Appendix A: Mathematical approach to modelling sexual behaviour

The purpose of this appendix is to explain, in mathematical terms, the modelling of sexual behaviour. Table A.1 summarizes the index variables that are used throughout this appendix.

Table A.1: Index variables

Symbol	Definition	State space
i	Individual risk group	01 = virgin with propensity for concurrency 02 = virgin with no propensity for concurrency 1 = Non-virgin, with propensity for concurrency 2 = Non-virgin, with no propensity for concurrency 3 = Commercial sex worker (relevant to females only)
j	Risk group(s) of partner(s)	0 = no partner; 1 = 1 high risk partner; 2 = 1 low risk partner; 11 = 2 high risk partners; 12 = primary high risk & secondary low risk; 21 = primary low risk & secondary high risk; 22 = 2 low risk partners*
l	Relationship type	1 = short-term (non-marital) 2 = long-term (marital)†
x	Individual age group	10, 15, 20, …, 85
y	Partner age	10, 15, 20, …, 85
g	Sex	1 = male; 2 = female
s	HIV disease state	0 = uninfected; 1 = acute HIV; 2 = asymptomatic HIV; 3 = WHO clinical stage 3; 4 = AIDS; 5 = on HAART
t	Time	0 to 40 (in years from mid-1985)

HAART = highly active antiretroviral treatment.
* Where the individual is in a marital relationship with one partner and a non-marital relationship with another, the first index refers to the risk group of the spouse and the second refers to the risk group of the other partner. † Where the individual has two partners, this index refers to the nature of the primary partnership (the secondary relationship is always short-term). Where the individual has no partners, the index is omitted.

Symbols are defined as follows:
K = number of cycles per year, i.e. the frequency at which sexual behaviour variables are updated (the default value is 12, but K can be any integer greater than one)
π = gender equality factor (explained below)
$N^{s}_{g,i,j,l}(x,t)$ = number of individuals of sex *g*, risk group *i*, aged *x*, who are in HIV disease state *s*, in relationship type *l* with partner(s) in group(s) *j*, at time *t*

A.1 Rates at which short-term partnerships are formed

To determine the rates at which short-term partnerships are formed, it is necessary to define the following symbols:
$c^{s}_{g,i,j,l}(x)$ = rate at which individual wishes to form short-term partnerships if they are of sex *g*, risk group *i*, aged *x*, in HIV disease state *s*, in relationship type *l* with a partner in group *j* (*j* = 0, 1 or 2 only)
$\rho_{g,i,j}(t)$ = desired proportion of new short-term partners who are in risk group *j*, if individual is of sex *s* and in risk group *i* (*j* = 1 or 2 only)

$\rho_{g,i,j}(t)$ is calculated according to the following formula:

$$\rho_{g,i,j}(t) = (1-\varepsilon)\delta_{ij} + \varepsilon \frac{\sum_{u=0}^{2}\sum_{l}\sum_{y}\sum_{s=0}^{5} N_{g^*,j,u,l}^{s}(y,t) c_{g^*,j,u,l}^{s}(y)}{\sum_{v=1}^{2}\sum_{u=0}^{2}\sum_{l}\sum_{y}\sum_{s=0}^{5} N_{g^*,v,u,l}^{s}(y,t) c_{g^*,v,u,l}^{s}(y)}, \qquad \text{(A1)}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, g^* is the sex opposite to g, and ε is the degree of assortative mixing. The degree of assortative mixing can be any value on the interval [0, 1], with lower values of the parameter indicating greater tendency to form partnerships with individuals in the same sexual activity class.

In order to balance the male and female rates of short-term partnership formation, it is necessary to define the following parameter:
$B_{g,i,j}^{1}(t)$ = adjustment factor applied to the rate at which individuals of sex g, in risk group i, form short-term partnerships with individuals in risk group j (j = 1 or 2 only).

For women ($g = 2$):

$$B_{2,i,j}^{1}(t) = \frac{\left[\pi\sum_{s=0}^{5}\sum_{u}\sum_{v}\sum_{y} N_{2,i,u,v}^{s}(y,t)c_{2,i,u,v}^{s}(y)\rho_{2,i,j}(t) + (1-\pi)\sum_{s=0}^{5}\sum_{u}\sum_{v}\sum_{y} N_{1,j,u,v}^{s}(y,t)c_{1,j,u,v}^{s}(y)\rho_{1,j,i}(t)\right]}{\sum_{s=0}^{5}\sum_{u}\sum_{v}\sum_{y} N_{2,i,u,v}^{s}(y,t)c_{2,i,u,v}^{s}(y)\rho_{2,i,j}(t)}. \qquad \text{(A2)}$$

The gender equality factor (π) can assume any value on the interval [0, 1]. The closer π is to 1, the closer the adjustment factor $B_{2,i,j}^{1}(t)$ is to 1. A similar formula is used to define the adjustment factor for males ($B_{1,i,j}^{1}(t)$), which approaches 1 as π approaches 0. The gender equality factor thus determines the extent to which rates of partnership formation are driven by male and female desires. An equality factor of 0.5 (the default value) would imply that men and women have equal control over the formation of partnerships, while an equality factor of 0 would effectively imply that women are coerced into sexual relationships and have no control over the rate at which they enter partnerships.

The independent probability of forming a new short-term partnership with a partner in risk group u, over the time period $[t, t+1/K)$, is calculated as

$$1-\exp\left(-c_{g,i,j,l}^{s}(x)B_{g,i,u}^{1}(t)\rho_{g,i,u}(t)/K\right) \qquad \text{(A3)}$$

for j = 0, 1 or 2, and is set to 0 for all other values of j (since it is assumed that an individual who already has two partners cannot form further partnerships).

A.2 Rates at which long-term (marital) partnerships are formed

Analogous to the previous section, the following variables are defined:
$m^s_{g,i,j}(x)$ = rate at which an individual wishes to marry a short-term partner in group *j* (*j* = 1 or 2 only) if they are of sex *g*, risk group *i*, aged *x*, in HIV disease state *s*
$B^2_{g,i,j}(t)$ = adjustment factor applied to the rate at which individuals of sex *g*, in risk group *i*, marry individuals in risk group *j* if they are currently in a short-term partnership with them (*j* = 1 or 2 only)

Similarly to the previous section, $B^2_{g,i,j}(t)$ is calculated for women as:

$$B^2_{2,i,j}(t) = \frac{\left[\begin{array}{l}\pi\sum_{s=0}^{5}\sum_u\sum_y N^s_{2,i,u,1}(y,t)m^s_{2,i,j}(y)I_{u,1}(j,1) \\ +(1-\pi)\sum_{s=0}^{5}\sum_u\sum_y N^s_{1,j,u,1}(y,t)m^s_{1,j,i}(y)I_{u,1}(i,1)\end{array}\right]}{\sum_{s=0}^{5}\sum_u\sum_y N^s_{2,i,u,1}(y,t)m^s_{2,i,j}(y)I_{u,1}(j,1)}, \qquad \text{(A4)}$$

where $I_{u,v}(j,l)$ is the number of partners of type *l* in risk group *j*, if the individual has partner(s) in group(s) *u*, with the primary partnership being of type *v*. It is thus assumed that the probability of marriage to a particular partner is independent of any other short-term partnerships the individual is in.

The independent probability that an individual in a short-term relationship with a partner in group *j*, at time *t*, becomes married to that partner over the time period $[t, t+1/K)$, is then

$$1-\exp\left(-m^s_{g,i,j}(x)B^2_{g,i,j}(t)/K\right), \qquad \text{(A5)}$$

provided the individual is not already married to another partner.

A.3 Rates at which partnerships are terminated

Partnerships can be terminated through death of the partner, through divorce (in the case of marital relationships) or through 'break up' (in the case of short-term relationships). In order to calculate rates of termination, it is necessary to define the following variables:
$D_{g,l}(x)$ = annual rate at which partnerships of type *l* dissolve, among individuals aged *x*, of gender *g* (ignoring mortality)
$f_g(y \mid x)$ = proportion of partners in age band *y*, if individual is of gender *g* and in age band *x*
$\mu^s_g(x,t)$ = force of mortality at time *t*, in individuals aged *x*, of gender *g*, who are in HIV disease state *s*

For a man who is of age *x*, in group *i*, in relationship type *l* with a partner in group *j* at time *t*, the independent probability of the relationship being terminated over the time period $\left[t, t+1/K\right)$ is

$$1-\exp\left(-\frac{1}{K}\left(D_{1,l}(x)+\sum_{y} f_1(y\mid x)\frac{\sum_{s=0}^{5} N_{2,j,\bullet}^{s}(y,t)\mu_2^{s}(y,t)}{\sum_{s=0}^{5} N_{2,j,\bullet}^{s}(y,t)}\right)\right), \qquad \text{(A6)}$$

where $N_{g,j,\bullet}^{s}(y,t)$ represents the sum across all behavioural states involving a relationship type *l* with a partner in group *i*. A similar formula is used to determine the probability that a woman's partnership is terminated.

A.4 Rates at which women become sex workers

In order to calculate the rate at which women become sex workers, it is necessary to specify the following variables:
C = average annual number of sex acts a sex worker has with clients
$w_{i,j,l}(x)$ = rate at which men in group *i*, aged *x*, visit sex workers when in relationship type *l* with partner(s) in group(s) *j* ($w_{i,j,l}(x)$ = 0 if $i \neq 1$)
$W^{s}(x)$ = factor by which the rate of recruitment into the 'sex worker' group is multiplied when the woman is of age *x* and in HIV disease state *s*
$\Delta_c\left(t, t+1/K\right)$ = the number of new sex workers required over the period $\left[t, t+1/K\right)$ in order to satisfy male demand

The variable $\Delta_c\left(t, t+1/K\right)$ is calculated as

$$\frac{1}{C}\sum_{s=0}^{5}\sum_{j}\sum_{l}\sum_{x} N_{1,1,j,l}^{s}(x,t)w_{1,j,l}(x)-\sum_{s=0}^{5}\sum_{x} N_{2,3,0}^{s}(x,t). \qquad \text{(A7)}$$

The independent probability that a woman in the 'high risk' group, who has no partners and is aged *x* and in HIV disease state *s* at time *t*, becomes a sex worker over the period $\left[t, t+1/K\right)$ is calculated as

$$\frac{\Delta_c\left(t, t+1/K\right)W^{s}(x)}{\sum_{s=0}^{5}\sum_{u} N_{2,1,0}^{s}(u,t)W^{s}(u)}. \qquad \text{(A8)}$$

A.5 Rates at which youth become sexually experienced

The independent probability that a male virgin aged *x* at time *t*, with a propensity for commercial sex and concurrent relationships, has his first sexual encounter over the time interval $\left[t, t+1/K\right)$ is

$$1 - \exp\left(-c^{s}_{1,01,0}(x)/K\right). \tag{A9}$$

This expression is a function of *s*, the HIV disease state of the male, though the only HIV-infected virgins would be those infected through mother-to-child transmission. Similar equations are used to determine the rates at which females and males in the 'low risk' group become sexually experienced.

A.6 Calculating movements between behavioural states

In order to calculate the movements between behavioural states over the $[t, t+1/K)$ time interval, it is necessary to convert the independent probabilities described previously into dependent probabilities, i.e. probabilities net of competing decrements. Suppose that there are *m* possible movements out of state S_v (excluding death and ignoring changes in *x* and *s*), and the potential states to which an individual can move are denoted $S_1, S_2, S_3, ..., S_m$. Suppose that the independent probability of moving from state S_v to state S_u is q_u over the $[t, t+1/K)$ time period. If it is assumed that in the absence of other decrements, the times of transition from state S_v to state S_u would be uniformly distributed over the interval $[t, t+1/K)$, then it can be shown (Neill 1977) that the dependent probability of movement from state S_v to state S_u is

$$\begin{aligned}(aq)_u &= q_u\left(1 - \frac{1}{2}\sum_{i=1}^{m} q_i + \frac{1}{3}\sum_{i=1}^{m}\sum_{j>i} q_i q_j - \frac{1}{4}\sum_{i=1}^{m}\sum_{j>i}\sum_{k>j} q_i q_j q_k + ...\right) \\ &\quad \times \exp\left(-\mu^{s}_{g}(x,t)/K\right)\end{aligned} \tag{A10}$$

Since the mortality rate is assumed to be independent of changes between states over the $[t, t+1/K)$ time interval (ignoring changes in *x* and *s*), the dependent rate of mortality is effectively the same as the independent rate of mortality. The dependent probability of remaining in state S_v over the interval $[t, t+1/K)$ is

$$\begin{aligned}(ap)_v &= \exp\left(-\mu^{s}_{g}(x,t)/K\right)\prod_{u=1}^{m}\left(1 - q_u\right) \\ &= \exp\left(-\mu^{s}_{g}(x,t)/K\right) - \sum_{u=1}^{m}(aq)_u\end{aligned} \tag{A11}$$

The probabilities of movement between the different behavioural states are the $(ap)_v$ and $(aq)_u$ values.

The assumption that the times of transition from state S_v to state S_u are uniformly distributed over the interval $[t, t+1/K)$ can be problematic in certain situations, and more accurate methods to estimate the transition probability matrix elements from the transition intensities have been proposed (Hoem and Funck Jensen 1982). However, these methods involve solving Kolmogorov differential equations or using exponential series expansions of transition intensity matrices, and such methods are computationally more demanding. Since the transition probability matrices have to be updated at monthly intervals for every age, sex, risk group and sexual activity state represented in Figure 1, the additional computational burden would be significant.

Appendix B: Likelihood definition for HIV prevalence data

This appendix describes the method used to define the likelihood in respect of the HIV prevalence data in the antenatal clinic surveys and the likelihood in respect of the HIV prevalence data collected in the 2005 HSRC household survey. The method used to define the likelihood function in respect of the sexual behaviour data is described in section 2.4 of the paper.

Suppose that $H_{x,t}(\boldsymbol{\varphi})$ represents the prevalence of HIV that we would expect to measure in pregnant women aged x to $x + 4$, in year t, based on the model predictions when the input parameters are represented by vector $\boldsymbol{\varphi}$.[1] Further suppose that the corresponding prevalence of HIV measured in the survey is $\theta_{x,t}$. It is assumed that if $\boldsymbol{\varphi}$ is the true set of parameter values, then the difference between the logit-transformed model prediction and the logit-transformed observed prevalence is normally distributed. The mean of this normal distribution represents the extent of antenatal bias. The variance of the distribution is assumed to be composed of a 'survey error' term, representing the uncertainty around the survey estimate due to binomial variation and cluster variation in the survey, and a 'model error' term. More formally, it is assumed that

$$\log\left(\frac{\theta_{x,t}}{1-\theta_{x,t}}\right) = \log\left(\frac{H_{x,t}(\boldsymbol{\varphi})}{1-H_{x,t}(\boldsymbol{\varphi})}\right) + a + \varpi_{x,t} + \eta_{x,t}\,, \qquad \text{(B1)}$$

where a is the antenatal bias parameter, $\varpi_{x,t} \sim N\left(0, \sigma_{\varpi}^2\right)$ and $\eta_{x,t} \sim N\left(0, \sigma_{x,t}^2\right)$. The latter two terms represent the model error and the survey error respectively. The logit transformations ensure that the error terms are closer to normality and the model error terms are roughly independent of the level of HIV prevalence. For a given parameter combination $\boldsymbol{\varphi}$, the antenatal bias parameter is estimated using the formula

$$\hat{a} = \frac{1}{N_A}\sum_x\sum_t\left(\log\left(\frac{\theta_{x,t}}{1-\theta_{x,t}}\right) - \log\left(\frac{H_{x,t}(\boldsymbol{\varphi})}{1-H_{x,t}(\boldsymbol{\varphi})}\right)\right), \qquad \text{(B2)}$$

where N_A is the number of antenatal prevalence estimates to which the model is calibrated. The $\sigma_{x,t}^2$ values are estimated from the 95% confidence intervals that have been published for the various survey estimates. Once these have been obtained, the σ_{ϖ}^2 parameter is estimated using the formula

$$\hat{\sigma}_{\varpi}^2 = \frac{1}{N_A}\sum_x\sum_t\left(\log\left(\frac{\theta_{x,t}}{1-\theta_{x,t}}\right) - \log\left(\frac{H_{x,t}(\boldsymbol{\varphi})}{1-H_{x,t}(\boldsymbol{\varphi})}\right) - \hat{a}\right)^2 - \sigma_{x,t}^2\,. \qquad \text{(B3)}$$

[1] The model estimate of HIV prevalence in pregnant women aged x to $x + 4$ is calculated as the HIV prevalence in sexually experienced women after weighting the numbers in the different HIV disease states by the relative levels of fertility in the different states.

The likelihood is then calculated based on the assumption that the error terms are normally distributed:

$$\prod_{x} \prod_{t=1997}^{2005} \left(2\pi\left(\hat{\sigma}_{\varpi}^2 + \sigma_{x,t}^2\right)\right)^{-0.5} \exp\left(-\frac{\left(\text{logit}\left(\theta_{x,t}\right) - \text{logit}\left(H_{x,t}(\varphi)\right) - \hat{a}\right)^2}{2\left(\hat{\sigma}_{\varpi}^2 + \sigma_{x,t}^2\right)}\right). \quad \text{(B4)}$$

The same approach to defining the likelihood function is used with the HSRC HIV prevalence data, except that the bias term (a) and model error term (ϖ) are both omitted. The model error term is omitted because the 95% confidence intervals around the HIV prevalence estimates are very wide, and introducing a model error term therefore reduces the weight given to the HSRC data to unreasonably low levels. The omission of the bias term is consistent with the approach adopted in other uncertainty analyses of HIV data in developing countries, in which it is assumed that household prevalence data provide an unbiased estimate of HIV prevalence in the general population.

References

Hoem J. M. and Funck Jensen U. (1982). Multistate life table methodology: a probabilist critique. In K. C. Land and A. Rogers, Eds. *Multidimensional Mathematical Demography*. New York: Academic Press**:** 155-264.

Neill A. (1977). Life Contingencies. Oxford: Heinemann Professional Publishing