Flexible Bayesian Models of Demographic and Health Indicators

Herb Susmann
School of Biostatistics, Department of Population Health
NYU Grossman School of Medicine

May 8, 2025

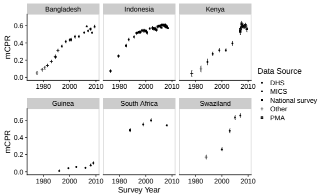
Research support: Gates Foundation (INV-00844); NIH (R01 HD070936).

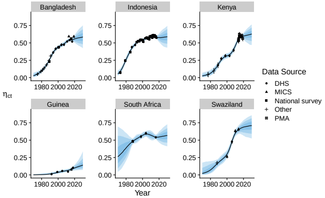
Who I am

- PhD at University of Massachusetts Amherst with Leontine Alkema
  - Bayesian modeling applied to estimating and projecting global family planning indicators
- One year research visit at MAP5 (Université de Paris Cité) and post-doc at CEREMADE (Université Paris Dauphine)
- Ongoing work with Adrian Raftery
  - Refugee and asylum seeker migration
- Currently post-doc at NYU Grossman School of Medicine
- `herbsusmann.com`

Outline

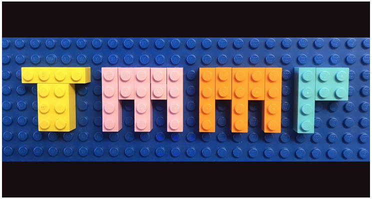
- Goal: estimate and project demographic and health indicators in multiple populations.
    - Example: analysis of indicators in the global indicator framework for the Sustainable Development Goals.
- Data may be from multiple sources, noisy, sparse, ...
- Statistical models (often Bayesian!) are needed to generate *probabilistic* estimates and projections from available data.
- **This talk:** a Bayesian modeling framework for demographic and health indicators, and how we've used it to build models.





Modeling Framework

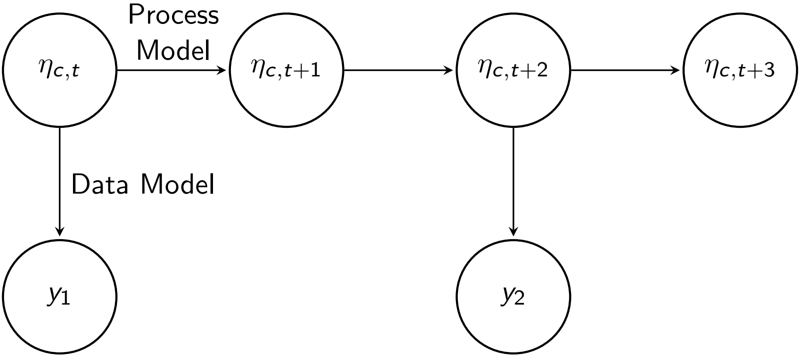
- Many statistical models have been created to provide estimates and projections, but...
  - comparing models can be difficult.
  - building a new model requires starting from scratch.
- Overarching model class: **Temporal Models for Multiple Populations** (Susmann et al. International Statistical Review 2022).



Temporal Models for Multiple Populations

- True value of indicator: $\eta_{c,t}$ for $c = 1, \ldots, C$, $t = 1, \ldots T$.
- *Process model* describes evolution of $\eta_{c,t}$.
  - Covariates
  - Systematic trends
- Observed data y_i, with associated properties $c[i]$, $t[i]$, $s[i]$, ...
- *Data model* describes relationship between y_i and $\eta_{c[i],t[i]}$.

Temporal Models for Multiple Populations



Process Model

$$g_1(\eta_{c,t}) = \underbrace{g_2(X_{c,t}, \beta_c)}_{\text{covariate}} + \underbrace{g_3(t, \eta_{c,s\neq t}, \alpha_c)}_{\text{systematic}} + \underbrace{a_{c,t}}_{\text{offset}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}}$$

Covariate component

$$g_1(\eta_{c,t}) = \underbrace{g_2(X_{c,t}, \beta_c)}_{\text{covariate}} + \underbrace{g_3(t, \eta_{c,s\neq t}, \alpha_c)}_{\text{systematic}} + \underbrace{a_{c,t}}_{\text{offset}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}}$$

- Regression function for incorporating covariates.

Systematic component

$$g_1(\eta_{c,t}) = \underbrace{g_2(X_{c,t}, \beta_c)}_{\text{covariate}} + \underbrace{g_3(t, \eta_{c,s\neq t}, \boldsymbol{\alpha}_c)}_{\text{systematic}} + \underbrace{a_{c,t}}_{\text{offset}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}}$$

- Parametric function for modeling systematic temporal trends.
- Example: modeling the rate of change in adoption of modern family planning as following logistic growth (Cahill et al. Lancet 2018).

Offset

$$g_1(\eta_{c,t}) = \underbrace{g_2(X_{c,t}, \beta_c)}_{\text{covariate}} + \underbrace{g_3(t, \eta_{c,s \neq t}, \alpha_c)}_{\text{systematic}} + \underbrace{a_{c,t}}_{\text{offset}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}}$$

- The offset term incorporates external information, for example from a separate modeling step.

Smoothing component

$$g_1(\eta_{c,t}) = \underbrace{g_2(X_{c,t}, \beta_c)}_{\text{covariate}} + \underbrace{g_3(t, \eta_{c,s\neq t}, \alpha_c)}_{\text{systematic}} + \underbrace{a_{c,t}}_{\text{offset}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}}$$

- The smoothing component allows data-driven deviations from the other components, while still enforcing smoothness.
- Many choices B-splines, Gaussian processes, AR(p), RW(p), spatial smoothing (ICAR), ...

Additional Examples

Paper includes additional examples of existing models that fall into the TMPP framework:

- Under-5 Mortality

  (Alkema and New AOAS 2014, Dicker et al. Lancet 2017)

- Family Planning

  (Cahill et al. Lancet 2018)

- Neonatal Mortality

  (Alexander and Alkema Demographic Research 2018)

- Maternal Mortality

  (Alkema et al. AOAS 2017)

- Subnational Mortality

  (Alexander et al. Demography 2017)

	GBD	B3
$\eta_{c,t}$	crisis-free U5MR	crisis-free U5MR
$g_1(\cdot)$	$\log_{10}$	log
Process model formula	$g_1(\eta_{c,t}) = g_2(\boldsymbol{X}_{c,t}, \boldsymbol{\beta}_c) + a_{c,t} + \epsilon_{c,t}$	$g_1(\eta_{c,t}) = g_3(t, \boldsymbol{\alpha}_c) + \epsilon_{c,t}$
Covariate Component		
$g_2(\cdot)$	non-linear regression formula (Equation 14)	·
Covariates	LDI, EDU, HIV	·
Systematic Component		
$g_3(\cdot)$	·	$\alpha_{c,0} + \alpha_{c,1}(t - t_c^*)$, with $t_c^* \approx$ middle of observation period
$\boldsymbol{\alpha}_c$	·	intercept $\alpha_{c,0}$ and slope $\alpha_{c,1}$
Offsets		
$a_{c,t}$	offsets obtained from smoothed residuals of a mixed-effects regression model fit	·
Smoothing Component $\boldsymbol{\epsilon}_c = \boldsymbol{B}_c\boldsymbol{\delta}_c$		
$\boldsymbol{B}$	$\boldsymbol{B} = \boldsymbol{I}$	$B_{c,k}$ = cubic B-splines, knots every 2.5 years
$s(t_1, t_2)$	Matérn	indep. $s(t_1, t_2) = \sigma_{\tau,c}^2 1(t_1 = t_2)$
r	0	2
$\mathcal{K}_{d,c}$	·	$\mathcal{K}_{0,c} = \{k^*\}, \mathcal{K}_{1,c} = \{2, \cdots, K_c\}$
Projections (if not defaulting to estimation model)		
Projections	·	logarithmic pooling approach: for projections, $\Delta_2\delta_{c,k} \sim N(\Gamma_{c,k}, \Theta_{c,k})$, $\Gamma_{c,k} = W \cdot G + (1-W) \cdot \Delta_2\delta_{c,k-1}$, $\Theta_{c,k} = W \cdot V + (1-W) \cdot \Theta_{c,k-1}$.

Comparison of two models for Under-5 Mortality

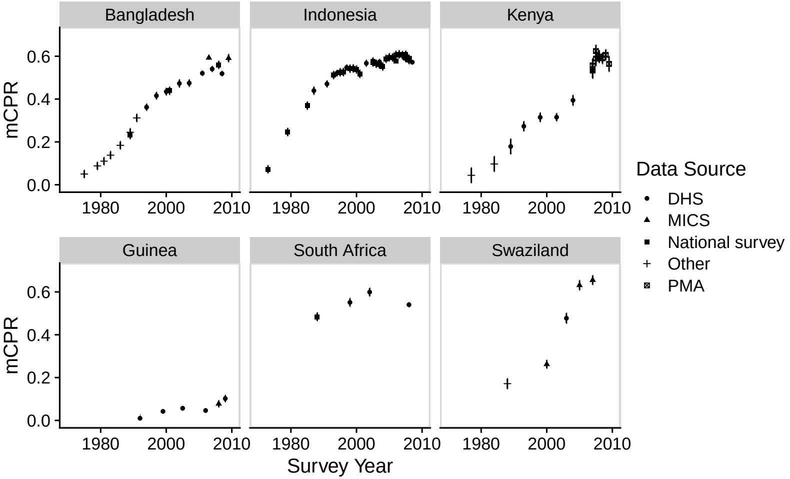
Vignette 1: Transitions

- Some indicators have been observed to evolve similarly across populations.
  - They tend to follow a *transition* between stable states.
- Classic example: demographic transition.
  - Transition from high total fertility rate and high under-5 mortality to low fertility, low mortality.
- Existing statistical models for estimating and projecting trends in these indicators draw on these patterns.
- We propose a new type of model, called *B-spline Transition Models*, for flexibly estimating indicators that follow transitions. (Susmann and Alkema JRSS-C 2025).

Case Study

- Modern Contraceptive Prevalence Rate (mCPR) for married or in-union women: proportion of married or in-union women of reproductive age using (or with partner using) a modern contraceptive method.
- Existing model: Family Planning Estimation Model (FPEM, Cahill et al. 2018).
- Goal: estimate and project mCPR in countries from 1970-2030.
- Dataset aggregated by United Nations Population Division (UNPD) from surveys conducted by governments or international organizations.

Data


Bangladesh
Indonesia
Kenya
Guinea
South Africa
Swaziland
mCPR
0.0
0.2
0.4
0.6
1980
2000
2010
Survey Year
Data Source
DHS
MICS
National survey
Other
PMA

Transition Models

- *Transition Models* have a process model given by

$$g_1(\eta_{c,t}) = \underbrace{g_3(t, \boldsymbol{\eta}_{c,s\neq t}, \boldsymbol{\alpha}_c)}_{\text{systematic}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}} .$$

- The systematic component has the following form:

$$g_3(t, \boldsymbol{\eta}_{c,s\neq t}, \boldsymbol{\alpha}_c) = \begin{cases} \Omega_c, & t = t_c^*, \\ g_1(\eta_{c,t-1}) + f(\eta_{c,t-1}, P_c, \beta_c), & t > t_c^*, \\ g_1(\eta_{c,t+1}) - f(\eta_{c,t+1}, P_c, \beta_c), & t < t_c^*, \end{cases}$$

where $\boldsymbol{\alpha}_c = \{\Omega_c, P_c, \beta_c\}$.

- The function f is called the *transition function*.

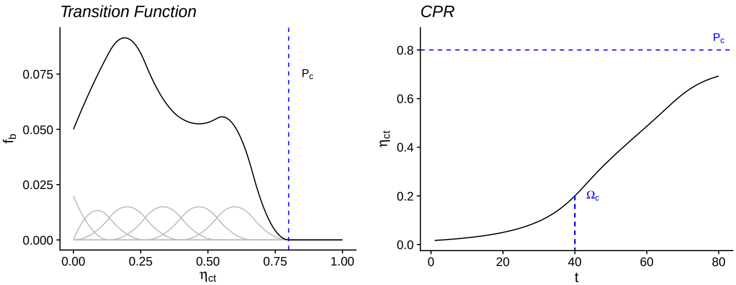
B-spline Transition Model

- Define a transition function f_b as:

$$f_b(\eta_{c,t}, P_c, \beta_c) = \sum_{j=1}^{J} \underbrace{h_j(\beta_{c,j})}_{\text{coefficient}} \cdot \underbrace{B_j(\eta_{c,t}/P_c)}_{\text{basis function}},$$

  where P_c is an asymptote parameter.
- Flexibility of f_b can be tuned through the spline degree and number and positioning of knots.

Example B-spline Transition Function



Smoothing component

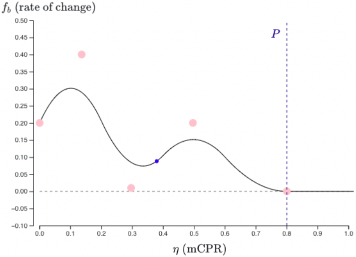
- Recall the process model has two components:

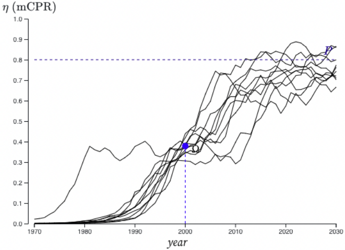
$$g_1(\eta_{c,t}) = \underbrace{g_3(t, \boldsymbol{\eta}_{c,s\neq t}, \boldsymbol{\alpha}_c)}_{\text{systematic}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}} .$$

- Smoothing component: AR(1) process of the form

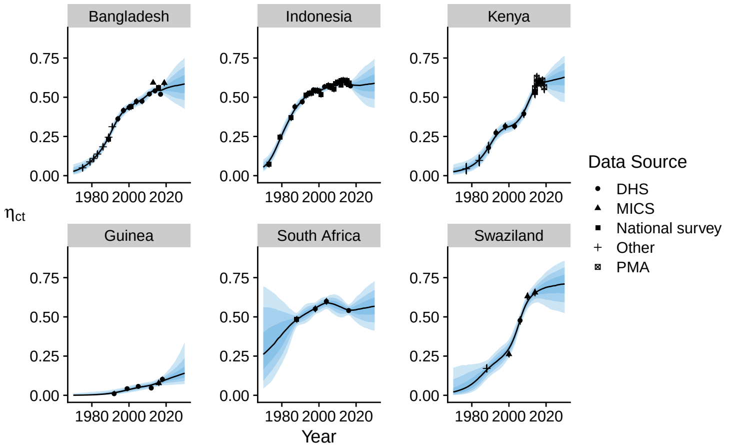
$$\epsilon_{c,t} | \epsilon_{c,t-1}, \tau, \rho \sim N(\rho \times \epsilon_{c,t-1}, \tau^2)$$

Smoothing component

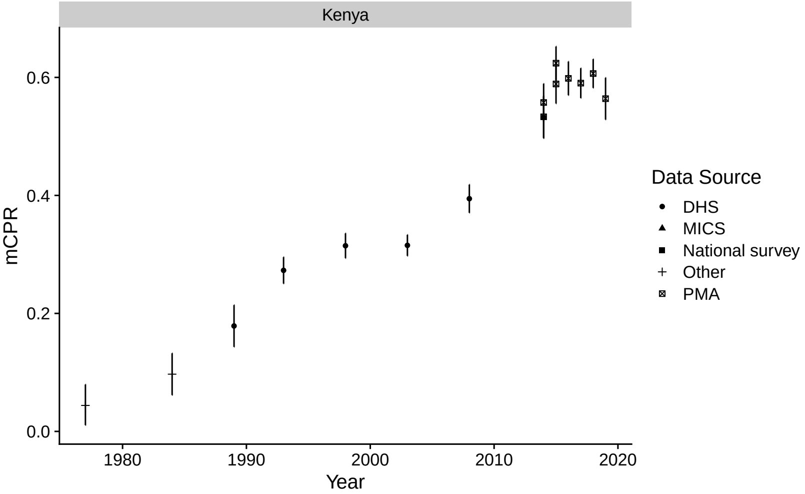




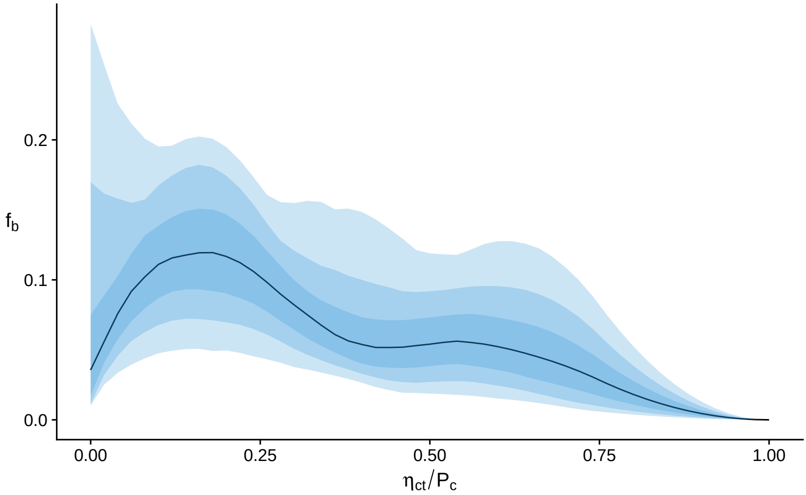
Illustrative Fits



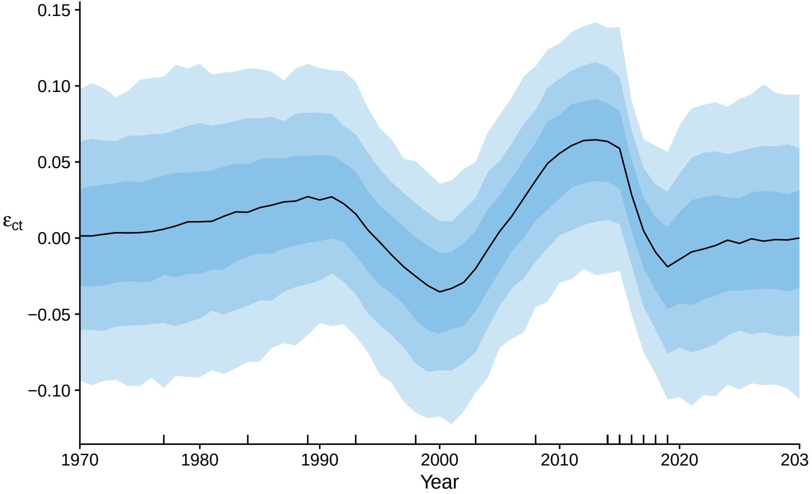
Kenya Data


Kenya
mCPR
0.0
0.2
0.4
0.6
1980
1990
2000
2010
2020
Year
Data Source
DHS
MICS
National survey
Other
PMA

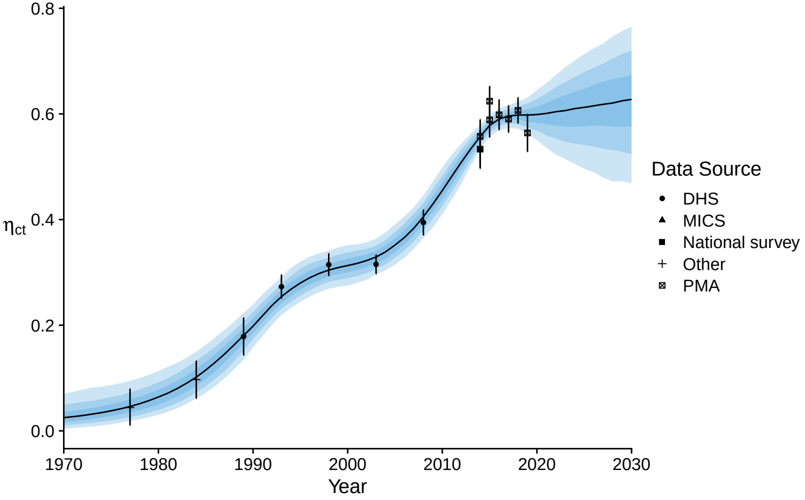
Kenya Transition Function



Kenya Smoothing Component



Kenya mCPR Estimates



Family Planning Estimation Tool (FPET)

- FPET: used by countries to produce estimates of family planning indicators such as mCPR.
- Results used within countries and in the FP2030 global initiative.
- Overview: *Statistical Demography Meets Ministry of Health: The Case of the Family Planning Estimation Tool* (Alkema et al. ArXiv 2024)

FPET training by Avenir Health, Kenya 2024

Vignette 2: Shocks

- Many statistical models assume *smoothness* of the data.
- Statistical models that assume smoothness typically will not perform well when fit to data that exhibit shocks.
- We propose using **Bayesian shrinkage priors** as a practical way to build statistical models robust to shocks.

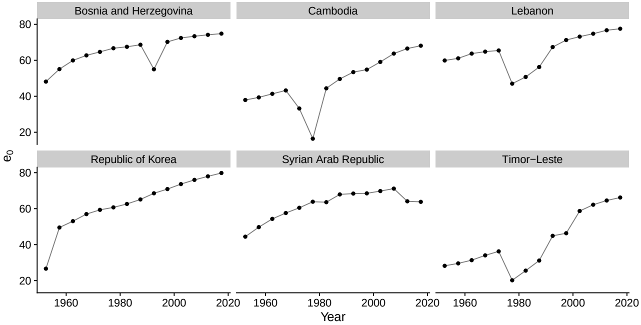


Figure: Male period life expectancy at birth (e_0) for six countries exhibiting shocks. Data: UN World Population Prospects, 2022 revision.

Smooth Transition Model

- Let $\eta_{c,t}$ be the true male period life expectancy at birth in country c and time t.
- Model change in $\eta_{c,t}$ as:

$$\eta_{c,t} = \underbrace{\eta_{c,t-1} + f(\eta_{c,t-1}, \beta_c)}_{systematic} + \underbrace{\epsilon_{c,t}}_{smoothing} .$$

- We model f using a B-spline transition model.
- Deviations typically modeled as ARIMA process; following Raftery et al. we use white noise for e_0:

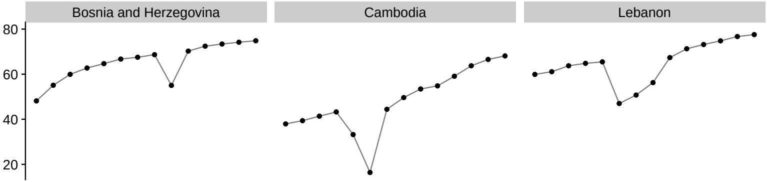
$$\epsilon_{c,t} | \tau_\epsilon \sim N(0, \tau_\epsilon^2).$$

Transition Model with Shocks

- Proposal: add an additional term to the process model to handle shocks.

$$\eta_{c,t} = \underbrace{\eta_{c,t-1} + f(\eta_{c,t-1}, \beta_c)}_{\text{systematic}} - \underbrace{\delta_{c,t}}_{\text{shock}} + \underbrace{\epsilon_{c,t}}_{\text{smoothing}},$$

  where $\delta_{c,t} > 0$.
- We call $\delta_{c,t}$ the *shock term*.
- A-priori we do not think that $\delta_{c,t}$ will be large for most country-years.

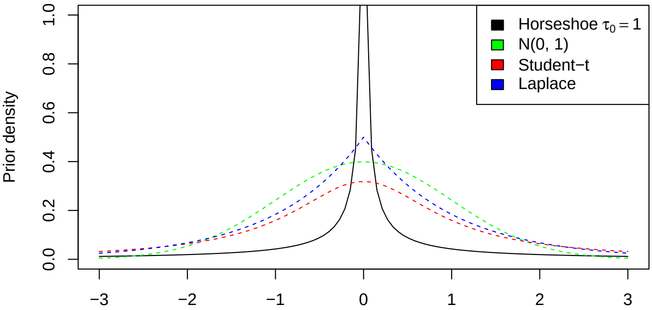


The horseshoe prior

- The *horseshoe prior* (Carvalho 2009 PMLR) is given by

$$\delta_{c,t} \mid \tau, \gamma_{c,t} \sim N(0, \tau^2\gamma_{c,t}^2)$$
$$\gamma_{c,t} \sim C^+(0,1).$$

- Global scale parameter $\tau > 0$ shrinks all shocks to zero.
- Local scale parameters $\gamma_{c,t} > 0$ allow some shocks to escape shrinkage.



Estimated shocks

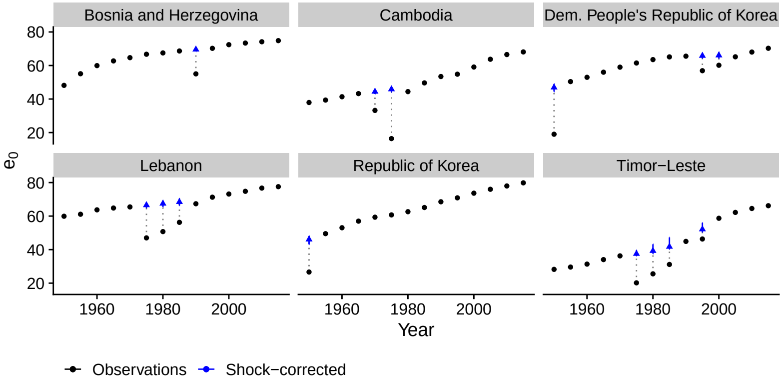


Figure: Six countries with the largest estimated detected shocks. Shocks are illustrated by plotting "shock-corrected" estimates, given by the observed e_0 minus the shock $\delta_{c,t}$

Example e_0 projections

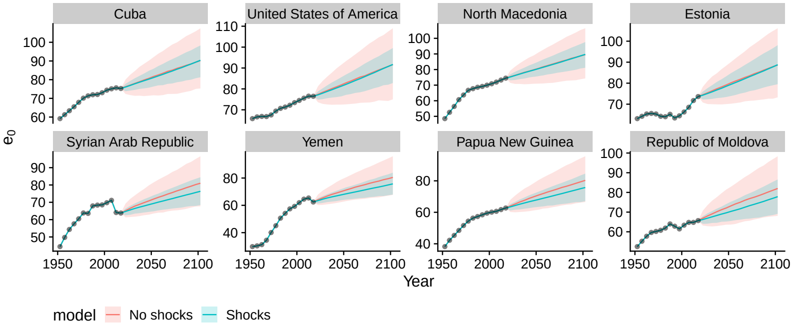


Figure: Projections of e_0 from the model with and without shocks, for countries with the smallest (top row) and largest (bottom row) differences in posterior median projected e_0 in 2095-2100.

Adding shocks has little effect on median projections, and reduces projection uncertainty

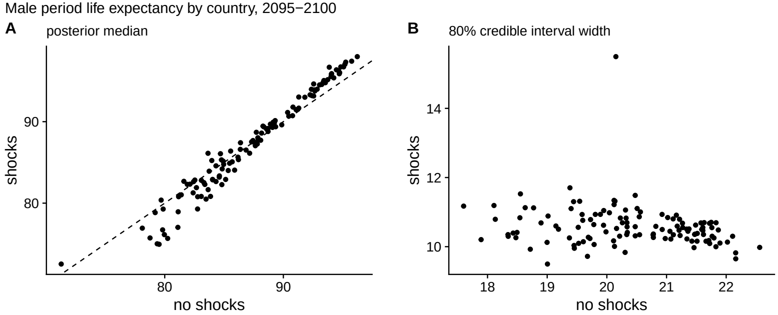


Figure: Posterior medians (A) and 80% projection interval widths (B) for male period life expectancy at birth by country in 2095-2100 for the model with and without shocks included.

Another application: adding shocks improves historical estimates

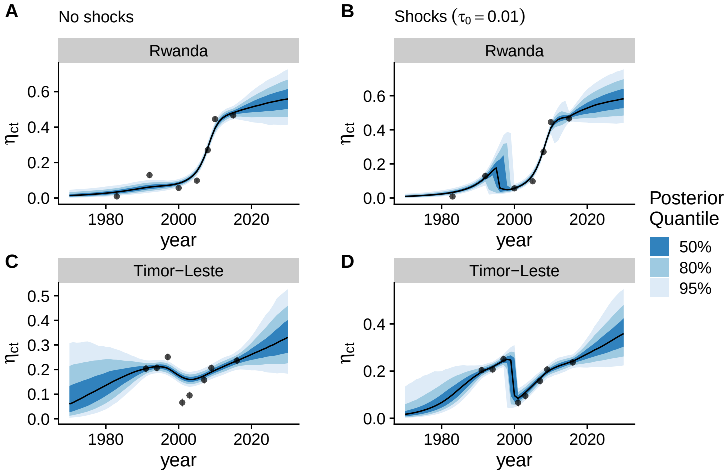


Figure: Modern Contraceptive Use Rate (mCPR) estimates with and without shocks.

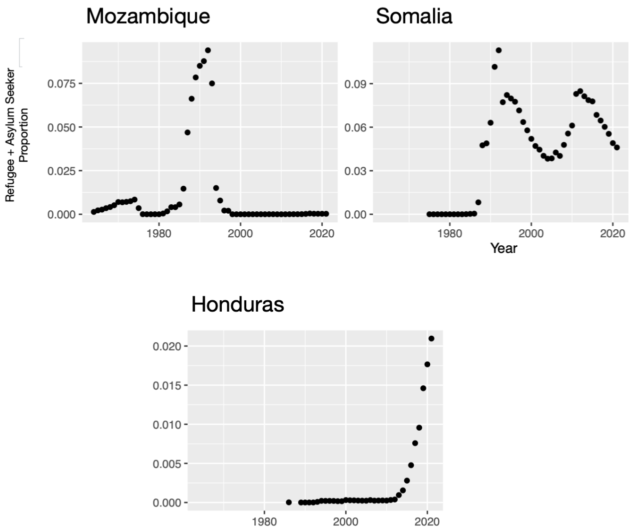
Vignette 3: Migration

- Migration is a key input to population projections.
- Projecting future migration patterns is difficult, especially for refugee and asylum seeker populations.
  - Current rule of thumb: 2/3 of refugees will return to country of origin within 5 years (UN WPP 2022 Methodology).
- We propose a modeling pipeline for projecting refugee and asylum seeker populations by country of origin. (Susmann and Raftery Demography 2025, forthcoming).

Data

- Yearly data on individuals classified as refugees and asylum seekers are sourced from from the United Nations High Commissioner on Human Rights (UNHCR 2021).
- Who counts as a refugee or asylum seeker is a fraught political and legal question. We rely on definitions used by UNHCR.
- We model refugee and asylum seeker counts as a proportion of their origin country's population.

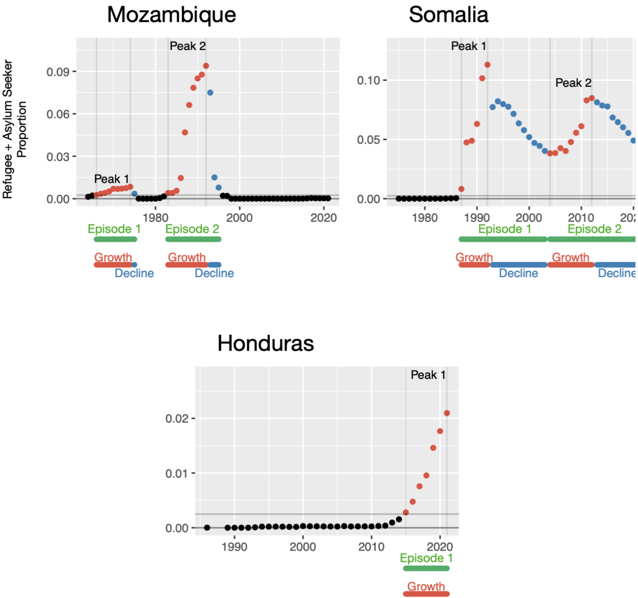
Data


Mozambique
Somalia
Honduras
Refugee + Asylum Seeker Proportion
Year
0.000
0.025
0.050
0.075
0.00
0.03
0.06
0.09
0.005
0.010
0.015
0.020
1980
2000
2020

Segmentation

- Conceptual model: refugee and asylum seeker populations follow **growth and decline phases**, separated by a **peak**.
- We call one of these cycles a "growth/decline period" or, simply, an "episode".
- The data from each origin country are **segmented** into separate episodes by a set of deterministic rules.

Segmentation



Interrupted Logistic Model

- Let $\mu_{c,t}$ be the refugee/asylum seeker proportion at time t from country of origin c.
- Rate of change in $\mu_{c,t}$ is modeled by a *logistic rate function* during the growth and decline phases.
- For example, during the growth phase:

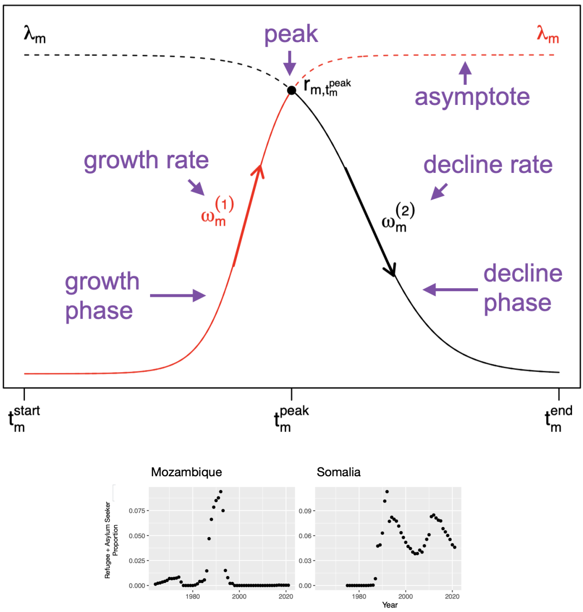
$$\mu_{c,t} = \underbrace{\mu_{c,t-1} + f(\mu_{c,t-1}, \omega, \lambda)}_{\textit{systematic}} + \underbrace{\epsilon_{c,t}}_{\textit{smoothing}}$$

where

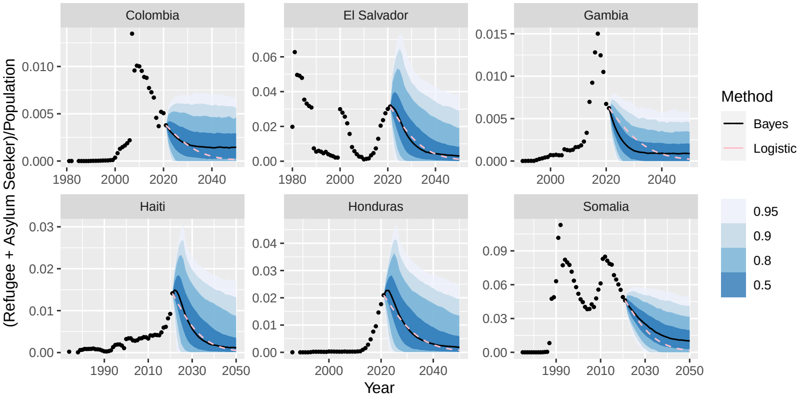
$$f(\mu_{c,t-1}, \omega, \lambda) = \begin{cases} \lambda - \mu_{c,t-1}, & \mu_{c,t-1} > \lambda, \\ \omega \cdot \mu_{c,t-1} \left(1 - \frac{\mu_{c,t-1}}{\lambda}\right), & \text{otherwise} \end{cases}$$

for asymptote λ and growth rate ω.

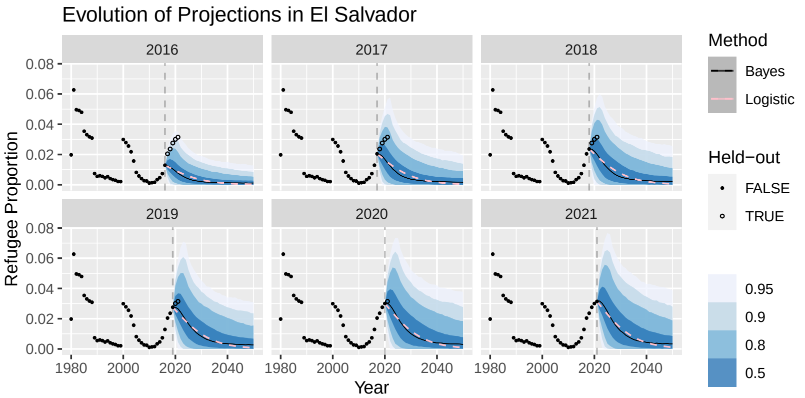
Interrupted Logistic Model



Projections in six ongoing episodes



Projections update as new data are available



Conclusion

- Many models of demographic and health indicators have a similar structure.
  - Proposed model class: Temporal Models for Multiple Populations (TMMPs).
- We have found this model class to be a useful starting point for solving real-world modeling challenges.
  - Mix/match/reuse existing components, or invent new ones if needed.
- Three vignettes
  - **Transitions:** a flexible approach for modeling indicators that follow a smooth *transition*.
  - **Shocks:** a principled Bayesian method for modeling indicators with *shocks*.
  - **Migration:** a novel interrupted logistic model for projecting refugee and asylum seeker populations.
- *My question for you:* what data and modeling challenges are you facing?

Resources

- Temporal Models for Multiple Populations
  - Paper: Susmann, Alexander, and Alkema International Statistical Review 2021.
  - Slides: `http://herbsusmann.com/paa2021`.
- Flexible B-spline Transition Models
  - Paper: Susmann and Alkema JRSS-C 2025.
  - Slides: `http://herbsusmann.com/epc2022/`.
- Shrinkage Priors for Modeling Shocks
  - Preprint: Susmann and Alkema ArXiv:2410.09217 2024.
  - Slides: `https://herbsusmann.com/paa2025/`.
- Refugee and Asylum Seeker Population Projections
  - Preprint: Susmann and Raftery arXiv:2405.06857 2024.
  - Paper: Susmann and Raftery Demography 2025 (forthcoming).
- Full list on `herbsusmann.com`.

Appendix: Model performs in validations vs benchmark

			MAE		ME	
Cutoff	Target	*n*	Bayes	Logistic	Bayes	Benchmark
1 year ahead						
2016	2017	25	**0.73**	0.82	**0.62**	0.75
2017	2018	26	**0.20**	0.28	**0.05**	0.23
2018	2019	26	**0.13**	0.20	**-0.03**	0.18
2019	2020	28	**0.15**	0.27	**0.01**	0.22
2020	2021	29	**0.26**	0.43	**0.20**	0.40
5 year ahead						
2011	2016	16	**1.46**	1.58	0.89	**0.79**
2016	2021	22	1.85	**1.76**	**1.67**	1.70
10 year ahead						
2011	2021	13	**2.26**	2.28	1.88	**1.77**

Appendix: validation coverage results

			Coverage		
Cutoff	Target	*n*	80%	90%	95%
1 year ahead					
2016	2017	25	88.0%	88.0%	88.0%
2017	2018	26	100.0%	100.0%	100.0%
2018	2019	26	96.2%	100.0%	100.0%
2019	2020	28	96.4%	96.4%	96.4%
2020	2021	29	89.7%	89.7%	93.1%
Average			94.0%	94.8%	95.5%
5 year ahead					
2011	2016	16	75.0%	81.2%	87.5%
2016	2021	22	81.8%	86.4%	100.0%
10 year ahead					
2011	2021	13	76.9%	84.6%	84.6%