Appendix to “Integrated nested Laplace approximations for extended latent Gaussian models with application to the Naomi HIV model”

Corresponding author: Adam Howes (ath19@ic.ac.uk)

Contents

1 Complete Naomi model description

1.1 Process specification

1.1.1 Notation

- District x
  - Located in Spectrum region file R_x
- Sex $s \in \{\mathrm{F}, \mathrm{M}\}$
- Age $a \in \{$0-5, 5-10, $\ldots$, 75-80, 80+$\}$
- Time $t \in \{T_1, T_2, T_3\}$
  - T_1: Most recent national household survey with HIV testing
  - T_2: Current time period at which to generate estimates
  - T_3: Short-term project period (typically 9-12 months)
  - Note $T_{1:3}$ specified to nearest calendar quarter
- Population size $N_{x,s,a,t}$
- HIV prevalence $\rho_{x,s,a,t}$
- ART coverage $\alpha_{x,s,a,t}$
- Annual HIV incidence rate $\lambda_{x,s,a,t}$
- $u \sim \mathrm{ICAR}(\sigma)$ refers to the intrinsic conditional auto-regressive model (ICAR) with marginal standard deviation $\sigma > 0$.
- $u \sim \mathrm{BYM2}(\sigma, \phi)$ refers to a reparameterised Besag-York-Mollie model (BYM2) with marginal standard deviation $\sigma > 0$ comprised of a spatially structured ICAR component with proportion $\phi \in (0, 1)$ and spatially unstructured IID component with proportion $1 - \phi$.
- For both the ICAR and BYM2 models, we follow recommendations on scaling, disconnected adjacency graph components, and islands.

1.1.2 HIV prevalence and ART coverage at T_1

1.1.2.1 HIV prevalence

HIV prevalence modelled by

$$\mathrm{logit}(\rho_{x,s,a,T_1}) = \beta_0^\rho + \beta_S^{\rho,s=\mathrm{M}} + u_a^\rho + u_a^{\rho,s=\mathrm{M}} + u_x^\rho + u_x^{\rho,s=\mathrm{M}} + u_x^{\rho,a<15} + \eta_{R_x,s,a}^\rho$$

where:

- β_0^ρ is the intercept
- $\beta_s^{\rho,s=\mathrm{M}}$ is the difference in logit prevalence for men compared to women
- $u_a^\rho \sim \mathrm{AR1}(\sigma_A^\rho, \phi_A^\rho)$ are age random effects for women
- $u_a^{\rho,s=\mathrm{M}}$ are age random effects for the difference in logit prevalence for men compared to women age a
- $u_x^\rho \sim \mathrm{BYM2}(\sigma_X^\rho, \phi_X^\rho)$ are spatial random effects for women
- $u_x^{\rho,s=\mathrm{M}} \sim \mathrm{BYM2}(\sigma_{XS}^\rho, \phi_{XS}^\rho)$ are spatial random effects for the difference in logit prevalence for men compared to women in district x
- $u_x^{\rho,a<15} \sim \mathrm{ICAR}(0, \sigma_{XA}^\rho)$ are spatial random effects for the ratio of paediatric prevalence to adult women prevalence
- $\eta_{R_x,s,a}^\rho$ are a fixed offset specifying assumed odds ratios for prevalence outside the age ranges for which data are available

For prior distributions, we use:

- $\mathcal{N}(0, 5)$ for all fixed effects $(\beta_0^\rho, \beta_s^{\rho,s=\mathrm{M}})$
- $\mathcal{N}^+(0, 2.5)$ for all standard deviation terms
- $\mathcal{U}(-1, 1)$ for all AR1 correlation parameters
- $\mathrm{Beta}(0.5, 0.5)$ for all BYM2 proportion parameters

1.1.2.2 ART coverage

ART coverage modelled by

$$\mathrm{logit}(\alpha_{x,s,a,T_1}) = \beta_0^\alpha + \beta_S^{\alpha,s=\mathrm{M}} + u_a^\alpha + u_a^{\alpha,s=\mathrm{M}} + u_x^\alpha + u_x^{\alpha,s=\mathrm{M}} + u_x^{\alpha,a<15} + \eta_{R_x,s,a}^\alpha$$

where terms are analogous to the HIV prevalence model.

1.1.2.3 HIV incidence rate

HIV incidence rate is modelled by

$$\log(\lambda_{x,s,a,t}) = \beta_0^\lambda + \beta_S^{\lambda,s=\text{M}} + \log(\rho_{x,t}^{\text{15-49}}) + \log(1 - \omega \cdot \alpha_{x,t}^{\text{15-49}}) + u_x^\lambda + \eta_{R_x,s,a,t}^\lambda$$

where:

- β_0^λ is the intercept, which is proportional to the average HIV transmission rate for untreated HIV positive adults
- $\beta_S^{\lambda,s=\text{M}}$ is the log incidence rate ratio for men compared to women
- $\rho_{x,t}^{\text{15-49}}$ is the HIV prevalence among adults 15-49 calculated by

$$\rho_{x,t}^{\text{15-49}} = \frac{\sum_{s \in \{\text{F},\text{M}\}} \sum_{a=15}^{49} N_{x,s,a,t} \cdot \rho_{x,s,a,t}}{\sum_{s \in \{\text{F},\text{M}\}} \sum_{a=15}^{49} N_{x,s,a,t}}$$

- $\alpha_{x,t}^{\text{15-49}}$ is the ART coverage among adults 15-49 calculated by

$$\alpha_{x,t}^{\text{15-49}} = \frac{\sum_{s \in \{\text{F},\text{M}\}} \sum_{a=15}^{49} N_{x,s,a,t} \cdot \rho_{x,s,a,t} \cdot \alpha_{x,s,a,t}}{\sum_{s \in \{\text{F},\text{M}\}} \sum_{a=15}^{49} N_{x,s,a,t} \cdot \rho_{x,s,a,t}}$$

- ω is the average reduction in HIV transmission rate per 1% increase in population ART coverage and is fixed at $\omega = 0.7$ from the EPP model
- $u_x^\lambda \sim \mathcal{N}(0, \sigma^\lambda)$ with $\sigma^\lambda \sim \mathcal{N}^+(0,1)$ are IID spatial random effects
- $\eta_{R_x,s,a,t}^\lambda$ specify log incidence rate ratios by sex and age group calculated from Spectrum model output

1.1.3 Short-term projection from $T_1 \to T_2$ and $T_2 \to T_3$

1.1.3.1 HIV prevalence

- HIV population projected from $T_1 \to T_2$ based on survival of PLHIV from Spectrum output, ageing from one five-year age group to the next, and the addition of new HIV infections calculated from district-level HIV incidence rate and survival and ageing after infection
- Let $H_{x,s,a,T_1} = N_{x,s,a,T_1} \cdot \rho_{x,s,a,T_1}$ be the number of PLHIV at T_1
- For age groups $a \geq 5$ the number of PLHIV at T_2 is modelled by

$$\begin{aligned} H_{x,s,a,T_2} = {} & H_{x,s,a,T_1} \cdot S_{R_x,s,a\to a,T_1} + H_{x,s,a-5,T_1} \cdot S_{R_x,s,a-5\to a,T_1} \\ & + (1 - \exp(-\delta_{T_1} \cdot \lambda_{x,s,a,T_1})) \cdot (N_{x,s,a,T_1} - H_{x,s,a,T_1}) \cdot L_{R_x,s,a\to a,T_1} \\ & + (1 - \exp(-\delta_{T_1} \cdot \lambda_{x,s,a-5,T_1})) \cdot (N_{x,s,a-5,T_1} - H_{x,s,a-5,T_1}) \cdot L_{R_x,s,a-5\to a,T_1} \end{aligned}$$

- $S_{R_x,s,a\to a,T_1}$ is the ratio for the number of PLHIV surviving from T_1 to T_2 and remaining in age group a (i.e. not aging from a to $a+5$)
- $S_{R_x,s,a-5\to a,T_1}$ the survival ratio from T_1 to T_2 and aging from $a-5$ to a
- $(1-\exp(-\delta_{T_1} \cdot \lambda_{x,s,a,T_1}))$ is the probability of acquiring HIV between T_1 and T_2 of duration $\Delta_{T_1} = T_2 - T_1$ for the susceptible population $N_{x,s,a,T_1} - H_{x,s,a,T_1}$
- $L_{R_x,s,a\to a,T_1}$ is the ratio surviving and remaining in age group a for those infected between T_1 and T_2
- $L_{R_x,s,a-5\to a,T_1}$ is the ratio surviving and aging from $a-5$ to a for those infected between T_1 and T_2
- The ratios S and L are calculated from Specturm results by
  1. Disaggregating single-year or single-age Specturm model results for PLHIV at T_1 and T_2 and new infections between T_1 and T_2 to quarterly birth cohorts
  2. Subtracting the number of new infections within the cohort from the number of PLHIV at T_2 to calculate the number of surviving PLHIV in each cohort
  3. Aggregating the survivors by quarter-age cohorts to those who survived and aged from a to a' between T_1 and T_2, and calculating the ratio by dividing by the initial PLHIV H_{x,s,a,T_1}
- Variation in ART coverage or effectiveness is assumed not to affect survival

1.1.3.2 ART coverage Change in ART coverage is modelled by

$$\text{logit}(\alpha_{x,s,a,T_2}) = \text{logit}(\alpha_{x,s,a,T_1}) + \beta^{\alpha}_{T_2} + \beta^{\alpha}_{T_2,s=\text{M}} + u^{\alpha}_{x,T_2} + u^{\alpha}_{x,a<15,T_2} + \eta^{\alpha}_{R_x,s,a,T_2}$$

where

- $\eta^{\alpha}_{R_x,s,a,T_2}$ are offsets for the logit change in ART coverage between T_1 and T_2 from Spectrum region R_x
- $\beta^{\alpha}_{T_2}$ is the average change in logit ART coverage between T_1 and T_2 for adult women
- $\beta^{\alpha}_{T_2,s=\text{M}}$ is the difference in average change for men compared to women
- $u^{\alpha}_{x,T_2} \sim \mathcal{N}(0, \sigma^{\alpha}_{XT})$ are district random effects
- $u^{\alpha}_{x,a<15,T_2}$ is the difference in logit ART coverage for children compared to adults

ART coverage at T_3 is modelled by[1]

$$\text{logit}(\alpha_{x,s,a,T_2}) = \text{logit}(\alpha_{x,s,a,T_1}) + \eta^{\alpha}_{R_x,s,a,T_3}$$

1.1.4 ANC testing cascade

- For women 15-49 the predicted number of ANC clients $\Psi_{x,a,t}$ is a log-linear model

$$\log(\Psi_{x,a,t}) = \log(N_{x,\text{F},a,t}) + \psi_{R_x,a,t} + \beta^{\psi} + u^{\psi}_x$$

  - $N_{x,\text{F},a,t}$ are the female population sizes
  - $\psi_{R_x,a,t}$ are ASFR in Spectrum region R_x at time t
  - β^{ψ} are the log rate ratio for the number of ANC clients relative to the predicted fertility
  - $u^{\psi}_x \sim \mathcal{N}(0, \sigma^{\psi})$ are district random effects
- HIV prevalence $\rho^{\text{ANC}}_{x,a,t}$ and ART coverage $\alpha^{\text{ANC}}_{x,a,t}$ among pregnant women modelled with logit-linear models

$$\text{logit}(\rho^{\text{ANC}}_{x,a,t}) = \text{logit}(\rho_{x,F,a,t}) + \beta^{\rho^{\text{ANC}}} + \beta^{\rho^{\text{ANC}}}_{t\in(T_2,T_3\}} + u^{\rho^{\text{ANC}}}_x + u^{\rho^{\text{ANC}}}_{x,t\in\{T_2,T_3\}} + \eta^{\rho^{\text{ANC}}}_{R_x,a,t} \tag{1}$$

$$\text{logit}(\alpha^{\text{ANC}}_{x,a,t}) = \text{logit}(\alpha_{x,F,a,t}) + \beta^{\alpha^{\text{ANC}}} + \beta^{\alpha^{\text{ANC}}}_{t\in\{T_2,T_3\}} + u^{\alpha^{\text{ANC}}}_x + u^{\alpha^{\text{ANC}}}_{x,t\in\{T_2,T_3\}} + \eta^{\alpha^{\text{ANC}}}_{R_x,a,t} \tag{2}$$

  - $\eta^{\theta^{\text{ANC}}}_{R_x,a,t}$ for $\theta \in \{\rho, \alpha\}$ are offsets for the log fertility rate ratios for HIV positive women compared to HIV negative women and for women on ART to HIV positive women not on ART, calculated from Spectrum model outputs for region R_x
  - $\beta^{\theta^{\text{ANC}}}$ for $\theta \in \{\rho, \alpha\}$ are the average differences between population and ANC outcomes at T_1 after removing the offset
  - $\beta^{\theta^{\text{ANC}}}_{t\in(T_2,T_3\}}$ for $\theta \in \{\rho, \alpha\}$ are the change in average difference from T_1 to T_2
  - $u^{\theta^{\text{ANC}}}_x \sim \mathcal{N}(0, \sigma^{\theta^{\text{ANC}}}_X)$ for $\theta \in \{\rho, \alpha\}$ are district random effects
  - $u^{\theta^{\text{ANC}}}_{x,t\in\{T_2,T_3\}} \sim \mathcal{N}(0, \sigma^{\theta^{\text{ANC}}}_{XT})$ for $\theta \in \{\rho, \alpha\}$ capture the change in district random effects between times T_1 and T_2

1.1.5 ANC attendance

- Number of PLHIV on ART is $A_{x,s,a,t} = N_{x,s,a,t} \cdot \rho_{x,s,a,t} \cdot \alpha_{x,s,a,t}$
- Let $\gamma_{x,x',t} \in [0, 1]$ be the probability that a person of ART residing in district x recieves ART in district x' at time t
- We assume that $\gamma_{x,x',t} = 0$ for $x \notin \{x, \text{ne}(x)\}$ i.e. that individuals seek treatment only in their residing district and its neighbours $\text{ne}(x) = \{x' : x' \sim x\}$ where $\sim$ is an adjacency relation
  - $\sum_{x'\in\{x,\text{ne}(x)\}} \gamma_{x,x',t} = 1$

[1] Don't understand this.

- Model $\gamma_{x,x',t}$ with a multinomial logit model where $\tilde{\gamma}_{x,t}$ is[2] the log odds ratio of seeking ART in the district x' neighbouring x

$$\tilde{\gamma}_{x,t} = \tilde{\gamma}_0 + u_x^{\tilde{\gamma}} + u_x^{\tilde{\gamma},t=T_2} \tag{3}$$
$$\tilde{\gamma}_0 = -4 \tag{4}$$
$$u_x^{\tilde{\gamma}} \sim \mathcal{N}(0, \sigma_X^{\tilde{\gamma}}) \tag{5}$$
$$u_x^{\tilde{\gamma},t\in\{T_2,T_3\}} \sim \mathcal{N}(0, \sigma_{XT}^{\tilde{\gamma}}) \tag{6}$$

  - $\tilde{\gamma}_{x,t}$ is an intercept specifying the prior mean for the log odds of seeking ART in each neighbouring district compared to the home district
  - $u_x^{\tilde{\gamma}}$ are district random effects
  - $u_x^{\tilde{\gamma},t\in\{T_2,T_3\}}$ is[3] the change in district level log-odds ratios between T_1 and T_2
  - Note: only one value estimated for each district, implying no preference between neighbouring districts, and no sex-age specific modelling
- Calculate multinomial probabilities via softmax (n.b. this is the sigmoid function, the inverse of the softmax)

$$\gamma_{x,x',t} = \frac{1}{1 + \sum_{x^\star \sim x} \exp(\tilde{\gamma}_{x^\star,t})}$$

- The number of ART clients who reside in district x and obtain ART in district x' are $A_{x,x',s,a,t} = A_{x,s,a,t} \cdot \gamma_{x,x',t} = N_{x,s,a,t} \cdot \rho_{x,s,a,t} \cdot \alpha_{x,s,a,t} \cdot \gamma_{x,x',t}$
- Total attending ART facilities in district x' is $\tilde{A}_{x',s,a,t} = \sum_{x \sim x', x=x'} A_{x,x',s,a,t}$

1.1.6 Awareness of HIV status

- The proportion of HIV positive adults aware of their HIV status is $v_{x,s,a,t}$
- The proportion of untreated HIV positive adults aware of their HIV status is $\tilde{\nu}_{x,s,a,t}$ such that

$$v_{x,s,a,t} = \alpha_{x,s,a,t} + \frac{\tilde{v}_{R_x,s,a,t}}{1 - \alpha_{x,s,a,t}}$$

1.2 Likelihood specification

1.2.1 Notation

- Use $\{x\}$ to denote a collection of districts, $\{s\}$ a collection of sexes and $\{a\}$ a collection of ages for which an outcome is observed

1.2.2 Household survey data

- Let ν be a household survey occuring at T_1
- Furnishes weighted observations for HIV prevalence $\hat{\rho}_{\{x\},\{s\},\{a\},\nu}$, ART coverage $\hat{\alpha}_{\{x\},\{s\},\{a\},\nu}$ and proportion recently infected $\hat{\kappa}_{\{x\},\{s\},\{a\},\nu}$ with respective Kish effective sample sizes $M^{\hat{\theta}}_{\{x\},\{s\},\{a\},\nu}$ for $\theta \in \{\rho, \alpha, \kappa\}$
- The observed cases are $Y^{\hat{\theta}}_{\{x\},\{s\},\{a\},\nu} = M^{\hat{\theta}}_{\{x\},\{s\},\{a\},\nu} \cdot \hat{\theta}_{\{x\},\{s\},\{a\},\nu}$
- For HIV prevalence we use a binomial working likelihood

$$Y^{\hat{\rho}}_{\{x\},\{s\},\{a\},\nu} \sim \text{Bin}(M^{\hat{\rho}}_{\{x\},\{s\},\{a\},\nu}, \rho_{\{x\},\{s\},\{a\},T_1}) \tag{7}$$
$$\rho_{\{x\},\{s\},\{a\},t} = \frac{\sum_{x\in\{x\}} \sum_{s\in\{s\}} \sum_{a\in\{a\}} N_{x,s,a,t} \cdot \rho_{x,s,a,t}}{\sum_{x\in\{x\}} \sum_{s\in\{s\}} \sum_{a\in\{a\}} N_{x,s,a,t}} \tag{8}$$

[2] Confused by this notation.
[3] Shouldn't this just be $t = T_2$ here rather than $t \in \{T_2, T_3\}$?

- For ART coverage we also use a binomial working likelihood

$$Y^{\hat{\alpha}}_{\{x\},\{s\},\{a\},\nu} \sim \text{Bin}(M^{\hat{\alpha}}_{\{x\},\{s\},\{a\},\nu}, \alpha_{\{x\},\{s\},\{a\},T_1}) \tag{9}$$

$$\alpha_{\{x\},\{s\},\{a\},t} = \frac{\sum_{x\in\{x\}}\sum_{s\in\{s\}}\sum_{a\in\{a\}} N_{x,s,a,t}\cdot\rho_{x,s,a,t}\cdot\alpha_{x,s,a,t}}{\sum_{x\in\{x\}}\sum_{s\in\{s\}}\sum_{a\in\{a\}} N_{x,s,a,t}\cdot\rho_{x,s,a,t}} \tag{10}$$

- For recent infections we also use a binomial working likelihood

$$Y^{\hat{\kappa}}_{\{x\},\{s\},\{a\},\nu} \sim \text{Bin}(M^{\hat{\kappa}}_{\{x\},\{s\},\{a\},\nu}, \kappa_{\{x\},\{s\},\{a\},T_1}) \tag{11}$$

$$\kappa_{x,s,a,t} = 1 - \exp(-\lambda_{x,s,a,t}\cdot\frac{1-\rho_{x,s,a,t}}{\rho_{x,s,a,t}}\cdot(\Omega_T-\beta_T)-\beta_T) \tag{12}$$

$$\Omega_T \sim \mathcal{N}(\Omega_{T_0}, \sigma^{\Omega_T}) \tag{13}$$

$$\beta_T \sim \mathcal{N}(\beta_{T_0}, \sigma^{\beta_T}) \tag{14}$$

  - $\kappa_{x,s,a,t}$ are the predicted proportion recently infected among HIV positive persons
  - Ω_T is the mean duration of recent infection (MDRI)
    - Use an informative prior based on the characteristics of the recent infection testing algorithm (RITA)
    - For PHIA surveys[4] $\Omega_{T_0} = 130$ days and $\sigma^{\Omega_T} = 6.12$ days
  - β_T is the false recent ratio (FRR)
    - For PHIA surveys $\beta_{T_0} = 0.0$ and $\sigma^{\beta_T} = 0.0$.[5]

1.2.3 ANC testing data

- Include two years of ANC testing data: the year of the most recent survey $Y[T_1]$ and the current year $Y[T_2]$
- $W^{\text{ANC}}_{\{x\},Y[t]}$ are the number of ANC clients
- $X^{\text{ANC}}_{\{x\},Y[t]}$ are the number of ANC clients with ascertained status
- $Y^{\text{ANC}}_{\{x\},Y[t]}$ are the number of ANC clients with positive status (either known or tested)
- $Z^{\text{ANC}}_{\{x\},Y[t]}$ are the number of ANC clients already on ART prior to first ANC
- Sometimes ANC testing data are only available for part of a given year, for example only the first three quarters. Denote $M^{\text{ANC}}_{Y[t]} \in \{1,\ldots,12\}$ the number of months of reported data reflected in counts for year $Y[t]$
- Likelihood for number of ANC clients specified only[6] for year $Y[T_2]$ as

$$W^{\text{ANC}}_{\{x\}Y[T_2]} \sim \text{Pois}\left(\frac{M^{\text{ANC}}_{Y[T2]}}{12}\sum_{x\in\{x\}}\sum_{a\in\{15,\ldots45\}}\Psi_{x,a,T_2}\right)$$

- Observed number of HIV positive and already on ART among ANC clients at both $Y[T_1]$ and $Y[T_2]$ are modelled by

$$\begin{aligned} Y^{\text{ANC}}_{\{x\}Y[t]} &\sim \text{Bin}\left(X^{\text{ANC}}_{\{x\}Y[t]}, \rho^{\text{ANC}}_{\{x\},\{15,\ldots49\},t}\right) \\ Z^{\text{ANC}}_{\{x\}Y[t]} &\sim \text{Bin}\left(Y^{\text{ANC}}_{\{x\}Y[t]}, \alpha^{\text{ANC}}_{\{x\},\{15,\ldots49\},t}\right) \end{aligned}$$

where predicted prevalence and ART coverage are aggregated weighted by the predicted number of

[4] What about other surveys?
[5] Is this just assuming that there is no false recency? How is this assumption justified?
[6] Why?

pregnant women by age $\Psi_{x,a,t}$

$$\rho^{\text{ANC}}_{\{x\}\{a\},t} = \frac{\sum_{x\in\{x\}}\sum_{a\in\{a\}}\Psi_{x,a,t}\cdot\rho^{\text{ANC}}_{x,a,t}}{\sum_{x\in\{x\}}\sum_{a\in\{a\}}\Psi_{x,a,t}}$$

$$\alpha^{\text{ANC}}_{\{x\}\{a\},t} = \frac{\sum_{x\in\{x\}}\sum_{a\in\{a\}}\Psi_{x,a,t}\cdot\rho^{\text{ANC}}_{x,a,t}\cdot\alpha^{\text{ANC}}_{x,a,t}}{\sum_{x\in\{x\}}\sum_{a\in\{a\}}\Psi_{x,a,t}\cdot\rho^{\text{ANC}}_{x,a,t}}$$

1.2.4 Number receiving ART

- Let $\dot{A}_{\{x\},\{s\},\{a\},t}$ be data for the number receiving ART[7]

$$\dot{A}_{\{x\},\{s\},\{a\},t} = \sum_{s\in\{s\}}\sum_{a\in\{a\}}\sum_{x\in\{x\}}\sum_{x\sim x',x=x'}\dot{A}_{x',x,s,a,t}$$

- Model the unobserved numbers of ART clients travelling from x' to x as

$$\dot{A}_{x',x,s,a,t} \sim \text{Bin}(N_{x',s,a,t}, \pi_{x',x,s,a,t})$$

where $\pi_{x',x,s,a,t} = \rho_{x',s,a,t}\cdot\alpha_{x',s,a,t}\cdot\gamma_{x',s,a,t}$[8]
- This likelihood is approximated using a normal for the sum of binomials

$$\dot{A}_{\{x\},\{s\},\{a\},t} \sim \mathcal{N}(\tilde{)}A_{\{x\},\{s\},\{a\},t}, \sigma^{\tilde{A}}_{\{x\},\{s\},\{a\},t})$$

where
  - $\tilde{A}_{\{x\},\{s\}\{a\},t} = \sum_{s\in\{s\}}\sum_{a\in\{a\}}\sum_{x\in\{x\}}\sum_{x\sim x',x=x'} N_{x',s,a,t}\cdot\pi_{x',x,s,a,t}$
  - $\sigma^{\tilde{A}}_{\{x\},\{s\}\{a\},t} = \sqrt{\sum_{s\in\{s\}}\sum_{a\in\{a\}}\sum_{x\in\{x\}}\sum_{x\sim x',x=x'} N_{x',s,a,t}\cdot\pi_{x',x,s,a,t}\cdot(1-\pi_{x',x,s,a,t})}$

1.3 Identifiability constaints

- Depends on data availability. If data is missing, some parameters might be fixed to default values
- Survey data on ART coverage by age and sex not available then set $u^{\alpha}_{a} = 0$ and $u^{\alpha}_{a,s=\text{M}} = 0$ and use the average age/sex pattern of ART coverage from the Spectrum offset $\eta^{\alpha}_{R_x,s,a}$
- If there is no ART data at T_1 (either survey or ART programme) or no ART data at T_2 then change in ART coverage is unidentifiable and set $\beta^{\alpha}_{T_2} = \beta^{\alpha}_{T_2,s=\text{M}} = u^{\alpha}_{x,T_2} = u^{\alpha}_{x,a<15,T_2} = 0$ such that change in ART coverage is modelled by

$$\text{logit}(\alpha_{x,s,a,T_2}) = \text{logit}(\alpha_{x,s,a,T_1}) + \eta^{\alpha}_{R_x,s,a,T_2}$$

- If no ART data (survey or ART programme) are available at T_1 or T_2 but data on ART coverage among ANC clients are available, the level of ART coverage is not identifiable, but spatial variation is identifiable. In this instance, overall ART coverage is determined by the Spectrum offset, and only area random effects are estimated

$$\text{logit}\left(\alpha_{x,s,a,T_1}\right) = u^{\alpha}_{x} + \eta^{\alpha}_{R_x,s,a}$$
$$\text{logit}\left(\alpha_{x,s,a,T_2}\right) = \text{logit}\left(\alpha_{x,s,a,T_1}\right) + u^{\alpha}_{x,T_2} + \eta^{\alpha}_{R_{x,s,a,T_2}}$$

- If survey data on recent HIV infection are not included in the model, then $\beta^{\lambda}_{0} = \beta^{\lambda,s=\text{M}}_{S} = u^{\lambda}_{x} = 0$. The sex ratio for HIV incidence is determined by the sex incidence rate ratio from Spectrum in the same years and the incidence rate in all districts is modelled assuming the same average HIV transmission rate for untreated adults, but varies according to district estimates of HIV prevalence and ART coverage

References

[7] I'm confused about the notation here. In general would be good to define precisely what is meant by $\{x\}$. Presumably the LHS below is a set, yet the RHS is a sum over the sets so I think just a number – so how can the two be equal?

[8] Should this be $\gamma_{x',x,s,a,t}$ since it's people moving from x' to x (and so far no where in the RHS is there an x which is suspicious)