

# The effect of changing sexual activity on HIV prevalence

Michael Kremer<sup>\*</sup>, Charles Morcom<sup>1</sup>

*Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 01239, USA*

Received 5 June 1996; received in revised form 19 March 1998

---

## Abstract

In a one-sex preferred mixing model, reductions in the rate of partner change by those with low sexual activity increase the average probability of HIV infection in the remaining pool of available partners. This increases prevalence among people with high activity, and since high activity people disproportionately influence the spread of HIV, may increase long-run prevalence in the population as a whole. Calculations using the model and survey data on sexual activity indicate that in low prevalence populations, many people have low enough activity that reductions in their activity might increase the endemic steady-state prevalence. If these results prove robust in more realistic models, they would support the case for targeting public health messages urging reduced sexual activity to high activity people. © 1998 Elsevier Science Inc. All rights reserved.

---

## 1. Introduction

This paper shows that, in a simple single-sex preferred mixing model, increases in the frequency of partner change by low-activity people may reduce long-run HIV prevalence. To see the intuition, suppose that 9 partners per year were required for HIV to be endemic in a homogeneous population. Consider a population in which a small minority had 10 partners a year, and the majority had no partners at all. In this population, the disease would persist among the active minority. Suppose that the inactive majority decided to have 1 partner

---

<sup>\*</sup>Corresponding author. Tel.: +1-617 253 3504; fax: +1-617 253 6915; e-mail: kremer@mit.edu.

<sup>1</sup> Tel.: +1-718 596 3685; e-mail: cmorcom@alum.mit.edu.

each over their lifetimes, that the groups mixed randomly, and that the proportions of the two groups were such that on average the high-activity people might have 5 partnerships a year with the low-activity people and 5 a year with other high-activity people. In this case, the disease would die out, because half the new infections would occur among low-activity people who would not infect others.

To see the intuition slightly more mathematically, recall that, under a simple random mixing model, the disease dies out or is endemic according to whether the threshold function,  $R_0$ , is greater or less than 1, where

$$R_0 = \frac{\beta}{\delta} \left( \mu + \frac{\sigma^2}{\mu} \right);$$

$\mu$  and  $\sigma$  are the mean and standard deviation of the rate of partner change, and  $\beta$  and  $\delta$  are the birth and death rates [1].

This can be rewritten as

$$R_0 = \frac{\beta \sum \alpha_k i_k^2}{\delta \sum \alpha_k i_k},$$

where there are  $N$  groups in the population classified by the number of partners per year,  $i_k$ , and each group represents a portion  $\alpha_k$  of the population. Differentiating with respect to  $i_k$  shows that small *increases* in activity by individuals with activity less than  $\frac{1}{2}(\mu + \sigma^2/\mu)$  will *reduce*  $R_0$ . Consider a population on a knife-edge between endemic HIV and the disease dying out, so that  $R_0 = 1$ . If members of the population with activity less than  $\frac{1}{2}(\mu + \sigma^2/\mu)$  increase activity,  $R_0$  will *decrease*, and the disease will die out. In fact, a uniform small increase in activity by the entire population will reduce  $R_0$  if  $\sigma > \mu$  [2].

It turns out that these effects on the *stability* of the endemic steady-state are far too sensitive to the assumption of homogeneous mixing among sexual partners to be of empirical importance. As we show in this paper, however, similar effects may apply to the level of steady-state prevalence, even if the disease remains endemic: namely, an increase in activity by low-activity members of the population may decrease prevalence in the long run. These effects on prevalence are robust enough with respect to preferred mixing among sexual partners that they may be of some practical significance.

In particular, crude calibration of a simple single-sex model using survey data on sexual activity suggests that these counter-intuitive effects may be more than just a theoretical curiosity. We develop expressions for a cut-off level of sexual activity such that increases in activity in the groups below this level will reduce steady-state prevalence. More than 80% of the population in a comprehensive study of sexual activity in the UK had low enough activity that reductions in activity would increase steady-state prevalence in a standard single-sex susceptible-infected (SI) epidemiological model with preferred mixing. Simulations using this simple model suggest that if everybody who had 1 partner every

5 years reduced their frequency of partner change by 5%, steady-state prevalence would increase by 7% under random mixing. However, the simulations also indicate that for reasonable parameter values, increases in activity would not lead to the eradication of the disease.

There are several important caveats. First, reductions in activity by low-activity people could only increase steady-state prevalence in populations that have low steady-state prevalence given current activity levels and transmission probabilities. The counterintuitive effects discussed in this paper thus may be relevant for heterosexuals in developed countries, but not for homosexuals, heterosexuals in the highest-prevalence areas of Africa, or IV drug users. Second, increases in activity by low-activity people will *increase* prevalence temporarily. Third, we consider the effects of changes in activity by one group while keeping the activity levels of all other groups constant. The conclusions in this paper may be weakened if reductions in activity by low-activity people make it harder for high-activity people to find partners, and they consequently reduce their own activity. More generally, since the model abstracts from important features of the epidemic (for example, our model does not allow concurrent partnerships or different sexes), the results should be considered provisional.

To the extent that the results of this paper prove robust in more realistic models, though, they reinforce arguments for targeting public health messages urging reductions in the frequency of partner change to high-activity people. Because anyone increasing activity will, at least in the short term, increase his or her risk, we would absolutely never recommend public health messages designed to increase activity by lower activity groups: people have a reasonable right to expect that public health policy will not act directly to increase their individual risk, even for the sake of a long-term reduction in prevalence in the population as a whole.

A large literature examines the dynamics of sexually transmitted diseases under a variety of mixing patterns, and considers the effect of changes in activity [1,3–10]. Both the present study and that of Kremer [2,11] follow independent work by Whitaker and Rentin [12]. They show that in a two group example with random mixing, an increase in activity by the low-activity group may reduce steady state prevalence. They, however, explicitly disclaim empirical relevance of its model for AIDS. Our analysis differs from Whitaker and Rentin [12] in that it extends the results to a preferred mixing model with an arbitrary distribution of activity, and uses empirical data to show that these effects may be important in low-prevalence populations, but not in high-prevalence populations.

Kremer [11] analyzes the externalities from sexual activity using the economic concept of asymmetric information. This paper presents the results in purely epidemiological terms. It also further develops the theory and mathematical methods of Ref. [11] using a different approach to determine the cut-off levels of activity, and extending the theory to models with preferred mixing. Kremer

[2] considers how the argument in this paper is modified if high-activity people change their activity in response to changes in activity by low-activity people.

The paper is organized as follows: In Section 2, we define the model we use, and cite some results concerning its stability and the endemic steady state. In Section 3, we solve for the cut-off levels of activity below which increasing activity reduces steady-state prevalence among other members of the population, or prevalence in the population as a whole. In Section 4, we examine the special case of random mixing. In Section 5, we calculate the cut-off values of activity under a preferred mixing model using data on the frequency of partner change among heterosexuals in the UK. Section 6 discusses the time-path of prevalence in response to reductions in the rate of partner change. In Section 7, we discuss directions for future research and possible policy implications.

## 2. The SI model with preferred mixing

We use a simplified version of the preferred mixing SI model as presented by Jacquez et al. [13]. In this model, people are born and die according to a Poisson process with parameter  $\delta$ , independent of whether or not they are infected. There are  $N$  groups of people, classified by the number of sexual partners they have per year,  $i_k$ , where  $k \in \{0, \dots, N-1\}$ , and each group represents a proportion  $\alpha_k$  of the total population. The mean sexual activity per year is  $\mu = \sum_{k=0}^{N-1} \alpha_k i_k$ , and the variance of sexual activity is  $\sigma^2 = \sum_{k=0}^{N-1} \alpha_k (i_k - \mu)^2$ . The prevalence of HIV in each group is  $y_k$ . The overall, or average, prevalence,  $\bar{y}$ , and the activity weighted prevalence, or pool-risk,  $\lambda$ , are defined by

$$\bar{y} = \sum_{k=0}^{N-1} \alpha_k y_k, \quad \lambda = \sum_k w_k y_k, \quad (1)$$

where  $w_k = \alpha_k i_k / \mu$ .

The pool-risk,  $\lambda$ , is the probability that a random partnership will be with someone who is infected if partners are picked randomly from the pool, with the chance of picking someone proportional to their number of partnerships per year.

Under the preferred mixing model, people select partners randomly with probability  $1 - \gamma$ , and select partners within their own group with probability  $\gamma$ . When  $\gamma = 0$ , the model reduces to random mixing. If  $\gamma = 1$ , there is restricted mixing (people select partners only from their own activity group).

Those born are uninfected, and enter activity groups pro rata to existing group size. This ensures that the relative sizes of the groups,  $\{\alpha_k\}$ , are constant, which simplifies the analysis. If an infected person mixes with an uninfected person, the probability of transmission is  $\beta$ .

Consider the dynamics of the number infected in group  $k$  during a short time interval  $\Delta t$ . Infected people die at rate  $\delta y_k$ . The uninfected proportion

$1 - y_k$  have  $i_k \Delta t$  partners. They pick from within group  $k$  with probability  $\gamma$ . These partners have probability  $y_k$  of being infected, and have probability  $\beta$  per contact of infecting the uninfected partner. Uninfected individuals thus become infected at rate  $\gamma\beta i_k y_k (1 - y_k)$  from mixing with their own activity group. The infection rate due to mixing with the general population at random is, analogously,  $(1 - \gamma)\beta i_k \lambda (1 - y_k)$ . The dynamics of HIV infection in group  $k$  are, thus, described by

$$\begin{aligned}\dot{y}_k &= -\delta y_k + \gamma\beta i_k y_k (1 - y_k) + (1 - \gamma)\beta i_k \lambda (1 - y_k) \\ &= \delta[-y_k + \theta i_k (1 - y_k)(\gamma y_k + (1 - \gamma)\lambda)],\end{aligned}\quad (2)$$

where  $\theta = \beta/\delta$  is the expected number of infections that would be caused during the life of single infected person who has 1 uninfected partner per year.

The dynamics and stability of such models are relatively well understood. There is either a stable endemic steady state, or the disease dies out, depending on the system parameters. Jacquez et al. [13] shows in Appendix A that a threshold function  $G$  can be defined as

$$G(\gamma, \theta, \{i_k\}, \{\sigma_k\}) = \theta(1 - \gamma)\left(\mu + \frac{\sigma^2}{\mu}\right) + \theta\gamma \max_k \{i_k\}, \quad (3)$$

where  $\mu$  and  $\sigma^2$  are the mean and variance of the activity levels  $\{i_k\}$ . If  $G < 1$ , the disease dies out, and zero prevalence for all activity groups is the unique globally asymptotically stable steady state. If  $G > 1$ , then there is a unique locally asymptotically stable steady state in which the activity group prevalences are  $Y_k$ .  $Y_k$  are easily found by solving Eq. (2) with  $\dot{Y}_k = 0$ , since we are in steady state

$$Y_k = \theta i_k (1 - Y_k)(\gamma Y_k + (1 - \gamma)A), \quad (4)$$

where  $A$  is the pool-risk at the steady state. This defines  $Y_k$  implicitly as a function of  $A$ ,  $i_k$ ,  $\gamma$ , and  $\theta$ . Solving, we find that

$$Y_k = \frac{\left\{ [1 + \theta i_k (A(1 - \gamma) - \gamma)]^2 + 4i_k^2 \theta^2 \gamma (1 - \gamma) A \right\}^{1/2} - \theta i_k (A(1 - \gamma) - \gamma) - 1}{2\gamma \theta i_k}$$

if  $0 < \gamma < 1$ . (5)

and

$$Y_k = \frac{\theta i_k A}{1 + \theta i_k A}, \quad \text{if } \gamma = 0. \quad (6)$$

We use capital letters to refer to all quantities at the endemic steady state. Thus  $\mathbf{Y} = (Y_0, \dots, Y_{N-1})$  are the group prevalences,  $\bar{Y}$  is the average prevalence, and  $A$  is the pool-risk in steady state. Note that prevalence in any group in endemic steady state is only directly a function of the group's own activity, and the pool-risk,  $A$ , given  $\gamma$  and  $\theta$ .

Note that, although we will examine how changes in the activity levels  $\{i_k\}$  affect the *endemic* steady state in the rest of the paper, it is theoretically possible that changes in activity levels could wipe out the disease entirely in steady state. We consider this possibility numerically in more detail in Section 5, but it turns out that, for anything other than cases with almost perfectly random mixing ( $\gamma$  very close to 0), increases in activity cannot wipe out the disease. Since real populations are likely to mix far from randomly, we do not consider this possibility in any more detail.

### 3. The effects of changing activity on prevalence

In this section, we consider the effects on steady-state pool-risk,  $\lambda$ , and prevalence,  $\bar{Y}$ , of changing the activity level of one of the groups. To this end, we shall always assume that  $0 \leq \gamma < 1$ , so there is at least some cross-group mixing.<sup>2</sup> For now, we restrict attention to the case in which  $G > 1$ , so that the endemic steady state exists and is stable.

If people from groups with group prevalence less than  $\lambda$  increase their activity then, in the short run,  $\lambda$  will fall as the chance of meeting someone from a low activity group, who is less likely to be infected, increases. Under certain conditions, this effect persists in the steady state, as well, and  $\lambda$  and  $\bar{Y}$  may fall. In this section, we examine such effects in the model of Section 2. There will be two quantities of interest.

First, we define  $j_e$  to be the number of partners below which an *increase* in activity causes a *reduction* in steady-state pool-risk,  $\lambda$ . This definition is equivalent to saying that, for  $i_0 < j_e$ ,  $d\lambda/di_0 < 0$ . In the language of economics, people with activity below  $j_e$  create a positive *externality* by increasing their activity, because doing so reduces the chance that *other* people will be infected. Economists normally assume that people can weigh the costs and benefits of their actions for themselves, but that they will not adequately consider the effects of their actions on others. Thus under an extreme *laissez faire* view of the world, there would be a case for encouraging reductions in activity only for those with activity higher than  $j_e$  (we do not endorse this view). As we show below,  $j_e$  is always positive, so that groups with low enough activity could always reduce steady-state pool-risk by increasing activity a bit.

Second, we define  $j_l$  to be the cut-off level of activity below which an *increase* in activity by a small group leads to a *reduction* in the long-term prevalence,  $\bar{Y}$ . This definition is equivalent to saying that, for  $i_0 < j_l$ ,  $d\bar{Y}/di_0 < 0$ . We derive necessary and sufficient conditions for  $j_l$  to be positive, and we show that it

<sup>2</sup> If  $\gamma = 1$ , a person who increases activity will increase his/her own chance of infection, and will not affect anyone else's chance of infection outside his or her own group.

is always the case that  $j_l \leq j_e$ , since if increasing activity reduces  $\bar{Y}$ , it must also reduce  $A$ .

First, we prove a technical Lemma, giving expressions for some derivatives of expressions which will prove useful for the results of the rest of the section. Partial derivatives in this and later sections are partial derivatives of the expression (4) with respect to the variables  $A$  and  $i_k$ . Total derivatives take account of the fact that changing activity will also change  $A$ .

**Lemma 1.** *Considering  $Y_k$  as a function of  $i_k$  and  $A$ , the following are true for  $A, Y_k > 0$  and for all  $k$ :*

$$\frac{\partial Y_k}{\partial A} = \frac{(1-\gamma)(1-Y_k)Y_k}{(1-\gamma)A + \gamma Y_k^2} < \frac{Y_k}{A}, \quad (7)$$

$$0 < \sum_k w_k \left( 1 - \frac{\partial Y_k}{\partial A} \right), \quad (8)$$

$$i_k \frac{\partial Y_k}{\partial i_k} = \frac{(1-Y_k)((1-\gamma)A + \gamma Y_k)Y_k}{(1-\gamma)A + \gamma Y_k^2}, \quad \text{and} \quad (9)$$

$$\frac{\partial Y_k}{\partial i_k} = \theta \frac{(1-Y_k)^2((1-\gamma)A + \gamma Y_k)^2}{(1-\gamma)A + \gamma Y_k^2}. \quad (10)$$

**Proof.** From Eq. (4), at the steady state,  $Y_k = \theta i_k(1-Y_k)(\gamma Y_k + (1-\gamma)A)$ . For Eq. (7), differentiate this identity with respect to  $A$ , treating  $Y_k$  as an implicit function of  $A$  and  $i_k$ :

$$\frac{\partial Y_k}{\partial A} = -\frac{Y_k}{1-Y_k} \frac{\partial Y_k}{\partial A} + \left[ \gamma \frac{\partial Y_k}{\partial A} + (1-\gamma) \right] \frac{Y_k}{\gamma Y_k + (1-\gamma)A}.$$

Rearranging yields the equality in Eq. (7). For the inequality, note that  $1-Y_k < 1$ , and  $(1-\gamma)A + \gamma Y_k^2 > (1-\gamma)A$ , and the result follows. For Eq. (8), The inequality from Eq. (7) gives  $\sum w_k(\partial Y_k/\partial A) < (1/A) \sum w_k Y_k = 1$ , by definition of  $A$ . For Eq. (9), differentiate Eq. (4) with respect to  $i_k$ , multiply by  $i_k$ , and rearrange. For Eq. (10), note that  $i_k = Y_k/[\theta(1-Y_k)(\gamma Y_k + (1-\gamma)A)]$ , and substitute for  $i_k$  in Eq. (9).  $\square$

Without loss of generality, we shall consider the effects of changing the activity of the zero group. The cut-off level  $j_e$  below which increases in activity reduce steady-state pool-risk,  $A$ , must be such that  $dA/di_0 = 0$  when the zero group has activity  $j_e$ .

## Proposition 2.

$$\frac{dA}{di_0} = \frac{\alpha_0 \left\{ Y_0 - A + i_0 \frac{\partial Y_0}{\partial i_0} \right\}}{\mu \sum_k w_k \left( 1 - \frac{\partial Y_k}{\partial A} \right)} \quad (11)$$

$$= -\frac{\alpha_0 (Y_0^2(A - \gamma) - 2(1 - \gamma)AY_0 + (1 - \gamma)A^2)}{\mu ((1 - \gamma)A + \gamma Y_0^2) \sum_k w_k (1 - \frac{\partial Y_k}{\partial A})}. \quad (12)$$

**Proof.** The definition of pool-risk (1) applies just as well to  $A$ , steady-state pool-risk:

$$A = \sum_{k=0}^{N-1} w_k Y_k(\gamma, \theta, i_k, A).$$

Differentiating this with respect to  $i_0$ ,

$$\frac{dA}{di_0} \sum_{k=0}^{N-1} \left[ \frac{\partial w_k}{\partial i_0} Y_k + w_k \left( \frac{\partial Y_k}{\partial i_0} + \frac{\partial A}{\partial i_0} \frac{\partial Y_k}{\partial A} \right) \right].$$

From the definitions of  $w_k$  and  $\mu$  (Eq. (1)),

$$\frac{\partial w_k}{\partial i_0} = \frac{\partial}{\partial i_0} \left( \frac{i_k \alpha_k}{\mu} \right) = \frac{\alpha_k}{\mu} \delta_{0,k} - \frac{\alpha_0}{\mu} w_k,$$

where  $\delta_{0,k}$  is 1 when  $k=0$ , and zero otherwise. Then

$$\frac{dA}{di_0} = \frac{\alpha_0}{\mu} \left( Y_0 + i_0 \frac{\partial Y_0}{\partial i_0} - A \right) + \frac{dA}{di_0} \left( \sum_{k=0}^{N-1} w_k \frac{\partial Y_k}{\partial A} \right).$$

Rearranging yields Eq. (11). Substituting for  $i_0 \partial Y_0 / \partial i_0$  from Lemma 1, Eq. (9) gives Eq. (12).  $\square$

The intuition behind Eq. (11) is as follows: the first two terms in the parentheses in the numerator are the net short-term effect of the change in  $i_0$  on the pool-risk,  $A$ : if the zero group has lower prevalence than  $A$ , increasing activity will cause  $A$  to fall (in the short run,  $Y_0$  will not change). The third term is the long-run effect on  $A$  of the increase in the zero group's prevalence from its own increased activity. The denominator shows how the effect is magnified by the change in prevalence caused in the rest of the population by the change in zero group behavior.  $j_e$  is the activity level for which  $dA/di_0 = 0$ , when  $i_0 = j_e$ .

Consider Eq. (12). The denominator is always positive, by Lemma 1, so the sign is determined by the numerator, which is a quadratic in  $Y_0$ . We may prove the following:

**Proposition 3.**  $dA/di_0$  is strictly negative for  $Y_0 \in [0, Y_e)$ , zero at  $Y_0 = Y_e$ , and strictly positive for  $Y_0 \in (Y_e, 1]$ , where  $0 < Y_e < A$ , and

$$Y_e = \frac{A(1 - \gamma)}{A - \gamma} \left\{ 1 - \sqrt{\frac{1 - A}{1 - \gamma}} \right\}. \quad (13)$$

**Proof.** Consider the numerator of Eq. (12):

$$-Y_0^2(A - \gamma) + 2(1 - \gamma)AY_0 - (1 - \gamma)A^2.$$

Its value at  $Y_0 = 0$  is  $-(1 - \gamma)A^2$ , which is strictly negative. Its value at  $Y_0 = A$  is  $(1 - A)A^2$ , which is strictly positive. There are, therefore, an odd number of roots of odd multiplicity between 0 and  $A$ . Since the numerator of Eq. (12) has order at most two, there must be exactly one root in  $[0, A]$ . Solving for this root yields Eq. (13). Moreover, at  $Y_0 = 1$ , the value of the numerator is positive:  $(1 - A)[A(1 - \gamma) + \gamma]$ , which implies that there is no other root on the segment  $(Y_c, 1]$  and  $dA/di_0 > 0$  by continuity.  $\square$

Now, in steady state, Eq. (4) implies that

$$i_k = \frac{Y_k}{\theta(1 - Y_k)((1 - \gamma)A + \gamma Y_k)},$$

which is increasing in  $Y_k$  for fixed  $A$ . This means that the root of  $dA/di_0$ ,  $Y_c$ , corresponds to an activity level,  $j_c$ , so that

$$j_c = \frac{Y_c}{\theta(1 - Y_c)((1 - \gamma)A + \gamma Y_c)}.$$

Substituting the value of  $Y_c$  from Proposition 3 and re-expressing  $\theta$  as  $\beta/\delta$  gives the following.

**Theorem 4.** *If a group has an activity of less than  $j_c$  partners per year, then a small increase in activity by that group will reduce steady-state pool-risk,  $A$ , where*

$$j_c = \frac{\delta}{\beta(\gamma + (1 - \gamma)A)^2} \left( \gamma - (1 - \gamma)A + \sqrt{\frac{1 - \gamma}{1 - A}}(A - (1 - A)\gamma) \right). \quad (14)$$

and

$$0 < j_c < \frac{\delta}{\beta(1 - A)}. \quad (15)$$

**Proof.** As discussed above. The bounds on  $j_c$  come from  $0 < Y_c < A$ .  $\square$

We now turn our attention to how average prevalence,  $\bar{Y}$ , in the population as a whole is affected by changes in the activity of one group. It will be helpful, first, to define

$$H = \frac{1}{\mu\theta} \frac{\sum z_k \frac{\partial Y_k}{\partial A}}{(1 - \sum w_k \frac{\partial Y_k}{\partial A})}. \quad (16)$$

Note that, by inequality Eq. (8) in Lemma 1,  $H > 0$ .

**Proposition 5.**

$$\begin{aligned} \frac{d\bar{Y}}{di_0} &= \alpha_0 \left( \frac{\partial Y_0}{\partial i_0} + \frac{1}{\mu} \left( \frac{\sum_k \alpha_k \frac{\partial Y_k}{\partial A} \left( Y_0 + i_0 \frac{\partial Y_0}{\partial i_0} - A \right)}{\sum_k w_k \left( 1 - \frac{\partial Y_k}{\partial A} \right)} \right) \right) \\ &= \theta \alpha_0 \frac{(1 - Y_0)^2 [(1 - \gamma)A + \gamma Y_0]^2 - H [(A - \gamma)Y_0^2 - 2(1 - \gamma)AY_0 + (1 - \gamma)A^2]}{(1 - \gamma)A + \gamma Y_0^2}. \end{aligned} \quad (17)$$

**Proof.**

$$\frac{d}{di_0} \bar{Y} = \sum_k \alpha_k \frac{dY_k}{di_0} = \sum_k \alpha_k \left( \frac{\partial Y_k}{\partial A} \frac{dA}{di_0} + \frac{\partial Y_k}{\partial i_0} \right),$$

and substitute for  $(dA/di_0)$  from Eq. (11) and Eq. (12). This gives the first equation of the proposition.

The second comes from substituting for the partial derivatives of  $Y_0$  from Lemma 1 and rearranging.  $\square$

In this case, we cannot find a root exactly as we could for  $Y_e$ . We can, however, prove the following:

**Theorem 6.** *If, and only if,  $H > 1 - \gamma$ , then there exists  $Y_t \in (0, Y_e)$  such that a group with group prevalence below  $Y_t$  will decrease steady-state average prevalence,  $\bar{Y}$ , by increasing group activity.*

**Proof.** Consider the numerator of Eq. (17):

$$(1 - Y_0)^2 (A(1 - \gamma) + \gamma Y_0)^2 - H (Y_0^2 (A - \gamma) - 2(1 - \gamma)AY_0 + (1 - \gamma)A^2). \quad (18)$$

$Y_t$  exists if and only if this expression is negative for  $Y_0 < Y_t$ , and zero at  $Y_t$ . It is continuous in  $Y_0$ . Its value at zero is  $A^2(1 - \gamma)^2 - HA^2(1 - \gamma)$ . This is strictly negative if and only if  $H > 1 - \gamma$ .

Now, consider its value at  $Y_e$ .  $Y_e$  is a root of the coefficient of  $H$ , so the value of the whole expression at  $Y_e$  is  $(1 - Y_e)^2 (A(1 - \gamma) + \gamma Y_e)^2$ , which is always strictly positive.

If  $H > 1 - \gamma$ , by the intermediate value property, Eq. (18) must have at least one root of odd order in  $(0, Y_e)$ . Define  $Y_t$  to be the smallest of these. Then, for  $0 \leq Y_0 < Y_e$ ,  $d\bar{Y}/di_0$  is negative, and the theorem follows.  $\square$

Note that although  $H$  is a function of all  $Y_k$ s, it is a characteristic function of the system, so that we are allowed to examine the roots of Eq. (18) for a given fixed  $H$ .

The activity level  $j_t$  corresponding to  $Y_t$  is the threshold below which increases in activity reduce steady-state prevalence. It seems clear that once  $Y_0$  is above  $Y_t$ , increases in activity would lead to increases in overall steady-state prevalence, but we have not been able to prove that the expression (18) has no more roots between  $Y_t$  and  $Y_e$ , so it is conceivable that there could be an interval above  $Y_0 = Y_t$  where  $d\bar{Y}/di_0$  is negative. We have not been able to find any cases where this happens, and we would not expect to. The stronger statement, that the sign of  $d\bar{Y}/di_0$  is *increasing* in  $Y_0$  on  $[0, A]$  is true in the case of random mixing (see Section 4). This is sufficient to prove that  $d\bar{Y}/di_0$  has a maximum of one root in  $[0, A]$ . That this is true for  $\gamma = 0$  implies, by continuity, that it is true for  $\gamma$  small but positive.

Quartic equations are soluble in closed form. We could, therefore, obtain an exact closed-form solution for  $Y_t$ , where it exists, in terms of  $H$ ,  $A$ , and  $\gamma$ . Since  $j_t$  is related to  $Y_t$  by a quadratic in  $Y_e$ , we could then find  $j_t$  explicitly. In general, though,  $d\bar{Y}/di_0$  has no rational roots, and so any such expression would be too complicated to be useful. In the case where  $\gamma = 0$ , we are able to factorize expression (18) relatively easily.

#### 4. The special case of random mixing

We now specialize to the case of random mixing, in which  $\gamma = 0$ . In this case, we obtain stronger results about the existence and uniqueness of  $j_t$ , the activity level below which long-run average prevalence is reduced by increases in activity. We also obtain a simple closed form expression for  $j_t$  in terms of the activity weighted prevalence,  $A$ , and the system parameters.

It is well known that if  $\gamma = 0$ , the endemic steady state will exist and be locally asymptotically stable if and only if  $\mu + \sigma^2/\mu > 1$  [1]. Substituting  $\gamma = 0$  into the expressions in Section 3 yields the expression in Kremer [11]:

$$Y_e = 1 - \sqrt{1 - A}, \quad \text{and} \quad (19)$$

$$j_e = \frac{\delta}{\beta A} \left( \frac{1}{\sqrt{1 - A}} - 1 \right). \quad (20)$$

The expression for  $d\bar{Y}/di_0$  factorizes when  $\gamma = 0$ , so that we can solve for  $Y_t$  and hence  $j_t$  in closed form. We can also prove that  $Y_t$  is the unique point at which  $d\bar{Y}/di_0 = 0$ , if any such point exists.

**Proposition 7.** *If there is random mixing, so that  $\gamma = 0$ , and if  $0 < A < 1$ , then*

$$\frac{d\bar{Y}}{di_0} = -\alpha_0 \theta \left\{ (Y_0 - 1)^2 (H - A) - H(1 - A) \right\}. \quad (21)$$

*If  $H > 1$ , then  $\exists Y_t$  such that  $0 < Y_t < Y_e < A < 1$ , and such that  $d\bar{Y}/di_0$  is strictly negative for  $Y_0 \in [0, Y_t)$ , zero at  $Y_0 = Y_t$ , and strictly positive on  $Y_0 \in (Y_t, 1]$ , where*

$$Y_1 = 1 - \sqrt{\frac{H(1-A)}{H-A}}. \quad (22)$$

If  $H < 1$ , then  $d\bar{Y}/di_0$  is strictly positive on  $[0,1]$ .

**Proof.** Substitute  $\gamma=0$  in the expression of Proposition 5 and simplify for Eq. (21).

$(Y_0 - 1)^2$  is decreasing in  $Y_0$  for  $0 < Y_0 < 1$  and  $H > A$ , so that Eq. (21) is increasing in  $Y_0$ . At  $Y_0 = 0$ , it has value  $-\alpha_0\theta A(H-1)$ , which is negative if and only if  $H > 1$ .

When  $\gamma = 0$ ,  $Y_e = 1 - \sqrt{1-A}$  so that, at  $Y_0 = Y_e$ ,  $d\bar{Y}/di_0$  has value  $\theta\alpha_0 A(1-A)$ , which is strictly positive, since  $0 < A < 1$ .

$d\bar{Y}/di_0$  has, therefore, exactly one root between 0 and  $Y_e$ . This is  $Y_1$ . Solving the quadratic yields Eq. (22). Note that, if  $H > 1$ , it must also be true that  $H > A$ .

If  $H < 1$ ,  $d\bar{Y}/di_0$  is still monotone on  $Y_0 \in [0,1]$ . As above, the value at  $Y_0=0$  is positive. At  $Y_0=1$ , the value is  $\alpha_0\theta H(1-A)$ , which is also positive.  $d\bar{Y}/di_0$  must be strictly positive on  $[0,1]$ .  $\square$

If  $H > 1$ , then we may use Eq. (4) to find the value of activity,  $j_1$ , which corresponds to  $Y_1$ . This gives a level of activity,  $j_1$ , such that *increases* in activity by individuals with activity less than  $j_1$  reduce steady-state endemic prevalence,  $\bar{Y}$ , where:

**Theorem 8.** *If and only if  $H > 1$ , there exists a positive activity level,  $j_1 \in (0, j_c)$ , below which increases in activity lead to a fall in steady-state average prevalence,  $\bar{Y}$ , where*

$$j_1 = \frac{\delta}{\beta A} \left( \sqrt{\frac{H-A}{H(1-A)}} - 1 \right). \quad (23)$$

**Proof.** From Proposition 7,  $i_0 = Y_0/[0(1-Y_0)A]$ . Substitute  $Y_1$  for  $Y_0$ , and the result follows.

In the case where  $\gamma = 0$ , we may find an expression for  $H$  in terms of the activity levels, the group prevalences, and their moments. First:

**Lemma 9.** *If we define  $\sigma_Y^2 = \sum \alpha_k Y_k^2 - \bar{Y}^2$  to be the variance in prevalence among the groups, then the endemic steady-state group prevalences satisfy*

$$\bar{Y} = \mu\theta A(1-A), \quad (24)$$

$$\mu\theta A^2 = \sigma_Y^2 + \bar{Y}^2 + A\theta\mu \sum_k w_k Y_k^2. \quad (25)$$

**Proof.** Since  $Y_k = A\theta i_k / (1 + A\theta i_k)$ ,  $Y_k + \theta i_k Y_k A = \theta i_k A$ . Multiplying by  $\alpha_k$  and summing gives Eq. (24). Multiplying by  $\alpha_k Y_k$  and summing gives Eq. (25).  $\square$

**Lemma 10.**

$$H = \frac{A(\bar{Y} - \bar{Y}^2 - \sigma_Y^2)}{\mu\theta A^2 - \sigma_Y^2 - \bar{Y}^2}. \quad (26)$$

$H > 1$ , and so  $j_1 > 0$ , if and only if

$$\sigma_Y^2 > \bar{Y} \left[ \frac{A^2}{(1-A)^2} - \bar{Y} \right]. \quad (27)$$

**Proof.** When  $\gamma = 0$ ,  $\partial Y_k / \partial A = (1 - Y_k)Y_k / A$ , from Lemma 1. Then  $\sum \alpha_k \partial Y_k / \partial A = (1/A)(\bar{Y} - \bar{Y}^2 - \sigma_Y^2)$ ,  $\sum w_k (1 - \partial Y_k / \partial A) = \sum w_k Y_k^2 / A$ . Substitute these into the definition for  $H$ , and

$$H = \frac{A(\bar{Y} - \bar{Y}^2 - \sigma_Y^2)}{A\mu\theta \sum w_k Y_k^2}.$$

Using the second equation of Lemma 9, the expression for  $H$  results.

The bound on  $\sigma_Y^2$  comes from using the first equation of Lemma 9 to replace  $\mu\theta A$  with  $\bar{Y}/(1-A)$ .  $\square$

Substituting this expression for  $H$  in the formula for  $j_1$  of Eq. (23), we get the following theorem.

**Theorem 11.** *If and only if*

$$\sigma_Y^2 > \bar{Y} \left[ \frac{A^2}{(1-A)^2} - \bar{Y} \right], \quad (28)$$

then  $j_1$  is positive, and is given by

$$j_1 = \frac{\mu(1-A)}{\bar{Y}} \left\{ \sqrt{\frac{\bar{Y}(1-2A)}{(1-A)^2(\bar{Y}(1-\bar{Y}) - \sigma_Y^2)}} - 1 \right\}. \quad (29)$$

Note that because this theorem holds for  $\gamma = 0$ ,  $j_1$  must also exist and be positive for  $\gamma$  in some neighborhood of  $\gamma = 0$  by continuity, as mentioned in Section 3.

Theorem 11 implies that for any  $\bar{Y} < (3 - \sqrt{5})/2$ , it is possible to construct a population with steady-state prevalence of  $\bar{Y}$  such that  $j_1 > 0$ , so at least some groups would increase steady-state prevalence by reducing activity.

To see this, consider a homogenous population with prevalence  $\bar{Y}$ . In this case,  $\sigma_Y^2 = 0$  and  $A = \bar{Y}$ , so Eq. (28) reduces to  $\bar{Y}^2 - 3\bar{Y} + 1 > 0$ . The only root of this equation  $\in [0, 1]$  is  $(3 - \sqrt{5})/2$ .

In this case

$$j_i = i \frac{\sqrt{(1-\bar{Y})(1-2\bar{Y})} - (1-\bar{Y})^2}{\bar{Y}(1-\bar{Y})} > 0,$$

where  $i$  is the rate of partner change associated with prevalence  $\bar{Y}$ , as shown by Kremer [2]. (It may seem strange to refer to  $j_i$  in a homogenous population, but note that by continuity (28) will also be satisfied if the population is not entirely homogenous, but instead a sufficiently small group has lower activity, and the remainder of the population has activity sufficient for prevalence in the population as a whole to be  $\bar{Y}$ ).

Theorem 11 can also be used to derive conditions under which such counter-intuitive cases do not arise, and instead increases in activity increase prevalence.

**Proposition 12.**

1. If  $\bar{Y} > 1/2$ , then increases in activity always increase steady-state prevalence.
2. For any  $\bar{Y} < 1$ , there exists a population with steady-state prevalence  $\bar{Y}$  in which increases in activity by people of any activity level increase overall prevalence.

**Proof**<sup>3</sup>.

1. Since  $Y_k$  is in  $[0,1]$ , the variance in prevalence among groups is less than  $1/4$ .  $\bar{Y}$  is always less than  $A$ . Thus, by Lemma 10,

$$H < 1 \iff 4\bar{Y} \left( \frac{A^2}{(1-A)^2} - \bar{Y} \right) > 1. \quad (30)$$

Since  $\bar{Y} < A$ , this condition will hold if

$$4\bar{Y} \left( \frac{\bar{Y}^2}{(1-\bar{Y})^2} - \bar{Y} \right) > 1.$$

This last inequality holds if  $\bar{Y} > 1/2$ , and  $H < 1$  implies that increases in activity increase steady-state prevalence by Theorem 8.

2. By the first part of this proposition, we can restrict attention to the case  $\bar{Y} < 1/2$ . Take a two-group population with prevalence  $\bar{Y}$ . One group has activity 0, and comprises a proportion  $(1-\varepsilon)$  of the population. The other group has activity  $i$ , and comprises proportion  $\varepsilon$  of the population. Assume random mixing ( $\gamma=0$ ). Prevalence in the zero group must be 0. Prevalence in the active group will be  $Y_\varepsilon = 1 - (1/\theta i)$ . Overall prevalence must be  $\bar{Y} = \varepsilon Y_\varepsilon$ . Thus, to ensure that the population prevalence is  $\bar{Y}$ , it must be that

<sup>3</sup> With thanks to Ed Drozd and Andrei Sarychev.

$(\varepsilon(\theta i - 1)/\theta i) = \bar{Y}$ .  $\theta i > 1$  for the disease to be endemic. We express  $i$  as  $\varepsilon/[\theta(\varepsilon - \bar{Y})]$ . We wish to show that, for  $i$  sufficiently high,  $H < 1$  so that  $j_i$  is zero, by Theorem 8. For this population,  $\Lambda = Y_e$ ,  $\mu = \varepsilon i$ , and  $\sigma_Y^2 = (1 - \varepsilon)\varepsilon Y_e$ . Using the expression for  $H$  from Lemma 10 and noting that  $H > 0$  and  $0 < Y_e < 1$  implies that the denominator of the expression for  $H > 0$ ,

$$H = \frac{\varepsilon Y_e(1 - Y_e)}{(\theta i - \varepsilon)Y_e - (1 - \varepsilon)} < 1 \iff \varepsilon Y_e(1 - Y_e) < (\theta i - \varepsilon)Y_e - (1 - \varepsilon).$$

Substituting for  $i$  and  $Y_e$ , this is equivalent to

$$-2\varepsilon\bar{Y} + \bar{Y}^2 + \frac{\varepsilon}{(\varepsilon - \bar{Y})}\bar{Y} - \varepsilon + \varepsilon^2 > 0. \quad (31)$$

Substitute  $\Phi$  and  $\bar{Y}$  for  $\varepsilon$ , and so long as  $\Phi > 0$  this is equivalent to

$$\Phi^2(\Phi - 1) + \bar{Y}^2 > 0.$$

For  $y > 0$ , we can always find  $\Phi^* > 0$  such that this is true and such that  $\varepsilon^* = \bar{Y} + \Phi^* < 1$ . Set  $i^* = \varepsilon^*/\theta(\varepsilon^* - \bar{Y})$  to complete the proof by observing that  $\theta_i^* < 1$ : the endemic steady state is stable, and  $H < 1$ , as required.  $\square$

## 5. Simulations calibrated to data from UK heterosexuals

In this section, we apply the model to data on rates of partner change taken from The National Survey of Sexual Attitudes and Lifestyles (NATSSAL), a comprehensive survey of sexual behavior in the UK encompassing some 18 000 people. It found that the mean number of heterosexual partners in the last five years was 1.98 and the variance was 19.03 [14]. We use data from the heterosexual population, because the sample was too small to make reliable inferences about the homosexual population. Our model, though, is a single-sex model. The results in this section cannot, therefore, strictly apply to a heterosexual population. This section is designed to illustrate the point that the cut-off values  $j_e$  and  $j_i$  may be above the activity levels of a significant proportion of the population in which endemic steady-state prevalence is low. Since the properties of the single-sex model giving rise to the effects we discuss in this paper would be qualitatively similar to those of a two-sex model we believe that the qualitative results of this section should at least be indicative of what one would be likely to find using a two-sex model. A full theoretical and simulated analysis of a two-sex preferred-mixing model is, though, beyond the scope of this paper. (To see that the same basic intuition would apply in a two-sex model, note that Anderson and May [1] show that in a two-sex randomly mixing population, the disease will survive if  $(\beta c \beta' c')^{1/2}/\delta > 1$ , where  $\beta$  and  $\beta'$  are the male-to-female and female-to-male transmission probabilities, respectively,

Table 1  
Numbers of partners per year over the last five years

| Ptrs/yr | Percentage | Ptrs/yr | Percentage | Ptrs/yr | Percentage | Ptrs/yr | Percentage |
|---------|------------|---------|------------|---------|------------|---------|------------|
| 0       | 8.95       | 2.2     | 0.218      | 4.4     | 0.033      | 8       | 0.061      |
| 0.2     | 62.56      | 2.4     | 0.261      | 4.6     | 0.010      | 9       | 0.010      |
| 0.4     | 10.56      | 2.6     | 0.069      | 4.8     | 0.023      | 10      | 0.037      |
| 0.6     | 6.39       | 2.8     | 0.033      | 5       | 0.061      | 12      | 0.026      |
| 0.8     | 3.44       | 3       | 0.347      | 5.2     | 0.008      | 13      | 0.003      |
| 1       | 2.05       | 3.2     | 0.026      | 5.4     | 0.008      | 14      | 0.018      |
| 1.2     | 1.64       | 3.4     | 0.063      | 6       | 0.149      | 15      | 0.015      |
| 1.4     | 0.90       | 3.6     | 0.031      | 6.4     | 0.005      | 16      | 0.003      |
| 1.6     | 0.70       | 3.8     | 0.027      | 7       | 0.027      | 18      | 0.003      |
| 1.8     | 0.24       | 4       | 0.263      | 7.6     | 0.005      | 20      | 0.006      |
| 2       | 0.83       | 4.2     | 0.003      | 7.8     | 0.005      | 100     | 0.002      |

$\delta$  is the death rate, which is assumed to be the same for both men and women,  $c = \mu + (\sigma^2/\mu)$ , where  $\mu$  is the mean rate of partner change for men,  $\sigma$  is its standard deviation, and  $c'$  is defined analogously for women. Given this condition, straightforward differentiation indicates that increases in activity by men with activity less than  $(\mu + (\sigma^2/\mu))/2$  will reduce  $c$ , and that increases in activity by women with activity less than  $(\mu' + (\sigma'^2/\mu'))/2$  will reduce  $c'$ .)

Table 1 shows the distribution of numbers of heterosexual partners in the last five years.<sup>4</sup> In the simulations below, we assume that the annual rate of partner change is one-fifth the number of heterosexual partners over the last 5 years.<sup>5</sup> We aggregate the responses from men and women in NATSSAL into a single population.

For the simulations, the data from Table 1 were aggregated into 18 activity groups<sup>6</sup> in order to simplify the calculations. Taking the NATSSAL partner change rates as given, we then calculate  $\{Y_k\}$ ,  $\bar{Y}$ , and  $A$  as functions of  $\beta/\delta$  and  $\gamma$  by numerically solving the equations for the steady states of Eq. (2). Inv-

<sup>4</sup> The data in the NATSSAL study were weighted before analysis to correct for differing responses in different geographical regions, and for differential probability of selection of individuals living in households of different sizes. Details of this may be found in Ref. [14], p. 54 & 55.

<sup>5</sup> There is no unproblematic way of moving from the theoretical concept of the rate of partner change to empirical observations of the number of partners per period, since people may have several partners simultaneously, and may re-establish old partnerships [16]. We assume that those people who have one partner over a 5 year period have an average of 0.2 partners per year. If the people listed as having one partner every 5 years actually changed partners less frequently, as seems plausible, the variance of sexual activity would be even larger. This would exacerbate some of the counterintuitive effects discussed in this paper.

<sup>6</sup> Groups as for Table 1 up to 1.6 partners per year, then groups: 1.8–2.2 ptrs. per year, 2.4–3.8, 4–5.4, 6–6.4, 7–7.8, 8–9, 10–13, 14–15, 16–20, and above 20. The group activities of these groups were taken to be the mean activities of the aggregated groups.

erting these numerically derived functions allows us to express  $Y_k$ ,  $\beta/\delta$ , and  $A$  as functions of  $\bar{Y}$  and  $\gamma$ , the overall prevalence and the degree of preferential partner mixing, respectively. We then calculate  $j_e$  and  $j_i$  using the methods of Section 3 and relate them to the NATSSAL distribution to get an idea of what percentage of the population is likely to increase prevalence or pool-risk by reducing activity under various assumptions about long-run endemic prevalence and mixing patterns.

The results are summarized in Table 2 and Fig. 1. The first entry in each cell of Table 2 shows  $j_e$ , the number of partners below which reductions in the number of partners will cause an increase in steady-state pool-risk,  $A$ , for particular values of  $\bar{Y}$ , steady-state prevalence, and  $\gamma$ , the degree of assortativeness in mixing. The second entry in each cell of Table 2 shows  $j_i$ , the cut-off number of partners below which reductions in the number of partners will increase steady-state prevalence in the population as a whole. The figures in parentheses show the proportion of the population with less than the cut-off frequency of partner change in the NATSSAL sample. Thus, for example, if the distribution of rates of partner change were as given in the NATSSAL sample,  $\gamma$  were 0.5, and the transmission rate were such that steady-state prevalence was 0.5%, then the 97% of the population with less than 1.8 partners per year would increase prevalence among others by reducing their rate of partner change. The 88% of the population with less than 0.67 partners per year would increase

Table 2  
 $j_e$  and  $j_i$  by steady-state prevalence and assortativeness in mixing

| $Y$ (%) | $\gamma = 0$                         | $\gamma = 0.25$ | $\gamma = 0.5$ | $\gamma = 0.75$ | $\gamma = 0.95$ | $\text{Lim}_{\gamma \rightarrow 1}$ |
|---------|--------------------------------------|-----------------|----------------|-----------------|-----------------|-------------------------------------|
| 0.5     | 0.91 <sup>a</sup> (92%) <sup>c</sup> | 1.43 (96%)      | 1.80 (97%)     | 2.08 (98%)      | 2.28 (98%)      | 2.57 (99%)                          |
|         | 0.79 <sup>b</sup> (88%) <sup>c</sup> | 0.78 (88%)      | 0.68 (88%)     | 0.59 (82%)      | 0.67 (88%)      | —                                   |
| 1       | 0.83 (92%)                           | 1.16 (94%)      | 1.38 (96%)     | 1.54 (96%)      | 1.61 (97%)      | 1.75 (97%)                          |
|         | 0.68 (88%)                           | 0.61 (88%)      | 0.51 (82%)     | 0.43 (82%)      | 0.45 (82%)      | —                                   |
| 2       | 0.74 (88%)                           | 0.93 (92%)      | 1.05 (94%)     | 1.11 (94%)      | 1.12 (94%)      | 1.17 (94%)                          |
|         | 0.54 (82%)                           | 0.44 (82%)      | 0.34 (72%)     | 0.28 (72%)      | 0.37 (72%)      | —                                   |
| 5       | 0.59 (82%)                           | 0.66 (88%)      | 0.70 (88%)     | 0.71 (88%)      | 0.67 (88%)      | 0.67 (88%)                          |
|         | 0.33 (72%)                           | 0.22 (72%)      | 0.15 (9%)      | 0.11 (9%)       | 0.13 (9%)       | —                                   |
| 10      | 0.47 (82%)                           | 0.50 (82%)      | 0.50 (82%)     | 0.48 (82%)      | 0.43 (82%)      | 0.39 (72%)                          |
|         | 0.13 (9%)                            | 0.05 (9%)       | 0.007(9%)      | 0.001(9%)       | 0.04 (9%)       | —                                   |
| 20      | 0.35 (72%)                           | 0.35 (72%)      | 0.34 (72%)     | 0.31 (72%)      | 0.26 (72%)      | 0.20 (9%)                           |
|         | 0 (0%)                               | 0 (0%)          | 0 (0%)         | 0 (0%)          | 0 (0%)          | —                                   |
| 30      | 0.28 (72%)                           | 0.28 (72%)      | 0.26 (72%)     | 0.24 (72%)      | 0.20 (9%)       | 0.17 (9%)                           |
|         | 0 (0%)                               | 0 (0%)          | 0 (0%)         | 0 (0%)          | 0 (0%)          | —                                   |
| 50      | 0.20 (9%)                            | 0.19 (9%)       | 0.17 (9%)      | 0.16 (9%)       | 0.14 (9%)       | 0.11 (9%)                           |
|         | 0 (0%)                               | 0 (0%)          | 0 (0%)         | 0 (0%)          | 0 (0%)          | —                                   |

<sup>a</sup>  $j_e$  is first entry.

<sup>b</sup>  $j_i$  is second entry.

<sup>c</sup> Percentage of population below cut-off value in parentheses.

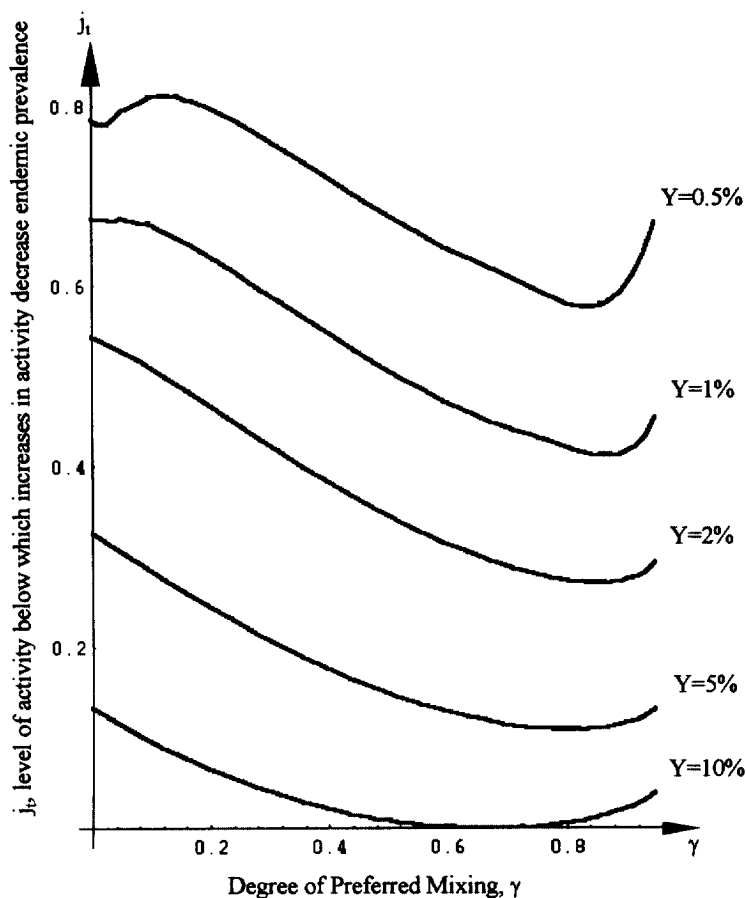


Fig. 1.  $j_i$  as a function of  $\gamma$ .

steady-state prevalence by reducing their rate of partner change. Thus, even assuming a steady-state prevalence among UK heterosexuals of 1% (which is likely to be a high estimate), more than 80% of the population have low enough sexual activity that they would reduce the overall prevalence in the long run by increasing activity. More than 90% would increase steady-state prevalence among others by decreasing their activity.

Preferred mixing does not necessarily mitigate the counterintuitive effects discussed in this paper, and in some cases can even exacerbate them. In low-prevalence populations, the proportion of the population that increases steady-state pool-risk,  $\Delta$ , by reducing sexual activity actually increases with  $\gamma$ , the degree of preferred mixing. The proportion of the population that increases total steady-state prevalence by reducing sexual activity does not change

monotonically with  $\gamma$ , the proportion of sexual activity which is within groups (see Fig. 1). For all values of  $\gamma$  in the table, though,  $j_t$  is greater than the number of partners of 80% of the population, assuming steady-state prevalence is less than 1%.<sup>7</sup>

To understand the intuition for why  $j_e$  and  $j_t$  may increase with  $\gamma$ , note that as  $\gamma$  rises, prevalence falls among low-activity people. Thus when low-activity people pick partners from the general pool they cause greater reductions in pool-risk. Moreover, as  $\gamma$  rises, prevalence rises less steeply in the number of partners for those with few partners. Thus low-activity people will increase their own probability of infection by a smaller amount if they increase their activity. These effects may cause increases in  $\gamma$  to increase  $j_e$  and  $j_t$ . On the other hand, as  $\gamma$  rises, people are less and less likely to pick partners from the general pool, and this effect will cause  $j_t$  to fall with  $\gamma$ .

It is less likely that many people could reduce steady-state prevalence by increasing activity in a higher prevalence population. Table 2 shows that if rates of partner change are similar, but transmission risk and steady-state prevalence are higher,  $j_t$  and  $j_e$  are smaller, and a much smaller percentage of the population has activity less than  $j_e$  or  $j_t$ . Thus, in the highest risk areas of Africa, where prevalence among adults is as high as one third, it seems certain that low-activity people would reduce total steady-state prevalence by having fewer partners. The situation is probably the same amongst homosexuals in large urban areas of developed countries. Similarly, prevalence among IV-drug users is high enough that reductions in needle-sharing by infrequent users are unlikely to increase total prevalence.<sup>8</sup>

If low activity groups increase their activity, while the mean activity will increase, the variance of activity will decrease, at least for moderate increases in activity. A look at the threshold function,  $G$ , in Eq. (3) suggests that, under some circumstances if  $\gamma$  is small enough, the reduction in variance caused by an increase in activity could actually be enough to eradicate completely the disease by causing  $G$  to fall below 1 and the endemic steady state to become unstable. Kremer [11] discusses this in detail for populations with random mixing. If we consider that case in which everyone below a certain activity level increases their activity to that level, for the NATSSAL distribution,  $G$  is minimized if everyone with activity less than 0.76 partners per year raises their activity to that level. With this new distribution of activity, the most  $\beta/\delta$  can be and have  $G < 1$  is 0.65 if mixing is random ( $\gamma = 0$ ). This maximal value of  $\beta/\delta$  falls very

<sup>7</sup> In these simulations,  $j_t$  is higher for high or low  $\gamma$  than for moderate  $\gamma$ . This is not true for a general distribution of rates of partner change, however.

<sup>8</sup> A sample of IV drug users in Thailand showed that 43% were HIV positive [17]. In Argentina, Brazil, and Uruguay, HIV prevalence among IV drug users is more than 50% in some communities [18].

rapidly in  $\gamma$  to 0.01 when  $\gamma$  approaches 1. With random mixing,  $\beta/\delta$  must be about 0.56 to give 0.5% prevalence, and 0.62 to give 1% prevalence. This calibrated level of  $\beta/\delta$  falls more slowly with preferred mixing. This means that, while  $\beta/\delta$  is low enough that the disease would possibly be wiped out if mixing were random, and people increased activity as discussed above, the result is extremely sensitive to  $\gamma$ . In fact, for the case where long run prevalence is 0.5%,  $\gamma$  could be no more than 0.003 for an activity increase to wipe out the disease. This suggests that increased activity would never wipe out the disease, as any realistic model would have considerably less than perfectly random mixing.

## 6. Dynamics

This paper mostly focuses on comparing steady states. Even if increases in rates of partner change cause decreases in long-run prevalence, they must, in the short run, increase prevalence. It is, therefore, also useful to briefly examine the time-path of prevalence in response to reductions in the frequency of partner change.

The transition period required before prevalence increases in response to a reduction in activity is fairly long under the simple SI model used in this paper, but much shorter under more realistic models. Table 3 shows the dynamics of prevalence in response to reduced activity under the simple SI model, a more realistic model with AIDS-induced mortality,<sup>9</sup> and a still more realistic model in which infectiousness is higher during the first few months of infection before the immune system has responded.<sup>10</sup> The dynamics are much faster in the more realistic models because the people infected immediately after the change in activity are not likely to continue to infect others for long. The simulations examine the impact of a reduction in activity to 0.16 partners per year by all people with 0.2 partner per year. In all three simulations, initial prevalence is

<sup>9</sup> The death rate from causes other than AIDS is assumed to be 0.03, and the death rate from AIDS is 0.1. The model requires a higher transmission rate to match any given steady-state prevalence. It also requires that a higher proportion of the population be born into highly active groups in order to match the observed proportion of high-activity groups in the population. It does not allow for the effect of AIDS-induced mortality on the birth rate, as would be appropriate in studies of aggregate population dynamics in some high prevalence African countries.

<sup>10</sup> Jacquez et al. [19] find that transmission rates are high in the first few months after infection, before the immune system has responded, and again in the final stage of the disease when AIDS has developed and the immune system has been overwhelmed. Transition probabilities between stages of the disease in the simulations reported here are taken from Ref. [19], and converted to Poisson hazard rates. The hazard rates for progression into the next stage are 0.970 and 0.119 in the first and second stage respectively. The death rate in the final stage of infection is 0.53. The transmission probability is set at 0.01 in the final stage, and 0.001 in the second stage. In the first stage, it is calibrated to match the desired prevalence, given  $\gamma$ .

Table 3  
Dynamics under different models

| Model   | Basic SI | AIDS mortality | Varying $\beta$ |
|---|----------|----------------|-----------------|
| Years until incidence = Initial incidence             | 5.4      | 1.4            | 0.14            |
| Increase in long run prevalence (%)                   | 30.9     | 16.5           | 15.0            |
| Years until prevalence is halfway to new steady state | 185      | 30             | 5               |
| Initial change in incidence (%)                       | -0.81    | -0.74          | -0.74           |

*Note:* In all three models prevalence is set at 0.005, and the group with 1 partner per 5 years reduces its activity to 0.8 partners per 5 years. The simulation with varying  $\beta$  requires three times as many state variables as groups. In order to keep the number of state variables manageable, the simulations in this table therefore aggregated everyone with more than 8 partners per year into a single group. The number of partners in this group was chosen not to be the average in the group, but to keep  $\mu - \sigma^2/\mu$  in the simulation equal to  $\mu + \sigma^2/\mu$  in the original data.

one half of 1%. The first row of Table 3 shows that incidence returns to its original level after 5.4 years in the basic model, after 1.4 years in a model that incorporates the effect of the disease on mortality, and after only two months if infectiousness depends on the stage of infection.<sup>11</sup> The second row of Table 3 shows that steady-state prevalence increases in response to the reduction in activity under all three models, but the effect is greatest under the basic model. The third row shows the number of years required for prevalence to fall halfway from its initial level to its new steady-state level.<sup>12</sup> The last row shows the initial percentage change in incidence in response to the reduction in activity. Note that in all three models, the initial reduction in incidence is negligible compared to the steady-state increase.

Although the dynamics are affected most strongly by AIDS-induced mortality and by varying infectiousness with the stage of the disease, Table 4 shows the dynamics are slower the greater are  $\gamma$ , the degree of assortativeness in mixing,  $\bar{Y}$ , the steady-state prevalence, and  $i$ , the number of partners in the group changing its activity. Both  $j_c$  and  $j_t$  seem to be lower under models that allow for mortality effects of the disease and for infectiousness to vary with the stage of infection.

Note that everyone with activity less than  $j_c$ , the cut-off for increasing steady-state prevalence among others in the long run by reducing activity, will also increase prevalence among others in the short run by reducing activity. In addition, all those from groups with prevalence between  $Y_c$  and  $A$  will reduce  $\lambda$  in the short run, but not in the long run, by reducing activity.

<sup>11</sup> Prevalence takes approximately twice as long to return to its original level.

<sup>12</sup> In the model in which infectiousness varies with the stage of infection, prevalence initially declines in response to a reduction in activity and then overshoots its steady-state value, before declining to a new steady-state value above its original level. The time until 50% of the steady-state change is attained therefore underemphasizes the costs of reductions in activity.

Table 4

Months until incidence rises to its original level in response to a reduction in activity

| $Y$ (%) | $\gamma = 0$ | $\gamma = 0.5$ | $\gamma = 0.9$ |
|---------|--------------|----------------|----------------|
| 0.5     | 4.2          | 10.4           | 23.8           |
|         | 1.8          | 5.2            | 9.6            |
| 1       | 4.7          | —              | —              |
|         | 2.2          | 6.6            | 12.7           |
| 2       | 7.3          | —              | —              |
|         | 3.2          | 9.1            | 18.7           |
| 5       | —            | —              | —              |
|         | 6.0          | —              | —              |

Top entry in each cell assumes a reduction from 0.4 partners per year to 0.3 partners per year. Bottom entry in each cell assumes a reduction from 0.2 partners per year to 0.1 partners per year. Dashes indicate that  $j_i$  is less than the number of partners. Infectiousness is assumed to vary with the stage of infection as in Ref. [19].

## 7. Directions for future research

This paper has shown that, in a preferred mixing, single-sex model, reductions in the frequency of partner change by low-activity people may increase the long-run prevalence of HIV/AIDS in populations that would have low steady-state prevalence given current activity levels. Given the limitations of the data, the simplifying assumptions of the model, and the fact that the model examines a homosexual population while the data are from heterosexuals, extreme caution should be used before applying these results to the real world. However, to the extent that these results are confirmed in future research, they reinforce arguments that public health messages urging reduced activity should be targeted to high-activity people, and should emphasize condom use, rather than abstinence.

Public health messages can be targeted both through their content, and through the choice of advertising media. For example, the 'Get high, Get stupid, Get AIDS' campaign warning people about the links between substance abuse, unprotected sex, and AIDS may have targeted high-activity people more than the mass mailing of AIDS-prevention literature to all US households in the early days of the epidemic.

We are not suggesting that public health officials encourage people to have more partners, since anyone who followed such advice would face a higher probability of infection, and people have an expectation that public health officials will inform them about how to protect themselves from health risks.

More research is needed to examine the robustness of the results. For example, further work is necessary to see if these results are robust when differences between the sexes, the age-structure of mixing, the process of partnership-for-

mation and dissolution, more general mixing patterns, time-varying infectiousness, and mortality effects of the disease are explicitly modeled.

Most important, this paper has examined the consequences of changes in the frequency of partner change by low-activity people, holding constant the number of partners of others. In fact, since reductions in the number of partners by low-activity people increase prevalence in the pool of available partners, they may lead to further reductions in activity. Kremer [11] explores a model in which people choose an activity level depending on prevalence in the pool of available partners, and shows that similar results obtain, but that there may be multiple equilibria, as in Akerlof's model of the market for lemons [15].

Reductions in activity by low-activity people could also directly cause high-activity people to reduce their activity by making it harder for them to find additional partners. The model in this paper would be applicable to an environment in which people met sexual partners in a bar in which one could always find a partner. One could also imagine a 'dating' model, in which people went on dates and decided whether or not to have sex, and there was a limit of one date per day. If the low-activity people decided to have sex on fewer dates, the high-activity people would automatically also have sex on fewer dates. To the extent that this 'dating model' is correct, and reductions in activity by low-activity people cause high-activity people to reduce activity rather than to mix with each other, increases in activity by low-activity people will be less likely to reduce steady-state prevalence. However, while the date model may be a good model for low-activity people, it may not be a good model for the tail of the distribution with extremely high activity, which disproportionately influences the spread of the disease. This group is not likely to continue dating without sex, but instead to seek other sexual partners.

If high-activity people respond to a potential partner's abstinence by seeking a new partner, but respond to a potential partner's preference for condom use by agreeing to use a condom, then public health messages directed to low-activity people urging abstinence could actually increase prevalence, but messages urging condom use could reduce prevalence.

## **Acknowledgements**

We thank Roy Anderson, Marie Claude Boilly, Gary Becker, Peter Diamond, Geoffrey Garnett, Sunetra Gupta, Anne Johnson, Ed Kaplan, Tomas Philipson, Jane Wadsworth, and two anonymous referees for comments and discussion; and Anne Johnson and Jane Wadsworth for generously providing data. We are particularly grateful to Ed Drozd, Ted Miguel, Cesaltina Pires, and Andrei Sarychev for excellent research assistance.

## References

- [1] R.M. Anderson, R.M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University, New York, 1991.
- [2] M. Kremer, Integrating behavioural choice into epidemiological models of AIDS, *Q. J. Econom.* 111 (2) (1996) 549.
- [3] R.M. Anderson, R.M. May, M.C. Boilly, G.P. Garnett, J.T. Rowley, The spread of HIV-1 in Africa: Sexual contact patterns and the predicted demographic impact of AIDS, *Nature* 352 (6336) (1991) 581.
- [4] S. Gupta, R.M. Anderson, R.M. May, Networks of sexual contacts: Implications for the pattern of spread of HIV, *AIDS* 3 (1989) 1.
- [5] H.W. Hethcote, Modeling heterogeneous mixing in infectious disease dynamics, in: *Models for Infectious Human Diseases*, V. Ishman, G.F.H. Medley, Eds., Cambridge University, Cambridge, 1996, p. 215.
- [6] H.W. Hethcote, J.A. Yorke, *Gonorrhea: Transmission dynamics and control*, Springer Lecture Notes in Biomath., vol. 56, Springer, Berlin, 1984, p. 1.
- [7] E.H. Kaplan, Modeling HIV infectivity: Must sex acts be counted?, *J. Acquired Immune Deficiency Syndromes* 3 (1990) 55.
- [8] E.H. Kaplan, P.C. Cramton, A.D. Paltiel, Nonrandom mixing models of HIV transmission, in: C. Castillo-Chavez (Ed.), *Mathematical Statistical Approaches to AIDS Epidemiology*, Springer Lecture Notes in Biomath., vol. 83, Springer, Berlin, 1990, p. 218.
- [9] R.M. May, R.M. Anderson, Transmission dynamics of HIV infection, *Nature* 326 (1987) 137.
- [10] Mead Over, Peter Piot, HIV infection sexually transmitted diseases, in: D.T. Jamison, W.H. Mosley (Eds.), *Disease Control Priorities in Developing Countries*, Oxford University, New York, 1992.
- [11] M. Kremer, AIDS and Imperfect Signals of Risk, unpublished notes.
- [12] L. Whitaker, A.M. Rentin, A theoretical problem of interpreting the recently reported increase of homosexual gonorrhea, *Eur. J. Epidemiology* 8 (2) (1992) 187.
- [13] J.A. Jacquez, C. P Simon, J. Koopman, L. Sattenspiel, T. Perry, Modelling analysing HIV transmission: The effect of contact patterns, *Math. Biosci.* 92 (1988) 119.
- [14] A.M. Johnson, J. Wadsworth, K. Wellings, J. Field, *Sexual Attitudes Lifestyles*, Blackwell Scientific, Oxford, 1993.
- [15] G. Akerlof, The market for lemons: Qualitative uncertainty the market mechanism, *Q. J. Econom.* 84 (1970) 488.
- [16] A. Johnson, Sources and use of empirical observations to characterize networks of sexual behavior, University College London, unpublished mimeo.
- [17] World Health Organization, *AIDS: Images of the epidemic*, World Health Organization, Geneva (1994).
- [18] Pan-American health organization, *Annual Report of the director*, Pan-American health organization, Washington (1992).
- [19] J.A. Jacquez, C.P. Simon, J. Koopman, I. Longini, The role of the primary infection in the epidemics of HIV infection in gay cohorts (mimeo), University of Michigan, 1994.