

SMR.780 - 43

FOURTH AUTUMN COURSE ON MATHEMATICAL ECOLOGY

(24 October - 11 November 1994)

"Like-with-Like Preference and Sexual Mixing Models"

Carlos Castillo-Chavez
Biometrics Unit
Cornell University
Ithaca, NY 14853-7801
U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

Like-with-Like Preference and Sexual Mixing Models

STEPHEN P. BLYTHE

Department of Physics and Applied Physics, University of Strathclyde, John Anderson Building, Glasgow, G4 0NG Scotland

AND

CARLOS CASTILLO-CHAVEZ

Biometrics Unit/Center for Applied Mathematics, Cornell University, Ithaca, New York 14853-7801

Received 8 October 1988; revised 9 March 1989

ABSTRACT

Two new general methods for incorporating like-with-like preference into one-sex mixing models in epidemiology are presented. The first is a generalization of the preferred mixing equation, while the second comprises a transformation of a general preference function for partners of similar sexual activity levels. Both methods satisfy the constraints implicit in a mixing model. The behavior of the transformation preference method is illustrated, and it is compared with the standard proportionate mixing model.

INTRODUCTION

In models of the dynamics of sexually transmitted diseases (STDs) within populations with heterogeneous sexual activity, it is necessary to specify the contact preference (who mixes with whom). Thus, for each level of sexual activity (number of new partners per unit time) we must know the fraction of partners coming from all other levels of activity. For practical modeling purposes, we require some function of activity that not only makes analysis straightforward but also is a reasonably accurate characterization of observed mixing patterns. Until recently the proportionate mixing model—equation (1), below—was the most common description of the mixing process available in analytic form, although arbitrary rules may, of course, be applied in stochastic simulations of the interaction of individuals. Proportionate mixing has been used extensively in a variety of situations by Barbour [3], Nold [15], Anderson and May [1], Dietz and Schenzle [7], Anderson and Grenfell [2], Hethcote and Van Ark [8], and Castillo-Chavez et al. [5, 6]. While this model has also proved useful in the study of the

epidemiology of STDs (see Hethcote and Yorke [9] for an outstanding example), it has become clear in recent years that more realistic mixing models are required for any detailed understanding of the transmission dynamics of HIV-1 and for any study of the possible value of various control measures.

At present there are few robust data on contact preference in a given community, although estimates of the distribution of *numbers* of sexual partners have been derived from various surveys. In the absence of detailed mixing information, models of HIV-1 transmission must take into account as many mixing patterns as possible if the impact of any preferential mixing in a given population is to be better understood.

Preliminary work on preferential mixing was done by Nold [15] and Hethcote and Yorke [9]. Their idea, as Nold [15] says, was to "supplement the proportionate mixing model with one which allows for more social or geographic separation of groups." In the development of their two-sex group model for the transmission of gonorrhea for very active and active subpopulations, Hethcote and Yorke [9] note that it may be that "a very active person may be more likely to have an encounter with a very active person" and therefore for the extreme case in which only like-with-like people mix we have "proportionate mixing within the very active sub-population and within the active subpopulation, but there is no interaction between these sub-populations." Hethcote and Yorke conclude that "actual mixing is probably somewhere in between the extremes of proportional mixing in the entire subpopulation and proportionate mixing in the activity levels." Earlier, Nold [15] had reached the same conclusion and introduced a mixing matrix M that is $1-s$ times the mixing matrix for proportionate mixing plus s times the matrix for proportionate mixing in the activity levels (i.e., within subpopulations). The parameter s was a measure of the separation between groups. Hethcote and Yorke called it the "selectivity constant."

Recently Sattenspiel [16, 17] questioned the use of proportionate mixing in dynamic models for the spread of disease in structured populations. She emphasizes those diseases for which the geographic and social structure plays an important role. Sattenspiel was the first to provide a *very general* formulation that allowed for truly distinct levels of interactions between individuals. More specifically, her framework allowed two distinct levels of random mixing: (1) nonsocial vs. social behaviors, with the condition that nonsocial individuals had within-group mixing only, and (2) two types of intragroup mixing, one involving individuals who interact only with neighbors (the local groups), and the second type involving individuals who interact randomly with all groups. Although her formulation was motivated by her work on the spread of hepatitis A among preschool children, Sattenspiel is aware (see [16]) that her framework could be applied to a variety of situations including the spread of STDs. Later, with Simon [18], she refined the mathematical analysis of their n -group model.

Hyman and Stanley [10, 11] have examined some approximations to like-with-like mixing; in addition, their simulations have illustrated the possible dangers of using proportionate mixing. Recently, Hyman and Stanley [11] have developed a model where the preference of half of the population may be specified by the modeler and the other half is defined by the preferences of the first, and Jacquez, Koopman, Sattenspiel, and Simon have explored a form of mixing slightly more general than Nold's that they called preferred mixing (see Sattenspiel et al. [19], Jacquez et al. [13], and Koopman et al. [12]). All of these approaches represent important and valuable contributions to the study of STD epidemiology.

In this paper we generalize the previous mixing framework by extending it to continuously distributed characteristics. In addition, we introduce two new solutions that satisfy the necessary mixing constraints: (1) generalized preferred mixing and (2) neighborhood mixing. A numerical example is provided to illustrate the effect of like-with-like mixing through the use of an appropriate chosen preference function. For further examples and extensions of this approach, see Blythe and Castillo-Chavez [4].

MIXING FUNCTIONS

In any one-sex model with heterogeneous sexual activity we have the mixing function $\rho(s, r)$, and hence $\int_r^{r+\Delta r} \rho(s, u) du$ specifies the fraction of partners that a person with activity s has among individuals with activities in the activity interval $[r, r + \Delta r]$. There are three constraints that $\rho(s, r)$ must satisfy for all s and r :

- (i) $\rho(s, r) \geq 0$,
- (ii) $\int_0^\infty \rho(s, r) dr = 1$,
- (iii) $\rho(s, r)sN(s) = \rho(r, s)rN(r)$,

where $N(x)$ is the number of people in the population with activity x . This is, of course, a function of time; however, we have suppressed time notation because (i)–(iii) must be true at *all* times. Conditions (i) and (ii) arise because $\rho(s, r)$ is in effect a probability density function, while condition (iii) expresses the requirement that the total number of partnerships of s -people with r -people must equal the total number of partnerships of r -people with s -people. We observe that if we look for solutions of the form $rN(r)A(s, r)$, then condition (iii) implies that $A(s, r) = A(r, s)$. These constraints are simple and obvious, but it is not easy to find functional forms for $\rho(s, r)$ that satisfy them simultaneously for all s , r , and time t .

We express the standard mixing model for proportionate mixing as

$$\rho(s, r) = \frac{rN(r)}{\int_0^\infty uN(u) du}. \quad (1)$$

Here $\rho(s, r)$ is actually independent of s and may be interpreted as saying

that the fraction of partners taken by any individual in the population from individuals with activity r is proportional to the total number of partnerships formed by all r -people and clearly satisfies (i)–(iii).

A preferred mixing function is an extension of Equation (1) to include a preference of individuals for partners with exactly the same activity level. In the continuous variables r and s used here, Nold's preferred mixing becomes

$$\rho(s, r) = (1 - \alpha) \frac{rN(r)}{\int_0^\infty uN(u) du} + \alpha\delta(s - r), \quad (2)$$

where $\delta(s - r)$ is a Dirac delta function and the constant α represents the bias toward partners of exactly the same activity. We define $\phi(s, r) = \alpha\delta(s - r)$ and call it the preference function. A discrete version of this model has been used recently by Jacquez et al. [13]. Although very useful for modeling purposes, and sufficient to demonstrate that even a small bias toward like-with-like can have a profound effect on epidemiological patterns. Equation (2) is rather restricted as a general model of preference.

A more general alternative to proportionate mixing has been derived by Hyman and Stanley [11] and in its continuous form can be expressed as

$$\rho(s, r) = \begin{cases} \rho(r, s) \frac{rN(r)}{sN(s)}, & r < s \\ \frac{f(s, r)rN(r)}{\int_s^\infty f(s, u)uN(u) du} \left(1 - \int_0^s \rho(s, u) du\right), & r > s, \end{cases} \quad (3)$$

where $\rho(r, s)$ for $r < s$ is arbitrarily specified by the modeler to suit available data and the rest of the values are derived from this constraint. The function $f(s, r)$ appears to be arbitrary and may be used to fine-tune the behavior of $\rho(s, r)$ to the modeler's needs. It can be shown that Equation (3) satisfies (i)–(iii).

We now introduce two new mixing functions that satisfy constraints (i)–(iii).

GENERALIZED PREFERRED MIXING

The first mixing model is a direct generalization of Nold's additive equation, Equation (2). A generalized mixing function that allows preferences for partners with activities that are arbitrary multiples of one's own and that is based on Equation (2) is provided in this section. An alternative version is provided in the next section.

To describe the mixing function of this section, we require the use of $2m+1$ delta functions with weights $\{\alpha_i\}$, describing the preference of individuals with activity s for individuals with activity $s/a_1, s/a_2, \dots, s/a_{2m+1}$. We assume that the sequences $\{\alpha_i\}$ and $\{a_i\}$ satisfy the following conditions:

- (a) $a_{m+1+j} = \frac{1}{a_{m+1-j}}, \quad j=1, 2, \dots, m-1; a_{m+1}=1,$
- (b) $\alpha_{m+1+j} = \alpha_{m+1-j}, \quad j=1, 2, \dots, m-1,$
- (c) $\sum_{i=1}^{2m+1} \alpha_i = 1.$

We now define our preference function $\phi(s, r)$ by the following symmetric expression on s and r ,

$$\phi(s, r) = \sum_{i=1}^{2m+1} \alpha_i \delta\left(s - \frac{r}{a_i}\right), \quad (4)$$

and observe that assumptions (a) and (b) imply that

$$\phi(s, r) = \alpha_{m+1} \delta(s - r) + \sum_{i=1}^m \alpha_i [\delta(s - a_i r) + \delta(a_i s - r)]. \quad (5)$$

In addition, we make the following assumptions and definitions:

- (d) $\theta(x) = xN(x),$
- (e) $\Lambda = \sup_x xN(x); \quad \Lambda < \infty,$
- (f) $Q(s) = \int_0^\infty \phi(s, r) \theta(r) dr = \alpha_{m+1} \theta(s) + \sum_{i=1}^m \alpha_i \left[\theta\left(\frac{s}{a_i}\right) + \theta(a_i s) \right],$
- (g) $W = \int_0^\infty \theta(u) du - \frac{1}{\Lambda} \int_0^\infty \theta(u) Q(u) du.$

Using these definitions and assumptions, we proceed to define the mixing function $\rho(s, r)$.

$$\rho(s, r) = \frac{\theta(r)[1 - Q(s)/\Lambda][1 - Q(r)/\Lambda]}{W} + \frac{\theta(r)}{\Lambda} \phi(s, r). \quad (6)$$

It is now straightforward to check that $\rho(s, r)$ satisfies conditions (i) and (ii). Since $\phi(s, r) = \phi(r, s)$, condition (iii) is satisfied. For like-with-like preference we could further assume that $a_{m+j} > a_{m+1-j+1}$ and that $\alpha_{m+j} > \alpha_{m+1+j+1}$, for $j=1, 2, \dots, m-1$, with weights $\{\alpha_i\}$ with a maximum at α_1 . It may in practice be a serious deficiency that there are "gaps" in the preference function between the arbitrarily chosen positions of the delta functions. Nonetheless, Equation (6) may be useful for preliminary investigations of a like-with-like preference distributed around $s = r$.

A NEIGHBORHOOD MIXING FUNCTION

Instead of the delta-function model of Equation (6), we should like to be free to specify like-with-like preference by some arbitrary function with well-understood properties. In particular we wish to use "neighborhood" functions that express preference as a continuous function with a single peak at $r = s$, falling off to either side. We know of no such functions that may be used directly, satisfying (i)–(iii). Equation (6) provides a clue as to how one might make use of an arbitrary continuous function, $\phi(s, r)$, as our *preference* function. We must ask: What transformation of the function $\phi(s, r)$

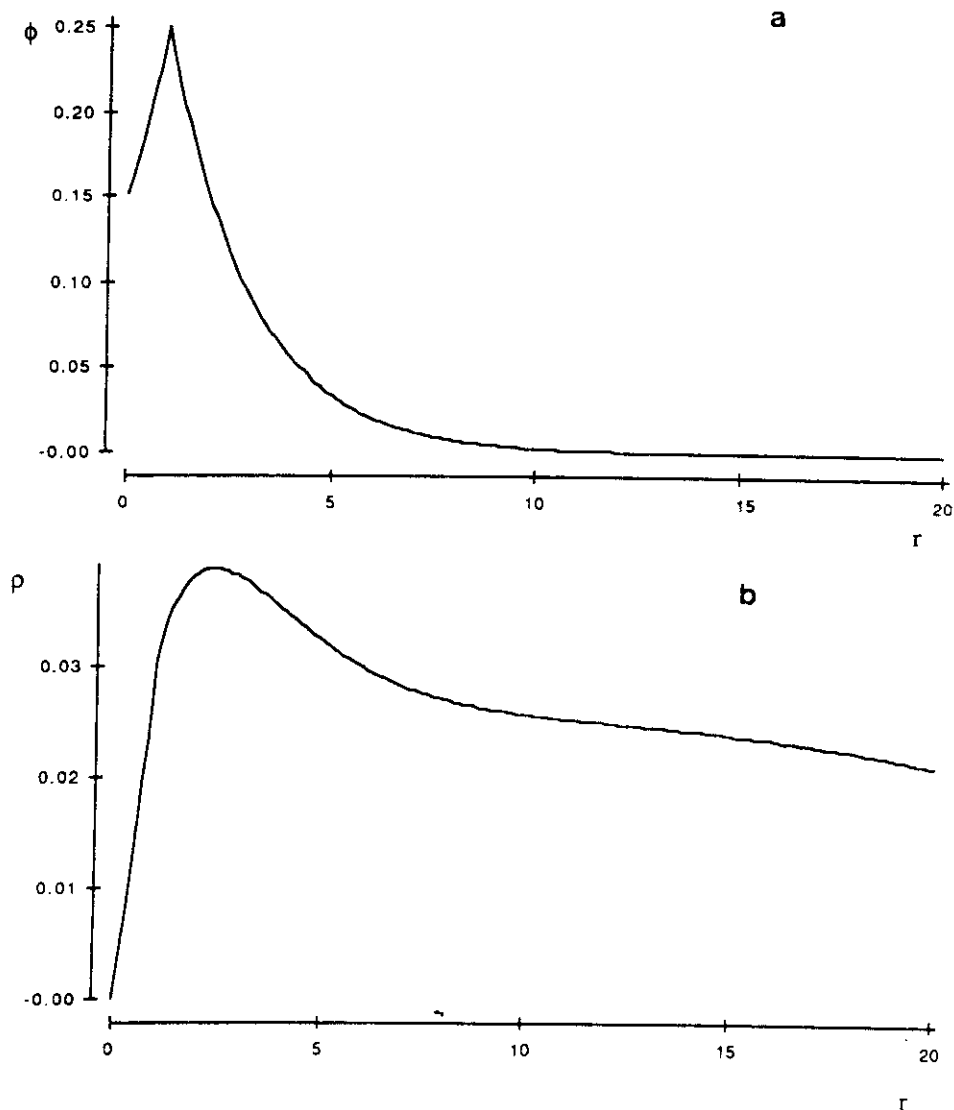


FIG. 1. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.1$, $c = 0.5$, and $s = 1.0$.

satisfies (i)–(iii)? If we restrict our choice of ϕ to functions with the property $\phi(s-r) = \phi(r-s)$ and state that

$$\int_{-\infty}^{\infty} \phi(y) dy = 1,$$

then we find that the transformation

$$\rho(s, r) = \frac{rN(r)P(r)P(s)}{\int_0^{\infty} uN(u)P(u) du} + \frac{rN(r)}{A} \phi(s-r) \quad (7)$$

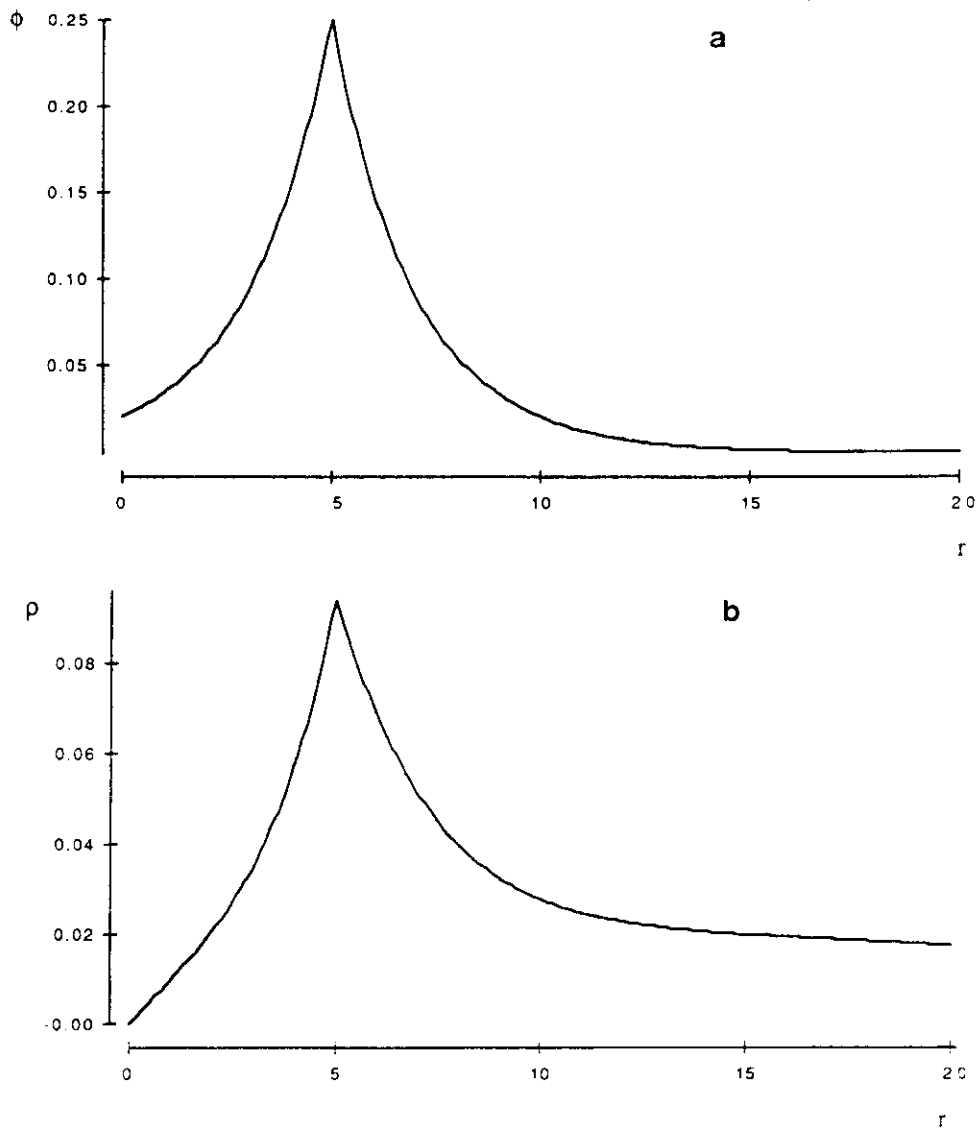


FIG. 2. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.1$, $c = 0.5$, and $s = 5.0$.

satisfies (i)–(iii). In (7),

$$P(x) = 1 - \frac{1}{A} \int_0^\infty u N(u) \phi(x-u) du \quad (8)$$

and A is a constant. We consider this constant further below. It is trivial to show that Equation (7) satisfies all the mixing constraints: the value of A must be large enough to give $P(x) > 0$ for all x , which in turn is sufficient to satisfy (i). If we assume that $\phi(s-r)$ is bounded by a constant M , that is,

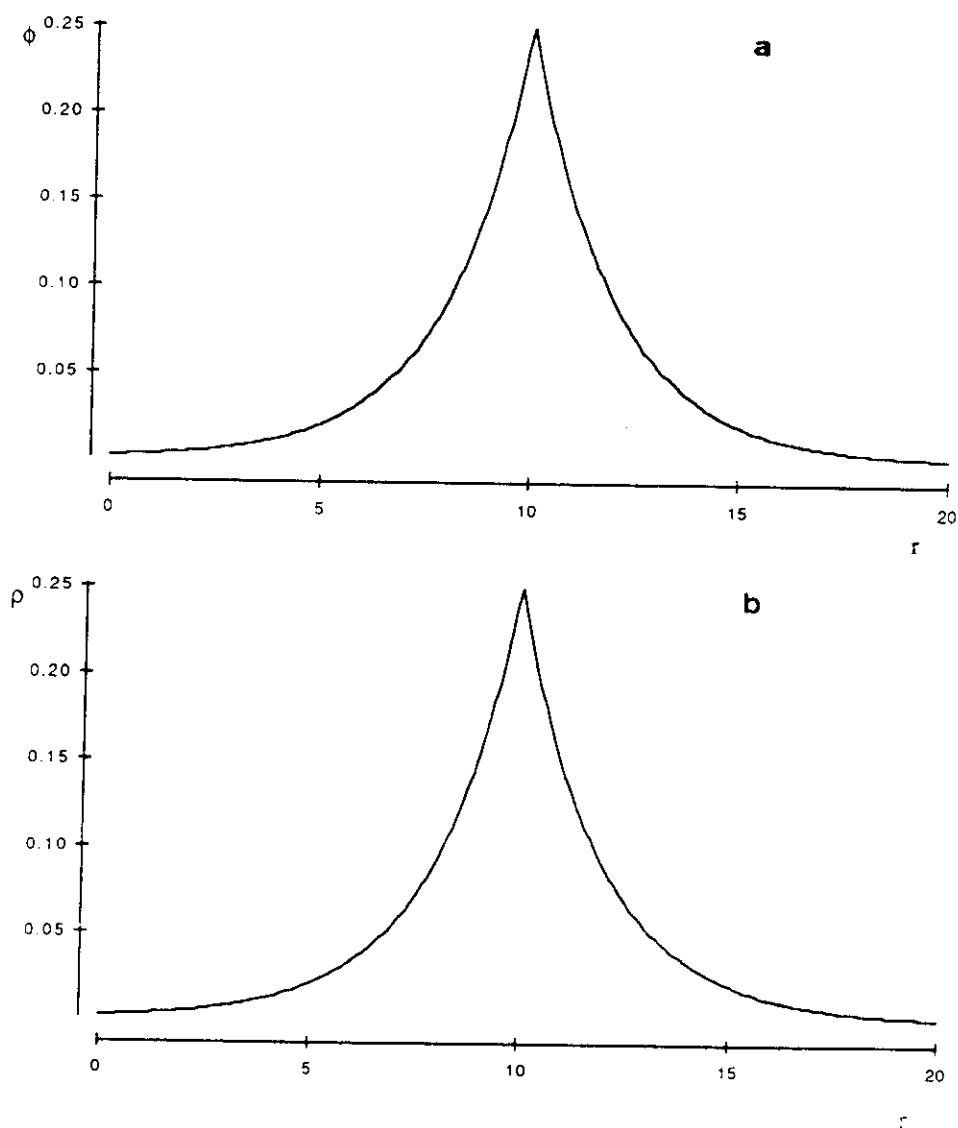


FIG. 3. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.1$, $c = 0.5$, and $s = 10.0$.

$\sup_z \phi(z) = M$, $z = s - r$, then the choice

$$A = M \int_0^\infty u N(u) du \quad (9)$$

is sufficient; for a delta function, $P(x)$ involves point values rather than integrals. For the special case of the previous section, since A roughly corresponds to Λ , $A > \max_x xN(x)$ is sufficient.

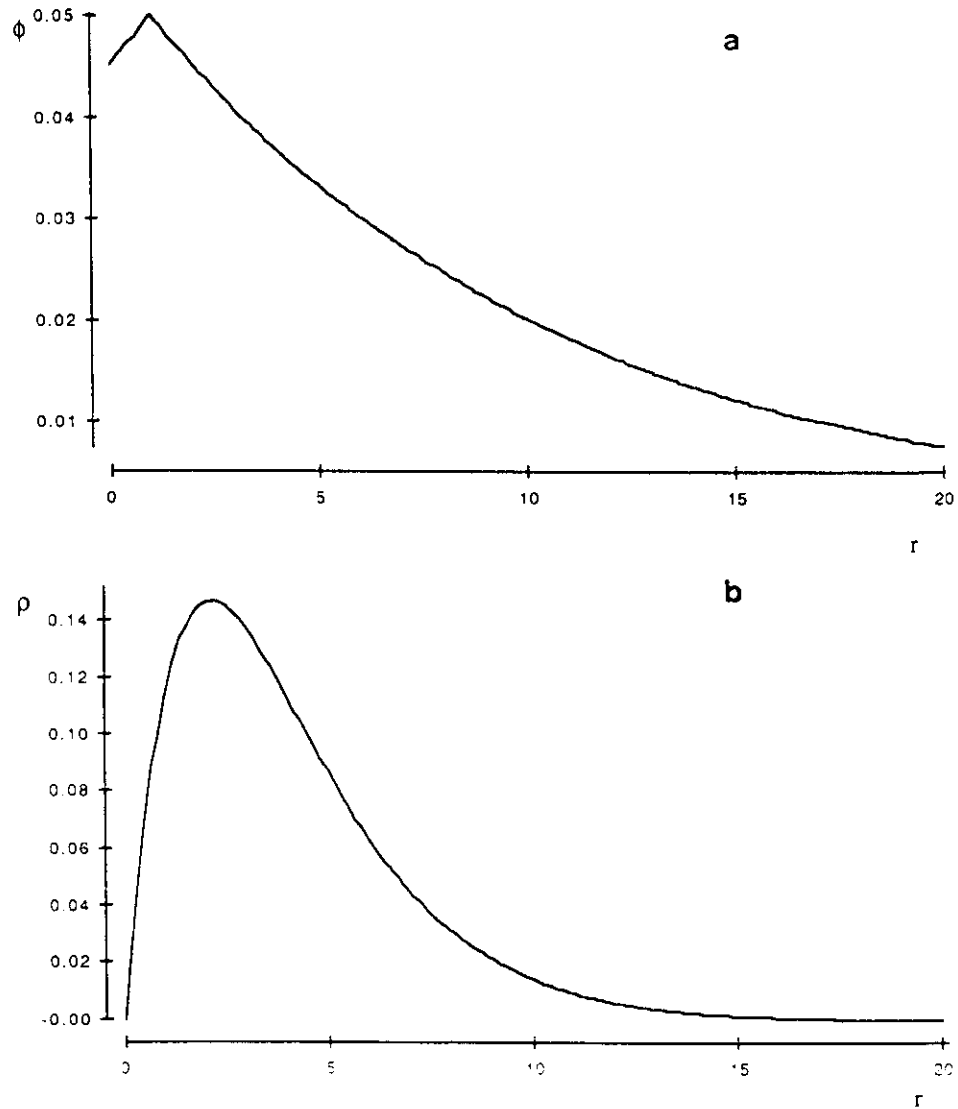


FIG. 4. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.5$, $c = 0.1$, and $s = 1.0$.

AN EXAMPLE

In this section we consider a simple example for which $\rho(s, r)$ can be calculated. We are not concerned here with a time-varying activity distribution (which would be the case in a real application or a dynamic model), and we choose the convenient exponential form

$$N(s) = Lke^{-ks}, \quad (10)$$

where $N(s)$ is the distribution of sexual activity in the population, L is the

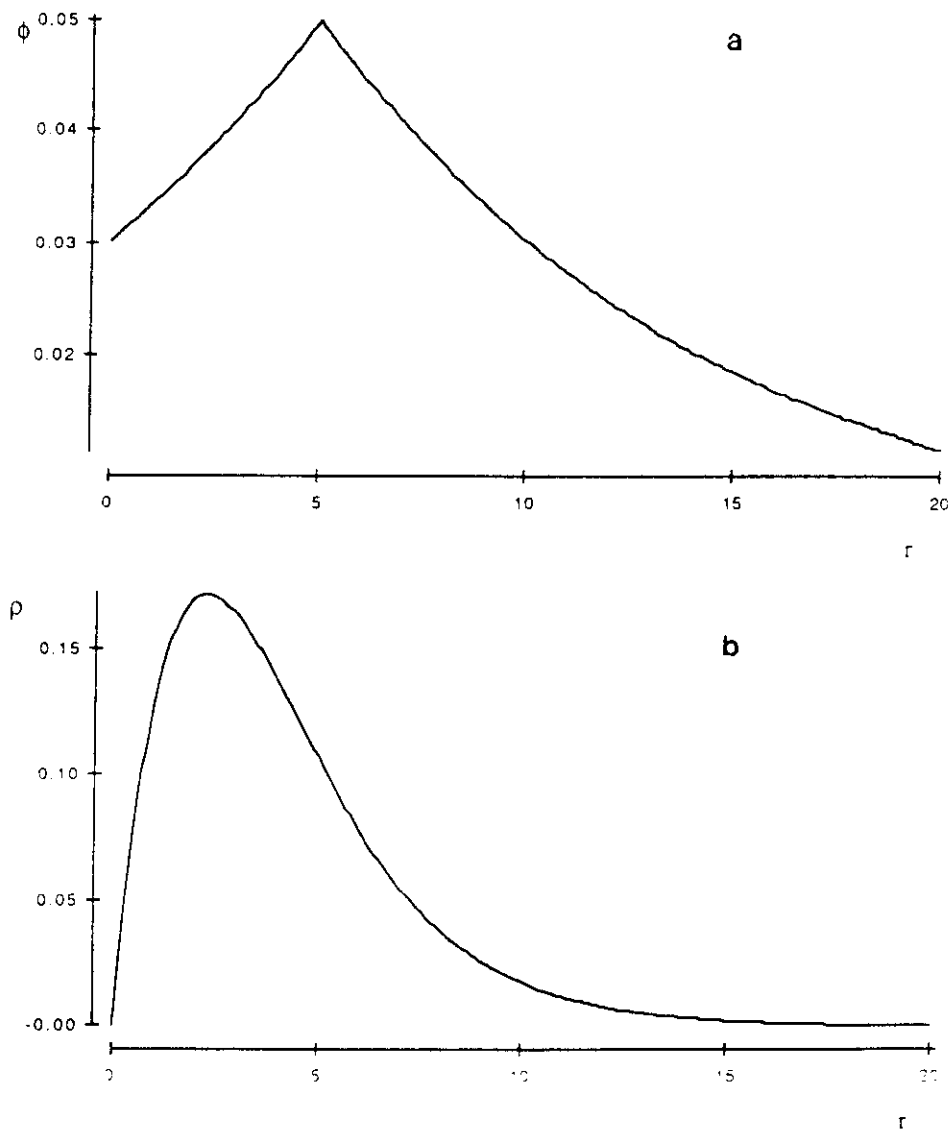


FIG. 5. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.5$, $\epsilon = 0.1$, and $\lambda = 5.0$.

total population size, and $1/k$ is the mean sexual activity. For the neighborhood preference functions $\phi(s, r)$, we choose

$$\phi(s, r) = \frac{c}{2} e^{-c|s-r|}, \quad (11)$$

which becomes more sharply peaked as c increases. In Figures 1-10 we take values of c in the interval $(0, 2]$; then choosing $A = L/k$ is sufficient. It is trivial to calculate the expression for $P(x)$ and $\rho(s, r)$ given Equations (10)

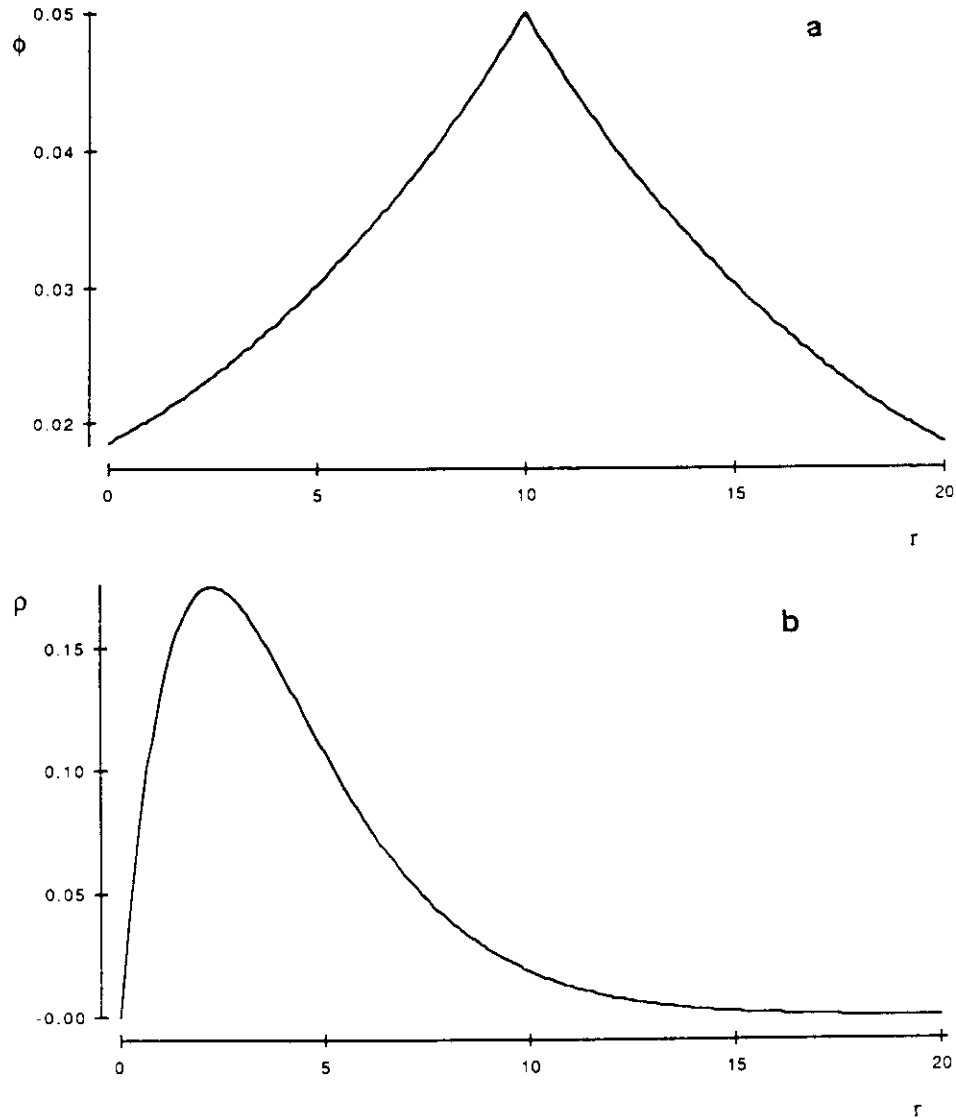


FIG. 6. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.5$, $c = 0.1$, and $s = 10.0$.

and (11), and in the figures we present some illustrative examples. We have graphed $\rho(s, r)$ as a function of r for different values of s and for a variety of values of c and k , with $A = N/k$.

In Figures 1–3 we illustrate $\rho(s, r)$ for $k = 0.1$ and $c = 0.5$, and $s = 1.0, 5.0$, and 10.0 , respectively. In this case $\rho(s, r)$ retains the sharply peaked form of $\phi(s, r)$ except when s is small, in which case $\rho(s, r)$ is much smoother. This case corresponds to a very narrow neighborhood function, with 50% of the area under $\phi(s, r)$ lying in the interval $r = s \pm 2 \ln 2$, and a large average activity: $1/k = 10.0$ partners per unit time.

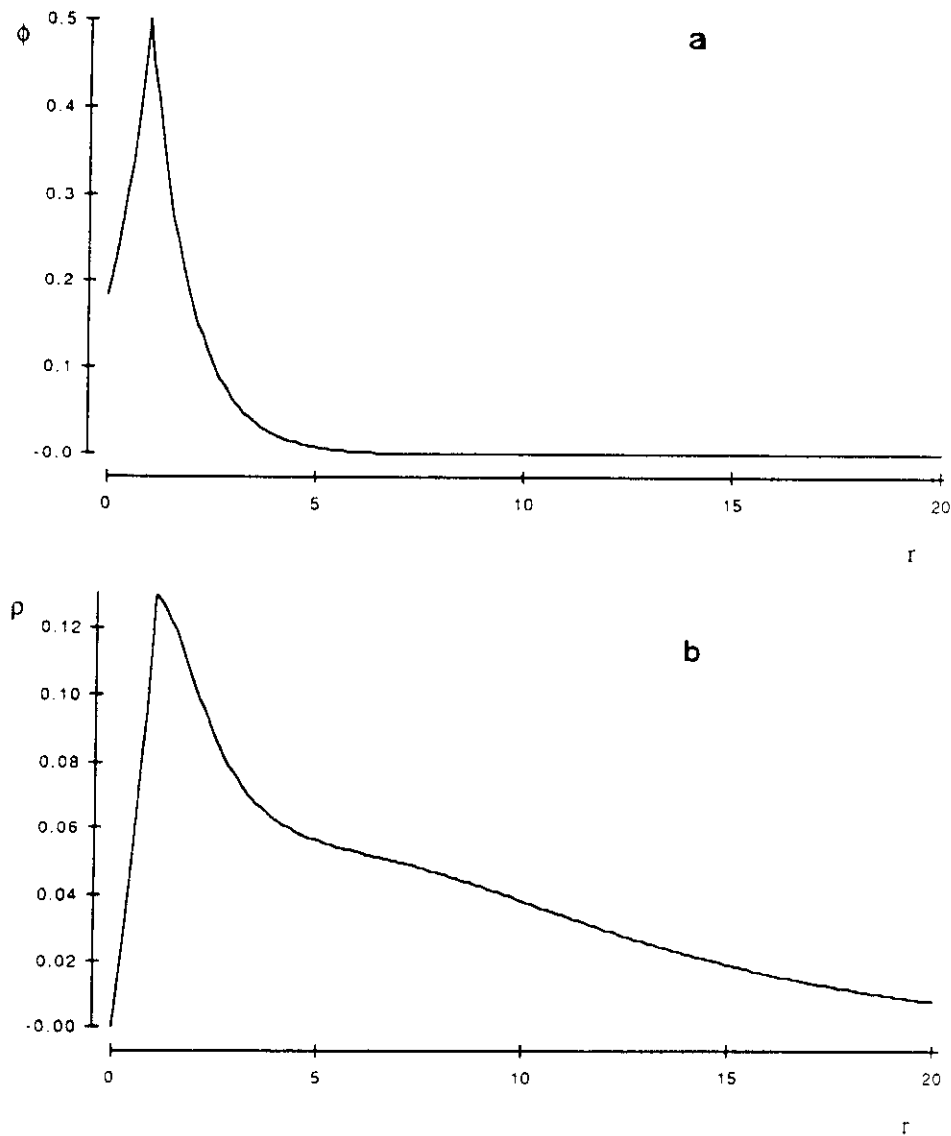


FIG. 7. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.25$, $c = 1.0$, and $s = 1.0$.

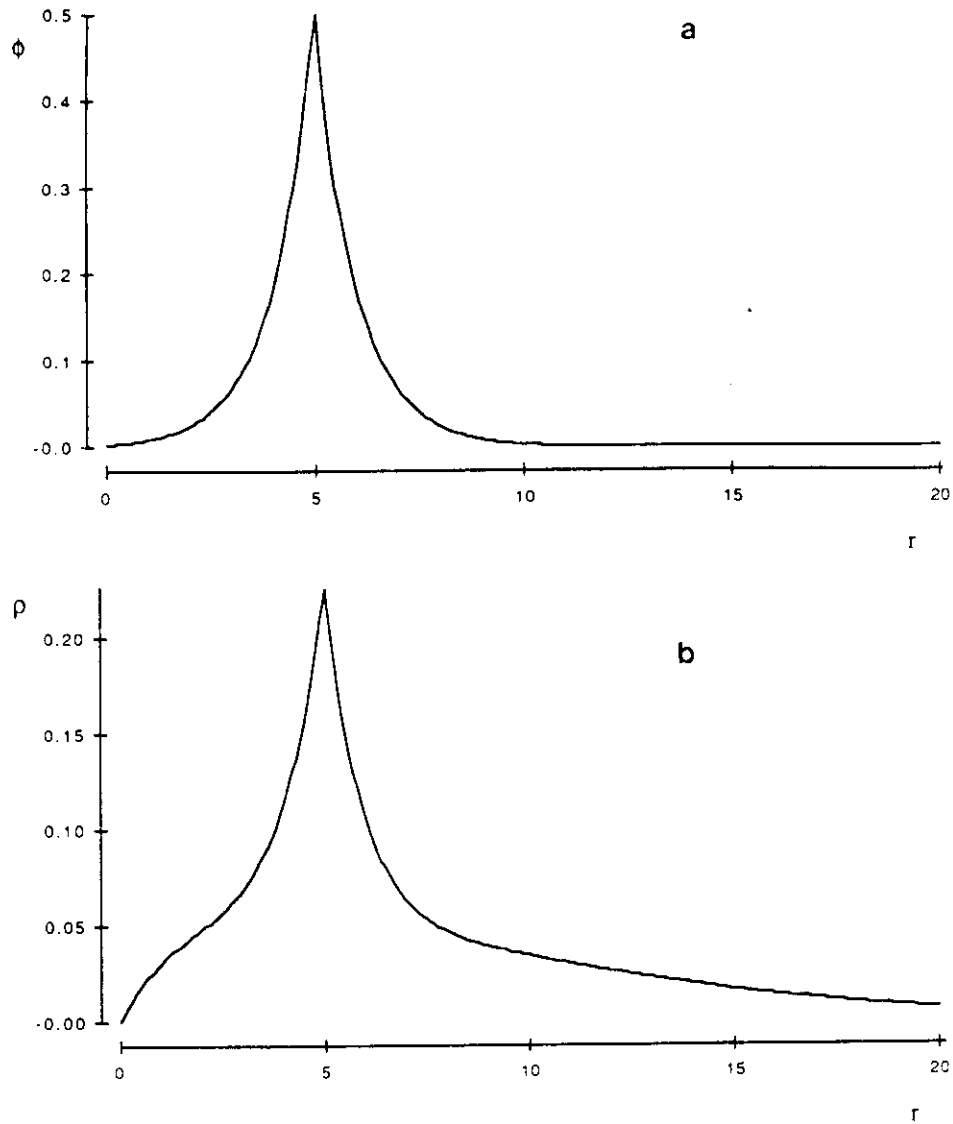


FIG. 8. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.25$, $c = 1.0$, and $s = 5.0$.

In Figures 4–6 we illustrate $\rho(s, r)$ for $k = 0.5$ and $c = 0.1$, with the same range of s values. In this case the neighborhood function is very broad and contributes very little to the shape of $\rho(s, r)$, which always behaves as $rN(r)$ (that is, like proportionate mixing).

In Figures 7–10 we illustrate the case $k = 0.25$ and $c = 1.0$ for $s = 1.0, 5.0, 10.0$, and 20.0 , respectively. Although here the neighborhood function is narrow, the mean sexual activity is small and the interplay between $N(r)$ and $\rho(s, r)$ is complicated. The essential form of $\rho(s, r)$ is a mixture of proportionate and like-with-like mixing. At a small s ($< 1/k$, Figure 7),

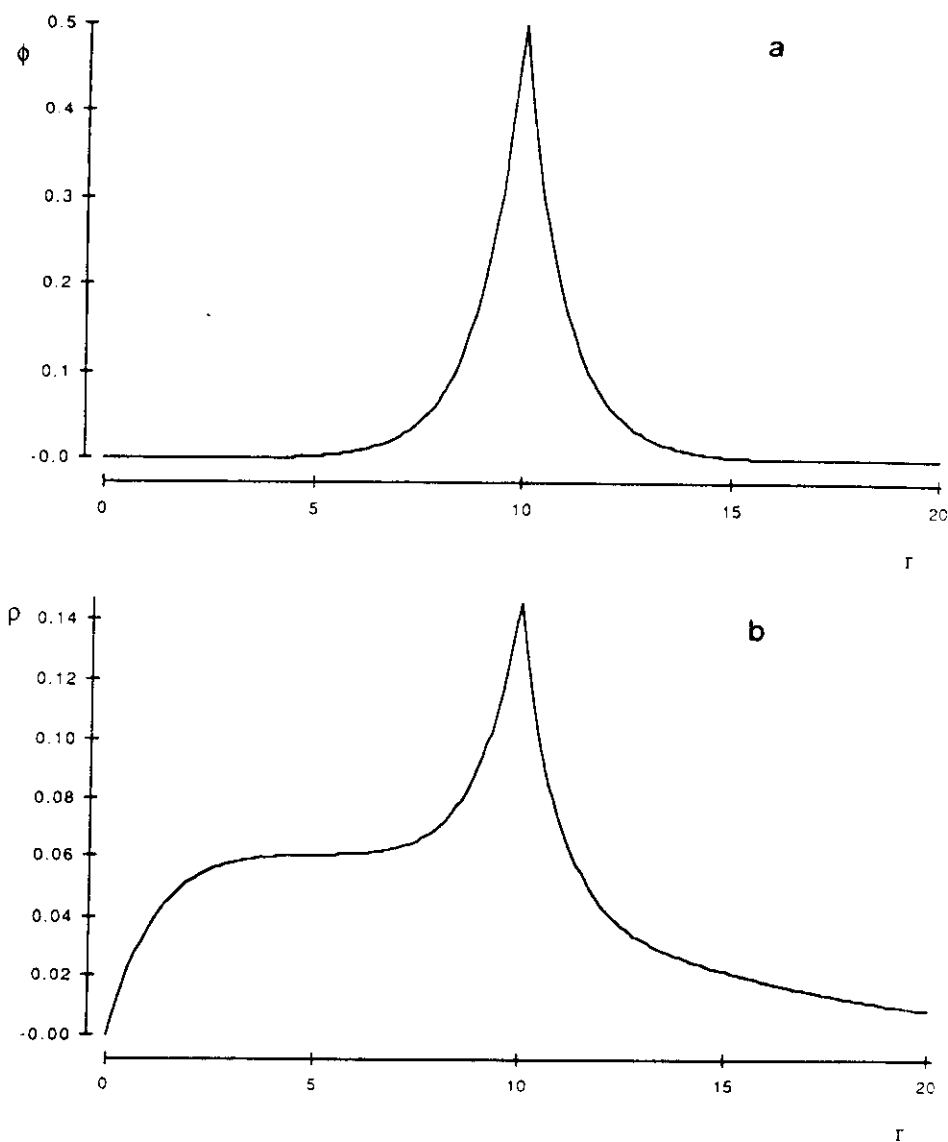


FIG. 9. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.25$, $c = 1.0$, and $s = 10.0$.

$\rho(s, r)$ is very much like $\phi(s, r)$ but with a more pronounced tail. As s increases (Figures 8–10), the component due to $\phi(s, r)$ decreases, until by the time $s = 20.0$ proportionate mixing is predominant.

We remark that the fidelity of the transformation $\rho(s, r)$ to the underlying neighborhood function $\phi(s, r)$, given Equations (10) and (11), depends upon the width of ϕ , the mean activity $1/k$, and the value of s in relation to $1/k$.

We observe that the above simple example fares well when dealing with those reported in the literature. To illustrate this point, we have superim-

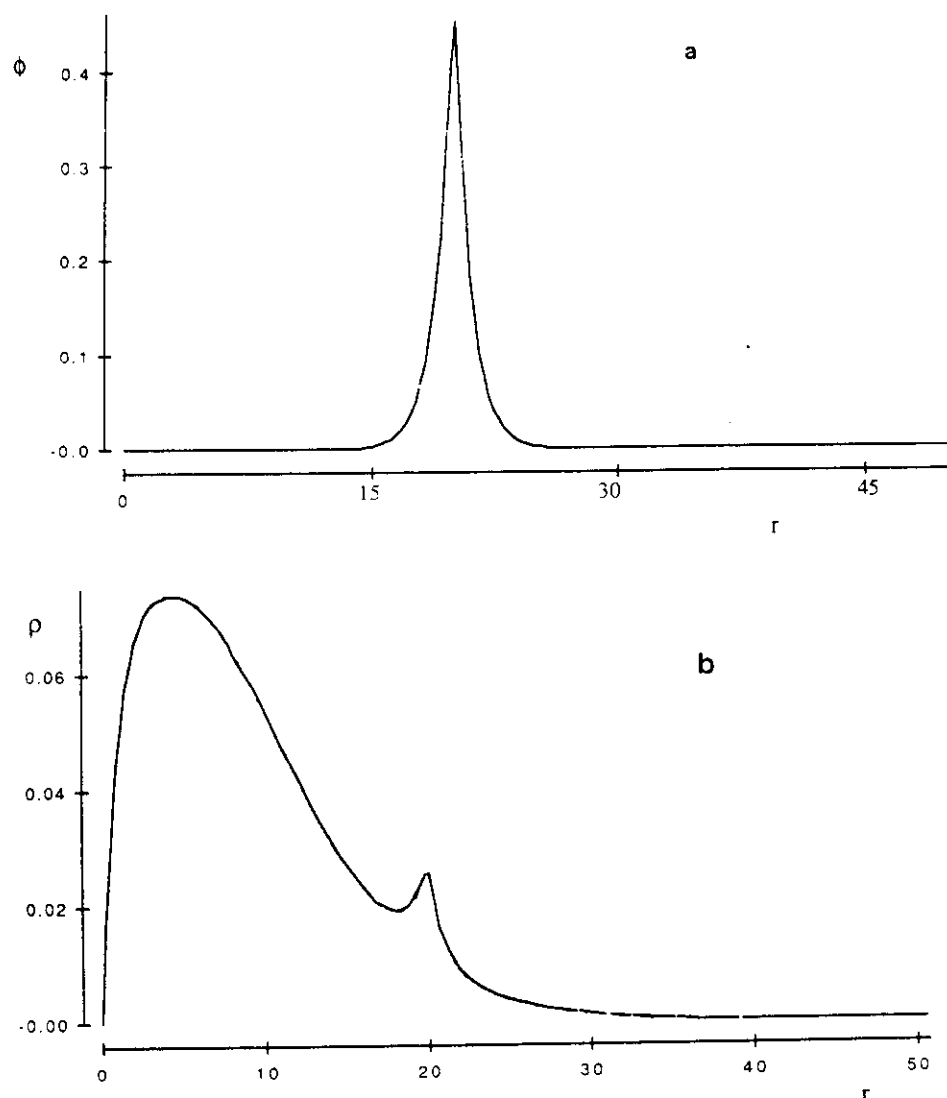


FIG. 10. Behavior of (a) $\phi(s, r)$ and (b) $\rho(s, r)$ for $k = 0.25$, $c = 1.0$, and $s = 10.0$.

posed (in Figure 11) three curves. They include the exponential function $N(x)$ of this paper with a mean of about two new partners per month, Hyman and Stanley's cubic (see Hyman and Stanley [10]), and data of Anderson and May (1987, as reported in Hyman and Stanley [10]).

These numerical simulations have been repeated using a "Gaussian" preference function, and the results have been consistent with those reported in this paper; however (not surprisingly), those involving different functional forms have produced significantly different results (see Castillo-Chavez and Blythe [4]).

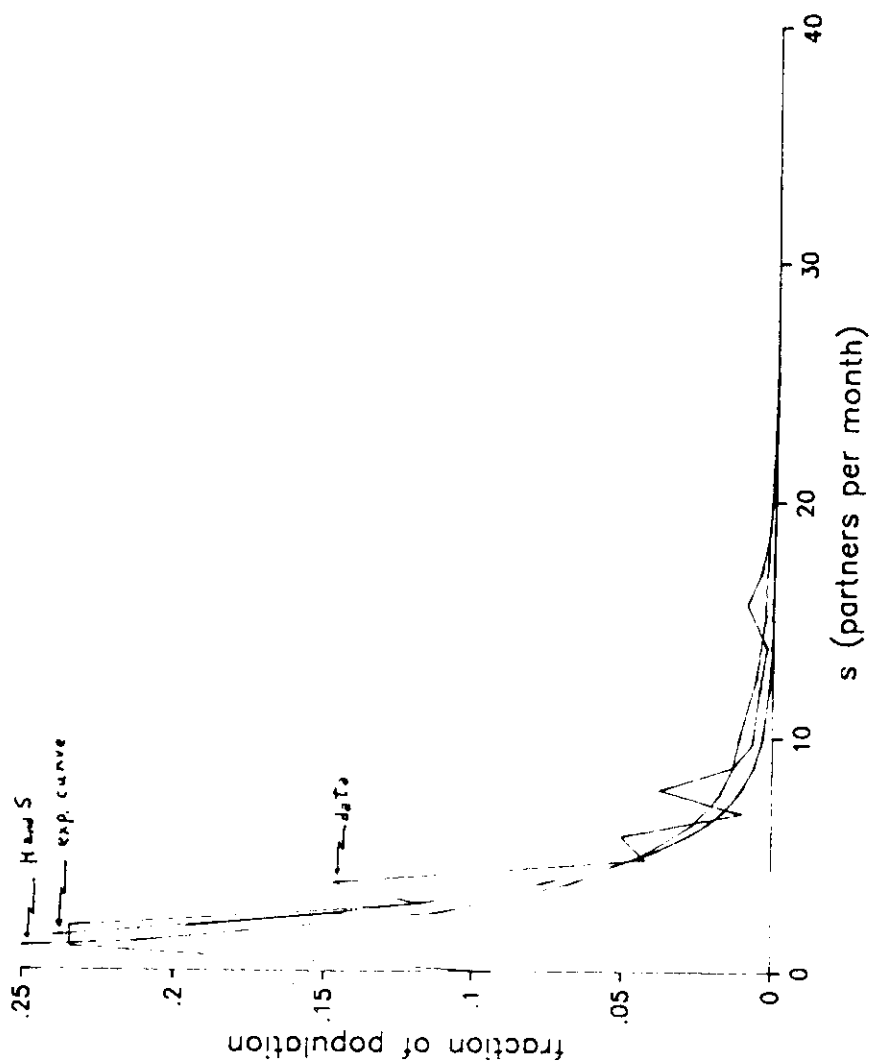


Fig. 11. Sexual activity distribution (partners per month) of homosexual men attending STD clinics in London. Data curve from Carne and Weller. Other curves: fit of inverse quartic by Hyman and Stanley (H and S curve, 1988), and an exponential curve with mean of two partners per month.

CONCLUSION

We have presented two new like-with-like mixing functions, one based on proportionate mixing biased at m values of the ratio s/r and the other based on a transformation of a general neighborhood function $\phi(s, r)$. A simple example for a static population indicates that the second mixing function behaves like the neighborhood function, provided that the latter is sharply peaked and the mean activity in the population is relatively high. In other cases proportionate mixing may be regained, with or without a level of bias toward like-with-like preference. These results support some of the numerical experiments of Hyman and Stanley [10, 11] regarding the role of the width (variance) of the neighborhood preference function and its relationship to proportionate mixing.

Much work remains to be performed before we have a complete understanding of the transformation method for an arbitrary neighborhood function, and the behavior of this $\rho(s, r)$ in a fully dynamic epidemiological model must be investigated. For further results the interested reader is referred to Castillo-Chavez and Blythe [4]. Finally, we speculate that if estimates for $N(s)$ (the activity distribution in the population) and $\phi(s, r)$ (tendency for like-with-like mixing) can be obtained from survey results, then examination of the transformation $\rho(s, r)$ of Equation (7) may be able to tell us whether or not the like-with-like preference is important in a given population and thus whether a proportionate mixing description is adequate or a more complicated model is required.

We thank S. A. Levin, J. M. Hyman, E. A. Stanley, and S. A. Colgate for their stimulating conversations; John Jacquez, C. E. McCulloch, and an anonymous referee for their comments; and L. Sattenspiel for her provocative ideas. This research has been partially supported by Hatch project NYC 151-409, USDA to C C-C.

REFERENCES

- 1 R. M. Anderson and R. M. May, Spatial, temporal, and genetic heterogeneity in host populations and the design of immunization programmes. *IMAJ. Math. Appl. Med. Biol.* 1:233-266 (1984).
- 2 R. M. Anderson and B. T. Grenfell, Quantitative investigations of different vaccination policies for the control of congenital rubella syndrome (CRS) in the United Kingdom. *J. Hyg. Camb.* 96:305-333 (1986).
- 3 A. D. Barbour, MacDonald's model and the transmission of bilharzia. *Trans. Roy. Soc. Trop. Med. Hyg.* 72:6-15 (1978).
- 4 C. Castillo-Chavez and S. P. Blythe, Mixing framework for social/social behavior, in *Mathematical and Statistical Approaches to AIDS Epidemiology*, C. Castillo-Chavez, ed., Lecture Notes in Biomathematics, Springer-Verlag (in press) (1989).

- 5 C. Castillo-Chavez, H. Hethcote, V. Andreasen, S. A. Levin, S. A., and W.-M. Liu. Cross-immunity in the dynamics of homogeneous and heterogeneous populations, in *Mathematical Ecology*, T. G. Hallam, L. G. Gross, and S. A. Levin, eds. World Scientific Publishing Co., Singapore, 1988. pp. 303–316.
- 6 C. Castillo-Chavez, H. Hethcote, V. Andreasen, S. A. Levin, S. A., and W.-M. Liu. Epidemiological models with age structure, proportionate mixing, and cross-immunity. *J. Math Biol.* (in press) (1989).
- 7 K. Dietz and D. Schenzle, Proportionate mixing models for age-dependent infection transmission, *J. Math. Biol.* 22:117–120 (1985).
- 8 H. W. Hethcote and J. W. Van Ark, Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation and immunization programs, *Math. Biosci.* 84:85–118 (1987).
- 9 H. W. Hethcote and J. A. Yorke, *Gonorrhea, Transmission Dynamics, and Control*, Lecture Notes in Biomathematics 56, Springer-Verlag, New York, 1984.
- 10 J. M. Hyman and E. A. Stanley, Using mathematical models to understand the AIDS epidemic, *Math. Biosci.* 90:415–473 (1988).
- 11 J. M. Hyman and E. A. Stanley, The effects of social mixing patterns on the spread of AIDS, in *Mathematical Approaches to Resource Management and Epidemiology*, C. Castillo-Chavez, S. A. Levin, and C. Shoemaker, eds., Lecture Notes in Biomathematics, Springer-Verlag, New York, (in press) (1989).
- 12 J. M. Koopman, C. P. Simon, J. Jacquez, L. Sattenspiel, and P. Taesung, Sexual partner selectiveness effects on homosexual HIV transmission dynamics, *J. AIDS* 1:486–504 (1988).
- 13 J. A. Jacquez, C. P. Simon, J. Koopman, L. Sattenspiel, and T. Perry, Modeling and analyzing HIV transmission: the effect of contact patterns, *Math. Biosci.* 92:119–199 (1988).
- 14 A. R. McLean and R. M. Anderson, Measles in developing countries. Part I. Epidemiological parameters and patterns, *Epidem. Inf.* 100:111–133 (1988).
- 15 A. Nold, Heterogeneity in disease-transmission modeling, *Math. Biosci.* 52:227–240 (1980).
- 16 L. Sattenspiel, Population structure and the spread of disease, *Hum. Biol.* 59:411–438 (1987).
- 17 L. Sattenspiel, Epidemics in nonrandomly mixing populations: a simulation, *Am. J. Phys. Anthropol.* 73:251–265 (1987).
- 18 L. Sattenspiel and C. P. Simon, The spread and persistence of infectious diseases in structured populations, *Math. Biosci.* (1988, in press).
- 19 L. Sattenspiel, J. Koopman, C. P. Simon, and J. Jacquez, The effects of population structure on the spread of HIV infection (ms.).

