{c,6},)$$

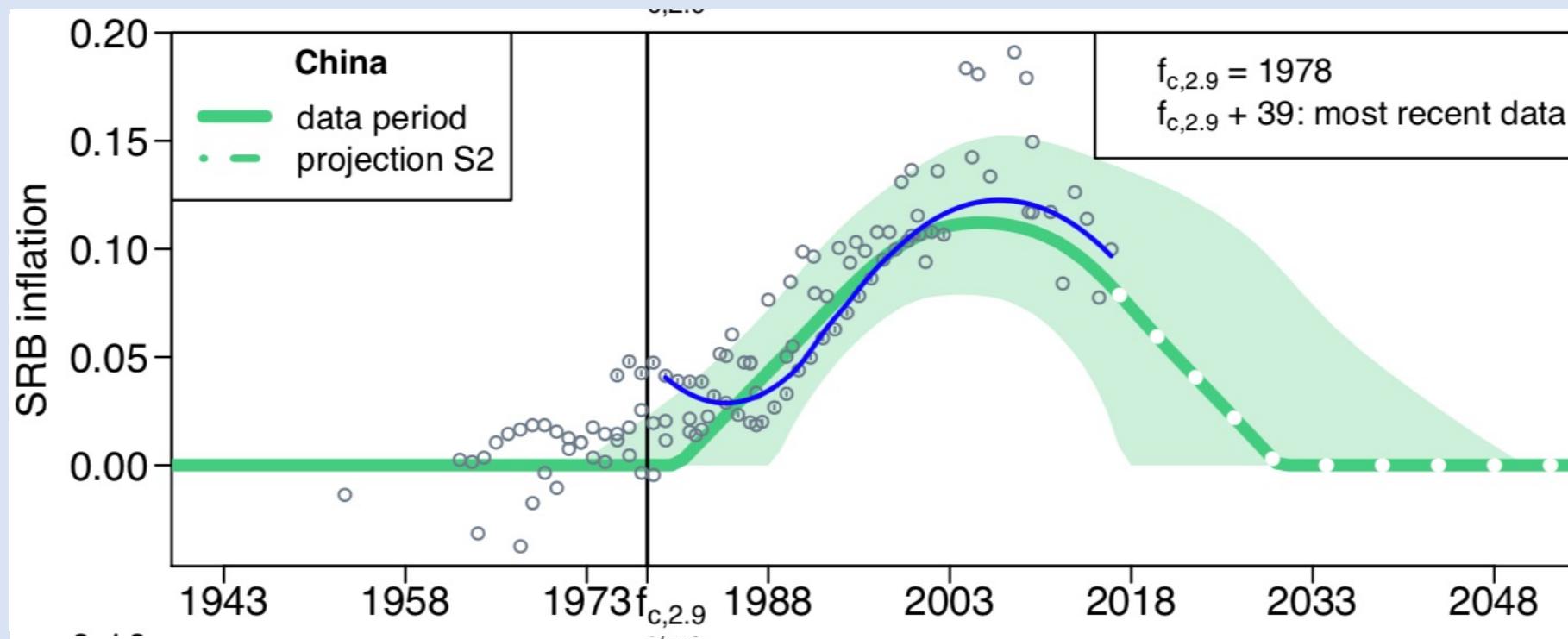
$f_{c,2.9}$: year in which TFR declines to 2.9

$f_{c,6}$: year in which TFR declines to 6

*TFR: total fertility rate

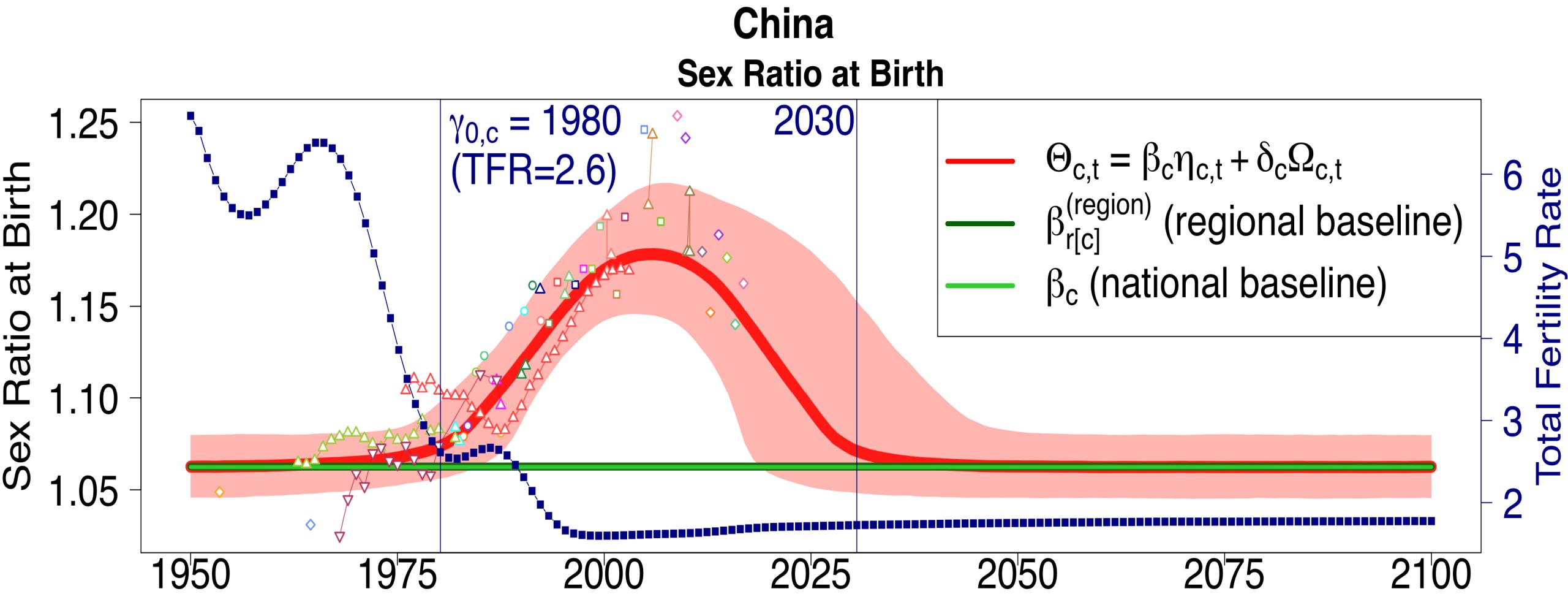
Inflation Model and Data

Parametric form of $\Omega_{c,t}$ captures the observed shape of inflated SRB



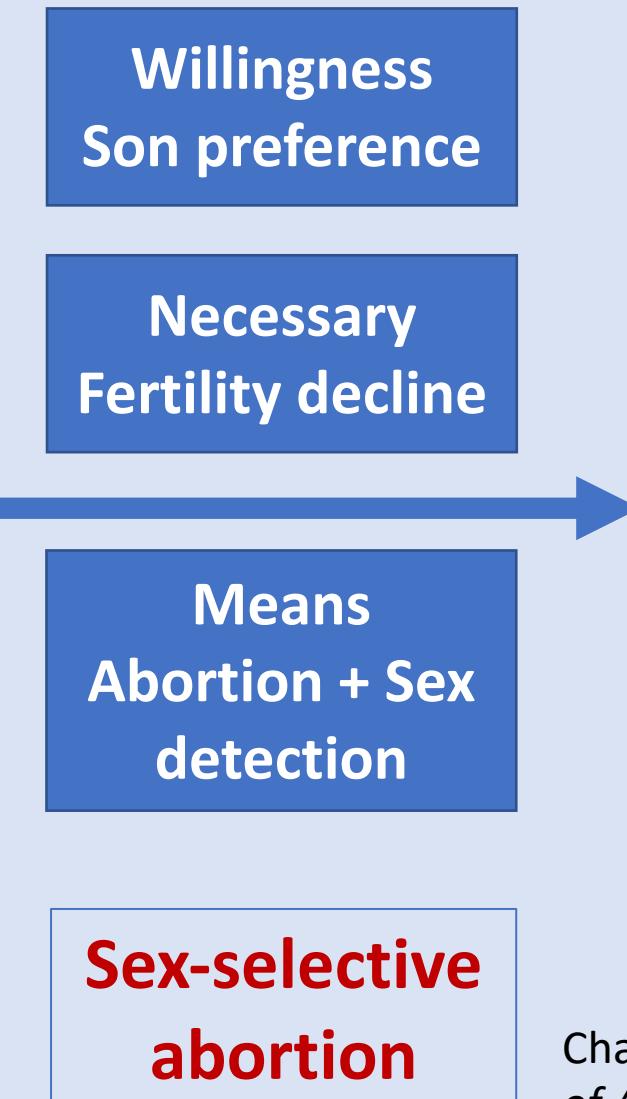
Data: $y_i - \hat{\beta}_c$

SRB Estimation and Projection Results for China



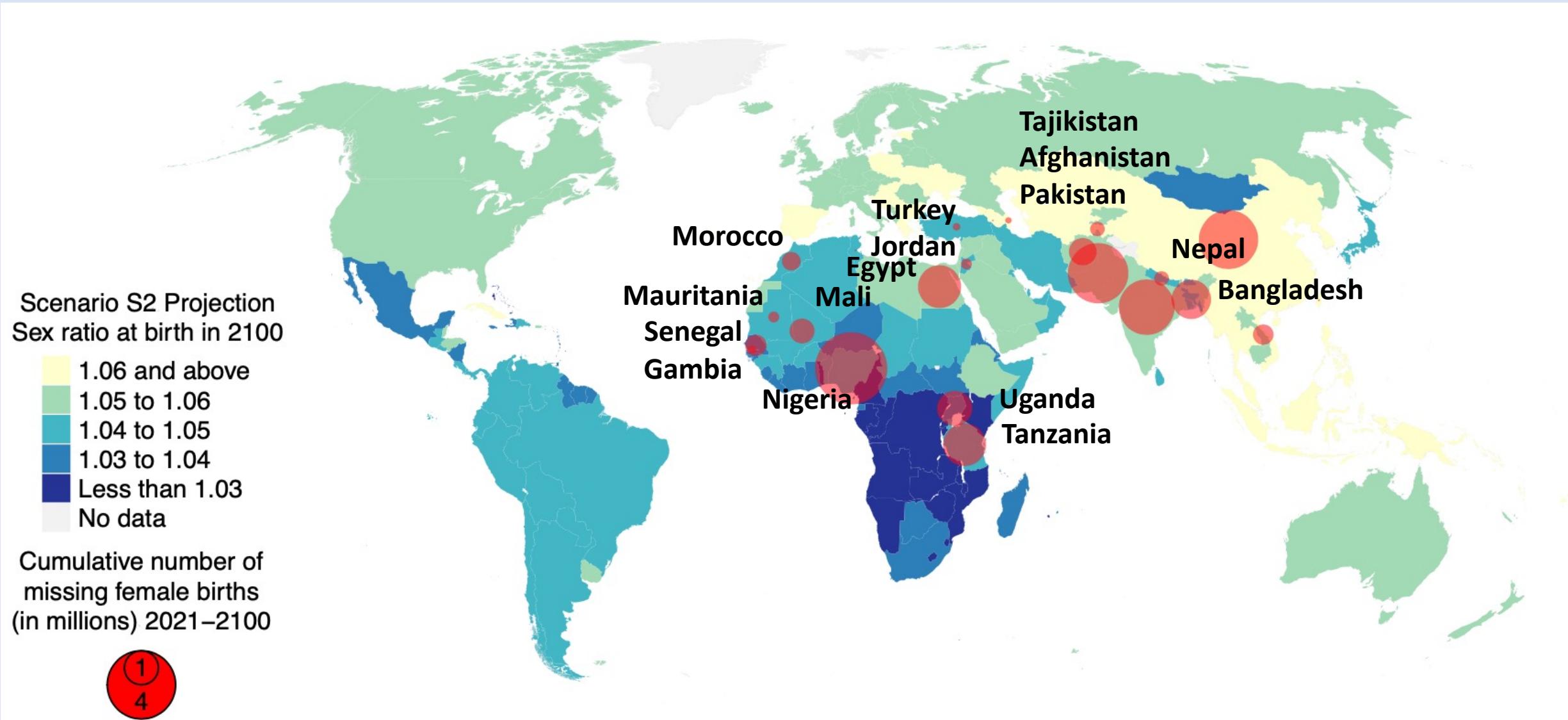
Scenario-Based SRB Projection till 2100

Some at-risk countries have normal SRB



- At-risk countries may have inflated SRB in the future
 - Mostly African countries
 - Scenario-based SRB projections:
 - No inflation $\delta_c = 0$
 $\Theta_{c,t} = \beta_c \eta_{c,t}$
 - With inflation $\delta_c = 1$
 $\Theta_{c,t} = \beta_c \eta_{c,t} + \Omega_{c,t}$

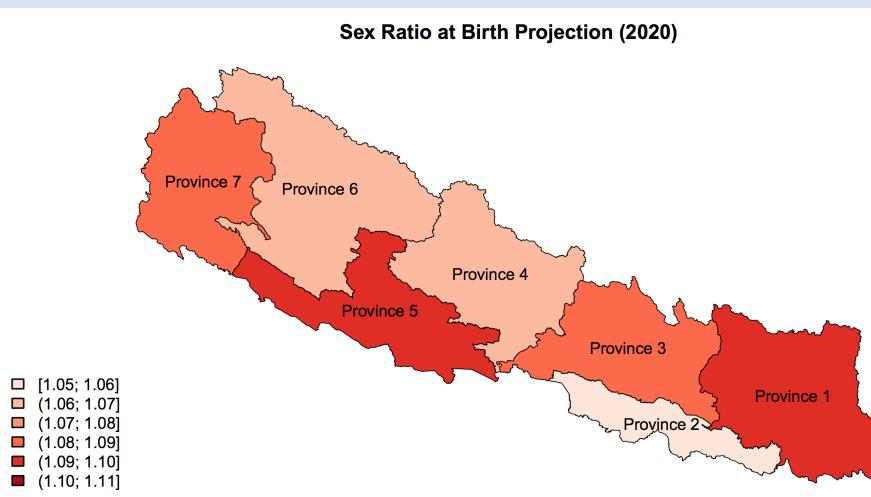
SRB & 22.1M Projected Missing Female Births till 2100



SRB Imbalance on Subnational Level in Asia

Bayesian hierarchical models with modifications can be used for estimating SRB inflation on subnational level.

Nepal



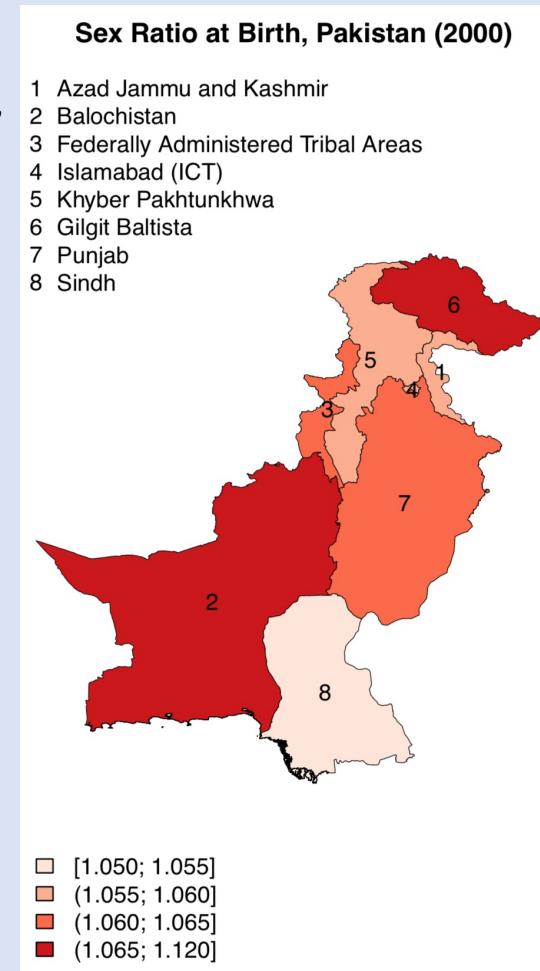
Chao, F., KC, S., & Ombao, H. (2022). *BMC Public Health*, 22(1), 1-15.

Vietnam



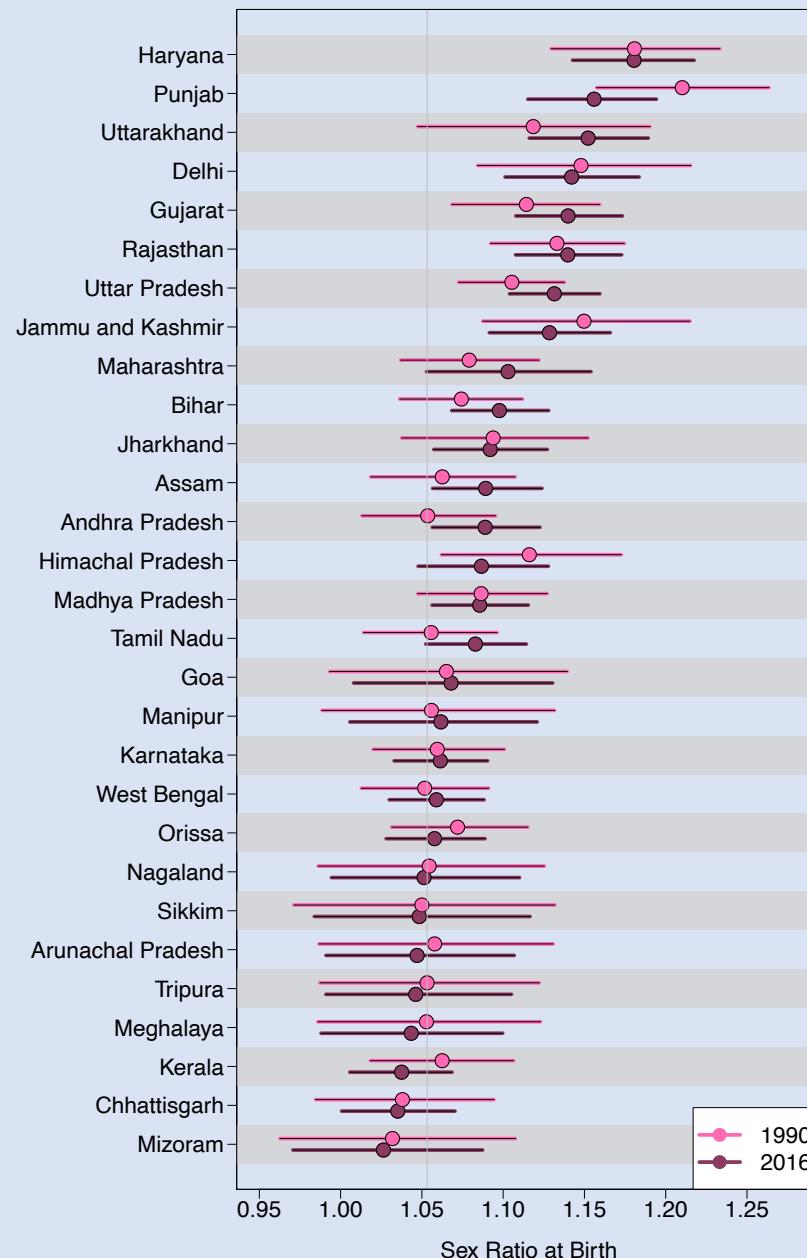
Chao, F.,
Guilmoto, C. Z.,
& Ombao, H.
(2021). *PLoS ONE*, 16(7),
e0253721.

Pakistan



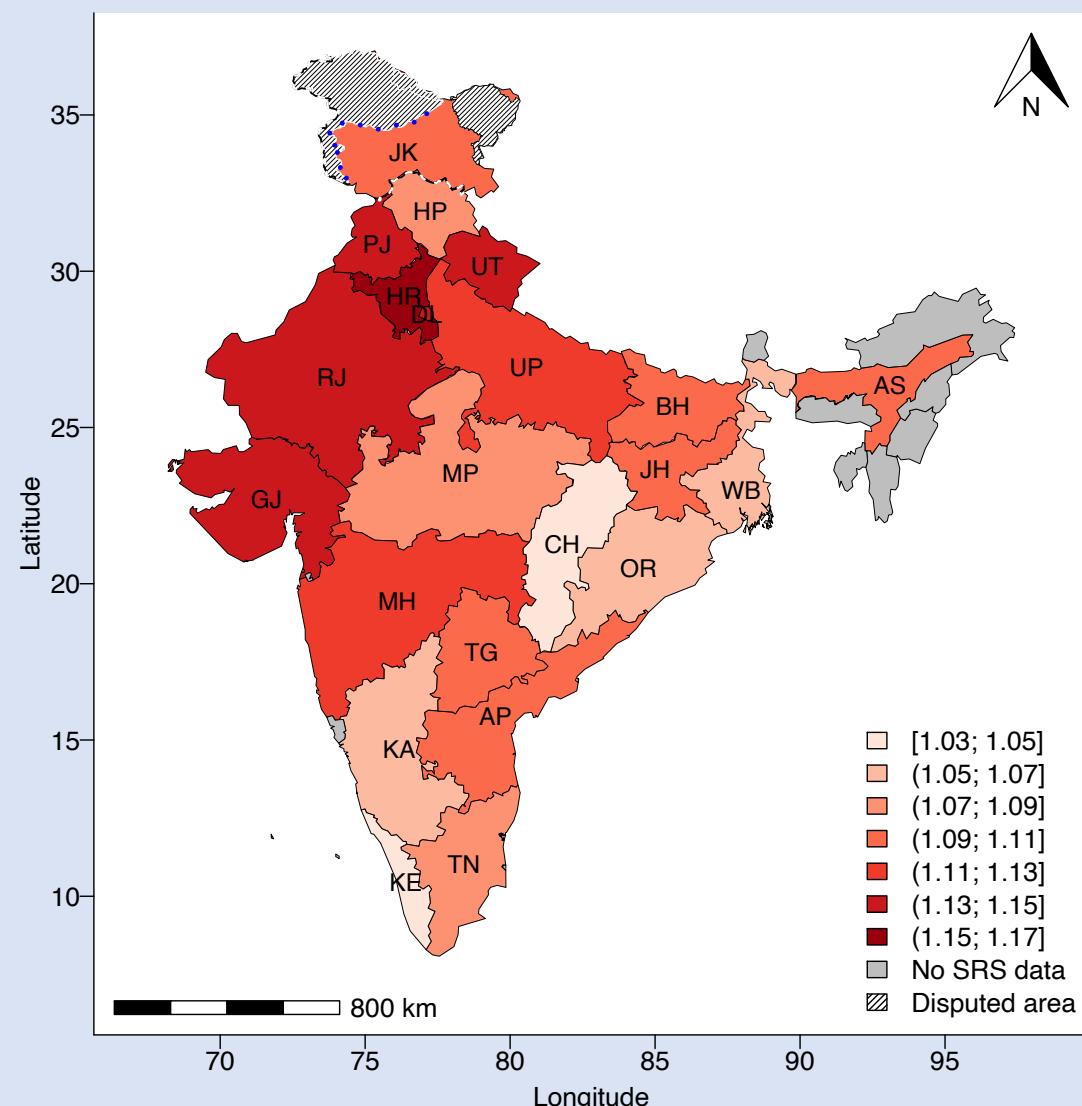
Chao, F.,
Wazir, M. A. &
Ombao, H.
Manuscript
under review.

SRB Imbalance on Subnational Level in India



Chao, F., &
Yadav, A. K.
(2019).
*Foundations of
Data
Science*, 1(2),
177.

Sex Ratio at Birth Projection (2030)



Research to the Real World

National SRB estimates* and scenario-based projections† have been used by the UNPD to inform policy makers and resource allocation.



**United
Nations**

Department of Economic and Social Affairs
Population Division

World Population Prospects 2022

UNPD World Population Prospects:
<https://population.un.org/wpp/>

*Chao, F., Gerland, P., Cook, A. R., & Alkema, L. (2019). *PNAS*, 116(19), 9303-9311.

†Chao, F., Gerland, P., Cook, A. R., Guilmoto, C. Z., & Alkema, L. (2021). *BMJ Global Health*, 6(8), e005516.

References

- 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 May 7;116(19):9303-11.
- Banerjea S. Three women: stories of Indian trafficked brides. *The Conversation*. 2020 Sep 30.
- Nga M. 40,800 female births doomed in Vietnam every year. *VNExpress*. 2020 Jun 19.
- Pune R. J. Why many Indian and Chinese men may need to delay marriage or remain bachelors. *The Hindu Business Line*. 2020 Jul 1.
- 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. *The Annals of Applied Statistics*. 2021 Sep;15(3):1499-528.
- 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 Aug 1;6(8):e005516.
- Bongaarts J. The implementation of preferences for male offspring. *Population and Development Review*. 2013 Jun;39(2):185-208.
- 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 Sep 1;2(9):e521-30.
- Chao F, KC S, Ombao H. Estimation and probabilistic projection of levels and trends in the sex ratio at birth in seven provinces of Nepal from 1980 to 2050: a Bayesian modeling approach. *BMC public health*. 2022 Dec;22(1):1-5.
- 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 Jul 14;16(7):e0253721.
- Chao F, Wazir MA, Ombao H. 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.
- 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.
- 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 Aug 19;15(8):e0236673.